TITLE PAGE

Classification

Social Sciences, Economic Sciences

Title

To mix or not to mix? Segregation and diffusion in heterogeneous groups

Authors

Segismundo S. Izquierdo[a*], Luis R. Izquierdo[b] and Dunia López-Pintado[c]

Affiliations

[a]Universidad de Valladolid. EII, paseo del cauce 59, 47011 Valladolid, Spain.

[b]Universidad de Burgos, Edificio A, Avda. Cantabria s/n, 09006 Burgos, Spain.

[c]Universidad Pablo de Olavide, Carretera de Utrera, Km. 1, 41013 Sevilla, Spain.

*Corresponding author: segis@eii.uva.es. +34 617 373043

Keywords

Diffusion, mixing, segregation, homophily, networks, SIS.

Abstract

The outbreak of epidemics, the rise of religious radicalization, or the motivational influence of fellow students in classrooms are some of the issues that can be described as diffusion processes in heterogeneous groups. Understanding the role that interaction patterns such as homophily, or segregation, play in the diffusion of certain traits or behaviors is a major challenge for contemporary societies. Here, we study the effects on diffusion processes of mixing (or segregating) two different groups – one group that is more sensitive or prone to "infection" or adoption, and the other which is more resistant –. In some cases we find non-monotonic effects of mixing, and Pareto inefficient segregation levels, e.g., situations where an increase in mixing can benefit *both* groups. These findings have fundamental consequences for the design of inclusion policies.

Significance Statement

Humans belong to different groups according to race, gender, age, abilities, preferences, etc. Most kinds of human interactions are biased in the sense that they can take place preferentially between individuals of the same group (homophily) or, conversely, between individuals of different groups (disassortative mixing). Social policies can be implemented to modify the level of mixing between different groups. We study the effect on diffusion processes -such as the adoption of a particular behavior or the spread of an illness- of mixing two groups with different propensities for adoption. We show that this effect is not always simple and there can be inefficient mixing levels, such that both groups can be better off by modifying the mixing level.

Text

Introduction

The adoption of new products, the spread of ideas, the transmission of diseases and other diffusion-dependent processes are shaped by the patterns of interaction between individuals with different susceptibilities to contagion or adoption (1–3). The influence of peers with different abilities or belonging to different groups also determines the adoption of behaviors in teams of workers in firms or students in classrooms (4, 5). Identifying the main features in the interaction structure that can foster or restrain the adoption of a given trait is a key consideration in network theory (6–15).

Here we study the variations on expected diffusion levels caused by a change in the degree of mixing or segregation between sensitive individuals (agents with a high propensity to adopt) and resistant individuals (agents with a low propensity to adopt). A particularly interesting case arises when group propensities are such that there are *Pareto inefficient mixing levels*, i.e. situations in which a change in the mixing level can improve the outcomes for both groups in the population.

Many social interactions exhibit significant homophily based on characteristics such as race, age, profession, etc. (16, 17). In contrast, other relationships (i.e., buyer-seller networks) are characterized by a high degree of heterophily or disassortative mixing. Regardless of the observed pattern of interaction in a population, a social planner could potentially modify this interaction by implementing specific goal-directed policies. Examples include health care programs to promote interactions between individuals of similar health characteristics (18), and compositional classroom designs to encourage school students with high academic ability to interact with students of lower ability (19, 20). The key question, then, is how changes in between-group interaction levels affect each group.

Qualitative results

We present a simple diffusion model which shows that the answer to the last question can be context dependent, consistent with the different evidence obtained in empirical and simulation studies (5, 19, 20). As a representative situation, imagine the case of an infectious disease spreading in a population composed of two distinct groups of agents (Fig. 1): a *sensitive group* (red) and a *resistant group* (blue). What are the consequences of increasing the level of mixing between the groups (i.e., the fraction of links between individuals belonging to different groups), while keeping constant the average level of individual interaction? One might intuitively predict that the infection levels in both groups should approximate, with an increase of infections in the resistant group and a decrease of infections in the sensitive group. Thus, all mixing levels would be Pareto efficient: the sensitive group would always benefit from higher between-group interactions, whereas the resistant group would always be harmed by it. However, we show that, while the first part of this intuition holds true (infection levels do become similar), the second part is not always true, since there can be *inefficient mixing levels*. In other words, we find situations for which an increase in between-group interaction leads to non-monotonic effects in one of the groups and, possibly, to a reduction of infections in *both* groups (Fig. 1, case B).

The underlying reason for this paradoxical effect is the feedback loop created between groups: from an initial stable situation corresponding to a given mixing level (e.g., 60% mixing in the scenario corresponding to case B in Fig. 1), increasing the interaction level between groups can be initially costly for the resistant group, which will initially meet more infected individuals. However, it can turn out to be beneficial for that same group once the returns from the positive effect induced on the other group are collected, and a new equilibrium is obtained in which both groups are better off.

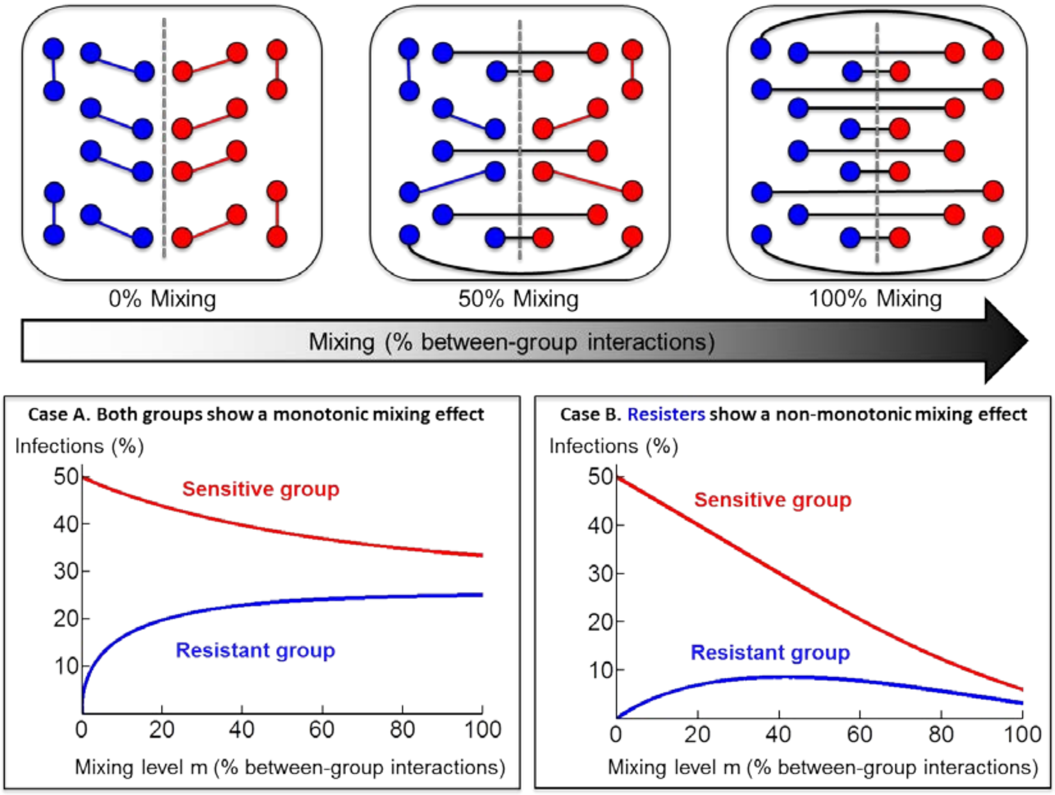

Fig. 1. Interaction structure and adoption levels as a function of mixing. Top, from left to right: segregated population case (m = 0), unbiasedly mixed population case (m = 0.5) and bipartite population case (m = 1). Bottom: Adoption levels in equilibrium for the resistant group (blue) and the sensitive group (red) as a function of the mixing level m for two different cases. In case B, the adoption level in the resistant group is a non-monotonic function of the mixing level, and Pareto inefficient mixing levels exist. Parameters $\{\lambda_1, \lambda_2\}$ for the SIS model: Case A: {1, 2}, Case B: {0.55, 2}.

The infection narrative constitutes a natural motivating example, but this framework can also be applied to the adoption of positive traits, and to questions such as: How does the market penetration of a new product depend on the degree of mixing between groups with different levels of willingness to adopt? How does the spread of a behavior like smoking depend on interaction patterns between groups with different propensities to smoke? Or, how does the internal organization of a classroom, where some students are easier to motivate than others, affect the overall level of motivation? In this last example, if we consider the related issue of ability sorting in schools – for which empirical findings are often controversial and sometimes even contradictory (5, 19, 20) –, our model suggests a way of reconciling apparently opposing recommendations within one single coherent explanation: increasing the level of mixing between students with high and low academic abilities will have a positive or negative effect depending on students' responsiveness to the state of partners.

Two-group SIS model

The simplest model we consider is based on the Susceptible-Infected-Susceptible (SIS) contagion framework used in epidemiology (21, 22), extended to a population composed of two groups of equal size (23–26): one of the groups is the resistant group and the other the sensitive group. Agents can be in one of two possible states: "susceptible" or "infected". In each time period, each agent interacts with another agent with probability $p > 0$ and, depending on the state of its partner, may become infected. Specifically, a susceptible agent in group $i \in \{1, 2\}$ becomes infected with probability $\upsilon_i > 0$ if it happens to interact with an infected agent. Otherwise, i.e. if the agent is already infected, this agent recovers and becomes susceptible again with probability $\delta_i > 0$. Let $m \in [0, 1]$ represent the mixing level, i.e., the probability that an interacting agent of one particular group meets an agent from the other group. Thus, m is the expected fraction of between-group interactions. As illustrated in Fig. 1, if $m = 0$, agents interact only with agents from their own group (this is the fully segregated society case), whereas if $m = 1$, agents interact only with agents from the other group (this is the fully bipartite society case).

We use a continuous-time mean-dynamic approximation to study the evolution of the adoption levels in each group. To do so, let ρ_i denote the fraction of adopters (infected individuals) in group i. The evolution of ρ_i in each group over time is described by the following non-linear system of differential equations:

$$\dot{\rho}_i = p\upsilon_i(1-\rho_i)[m\rho_j + (1-m)\rho_i] - \delta_i\rho_i$$

where $i, j \in \{1, 2\}$ and $i \neq j$. Note that, here, the parameters p, υ_i and δ_i can be interpreted as rates instead of probabilities. We can rewrite these equations as:

$$\dot{\rho}_i = \delta_i[\lambda_i(1-\rho_i)[m\rho_j + (1-m)\rho_i] - \rho_i]$$

where $\lambda_i = (p\, \upsilon_i / \delta_i)$ represents the *effective adoption rate for group i*. The equilibria and the long-run predictions of this model can be characterized exclusively by three parameters: λ_1, λ_2 and m. Depending on the values of these parameters, the diffusion dynamics either converge from any initial state to the situation where there is no diffusion in either group, or present a positive, *almost globally asymptotically stable state* to which the system converges from any *non-null* initial fraction of adopters (26).

The space of plausible effective adoption rates $(\lambda_1, \lambda_2) \in [0, \infty) \times [0, \infty)$ can be partitioned into different regions (see Fig.2) according to the effects that variations in the mixing level m cause on the positive equilibrium values for each group, and for the population average. In region A, the no-diffusion state is a global attractor. In region B, there is a threshold for the mixing level below which there is a positive, almost globally asymptotically stable state, leading to positive diffusion in both groups. Above this threshold, no diffusion occurs in the long term. In the remaining regions, the positive, almost globally asymptotically stable state exists for any non-null value of the mixing level $m > 0$.

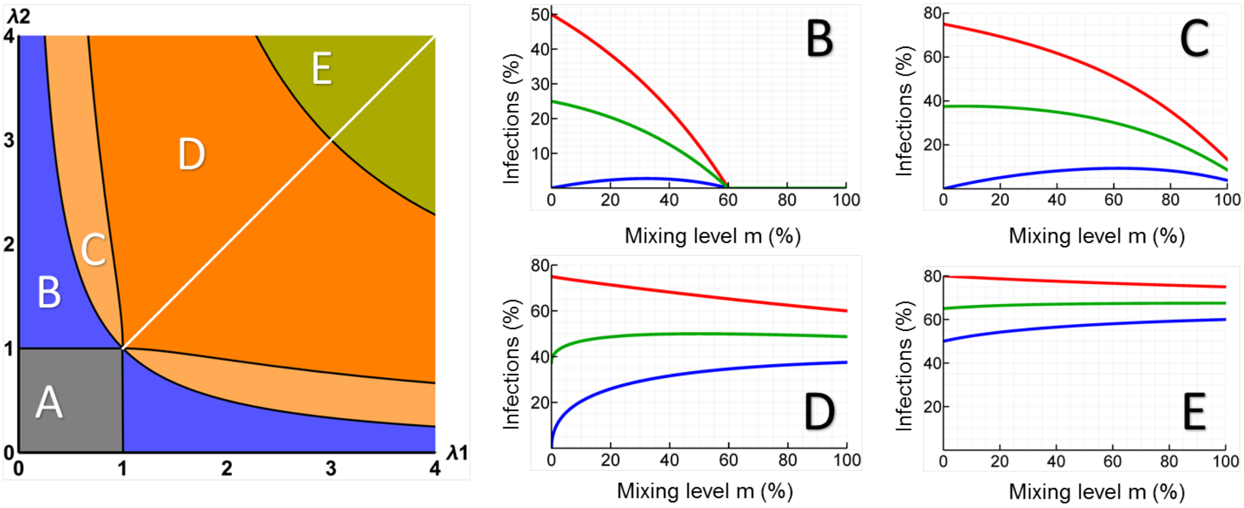

Fig. 2. Regions of effective adoption rate values (λ_1 and λ_2) corresponding to qualitatively different effects of the mixing level *m* on equilibrium diffusion levels. The main graph (left) illustrates the different regions in the $\{\lambda_1, \lambda_2\}$ plane. The accompanying graphs represent the stable equilibrium diffusion levels (blue for the resistant group, red for the sensitive group, and green for the average) as functions of *m* for $\{\lambda_1, \lambda_2\}$ values corresponding to regions B, C, D and E. In region A there is no diffusion, as also happens in region B for large *m*. In regions B and C, the diffusion level for the resistant group is a non-monotonic function of *m*, with an interior maximum. In regions C and D, the average diffusion level is also a non-monotonic function of *m*, with an interior maximum. In region E, the three diffusion levels are monotonic functions of *m*. The $\{\lambda_1, \lambda_2\}$ values selected for the graphs are: B: {0.25, 2}, C: {0.3, 4}, D: {1, 4}, and E: {2, 5}.

As mentioned above, the level of adoption in the resistant group (blue) in regions B and C is a non-monotonic function of *m*. This is a consequence of two opposing effects taking place simultaneously. First, the more intuitive effect, by mixing more with the sensitive group, which always has a larger number of adopters in any positive equilibrium, the number of adopters in the resistant group should increase. The second effect, however, is that as the sensitive group increases its between-group interactions, its adoption level decreases. This, in turn, creates a feedback loop effect on the resistant group such that, when the second effect is stronger than the first (i.e., when *m* is sufficiently high), the number of adopters in the resistant group decreases as mixing increases. Unlike that which occurs for the resistant group, the adoption level in the sensitive group always decreases with *m*. Consequently, Pareto inefficient mixing levels exist, wherein both groups would be better off at some other mixing level. For other combinations of effective adoption rates (regions D and E in Fig.2), all mixing levels are Pareto efficient; namely, an increase or decrease of mixing would always benefit one of the groups to the detriment of the other.

Discussion

Avoiding Pareto-dominated outcomes can be considered a generally desirable objective. In regions B and C (Fig. 1), this condition can rule out a wide range of non-efficient mixing levels but would generally not provide just one single optimal mixing level. A possible objective that a social planner could focus on would be to maximize the *average diffusion*. In such a case, and depending on the groups' propensities (SI), the optimal mixing level can range from the totally segregated case ($m = 0$) to the bipartite case ($m = 1$). An alternative objective worth exploring would be to reduce the diffusion difference between groups. To implement this, the bipartite case where all interactions are between groups (m = 1), is always the best structure (SI). However, it should be noted that this situation may be Pareto inefficient (see, e.g., the graph representative of region C in Fig. 1, assuming a positive trait).

This analysis relies on a deterministic mean-dynamic approximation of a stochastic process, which provides a good approximation to the actual stochastic dynamics occurring in finite populations when those populations are large (27, 28). Via simulation (SI), we show the robustness of the reported results to different population sizes, as well as to the inclusion of some within-group heterogeneity. More importantly, the SIS model of contagion analyzed here can be extended to better address some more general situations, such as peer effects in classrooms, where it seems more reasonable to assume that the probability of switching from infected (motivated student) to susceptible (non-motivated student) also depends on the current state of partners. The qualitative effects that we have described here for the multi-group SIS model, namely the existence of situations where the equilibrium diffusion levels are non-monotonic functions of the mixing level m, as well as the existence of Pareto-inefficient outcomes, hold in this more general case (see SI).

We conclude by emphasizing that selecting the best mixing level among heterogeneous groups depends not only on the desired objective, but typically on the effective adoption rates (or propensities) of each group, these being parameters that are well defined and potentially measurable. A 'one-fits-all' recommendation does not exist, meaning that the optimal policy could be very different for different contexts. Furthermore, the existence of inefficient mixing levels highlights the importance of estimating appropriately the relevant contagion parameters before embracing any particular policy.

Acknowledgments

Dunia López-Pintado gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation (ECO2011-22919).

References and Notes

1. Rogers EM (1995) *Diffusion of innovations* (The Free Press).
2. Granovetter M (1978) Threshold Models of Collective Behavior. *Am J Sociol* 83:489–515.
3. Jackson MO, Yariv L (2011) Diffusion, strategic interaction, and social structure. *Handbook of Social Economics* (North Holland Press).
4. Festinger L, Back K, Schachter S (1950) *Social pressures in informal groups: A study of human factors in housing* (Stanford University Press).
5. Gamoran A (1992) The Variable Effects of High School Tracking. *Am Sociol Rev* 57(6):812–828.
6. Jackson MO (2008) *Social and economic networks* (Princeton University Press).
7. Vega-Redondo F (2007) *Complex social networks* (Cambridge University Press).
8. Goyal S (2007) *Connections : an introduction to the network economy* (Princeton Univ. Press).

9. Watts DJ (2002) A simple model of global cascades on random networks. *Proc Natl Acad Sci U S A* 99(9).

10. Salganik MJ, Dodds PS, Watts DJ (2006) Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science (80-)* 311(5762):854–856.

11. Jackson MO, López-Pintado D (2013) Diffusion and contagion in networks with heterogeneous agents and homophily. *Netw Sci* 1(1):49–67.

12. López-Pintado D (2012) Influence networks. *Games Econ Behav* 75(2):776–787.

13. Jackson MO, Rogers BW (2007) Relating network structure to diffusion properties through stochastic dominance. *BE J Theor Econ* 7(1):1–13.

14. Goel S, Anderson A, Hofman J, Watts DJ (2016) The structural virality of online diffusion. *Manage Sci* 62(1).

15. Gómez S, Gómez-Gardeñes J, Moreno Y, Arenas A (2011) Nonperturbative heterogeneous mean-field approach to epidemic spreading in complex networks. *Phys Rev E - Stat Nonlinear, Soft Matter Phys* 84(3).

16. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a Feather: Homophily in Social Networks. *Annu Rev Sociol* 27(1):415–444.

17. Bramoullé Y, Currarini S, Jackson MO, Pin P, Rogers BW (2012) Homophily and long-run integration in social networks. *J Econ Theory* 147(5):1754–1786.

18. Centola D (2011) An Experimental Study of Homophily in the Adoption of Health Behavior. *Science (80-)* 334(6060):1269–1272.

19. Androushchak G, Poldin O, Yudkevich M (2013) Role of peers in student academic achievement in exogenously formed university groups. *Educ Stud* 39(5):568–581.

20. Hidalgo-Hidalgo M (2014) Tracking Can Be More Equitable Than Mixing. *Scand J Econ* 116(4):964–981.

21. Bailey N (1975) *The mathematical theory of infectious diseases and its applications* (Charles Griffin & Company Ltd.).

22. Pastor-Satorras R, Vespignani A (2001) Epidemic dynamics and endemic states in complex networks. *Phys Rev E* 63(6):66117.

23. Galeotti A, Rogers BW (2013) Strategic immunization and group structure. *Am Econ J Microeconomics* 5(2):1–32.

24. Reluga TC (2009) An SIS epidemiology game with two subpopulations. *J Biol Dyn* 3(5):515–531.

25. Gómez-Gardeñes J, Latora V, Moreno Y, Profumo E (2008) Spreading of sexually transmitted diseases in heterosexual populations. *Proc Natl Acad Sci U S A* 105(5).

26. Rass L, Radcliffe J (2000) Global Asymptotic Convergence Results for Multitype Models. *Int J Appl Math Comput Sci* Vol. 10(no 1):63–79.

27. Sandholm WH (2010) *Population games and evolutionary dynamics* (The MIT Press).

28. Izquierdo SS, Izquierdo LR (2013) Stochastic Approximation to Understand

Simple Simulation Models. *J Stat Phys* 151(1–2):254–276.
29. Lajmanovich A, Yorke JA (1976) A deterministic model for gonorrhea in a nonhomogeneous population. *Math Biosci* 28(3–4):221–236.

Supporting Information

A. Two-group SIS model. Results

The system of differential equations describing the evolution of adoption in each group $i \in \{1, 2\}$ over time is:

$$\dot{\rho}_i = p\upsilon_i(1-\rho_i)[m\rho_j + (1-m)\rho_i] - \delta_i\rho_i \tag{1}$$

where m is the mixing level, and where $j \in \{1,2\}, j \neq i$, indicates the other group. In terms of the effective adoption rates, Eq. 1 is:

$$\dot{\rho}_i = \delta_i[\lambda_i(1-\rho_i)[m\rho_j + (1-m)\rho_i] - \rho_i] \tag{2}$$

The stationary states of Eq. 1 are the pairs of values $(\rho_1, \rho_2) \in [0, 1]^2$ such that $\dot{\rho_1} = 0$ and $\dot{\rho_2} = 0$. We are interested in the dynamics and the stationary states of Eq. 1 for $0 < \lambda_1 < \lambda_2$ and $m \in (0,1)$ - these assumptions will be kept throughout the analysis, with the extreme cases $m = 0$, $m = 1$ and $\lambda_1 = \lambda_2$ discussed in the proofs section.

A1. Proposition 1

If $\lambda_2 > 1$ *and either* $\lambda_1\lambda_2 > 1$ *or* $m < \overline{m}$, where $\overline{m} = 1 - \frac{1-\lambda_1\lambda_2}{\lambda_2+\lambda_1-2\lambda_1\lambda_2}$, *there is a positive globally asymptotically stable state in* $[0, 1]^2 \backslash \{(0,0)\}$. *Otherwise the no-diffusion state is globally asymptotically stable in* $[0, 1]^2$.

As an immediate corollary, if $\lambda_2 \leq 1$ the no-diffusion state is globally asymptotically stable. As a note, the same threshold $\overline{m}$ can be found in (11) for local stability of the no-diffusion state.

Next we study the effect of the mixing level m on the stable diffusion levels, within the range of positive values of m where the stable positive equilibrium exists, i.e., for $\lambda_2 > 1$, $m \in (0,1]$ if $\lambda_1\lambda_2 > 1$ and $m \in (0, \overline{m})$ if $\lambda_1\lambda_2 \leq 1$.

Let ρ_1^E and ρ_2^E be the functions of m in the indicated range that provide the corresponding positive equilibrium value for each group. By analyzing the derivatives of ρ_1^E and ρ_2^E with respect to m we obtain the following results for the positive equilibrium diffusion levels.

A2. Proposition 2

The stable positive equilibrium diffusion level in the sensitive group ρ_2^E is a strictly decreasing function of the mixing level m.

Its maximum level, $\rho_2^E = 1 - \lambda_2^{-1}$, corresponds to $m = 0$. If $\lambda_1\lambda_2 > 1$, ρ_2^E obtains its minimum value $\rho_2^E = \frac{\lambda_1\lambda_2 - 1}{\lambda_1\lambda_2 + \lambda_1} > 0$ at $m = 1$. If $\lambda_1\lambda_2 \leq 1$, ρ_2^E decreases to $\lim_{m \to \overline{m}^-} \rho_2^E = 0$.

A3. Proposition 3

For adoption rates $\lambda_1 > \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1}$, which is the case if $\lambda_1 > 1$, the stable positive diffusion level in the resistant group ρ_1^E is a strictly increasing function of the mixing level $m \in (0,1)$. The minimum level for ρ_1^E, corresponding to $m = 0$, is $\lim_{m \to 0^+} \rho_1^E =$ $\text{Max}(0, 1 - \lambda_1^{-1})$.

For adoption rates $\lambda_1 < \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1} < 1$, the stable positive diffusion level in the resistant group ρ_1^E is a non-monotonic function of m which increases from $\lim_{m \to 0^+} \rho_1^E = 0$, obtains a maximum value for some interior value of m and then decreases either back to $\lim_{m \to \overline{m}^-} \rho_1^E = 0$ if $\lambda_1\lambda_2 \leq 1$, or to $\rho_1^E = \frac{\lambda_1\lambda_2 - 1}{\lambda_1\lambda_2 + \lambda_2} > 0$ (at $m = 1$) if $\lambda_1\lambda_2 > 1$.

Let $\rho^E = \frac{1}{2}(\rho_1^E + \rho_2^E)$ be the average diffusion level in the population in the stable positive equilibrium, which is a function of m in the range in which the positive equilibrium exists: $\lambda_2 > 1$, $m \in (0,1]$ if $\lambda_1\lambda_2 > 1$ and $m \in (0, \overline{m})$ if $\lambda_1\lambda_2 \leq 1$.

A4. Proposition 4

If $\lambda_1\lambda_2 \leq 1$, ρ^E is a decreasing function of m, with $lim_{m \to \overline{m}^-} \rho^E = 0$. If $\lambda_1^{3/4}(\lambda_2 + 1) \geq \lambda_2^{3/4}(\lambda_1 + 1)$, which requires $\lambda_1\lambda_2 > 1$ and $\lambda_2 > 3$, ρ^E is an increasing function of m.

Otherwise ρ^E is a non-monotonic function of m, first increasing and then decreasing, *which obtains an interior maximum. In this last case, the minimum average diffusion is obtained at* $m = 1$ *if* $\lambda_1 < \frac{\sqrt{\lambda_2}}{\lambda_s - \sqrt{\lambda_2} + 1}$, *or at* $m = 0$ *otherwise.*

A5. Proposition 5

The difference between the stable positive diffusion levels in the sensitive and resistant groups $(\rho_2^E - \rho_1^E)$ *is a strictly decreasing function of the mixing level* m*. This difference is always positive for* $(\rho_1^E, \rho_2^E) \neq (0,0)$.

B. Proofs and auxiliary results for the two-group SIS model

B1. Stationary states for $\mathbf{m = 0}$, $\mathbf{m = 1}$ or $\boldsymbol{\lambda_1 = \lambda_2}$

If m = 0, then each group is independent of the other and it is quick to check that the stationary states for group i are $\rho_i = 0$ and, if $\lambda_i > 1$, also $\rho_i = 1 - \lambda_i^{-1}$.
If m = 1 the stationary states can also be found directly by solving $\dot{\rho_1} = 0$ and $\dot{\rho_2} = 0$, leading to $(\rho_1, \rho_2) = (0,0)$ and, if $\lambda_1\lambda_2 > 1$, also $(\rho_1, \rho_2) = (\frac{\lambda_1\lambda_2 - 1}{\lambda_1\lambda_2 + \lambda_2}, \frac{\lambda_1\lambda_2 - 1}{\lambda_1\lambda_2 + \lambda_1})$.
For $\lambda_1 = \lambda_2$ the stationary states are $(\rho_1, \rho_2) = (0,0)$ and, if $\lambda_1 > 1$, also $\rho_1 = \rho_2 = 1 - \lambda_1^{-1}$. In this case the stationary states are not affected by the value of the mixing level $m \in (0,1]$.

B2. Proof of proposition 1

An analytical proof of proposition 1 can be derived from the multitype SIS model results in Rass and Radcliffe (26) or Lajmanovich and Yorke (29). Here we present an alternative approach that relies on graphical arguments, and which can be applied to some parameterizations of the extended model presented in section D.

By substituting the values $\rho_1 = 0$ and $\rho_2 = 0$ in Eq. 1, it is quick to check that the no-diffusion state $(\rho_1, \rho_2) = (0,0)$ is a stationary state. We will show that, under certain conditions, it is unique and globally asymptotically stable; otherwise, there is a unique positive stationary state which is almost globally asymptotically stable.

Apart from the no-diffusion state, we can restrict the analysis of stationary states to the region $(\rho_1, \rho_2) \in (0,1) \times (0,1)$: at any stationary state the diffusion levels in each group must be less than one (note in Eq. 1 that if $\rho_1 = 1$ then $\dot{\rho_1} < 0$) and the only stationary state in which the equilibrium value in some group is null is the no-diffusion state (note in Eq. 1 that, for $m > 0$, $\dot{\rho_1} = 0$ with $\rho_1 = 0$ imply $\rho_2 = 0$).

For $\rho_2 \in (0,1)$ and $m \in (0,1)$, $\dot{\rho_1}$ is a second-degree polynomial in ρ_1:

$$\dot{\rho_1} = \delta_1[-\lambda_1(1-m)\rho_1^2 + (\lambda_1(1-m+m\rho_2)-1)\rho_1 + \lambda_1 m\rho_2]$$

This polynomial is such that $\dot{\rho_1}(\rho_1 = 0) = \delta_1\lambda_1 m\rho_2 > 0$ and $\dot{\rho_1}(\rho_1 = 1) = -\delta_1 < 0$. Consequently, there is a unique value $\rho_1^* \in (0,1)$ such that $\dot{\rho_1}(\rho_1 = \rho_1^*) = 0$. Besides, $\dot{\rho_1} < 0$ if $\rho_1 \in (\rho_1^*, 1)$ and $\dot{\rho_1} > 0$ if $\rho_1 \in (0, \rho_1^*)$. The equation $\dot{\rho_1} = 0$ defines a "reaction" function $\rho_1^R: (0,1) \times (0,1) \to (0,1)$, $(\rho_2, m) \to \rho_1^R(\rho_2, m)$, which provides the corresponding equilibrium diffusion level in the resistant group and which, for any fixed value of m, can be shown to be a continuous strictly increasing and strictly concave function of ρ_2. [See B3]

Likewise, $\dot{\rho_2} = 0$ defines a reaction function $\rho_2^R: (0,1) \times (0,1) \to (0,1)$, $(\rho_1, m) \to \rho_2^R(\rho_1, m)$, which, for any fixed value of m, is a strictly increasing and strictly concave function of ρ_2.

For any fixed value of m, a positive stationary state of Eq. 1 is an interior crossing point of the graphs of ρ_1^R and ρ_2^R in the $\rho_1 - \rho_2$ plane in $(0,1) \times (0,1)$ - See Figure 1 and Figure 2-. Since these two continuous functions can be easily extended to the compact $[0,1] \times [0,1]$ and are increasing and concave with respect to their first argument, there is a unique positive stationary state of Eq. 1 if and only if:

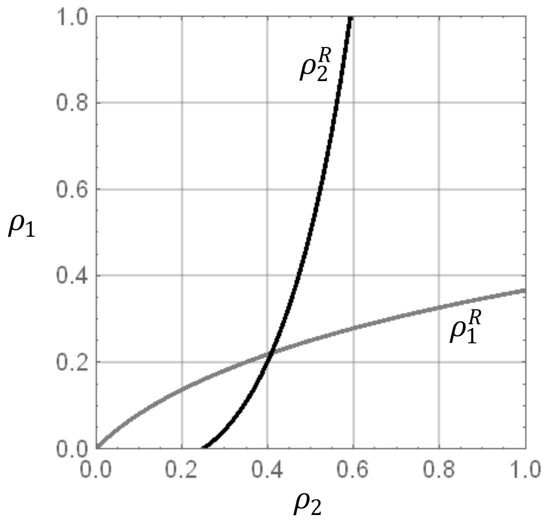

Fig. S1. Reaction curves which do not meet at the origin: $\lim_{\rho_1 \to 0^+} \rho_2^R > 0$

i) Either any of the graphs does not converge to the origin, which happens if $\lim_{\rho_1 \to 0^+} \rho_2^R > 0$ (as illustrated in Figure S1), or, otherwise,

ii) The graph of ρ_2^R is above the graph of ρ_1^R near the origin, which happens if the product of the slopes near the origin is greater than one, i.e., if $\lim\limits_{\rho_1 \to 0^+} \frac{\partial \rho_2^R}{\partial \rho_1} \lim\limits_{\rho_2 \to 0^+} \frac{\partial \rho_1^R}{\partial \rho_2} > 1$ (as illustrated in Figure S2).

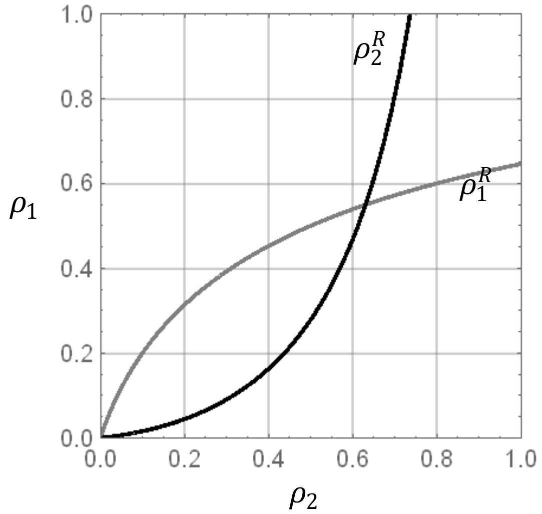

Fig S2. Reaction curves with slopes at the origin such that, besides the crossing point at the origin, there is an interior crossing point in $[\mathbf{0},\mathbf{1}] \times [\mathbf{0},\mathbf{1}]$: $\boldsymbol{\rho_1^R}$ is above $\boldsymbol{\rho_2^R}$ near the origin.

Almost global convergence of Eq. 1 to the positive stationary state when it exists, or, otherwise, global convergence of Eq. 1 to the no-diffusion state, follow from the analysis of the vector $(\dot{\rho_1}, \dot{\rho_2})$ in the different sub-regions of $(0,1) \times (0,1)$ obtained from the graphs of ρ_1^R and ρ_2^R, as indicated in Figure S3. Note that sub-region A in figure S3 does not exist when there is no interior stationary state, and then there is global convergence to the origin.

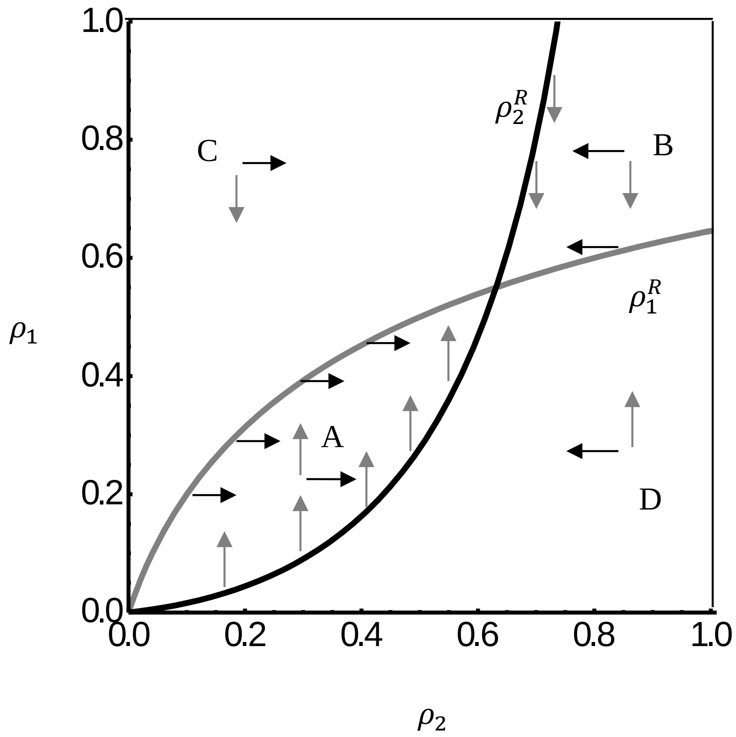

Fig. S3. Phase portrait schema for Eq. 1. The arrows next to the letters labelling each region and over the reaction curves indicate the positive components of the vector field at the interior of each region and at the reaction curves.

Considering that

$$\lim_{\rho_1 \to 0^+} \rho_2^R\,(\rho_1, m) = Max(0, \frac{\lambda_2(1-m)-1}{\lambda_2(1-m)}),$$

and, if $\lambda_2(1-m) < 1$,

$$\lim_{\rho_1 \to 0^+} \frac{\partial \rho_2^R}{\partial \rho_1} \lim_{\rho_2 \to 0^+} \frac{\partial \rho_1^R}{\partial \rho_2} = \frac{\lambda_1 m}{1-\lambda_1(1-m)} \frac{\lambda_2 m}{1-\lambda_s(1-m)},$$

the conditions for the existence of the almost globally asymptotically stable interior stationary state are $\lambda_2(1-m) \geq 1$ or $\frac{\lambda_1 m}{1-\lambda_1(1-m)} \frac{\lambda_2 m}{1-\lambda_2(1-m)} > 1$, which are equivalent (See B4) to $\lambda_2 > 1$ and either $\lambda_1\lambda_2 > 1$ or $m < \overline{m} = 1 - \frac{1-\lambda_1\lambda_2}{\lambda_2+\lambda_1-2\lambda_1\lambda_2}$.

B3. Strictly increasing and strictly concave reaction functions.

From (2), let $F_i \equiv \lambda_i(1-\rho_i)[m\rho_j + (1-m)\rho_i] - \rho_i$

Note that $\dot{\rho_i} = \delta_i F_i$. Solving $F_1 = 0$ for $(\rho_1, \rho_2) \in (0,1) \times (0,1)$ and $m \in (0,1)$ provides the solution function

$$\rho_1^R = \frac{\lambda_1(1-m+m\rho_2) - 1 + \sqrt{(\lambda_1(1-m+m\rho_2)-1)^2 + 4\lambda_1^2 m(1-m)\rho_2}}{2\lambda_1(1-m)}$$

An implicit differentiation of $F_1 = 0$ leads to

$$\frac{\partial \rho_1^R}{\partial \rho_2} = -\frac{\partial F_1/\partial \rho_2}{\partial F_1/\partial \rho_1} = \frac{\lambda_1 m(1-\rho_1)}{\lambda_1[m\rho_2 + (1-m)\rho_1 - (1-m)(1-\rho_1)] + 1}$$

which, considering $F_1 = 0$, simplifies to

$$\frac{\partial \rho_1^R}{\partial \rho_2} = \frac{\lambda_1 m(1-\rho_1)^2}{1-\lambda_1(1-m)(1-\rho_1)^2}$$

Taking into account $F_1 = 0$ again, the term $\lambda_1(1-m)(1-\rho_1)^2$ in the denominator is

$$\lambda_1(1-m)(1-\rho_1)^2 = \frac{(1-m)\rho_1}{m\rho_2 + (1-m)\rho_1}(1-\rho_1) < 1$$

Consequently, $1-\lambda_1(1-m)(1-\rho_1)^2 > 0$ and $\frac{\partial \rho_1^R}{\partial \rho_2} > 0$.

To show that $\frac{\partial^2 \rho_1^R}{\partial \rho_2^2} < 0$ note that the function $\frac{\partial \rho_1^R}{\partial \rho_2} = \frac{\lambda_1 m(1-\rho_1)^2}{1-\lambda_1(1-m)(1-\rho_1)^2}$ is increasing with respect to $(1-\rho_1)^2$, while $\frac{\partial (1-\rho_1^R)^2}{\partial \rho_2} < 0$.

The proof for $\frac{\partial \rho_2^R}{\partial \rho_1} > 0$ and $\frac{\partial^2 \rho_2^R}{\partial \rho_1^2} < 0$ proceeds analogously.

B4. Existence regions for the interior almost globally asymptotically stable stationary state.

For $m \in (0,1)$ and $0 < \lambda_1 < \lambda_2$, the region satisfying

$$\lambda_2(1-m) \geq 1 \tag{3}$$

or

$$\frac{\lambda_1 m}{1-\lambda_1(1-m)} \frac{\lambda_2 m}{1-\lambda_2(1-m)} > 1 \tag{4}$$

is the same region satisfying

$$\lambda_1\lambda_2 > 1 \tag{5}$$

or

$$\lambda_2 > 1 \text{ and } m < \overline{m} = 1 - \frac{1-\lambda_1\lambda_2}{\lambda_2 + \lambda_1 - 2\lambda_1\lambda_2}. \tag{6}$$

We will need some auxiliary results to prove this result. Note that, for $m \in (0{,}1)$ and $0 < \lambda_1 < \lambda_2$:

i) Any of the equations from Eq. 3 to Eq. 6 independently imply $\lambda_2 > 1$.

ii) For $\lambda_2 > 1$ and $\lambda_1\lambda_2 \leq 1$, Eq. 4 $\Leftrightarrow$ Eq. 6. It is quick to check that, for $\lambda_2 + \lambda_1 - 2\lambda_1\lambda_2 > 0$, Eq. 4 and Eq. 6 are the same condition. Besides, $\lambda_1\lambda_2 \leq 1$ implies $\lambda_2 + \lambda_1 - 2\lambda_1\lambda_2 > 0$, given that $\lambda_2 + \lambda_1 - 2\lambda_1\lambda_2 \geq \lambda_2 + \lambda_1 - 2\sqrt{\lambda_2\lambda_1} = (\sqrt{\lambda_2} - \sqrt{\lambda_1})^2 > 0$.

iii) If $\lambda_2 > 1$ and $\lambda_1\lambda_2 \leq 1$ then $0 \leq \frac{1-\lambda_2\lambda_1}{\lambda_2+\lambda_1-2\lambda_1\lambda_2} < \frac{1}{\lambda_2}$ and, consequently, $\overline{m} > 1 - \frac{1}{\lambda_2}$ and Eq. 3 $\Longrightarrow$ Eq. 6.

Here we have used the result $\lambda_2 + \lambda_1 - 2\lambda_1\lambda_2 > \lambda_2(1 - \lambda_1\lambda_2)$. To show this, note that $\lambda_2 > 1 (\lambda_2 - 1 > 0) 1 - 2\lambda_2 > -\lambda_2^2$. Consequently,

$\lambda_2 + \lambda_1 - 2\lambda_1\lambda_2 = \lambda_2 + \lambda_1(1 - 2\lambda_2) > \lambda_2 - \lambda_1\lambda_2^2 = \lambda_2(1 - \lambda_1\lambda_2)$.

Proof of the equivalence of regions:

Eq. 3 $\Longrightarrow$ (Eq. 5 or Eq. 6). Suppose Eq. 3 holds. Then either Eq. 5 or $\lambda_1\lambda_2 \leq 1$ with $\lambda_2 > 1$, in which case, because of iii), we obtain Eq. 6.

Eq. 4 $\Longrightarrow$ (Eq. 5 or Eq. 6). Suppose Eq. 4 holds. Then either Eq. 5 or $\lambda_1\lambda_2 \leq 1$ with $\lambda_2 > 1$, in which case, because of ii), we obtain Eq. 6.

Eq. 5 $\Longrightarrow$ (Eq. 3 or Eq. 4). Suppose Eq. 5 holds. Then either Eq. 3 or $\lambda_2(1 - m) < 1$, in which case Eq. 4 can be seen to be equivalent to

$$\lambda_1 > \frac{1 - \lambda_2(1 - m)}{\lambda_2 + (1 - m)(1 - 2\lambda_2)}$$

which holds because, for $\lambda_2 > 1$, we have

$$\frac{1 - \lambda_2(1 - m)}{\lambda_2 + (1 - m)(1 - 2\lambda_2)} < \frac{1}{\lambda_2},$$

given that $\lambda_2 + (1 - m)(1 - 2\lambda_2) > \lambda_2 - (1 - m)\lambda_2^2 > 0$.

Eq. 6 $\Longrightarrow$ (Eq. 3 or Eq. 4). Suppose Eq. 6 holds. Then either Eq. 5, in which case we have (Eq. 3 or Eq. 4), or $\lambda_1\lambda_2 \leq 1$, in which case, because of ii), we obtain Eq. 4.

B5. Lemma 1

Any stationary state $(\rho_1, \rho_2) \neq (0,0)$ *of Eq. 1 satisfies* $0 < \rho_1 < \rho_2 < 1$. *Furthermore,* $1 - \lambda_1^{-1} < \rho_1 < \rho_2 < 1 - \lambda_2^{-1}$.

Proof: It was shown before that $0 < \rho_1 < 1$ and $0 < \rho_2 < 1$. Then, from $F_1 = 0$ and $F_2 = 0$ we obtain

$$\lambda_1 < \lambda_2 \Rightarrow \frac{\rho_1}{(1-\rho_1)[m\rho_2 + (1-m)\rho_1]} < \frac{\rho_2}{(1-\rho_2)[m\rho_1 + (1-m)\rho_2]} \Rightarrow$$

$$\Rightarrow \frac{\rho_1^2}{m(1-\rho_1) + (1-m)\rho_1} < \frac{\rho_2^2}{m(1-\rho_2) + (1-m)\rho_2} \Rightarrow \rho_1 < \rho_2$$

where the last step follows from the fact that the function $f(x) = \frac{x^2}{m(1-x)+(1-m)x}$ is strictly increasing for $x > 0$.

Now, given that the reaction functions are increasing in their first argument, we have that, at a stationary state (ρ_1, ρ_2),

$$\rho_1 = \rho_1^R(\rho_2, m) > \rho_1^R(\rho_1, m) = 1 - \lambda_1^{-1}$$

$$\rho_2 = \rho_2^R(\rho_1, m) < \rho_2^R(\rho_2, m) = 1 - \lambda_2^{-1}$$

B6. Derivatives of ρ_1^E and ρ_2^E with respect to m.

These derivatives are defined in the region $\lambda_2 > 1$, with $m \in (0, 1)$ if $\lambda_1\lambda_2 > 1$, and $m \in (0,\ \overline{m})$ if $\lambda_1\lambda_2 \leq 1$.

Let $|J_{F,\rho}|$ be the determinant of the Jacobian matrix

$$J_{F,\rho} = \begin{pmatrix} \partial F_1/\partial \rho_1 & \partial F_1/\partial \rho_2 \\ \partial F_2/\partial \rho_1 & \partial F_2/\partial \rho_2 \end{pmatrix}$$

By the implicit function theorem, we have that, at any stationary point with $|J_{F,\rho}| \neq 0$,

$$\frac{\partial \rho_1^E}{\partial m} = |J_{F,\rho}|^{-1}\left(\frac{\partial F_1}{\partial \rho_2}\frac{\partial F_2}{\partial m} - \frac{\partial F_2}{\partial \rho_2}\frac{\partial F_1}{\partial m}\right), \text{ and}$$

$$\frac{\partial \rho_2^E}{\partial m} = |J_{F,\rho}|^{-1}\left(\frac{\partial F_2}{\partial \rho_1}\frac{\partial F_1}{\partial m} - \frac{\partial F_1}{\partial \rho_1}\frac{\partial F_2}{\partial m}\right)$$

where

$$\frac{\partial F_1}{\partial \rho_1} = \lambda_1[(1-m)(1-2\rho_1) - m\rho_2] - 1,$$

$$\frac{\partial F_1}{\partial \rho_2} = \lambda_1(1-\rho_1)m, \text{ and}$$

$$\frac{\partial F_1}{\partial m} = \lambda_1(1-\rho_1)(\rho_2-\rho_1),$$

with equivalent equations for the derivatives of F_2. Considering that, at a stationary state, $F_1 = 0$ and $F_2 = 0$, we also have

$$\frac{\partial F_1}{\partial \rho_1} = \frac{\lambda_1(1-\rho_1)^2(1-m)-1}{1-\rho_1}, \text{ and}$$

$$\frac{\partial F_2}{\partial \rho_2} = \frac{\lambda_2(1-\rho_2)^2(1-m)-1}{1-\rho_2}$$

Leading to

$$\frac{\partial \rho_1^E}{\partial m} = |J_{F,\rho}|^{-1}\lambda_1(1-\rho_1)(\rho_2-\rho_1)[-m\lambda_2(1-\rho_2) - \frac{\lambda_2(1-\rho_2)^2(1-m)-1}{1-\rho_2}], \text{ and}$$

$$\frac{\partial \rho_2^E}{\partial m} = |J_{F,\rho}|^{-1}\lambda_2(1-\rho_2)(\rho_2-\rho_1)[m\lambda_1(1-\rho_1) + \frac{\lambda_1(1-\rho_1)^2(1-m)-1}{1-\rho_1}]$$

with

$$|J_{F,\rho}| = \frac{\lambda_1(1-\rho_1)^2(1-m)-1}{1-\rho_1}\frac{\lambda_2(1-\rho_2)^2(1-m)-1}{1-\rho_2} - \lambda_1\lambda_2 m^2(1-\rho_1)(1-\rho_2)$$

Let us first show that $|J_{F,\rho}| \neq 0$ at the positive stationary state.

As shown in B3, the first term in $|J_{F,\rho}|$ is not null at a stationary state, so

$$|J_{F,\rho}| = 0 \Rightarrow 1 - \frac{\lambda_1 m(1-\rho_1)^2}{1-\lambda_1(1-m)(1-\rho_1)^2}\,\frac{\lambda_2 m(1-\rho_2)^2}{1-\lambda_2(1-m)(1-\rho_2)^2} = 0$$

But the product of the fractions in this expression is the product of the slopes of the reaction curves $\frac{\partial \rho_1^R}{\partial \rho_2}\frac{\partial \rho_2^R}{\partial \rho_1}$ which, following our previous discussion, must be less than one and positive at an interior crossing point. Consequently, $|J_{F,\rho}| > 0$ at the positive stationary state.

B7. Proof of proposition 2

Solving $\frac{\partial \rho_2^E}{\partial m} = 0$ with $F_1 = 0$ and $F_2 = 0$ provides a single solution at $m = \frac{\rho_1^2}{\rho_1 - \rho_2}$ which, by lemma 1, is negative, so $\frac{\partial \rho_2^E}{\partial m}$ does not change sign for $m \in (0, 1)$. It is then quick to check that $\frac{d\rho_2^E}{dm}$ is negative in the region of existence of the positive stationary equilibrium, i.e., for $m \in (0, 1)$ if $\lambda_1\lambda_2 > 1$, or $m \in (0, \overline{m})$ if $\lambda_1\lambda_2 \leq 1$.

B8. Proof of proposition 3

Solving $\frac{\partial \rho_1^E}{\partial m} = 0$ with $F_1 = 0$ and $F_2 = 0$ provides no solution for $m \in (0, 1)$ if $\lambda_1 \geq \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1}$.

Given that $\frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1} = \frac{\sqrt{\lambda_2}}{(\sqrt{\lambda_2} - 1)^2 + \sqrt{\lambda_2}} \leq 1$, we have that if $\lambda_1 > 1$ then $\lambda_1 > \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1}$.
Besides, for $\lambda_2 > 1$, $\lambda_1 > \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1} \Rightarrow \lambda_2\lambda_1 > 1$, and consequently the almost globally asymptotically stable positive equilibrium exists for $m \in [0, 1]$.
Considering that, if $\lambda_1 > \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1}$, $\frac{d\rho_1^E}{dm}$ does not change sign with $m \in (0, 1)$, it is quick to check that $\frac{d\rho_r^E}{dm}$ is positive for $m \in (0, 1)$.

B9. Proof of proposition 4

Solving $\frac{\partial \rho_1^E}{\partial m} = 0$ with $F_1 = 0$ and $F_2 = 0$ provides, for $\lambda_1 < \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1}$, a unique solution at $m = m_1 \in (0,1)$:

$$m_1 = \frac{2\lambda_1\left(\sqrt{\lambda_2} - 1\right)^2}{\lambda_1 + \lambda_2 + 2\lambda_1\lambda_2 - 4\lambda 1\sqrt{\lambda_2} - \sqrt{4\lambda_1^2\lambda_2\left(\sqrt{\lambda_2} - 1\right)^2 + \left(\lambda_1 + \lambda_2 - 2\lambda_1\sqrt{\lambda_2}\right)^2}}$$

such that $\frac{d\rho_1^E}{dm}$ is positive for $m \in (0, m_1)$ and negative for $m \in (m_1, 1)$ if $\lambda_1\lambda_2 > 1$, or for $m \in (m_1, \overline{m})$ if $\lambda_1\lambda_2 \leq 1$.
The condition $\lambda_1 < \frac{\sqrt{\lambda_2}}{\lambda_2 - \sqrt{\lambda_2} + 1}$ implies $\lambda_1 < 1$, so $\lim_{m \to 0^+} \rho_1^E = 0$.

B10. Proof of proposition 5

Solving $\frac{\partial \rho_2^E}{\partial m} + \frac{\partial \rho_1^E}{\partial m} = 0$ leads to $\lambda_1(1-\rho_1)^2 = \lambda_2(1-\rho_2)^2$. Together with $F_1 = 0$ and $F_2 = 0$, this provides the formula for the mixing level m_ρ that maximizes the average diffusion, from which the results follow:

$$m_\rho = \frac{1}{2\lambda_1(\lambda_1-\lambda_2)^2\lambda_2}\Big[2\lambda_1^2\lambda_2 + \lambda_1^3\lambda_2 - 4\lambda_1^{5/2}\lambda_2^{3/2} + 2\lambda_1\lambda_2^2 + 6\lambda_1^2\lambda_2^2 + 2\lambda_1^3\lambda_2^2 -$$

$$4\lambda_1^{3/2}\lambda_2^{5/2} + \lambda_1\lambda_2^3 + 2\lambda_1^2\lambda_2^3 - 2(\lambda_1\lambda_2)^{3/2} - 4(\lambda_1\lambda_2)^{5/2} - \sqrt{\lambda_1^5\lambda_2} - \sqrt{\lambda_1\lambda_2^5} +$$

$$\mathrm{Abs}\left[-1+\sqrt{\lambda_1\lambda_2}\right]\left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^2\sqrt{\lambda_1\lambda_2\left(\lambda_2^2+\lambda_1^2(1+4\lambda_2^2)+2\lambda_1\lambda_2\left(1-4\sqrt{\lambda_1\lambda_2}\right)\right)}\Big]$$

B11. Proof of proposition 6

Solving $\frac{\partial \rho_2^E}{\partial m} - \frac{\partial \rho_1^E}{\partial m} = 0$ for m, with $F_1 = 0$ and $F_2 = 0$, provides a unique positive solution at $m = \frac{\rho_2\rho_1}{\rho_2+\rho_1-1}\frac{\rho_2+\rho_1-2\rho_2\rho_1}{(\rho_2-\rho_1)^2}$. For $0 < \rho_i < 1$, this expression is greater than one: it can be seen that the first fraction $\frac{\rho_2\rho_1}{\rho_2+\rho_1-1}$ is greater than one from expanding $(1-\rho_1)(1-\rho_2) > 0$. The second fraction is also greater than one because $\rho_i^2 < \rho_i$. Consequently, $\frac{\partial \rho_2^E}{\partial m} - \frac{\partial \rho_1^E}{\partial m}$ does not change sign in the considered range of values of m where the positive equilibrium exists. It is then quick to check that the difference is strictly decreasing.

C Robustness

A computer program available at https://luis-r-izquierdo.github.io/micopro/ implements the two-group SIS model. It allows testing the validity of the mean dynamic approximation for different population sizes and with heterogeneous propensities, uniformly distributed in each group around the group's average value λ_i.

As an example, figure S4 shows the diffusion values obtained in each group and on average for a population of 1000 individuals with individual propensities in each group uniformly distributed in the range $\lambda_i \mp 0.3\,\lambda_i$, with $\lambda_1 = 0.25$ for the resistant group, and $\lambda_2 = 2$ for the susceptible group. The lines correspond to the equilibria obtained with the mean dynamic equations and each dot corresponds to an average across 10 independent samples of the time-averaged infection values obtained between time periods 10 000 and 11 000 (after the system has had time to evolve towards an equilibrium from the initial conditions). Even with the considered level of heterogeneity, the correspondence between the simulated values and the mean field is very precise but for the cases in which the expected equilibrium levels are positive but low (corresponding to mixing levels between 0.45 and 0.60 in the represented case), as some simulations will then reach the absorbing state in which there is no infection and will remain there. In every other case, we obtain dispersion ranges between samples which are less than 1% for the resistant group and less than 6% for the susceptible group. The initial conditions turn out not to be relevant (as expected), provided that the initial fractions of infected individuals are not too close to 0. The results in the graph were obtained with a 10% initial fraction of infected individuals.

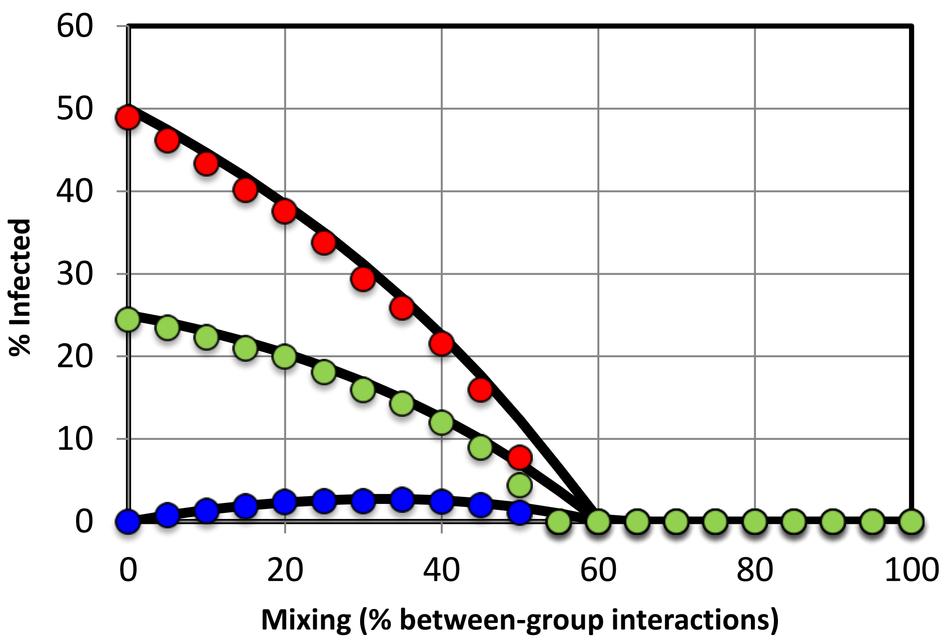

Fig. S4. Average fraction of infected individuals in each group (red and blue dots) and in the whole population (green dots) as a function of the mixing level, with heterogeneous individual propensities within groups. The dots correspond to simulations of the process. The lines correspond to the equilibria obtained from the mean dynamic equations using the average values for the propensities in each group: $\boldsymbol{\lambda_1 = 0.25}$ and $\boldsymbol{\lambda_2 = 2}$. The dispersion range of individual propensities within each group is ±30% of the average value, following a uniform distribution in that range.

D Extended model

Consider an extended model which includes as parameters for each population $i \in \{1, 2\}$ the following contagion and recovery rates, conditional on the current state of the interacting individual (or network partner):

- $\upsilon_{i|0}$: rate at which individuals in state 0 being matched to an individual in state 0 adopt state 1. In an infection model, this is the rate of infection when meeting a healthy partner. In a school motivation model, this is the rate at which students in motivation state 0 and with a partner in motivation state 0 adopt motivation state 1.
- $\upsilon_{i|1}$: rate at which individuals in state 0 being matched to an individual in state 1 adopt state 1. This is the rate of infection when meeting an infected partner.
- $\delta_{i|0}$: rate at which individuals in state 1 being matched to an individual in state 0 adopt state 0. This is the rate of recovery when having a healthy partner.
- $\delta_{i|1}$: rate at which individuals in state 1 being matched to an individual in state 1 adopt state 0. This is the rate of recovery when having an infected partner.

Leaving apart the limit cases in which some parameters are either 1 or null, we assume that $0 < \upsilon_{i|0} < \upsilon_{i|1} < 1$ and $0 < \delta_{i|1} < \delta_{i|0} < 1$.

The probability that an individual in group i interacts with an individual in state 1 is $I_i = (1 - m)\rho_i + m\rho_j$, with index j indicating the other group: $j \in \{1, 2\}, j \neq i$.

The system of differential equations describing the evolution of adoption in each group over time is:

$$\dot{\rho}_i = (1 - \rho_i)[\upsilon_{i|1} I_i + \upsilon_{i|0}(1 - I_i)] - \rho_i[\delta_{i|1} I_i + \delta_{i|0}(1 - I_i)]$$

For each value of m and $\rho_j \in [0, 1]$, $\dot{\rho}_i$ is a second degree polynomial in ρ_i such that

$$\dot{\rho}_i(\rho_i = 0) = \upsilon_{i|0} + m\rho_j(\upsilon_{i|1} - \upsilon_{i|0}) > 0, \text{ and}$$

$$\dot{\rho}_i(\rho_i = 1) = -[\delta_{i|1} + m(1 - \rho_j)(\delta_{i|0} - \delta_{i|1})] < 0.$$

Consequently, for each value of m and $\rho_j \in [0, 1]$ there is a unique and positive value of $\rho_i \in [0, 1]$ satisfying $\dot{\rho}_i = 0$, and this equation defines a "reaction function" $\rho_i^R: [0, 1] \times [0, 1] \to (0, 1)$, $(\rho_j, m) \to \rho_i^R(\rho_j, m)$, which provides the corresponding equilibrium diffusion level in group i.

It is not difficult to see that, for a fixed value of m, ρ_i^R is strictly increasing in ρ_j (this follows from the fact that $\frac{\partial \dot{\rho}_i}{\partial \rho_j} > 0$) and that, if $(v_{i|1} - v_{i|0}) > (\delta_{i|0} - \delta_{i|1})$, it is also strictly concave (strictly convex if the inequality is reversed). For the strictly concave-concave case, as well as for the strictly convex-convex case (which can be easily transformed to the concave-concave case via a change of variables), the same arguments used before can be applied to show that there is a unique interior globally asymptotically stable stationary state.

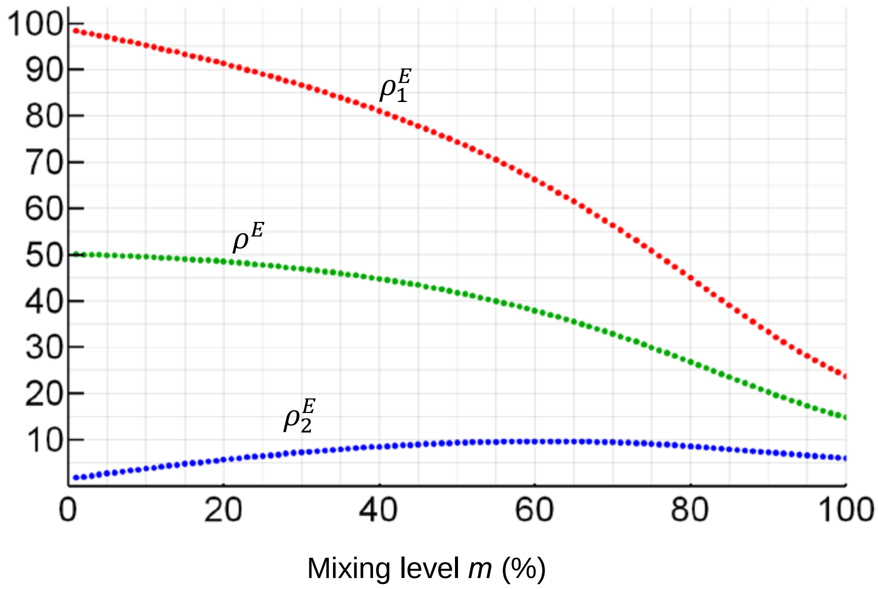

Fig. S5. Equilibrium diffusion values as a function of the mixing level in the extended model. Parameter values (%): $(\boldsymbol{v_{1|0}}, \boldsymbol{v_{1|1}}, \boldsymbol{\delta_{1|1}}, \boldsymbol{\delta_{1|0}}) = (\mathbf{1}, \mathbf{80}, \mathbf{1}, \mathbf{20})$; $(\boldsymbol{v_{2|0}}, \boldsymbol{v_{2|1}}, \boldsymbol{\delta_{2|1}}, \boldsymbol{\delta_{2|0}}) = (\mathbf{1}, \mathbf{15}, \mathbf{60}, \mathbf{70})$.

Figure S5 illustrates the equilibrium diffusion values (ρ_1^E, ρ_2^E) as a function of the mixing level m for a parameterization of this model. It shows that the same qualitative features that we discussed for the multi-group SIS model also apply to this extended model: in particular, the diffusion level in a group can be a non-monotonic function of the mixing level, and there can be Pareto-inefficient mixing levels. Again, for policy considerations, it also shows that the qualitative global effect of modifying the level of segregation or mixing between groups varies with the parameter values of the groups: given most reasonable sets of objectives, a modification of the segregation level in one direction (increase or decrease) could be beneficial or prejudicial depending on the status quo and on the parameter values of both groups.