

Testing, Voluntary Social Distancing and the Spread of an Infection

Daron Acemoglu^{*} Ali Makhdoumi[†] Azarakhsh Malekian[‡] Asuman Ozdaglar[§]

June 21, 2020

Abstract

We study the effects of testing policy on voluntary social distancing and the spread of an infection. Agents decide their social activity level, which determines a social network over which the virus spreads. Testing enables the isolation of infected individuals, slowing down infection. But greater testing also reduces voluntary social distancing or increases social activity, exacerbating the spread of the virus. We show that the effects of testing on infections is non-monotone. This non-monotonicity also implies that the optimal testing policy may leave some of the testing capacity of society unused.

1 Introduction

The COVID-19 pandemic has reignited interest in models of epidemics and their control. A point of broad agreement among different approaches is that ramping up testing capacity is one of the most effective ways of combating the pandemic (see e.g. [Searchinger et al. \[2020\]](#), [Brumfiel \[2020\]](#), and [BruSoe-Lin and Hecht \[2020\]](#)). One issue that has not received much attention, however, is whether and how different testing strategies will impact voluntary social distancing decisions of individuals.

In this paper, we develop a simple model to investigate the effects of testing on infections and provide insights on optimal testing strategies. We model social activity and voluntary distancing as a network formation problem and use a simple percolation process to represent the spread of a virus over the endogenous social network. Our choice for a model of percolation rather than the SIR (susceptible, infected and recovered) model more commonly used in the analysis of COVID-19 is motivated by two considerations. First, our percolation model enables a fairly tractable, albeit general, analysis of the spread of the virus and how it depends on the endogenous social network. Second, to the extent that the optimal social policy is to minimize infections until a vaccine and

^{*}Dept. of Economics, Massachusetts Institute of Technology. daron@mit.edu

[†]Fuqua School of Business, Duke University. ali.makhdoumi@duke.edu

[‡]Rotman School of Management, University of Toronto. azarakhsh.malekian@rotman.utoronto.ca

[§]Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. asuman@mit.edu

a cure arrive (rather than reaching herd immunity), a percolation model, even though it does not have the recovered individuals, will provide a good approximation to the outcomes under an SIR model.

Our main results confirm the major benefit society can reap from testing and isolating infected individuals, but also leads to two new insights. First, greater testing can lead to more social activity (less social distancing) and thus a denser social network, because when infected individuals are more likely to be identified and isolated, agents feel more secure to initiate contacts. We show that, for a nontrivial set of parameters, greater testing can increase infections. This happens, in particular, when the equilibrium involves some groups choosing an intermediate level of social activity because of their fear of infection. Second, we demonstrate that when the testing capacity of society is limited, optimal policy may involve leaving some of this capacity unused in order to avoid adverse effects on social distancing. This also implies that testing should be combined with mandatory social distancing measures to avoid these adverse behavioral effects.

More formally, our model consists of n individuals of two types, high-type agents with greater value from social activities and low-type agents with lower values (more than two types are considered as an extension). Each individual i chooses a social activity level $x_i \in [0, 1]$, and the vector of social activities defines a contact (social) network among individuals. Specifically, we assume that there will be contact between agents i and j with probability $\eta x_i x_j$, where $\eta < 1$ captures the probability of a match conditional on activities. The utility of each individual is her utility gain from social activities minus her infection probability. The virus spreads over the social network following a cascade process, whereby each infected and non-isolated individual transmits the infection to its neighbors with probability $\beta \in (0, 1]$. An infected individual does not transmit the virus after isolation, but we assume that transmission still occurs before testing detects and isolates an infected individual. More specifically, we assume that an infected individual who is detected transmits the virus with a smaller probability βp , where $p \in [0, 1]$.

Our analysis relies on a key lemma which establishes that the stochastic process for the spread of the virus satisfies natural monotonicity and concavity (submodularity) properties: more social activity leads to more infections but does so less when an individual himself has a high level of activity.

We first prove that, as often maintained in prior analyses, more testing reduces infections when we take activity levels (and thus the social network) as given. Our key results concern the case in which the social activity levels are endogenous. First, we show that when testing probabilities are large, both type of individuals choose maximal social activity (which can be interpreted as the same level of activity as the pre-pandemic period). In this region, the equilibrium behaves in an identical fashion to the exogenous social network case and testing always reduces infections. More interestingly, however, for lower testing probabilities, either high-value or low-value individuals choose intermediate levels of social activity and greater testing increases their infection probability.

We next turn to an analysis of optimal testing (where the objective is to maximize the value of

social activity minus the cost of infections). We assume that there is a fixed supply of tests and then investigate how these should be allocated. Because high-value agents are socially more active, they are more likely to be infected and thus they should be tested first. If there is sufficient testing capacity, it is optimal to test all agents. But most interestingly, we prove that for intermediate or low values of testing capacity, it is socially optimal not to use all available tests. The social planner should either test all high-value agents but not the low-value agents, so as to discourage them from social activity. Or she should have no testing, even though there is capacity to test some of the high-value agents. The intuition is again via the impact of testing on social activity: additional testing over this range would increase social activity so much that it would have no benefit in terms of containing the infection.

As already emphasized, these non-monotonicity results are due to the impact of testing on equilibrium social activity (social distancing). If the social planner can mandate social distancing, then it is straightforward to show that there is no non-monotonicity and it is always optimal to use all of the available testing capacity.

We also show that uniform testing policies, where tests are allocated without reference to the type of agents, are significantly worse than targeted testing policies.

In our baseline model, high-value agents should be tested first because they are more active and thus more likely to transmit the virus. In an extension where we allow individuals to choose type-specific social activity levels (e.g., how much to socialize with high-value and low-value agents), we show that not testing high-value agents may have an additional strategic benefit—it discourages low-value agents from socializing with high-value agents, slowing down the spread of the virus.

Related Literature

Our paper is related to three distinct literature. First, ours is a model of endogenous social network formation. Seminal papers in this area include [Jackson and Wolinsky \[1996\]](#), [Bala and Goyal \[2000\]](#), and [Currarini et al. \[2009\]](#) (see [Jackson \[2008\]](#) and [Vega-Redondo \[2007\]](#) for book-length treatments of issues of network formation and contagion in networks). Differently from the most common approach in this literature, which is to look at pairwise or coalition-wise stable outcomes, we adopt a non-cooperative approach to network formation (see also [Bollobás and Béla \[2001\]](#) and [Newman et al. \[2001\]](#)). In particular, we adopt a very simple (and to the best of our knowledge, novel) modeling strategy, where the probability of connection between two agents is proportional to the product of their levels of social activity. This leads to a tractable but rich set of interactions, in part because this structure makes activity levels both strategic substitutes and complements. [Goyal \[1993\]](#), [Bala and Goyal \[2000\]](#), and [Acemoglu et al. \[2017a\]](#) assume agents can unilaterally create directed links to others. In contrast, in our model the probability of a link between two agents depends on the action of both of them.

Second, there is a large literature spanning various areas of economics where precautionary tools increase risk-taking ([Peltzman \[1975\]](#)). For example, hydraulic breaks increase driving speed

and may have much lower effects on accidents than they would have done had driving behavior not responded (see e.g., [Lindgren and Stuart \[1980\]](#), [Crandall and Graham \[1984\]](#), and [Keeler \[1994\]](#)). Differently from other works in this vein, we showed that in our setting with social network interactions, the countervailing effect can outweigh the direct impacts and lead to higher infections from more testing.

Third, our paper is related to the recent literature on the effects of pandemics and especially of COVID-19. See, for example, [Kruse and Strack \[2020\]](#), [Atkeson \[2020\]](#), [Jones et al. \[2020\]](#), [Glover et al. \[2020\]](#), [Berger et al. \[2020\]](#), and [Birge et al. \[2020\]](#). More closely related are several papers endogenizing behavior and social distancing in the context of SIR models, such as [Leung et al. \[2018\]](#), [Toxvaerd \[2020\]](#), [Eichenbaum et al. \[2020\]](#), [Farboodi et al. \[2020\]](#), and [Maloney and Taskin \[2020\]](#). In particular, [Eichenbaum et al. \[2020\]](#) extend the canonical SIR model to study the interaction between economic decisions and pandemics, while [Leung et al. \[2018\]](#), [Toxvaerd \[2020\]](#), [Farboodi et al. \[2020\]](#), and [Maloney and Taskin \[2020\]](#) incorporate simple social distancing behavior in an otherwise standard SIR model.¹

Also closely related within this recent literature are a few recent papers studying the effects of lockdown and testing strategies. For example, [Alvarez et al. \[2020\]](#), [Acemoglu et al. \[2020\]](#), [Piguillem and Shi \[2020\]](#), and [Brotherhood et al. \[2020\]](#) look at the effects of different testing and isolation policies on the dynamics of infections and optimal lockdowns.

Various other issues such as estimation, testing, and control related to COVID-19 have also been studied. [Kaplan \[2020\]](#) considers a statistical model for estimating the effectiveness of isolation and quarantine. [Drakopoulos and Randhawa \[2020\]](#) study the allocation of inaccurate tests to strategic agents, and find the optimal policy for the social planner (see also [Ely et al. \[2020\]](#)) while [Kasy et al. \[2020\]](#) study the implications of false quarantine. Finally, [Drakopoulos et al. \[2017\]](#) study the open-loop control of epidemics on a network and shows how the network structure affects the amount of resources required to contains the spread of infection. To the best of our knowledge, no other paper has either provided a full characterization of social activity and how it depends on infection probabilities and testing strategies nor has pointed out the non-monotonic impacts of testing on infections.

The rest of the paper is organized as follows. In Section [2](#) we present our model, describing the formation of contact network and the stochastic process governing the spread of infection. We also show that for a given (exogenous) network, as expected, increasing the testing capability decreases the infection probability. In Section [3](#) we characterize the equilibrium outcome and show that increasing testing probabilities may adversely increase the (equilibrium) infection probability. In Section [4](#) we characterize the optimal testing policy and show it may be optimal to have underutilized testing capacity. Section [5](#) considers two extensions: multiple groups of agents and social activity differentially targeted towards different groups. Section [6](#) concludes, while the Appendix contains all the proofs and additional results.

¹Prior related work includes [Acemoglu et al. \[2016\]](#), who introduce precautionary behavior in the context of a virus spreading over a network, and [Morris \[2000\]](#), [Tardos and Wexler \[2007\]](#), [Blume et al. \[2011\]](#), [Acemoglu et al. \[2017a\]](#), [Elliott et al. \[2014\]](#), and [Acemoglu et al. \[2017b\]](#) who study contagion over financial networks.

2 The Environment

We study the spread of an infection among n individuals (also referred to as agents) represented by the set $\mathcal{V} = \{1, \dots, n\}$. Each agent $i \in \mathcal{V}$ decides about her level of *social activity* denoted by $x_i \in [0, 1]$. Higher social activity provides greater utility to agents but also leads to faster spread of the infection. Agents are heterogeneous in terms of their value of social interaction. In our main model, we assume that the agent's type (value of interaction) is either $v_H \in [0, 1]$ or $v_L \in [0, 1]$, where $v_L < v_H$. We also use \mathcal{H} and \mathcal{L} to denote individuals with high and low values, respectively, and r_H and r_L to denote the population fractions of high- and low-value agents. Section 5 extends our results to a setting with $m > 2$ types of agents.

An infection hits the society, starting from a random individual and then spreads to others through a stochastic process described in subsection 2.2. We analyze the implications of infection testing, represented by testing probabilities for two types of agents, $\alpha_L, \alpha_H \in [0, 1]$, and study the optimal *testing policy*, (α_L, α_H) , of a (benevolent) social planner. Each infected individual that is tested positive will be isolated from the rest of individuals. If an individual is infected and not isolated, she will expose her neighbors to the infection. If she is isolated, the individual still has a (smaller) chance to infect her neighbors and expose them to the infection.

We let $\mathbf{x} = (x_1, \dots, x_n)$ denote the social activity profile of all individuals. We also let \mathbf{x}_{-i} represent the social activity profile of all individuals except agent i . In what follows, for any vector $\mathbf{x} \in \mathbb{R}^n$ and set $S \subseteq \mathcal{V}$, \mathbf{x}_S denotes the elements of \mathbf{x} for the indices in S and \mathbf{x}_{-S} denote the elements of \mathbf{x} for the indices outside set S .

We next describe how social activity levels determine the (endogenous) social network in this community and how the infection spreads over this social network.

2.1 Network of Contacts

The social activity profile $\mathbf{x} = (x_1, \dots, x_n)$ of agents generates a social (contact) network in which agents i and j are connected with a probability that depends on x_i and x_j . In particular, let $\mathbf{G} = (\mathcal{V}, \mathbf{E})$ be a random network where $\mathbf{E} \in \{0, 1\}^{n \times n}$ denotes the random edges (thus, $\mathbf{E}_{ij} = \mathbf{E}_{ji}$), where \mathbf{E}_{ij} for $i, j \in \mathcal{V}$ are independent binary random variables with

$$\mathbb{P}[\mathbf{E}_{ij} = 1] = \eta x_i x_j,$$

for some $\eta \in (0, 1]$. This η captures the probability of match between two individuals conditional on activity levels. We denote a *realized network* by $G = (\mathcal{V}, E)$, where E_{ij} is a realization of \mathbf{E}_{ij} , and thus $E_{ij} = 1$ means that there is a link between agents i and j .

We next describe the stochastic process governing the spread of infection and define the infection probability of agents.

2.2 The Spread of Infection

Let us denote the neighbors of node i by $N(i) = \{j \in \mathcal{V} : E_{ij} = 1\}$. For a given network $G = (\mathcal{V}, E)$, there is a stochastic process that governs the spread of infection as follows. One of the individuals uniformly at random becomes infected, and the infection then spreads to others via a percolation process over the social network that is a generalization of the independent cascades model (see [Kempe et al. \[2015\]](#) for more details).

We can describe the dynamics of infection as follows: At time 0, one of the agents (chosen uniformly at random) denoted by $s \in \mathcal{V}$ gets infected. At any round $t \geq 0$ for any agent $i \in \mathcal{V}$ we let $d_i^{(t)}$ denote the number of neighbors of agent i that are infected and tested at time t and $\hat{d}_i^{(t)}$ denote the number of neighbors of agent i that are infected at time t but not tested. For each agent i , these two variables are initially 0 and then evolve over time as described next. At time 0 node s will be tested with probability α_i in which case at time 1 for all neighbors of agent s such as j we have $d_j^{(1)} = 1$. Here α_i is either α_H or α_L depending on whether agent s belongs to \mathcal{H} or \mathcal{L} . With probability $1 - \alpha_i$, however, agent s will not be tested in which case at time 1 for all neighbors of agent s such as j we have $\hat{d}_j^{(1)} = 1$. If an infected node is not tested (and therefore not isolated), she will be *active* for one round and transmits infection to its neighbors with *transmission probability* $\beta \in (0, 1]$. If an infected node is tested, she can be isolated and prevented from transmitting the virus. However, before isolation takes place, some social contacts will occur and lead to the spread of the infection. We represent this possibility with a smaller probability of infection, βp , where $p \in [0, 1)$. From active agents, the infection simultaneously and independently transmits to each of their uninfected neighbors. If an uninfected agent is a neighbor to multiple infected individuals (i.e., there exists at least two i and i' such that $j \in N(i)$ and $j \in N(i')$), then the infection is transmitted to agent j in an order-independent fashion. This implies, for example, that if j is uninfected and a neighbor to two active agents that are not tested, then j becomes infected with probability $1 - (1 - \beta)^2$. Note that if a neighbor of an active node does not become infected at time t , then it will never again become infected via that node. Given this process, for each node i , the probability of getting infected at time $t + 1$ becomes

$$1 - (1 - \beta)^{\hat{d}_i^t} (1 - \beta p)^{d_i^t},$$

where, we recall that \hat{d}_i^t is the number of i 's neighbors that are infected and not tested at time t and d_i^t is the number of i 's neighbors that are infected and tested at time t .

Definition 1. For any agent $i \in \mathcal{V}$, network $G = (\mathcal{V}, E)$, and testing policy (α_L, α_H) we let $\mathbb{P}_i^{\text{inf}}(G, \alpha_L, \alpha_H)$ denote the probability of infection reaching agent i . This probability is over the randomness in the source of infection, the randomness in testing, and the randomness in the stochastic process described above.

We next illustrate the stochastic process and the infection probability by means of an example.

Example 1.

We consider a setting with three agents $\{1, 2, 3\}$ who are fully connected to each other with testing probabilities $\alpha_H = \alpha_L = \alpha$ and find the infection probability of agent 3. We list the three cases for the source of infection and find this probability:

- With probability $1/3$, the infection hits agent 3: In this case, agent 3 gets infected.
- With probability $1/3$, the infection hits agent 1: In this case, there are two possibilities:
 - (i) Node 1 will get tested whose probability is α . Here there are two cases. In the first case, the infection reaches node 3 directly through the edge between nodes 1 and 3 with probability βp (i.e., through the red solid path in Figure 1). In the second case, this edge is not active and the infection reaches node 3 through node 2 whose probability is $(1 - \alpha)\beta^2 p + \alpha\beta^2 p^2$ (i.e., through the blue dashed path in Figure 1). The first term is the probability of node 2 not being tested in which case the infection reaches node 3 with probability $\beta p \times \beta$ and the second term is the probability of node 2 being tested in which case the infection reaches node 3 with probability $\beta p \times \beta p$.
 - (ii) Node 1 will not get tested whose probability is $(1 - \alpha)$. Similarly, we can find the infection probability.

Therefore, the overall infection probability of node 3, if the infection hits agent 1, is

$$\alpha (\beta p + (1 - \beta p) ((1 - \alpha)\beta^2 p + \alpha\beta^2 p^2)) + (1 - \alpha) (\beta + (1 - \beta) ((1 - \alpha)\beta^2 + \alpha\beta^2 p)).$$

- With probability $1/3$, the infection hits agent 2: This case is identical to the previous one by swapping the role of nodes 1 and 2.

Putting these three cases together, the infection probability of agent 3, $\mathbb{P}_3^{\text{inf}}(G, \alpha_L = \alpha, \alpha_H = \alpha)$, is

$$\frac{1}{3} + \frac{2}{3} (\alpha (\beta p + (1 - \beta p) ((1 - \alpha)\beta^2 p + \alpha\beta^2 p^2)) + (1 - \alpha) (\beta + (1 - \beta) ((1 - \alpha)\beta^2 + \alpha\beta^2 p))).$$

Definition 2. For a given social activity profile \mathbf{x} and testing policy (α_L, α_H) , we denote the infection probability of individual i in the random network of contacts by $\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H)$, i.e.,

$$\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) = \mathbb{E}_{\mathbf{G}=(\mathcal{V}, E): E_{ij} \sim \text{Bernoulli}(\eta x_i x_j)} \left[\mathbb{P}_i^{\text{inf}}(\mathbf{G}, \alpha_L, \alpha_H) \right].$$

This probability is over the randomness in the formed network, the randomness over the source of infection, the random in testing, and the randomness in the stochastic process governing the spread of infection.

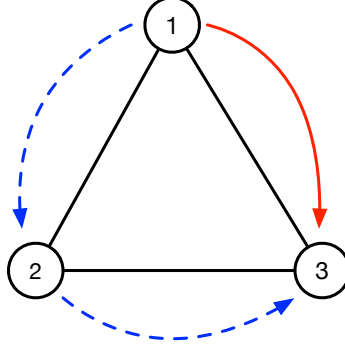


Figure 1: From agent 1 the infection can reach agent 3 from two paths: (i) directly via the edge between them whose probability depends on whether agent 1 will be tested or not and (ii) through agent 2 which, again, depends on whether agents 1 and 2 will be tested or not.

2.3 Utility of Agents and Solution Concept

The utility function of agent i is given by

$$u_i(\mathbf{x}, \alpha_L, \alpha_H) = v_i x_i - \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) - c (\alpha_L \mathbf{1}\{i \in \mathcal{L}\} + \alpha_H \mathbf{1}\{i \in \mathcal{H}\}), \quad (1)$$

where $v_i \in \{v_L, v_H\}$ and $c > 0$ is a (possibly small) cost of testing for an agent. The first term, $v_i x_i$, represents the utility gain from social activity, the second term, $-\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H)$, is the loss from getting infected, and the last term is the expected cost of getting tested.

As a solution concept, we use symmetric pure-strategy or mixed-strategy (Nash) equilibrium.

Definition 3. A pure-strategy social activity profile \mathbf{x}^e is a (pure-strategy) Nash equilibrium if

$$u_i(\mathbf{x}^e, \alpha_L, \alpha_H) \geq u_i((x_i, \mathbf{x}_{-i}^e), \alpha_L, \alpha_H), \quad \text{for all } i \in \mathcal{V}, x_i \in [0, 1].$$

A mixed-strategy social activity profile is of the form $\boldsymbol{\mu}^e = \prod_{i=1}^n \mu_i^e$ where μ_i^e is a probability distribution over $[0, 1]$. A mixed-strategy social activity profile is in equilibrium if

$$\mathbb{E}_{\mathbf{x}^e \sim \boldsymbol{\mu}^e} [u_i(\mathbf{x}^e, \alpha_L, \alpha_H)] \geq \mathbb{E}_{\mathbf{x}_{-i}^e \sim \boldsymbol{\mu}_{-i}^e} [u_i((x_i, \mathbf{x}_{-i}^e), \alpha_L, \alpha_H)], \quad \text{for all } i \in \mathcal{V}, x_i \in [0, 1].$$

A *symmetric* pure (mixed) strategy equilibrium is a pure (mixed) strategy equilibrium in which the decision of each agent i only depends on its value v_i and its infection probability and not on her identity.

2.4 Monotonicity and Concavity of Infection Probability

As illustrated in Example [1](#), the infection probability depends on the graph structure and testing policy in a complex way. Nevertheless, the next lemma shows that the stochastic process of the spread of the infection satisfies natural monotonicity and concavity properties. In what follows

for two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$, we write $\mathbf{a} \geq \mathbf{b}$ to denote $a_i \geq b_i$ for $i = 1, \dots, m$.

Lemma 1.

(a) For any agent $i \in \mathcal{V}$, we have

$$\mathbb{P}_i^{\text{inf}}(\hat{\mathbf{x}}, \alpha_L, \alpha_H) \geq \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H), \quad \text{for all } \hat{\mathbf{x}} \geq \mathbf{x}. \quad (2)$$

(b) For any agent $i \in \mathcal{V}$, and any social activity profile $\mathbf{x}_{-i} \in [0, 1]^{n-1}$, $\mathbb{P}_i^{\text{inf}}((x_i, \mathbf{x}_{-i}), \alpha_L, \alpha_H)$ is concave in x_i .

Part (a) of Lemma 1 is straightforward. It shows that the infection probability for an individual is increasing in the social activity levels of all agents in society, because higher social activity leads to a denser social network. Part (b), on the other hand, shows that this probability is concave in the individual's own social activity level, because additional social activity brings the virus to the individual only if her existing links did not do so already.

2.5 Exogenous Activities: Impact of Testing Policy on Infection Probability

We first analyze the case in which social activity levels are fixed and thus the social (contact) network G is given. The next proposition shows that more testing always reduces the spread of the infection in this case.

Proposition 1. For any network $G = (\mathcal{V}, E)$, the infection probability of each agent is decreasing in (α_L, α_H) . In particular, we have

$$\mathbb{P}_i^{\text{inf}}(G, \alpha'_L, \alpha'_H) \leq \mathbb{P}_i^{\text{inf}}(G, \alpha_L, \alpha_H), \quad \text{for all } i \in \mathcal{V}, (\alpha'_L, \alpha'_H) \geq (\alpha_L, \alpha_H).$$

Proposition 1 follows from a simple coupling argument, showing that when agents do not choose their social activity levels, greater testing probabilities (α_L, α_H) reduce the infection probability of all agents. This is intuitive: greater testing enables the detection and isolation of infected individuals, slowing down the spread of the virus.

We conclude this section by noting that all the results that will follow hold for any contagion process and infection probability as long as the infection probability has the properties stated in Lemma 1 and Proposition 1. Put differently, provided that for any $i \in \mathcal{V}$, $\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H)$ is increasing in \mathbf{x} , concave in x_i , and decreasing in (α_L, α_H) , the equilibrium characterization and the non-monotonicity of infection probability in testing policy, presented in the next section, hold.

3 Endogenous Activities

With endogenous networks each individual has a strategic decision to make: the social activity level which determines the number of her neighbors in the contact network.

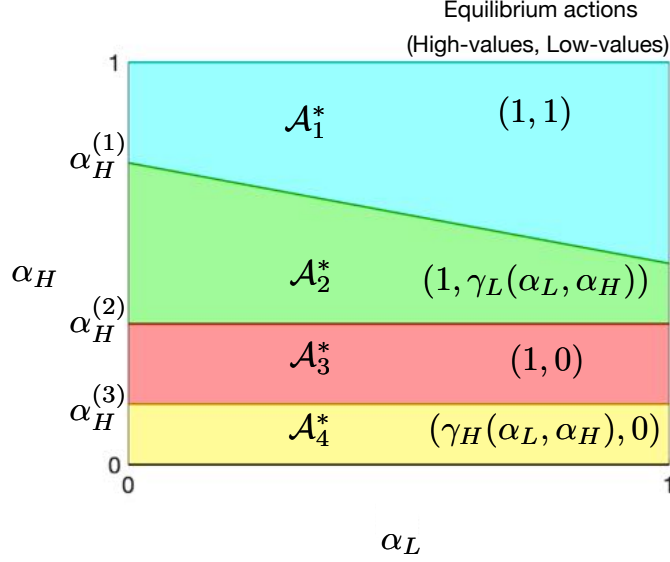


Figure 2: This graph depicts the regions $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$, and \mathcal{A}_4 (as $n \rightarrow \infty$) that determine the equilibrium in Proposition 2.

3.1 Equilibrium Characterization

We first introduce a few notations. We let $\mathbf{1}$ and $\mathbf{0}$ denote the vectors of all 1 and all 0, respectively where their dimension will be clear from the context. We also let l and h be a low- and high-value agent and define the following sets:

$$\begin{aligned}
\mathcal{A}_1 &= \{(\alpha_L, \alpha_H) \in [0, 1]^2 : \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{1}, \alpha_L, \alpha_H) \leq v_L + \frac{1}{n}, \\
&\quad \mathbb{P}_h^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{1}, \alpha_L, \alpha_H) \leq v_H + \frac{1}{n}\} \\
\mathcal{A}_2 &= \{(\alpha_L, \alpha_H) \in [0, 1]^2 : \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{1}, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}, \\
&\quad \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, x_l = 1, \mathbf{x}_{L \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) \leq v_L + \frac{1}{n}, \\
&\quad \mathbb{P}_h^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{0}, \alpha_L, \alpha_H) \leq v_H + \frac{1}{n}\} \\
\mathcal{A}_3 &= \{(\alpha_L, \alpha_H) \in [0, 1]^2 : \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{1}, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}, \\
&\quad \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, x_l = 1, \mathbf{x}_{L \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}, \\
&\quad \mathbb{P}_h^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{0}, \alpha_L, \alpha_H) \leq v_H + \frac{1}{n}\} \\
\mathcal{A}_4 &= \{(\alpha_L, \alpha_H) \in [0, 1]^2 : \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{1}, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}, \\
&\quad \mathbb{P}_h^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{0}, \alpha_L, \alpha_H) \geq v_H + \frac{1}{n}\}.
\end{aligned}$$

We recall that \mathcal{H} denotes the set of agents with a high value v_H for social activity and \mathcal{L} denotes

the set of agents with a low value v_L for social activity.

Proposition 2. *There exists $M \in \mathbb{N}_+$ and functions $\gamma_L : [0, 1]^2 \rightarrow [0, 1]$ and $\gamma_H : [0, 1]^2 \rightarrow [0, 1]$ such that for $n \geq M$, depending on the testing policy, (α_L, α_H) there are four possibilities for the equilibrium:*

- (a) *For $(\alpha_L, \alpha_H) \in \mathcal{A}_1$, there exists a unique symmetric equilibrium where $x_i^e = 1$ for all $i \in \mathcal{V}$.*
- (b) *For $(\alpha_L, \alpha_H) \in \mathcal{A}_2$, there exists a unique symmetric equilibrium where $x_i^e = 1$ for all $i \in \mathcal{H}$ and a mixed-strategy μ_j^e for all $j \in \mathcal{L}$ that puts probability $\gamma_L(\alpha_L, \alpha_H)$ on 1 and probability $1 - \gamma_L(\alpha_L, \alpha_H)$ on 0.*
- (c) *For $(\alpha_L, \alpha_H) \in \mathcal{A}_3$, there exists a unique symmetric equilibrium where $x_i^e = 1$ for all $i \in \mathcal{H}$ and $x_j^e = 0$ for all $j \in \mathcal{L}$.*
- (d) *For $(\alpha_L, \alpha_H) \in \mathcal{A}_4$, there exists a unique symmetric equilibrium where $x_j^e = 0$ for all $j \in \mathcal{L}$ and a mixed-strategy μ_i^e for all $i \in \mathcal{H}$ that puts probability $\gamma_H(\alpha_L, \alpha_H)$ on 1 and probability $1 - \gamma_H(\alpha_L, \alpha_H)$ on 0.*

Proposition 2 characterizes the equilibrium outcome for different testing policies (α_L, α_H) . For the case in which the number of agents in society is greater than a threshold, it divides testing policies into four regions, which lead to different types of equilibrium behavior. For example, when testing policy is in the set \mathcal{A}_1 , the infected are sufficiently likely to be identified and isolated that all individuals choose full social activity—as if social contacts did not increase their probability of infection. Outside of this region, individuals take precautionary action by reducing their social activity.² In the region \mathcal{A}_2 , high-value agents (who receive greater utility from social activity) still choose full activity, but now low-value agents mix between no activity and full activity.³ The other regions are defined similarly. In the proof of Proposition 2 we show that the four sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$, and \mathcal{A}_4 will cover all possible testing policies $(\alpha_L, \alpha_H) \in [0, 1]^2$ (for sufficiently large n).

Note finally that the sets \mathcal{A}_1 to \mathcal{A}_4 are simple functions of the model primitives. In fact, the next lemma shows that as $n \rightarrow \infty$, these sets have a very simple characterization. For this lemma, recall that r_H and r_L denotes the fractions of the two types of individuals.

Lemma 2. *There exist $\alpha_H^{(1)} \geq \alpha_H^{(2)} \geq \alpha_H^{(3)}$ such that as $n \rightarrow \infty$ the sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4$ converge to⁴*

$$\begin{aligned} \mathcal{A}_1^* &= \left\{ (\alpha_L, \alpha_H) \in [0, 1]^2 : \alpha_L r_L + \alpha_H r_H \geq \alpha_H^{(1)} r_H \right\}, \\ \mathcal{A}_2^* &= \left\{ (\alpha_L, \alpha_H) \in [0, 1]^2 : \alpha_L r_L + \alpha_H r_H \leq \alpha_H^{(1)} r_H, \alpha_H \geq \alpha_H^{(2)} \right\}, \\ \mathcal{A}_3^* &= \left\{ (\alpha_L, \alpha_H) \in [0, 1]^2 : \alpha_L r_L + \alpha_H r_H \leq \alpha_H^{(1)} r_H, \alpha_H^{(3)} \leq \alpha_H \leq \alpha_H^{(2)} \right\}, \text{ and} \end{aligned}$$

²In particular, in this region n must be large enough such that, when all agents are playing 1, a change in one agent's testing probability does not change the infection probability of other agents by more than $v_H - v_L$.

³The reason they do not choose an intermediate level of activity is that, given the concavity of the infection probability, mixing gives them higher utility.

⁴We say a sequence of sets $\{A\}_{n=1}^\infty$ converges to set A if for any $\epsilon > 0$, there exists $M \in \mathbb{N}_+$ such that for $n \geq M$, we have $A_n \subseteq A^{(\epsilon)}$ and $A \subseteq A_n^{(\epsilon)}$, where for any set B , $B^{(\epsilon)}$ denotes $\bigcup_{b \in B} \{x : \|x - b\|_2 \leq \epsilon\}$.

$$\mathcal{A}_4^* = \left\{ (\alpha_L, \alpha_H) \in [0, 1]^2 : \alpha_L r_L + \alpha_H r_H \leq \alpha_H^{(1)} r_H, \alpha_H \leq \alpha_H^{(3)} \right\},$$

respectively.

To obtain the intuition of the proof of Lemma 2, let us consider set \mathcal{A}_1 for which both type of agents are playing 1 and therefore their testing probability affect the infection probabilities of other agents. The convergence of the sets to these asymptotic objects follows from the fact that for large enough n , the number of tests will be concentrated around its mean which is $\alpha_L r_L + \alpha_H r_H$ and the infection probability is decreasing in the number of tests. The proof does not readily follow from law of large numbers because the infection probability is a complex non-linear function of the number of tests. To establish this result, we develop a “peeling argument” that uses the properties of the infection probability (such as concavity) together with a concentration bound. These four sets are depicted in Figure 2. The equilibrium action profiles of high- and low-value agents as a function of the pair (α_L, α_H) is also depicted in Figure 2.

In our baseline analysis we adopt symmetric equilibrium solution concept to simplify the notation (i.e., the strategy of an agent depends on the infection probability from her perspective and not her identity). In Appendix A.2 we show that the equilibrium characterization given in Proposition 2 is essentially unique. In particular, we show that the pure-strategy equilibrium in parts (a) and (c) are unique. We then characterize the asymmetric pure-strategy equilibrium (for parts (b) and (d)) and show that, for large n , the expected number of infected individuals in any asymmetric pure-strategy equilibrium is the same as the expected number of infected individuals in the symmetric equilibrium characterized in Proposition 2.

3.2 Impact of Testing Policy on Infection Probability

The next theorem presents one of our main results in this paper. It shows the non-monotonic impact of greater testing on infections.

Theorem 1. *There exists $M \in \mathbb{N}_+$ such that for $n \geq M$, in the unique symmetric equilibrium we have:*

- (a) *Higher (α_L, α_H) in the interior of \mathcal{A}_1 decreases the infection probabilities of both types of agents.*
- (b) *Higher (α_L, α_H) in the interior of \mathcal{A}_2 increases the infection probability of low-value agents and does not change the infection probability of high-value agents.*
- (c) *Higher (α_L, α_H) in the interior of \mathcal{A}_3 decreases the infection probabilities of both types of agents.*
- (d) *Higher (α_L, α_H) in the interior of \mathcal{A}_4 increases the infection probability of high-value agents and does not change the infection probability of low-value agents.*

Moreover, infection probabilities are continuous in (α_L, α_H) at the boundaries of the above sets.

Theorem 1 establishes that whenever we have a mixed-strategy equilibrium, as in regions \mathcal{A}_2 and \mathcal{A}_4 , the effects of greater testing on infections are non-monotonic. Intuitively, this is because

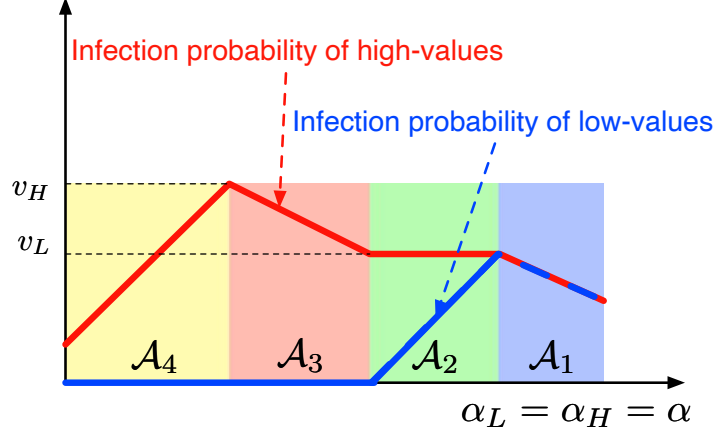


Figure 3: The schematic view of the infection probability of high- and low value agents, in the limit as $n \rightarrow \infty$, for uniform policy (i.e., $\alpha_H = \alpha_L = \alpha$) as α increases.

greater testing and isolation makes agents that are mixing wish to go to full activity. But when we are (and remain) in the interior of the sets \mathcal{A}_2 and \mathcal{A}_4 , the pure-strategy full activity level is not an equilibrium (either for the low-value or the high-value agents). Hence, equilibrium is restored by some more of the relevant agents choosing selectivity at the margin, which increases contacts and thus restores the incentives for mixing by increasing infection probabilities. In both cases, as Proposition 1 highlighted, with given activity levels, greater testing would have reduced the spread of the infection. The reason why the infection spreads more is because greater social activity levels make the social network denser.

In practice, it may not be possible to test different types of agents at different rates. If so, we would have to impose $\alpha_L = \alpha_H = \alpha$ for some $\alpha \in [0, 1]$. Figure 3 depicts the infection probability of both types as a function of α , illustrating that infection proclivities continue to be non-monotonic in testing probabilities (in this case α).

Theorem 1 again highlights the importance of the sets $\mathcal{A}_1, \dots, \mathcal{A}_4$ in our analysis. We next provide a comparative statics for these sets as η (the probability of match conditional on activities), β (the transmission rate of the infection), and p (where βp is the transmission rate of tested individuals) vary.

Lemma 3. *Let $\mathcal{A}_1(\eta, \beta, p)$ and $\mathcal{A}_4(\eta, \beta, p)$ denote the sets \mathcal{A}_1 and \mathcal{A}_4 as a function of the parameters η , β , and p . We have*

(a) *Higher η shrinks the set \mathcal{A}_1 and expands the set \mathcal{A}_4 . That is,*

$$\mathcal{A}_1(\eta', \beta, p) \subseteq \mathcal{A}_1(\eta, \beta, p) \text{ and } \mathcal{A}_4(\eta, \beta, p) \subseteq \mathcal{A}_4(\eta', \beta, p), \quad \text{for all } \eta' \geq \eta, \beta, p.$$

(b) *Higher β shrinks the set \mathcal{A}_1 and expands the set \mathcal{A}_4 . That is,*

$$\mathcal{A}_1(\eta, \beta', p) \subseteq \mathcal{A}_1(\eta, \beta, p) \text{ and } \mathcal{A}_4(\eta, \beta, p) \subseteq \mathcal{A}_4(\eta, \beta', p), \quad \text{for all } \beta' \geq \beta, \eta, p.$$

(c) Higher p shrinks the set \mathcal{A}_1 and expands the set \mathcal{A}_4 . That is,

$$\mathcal{A}_1(\eta, \beta, p') \subseteq \mathcal{A}_1(\eta, \beta, p) \text{ and } \mathcal{A}_4(\eta, \beta, p) \subseteq \mathcal{A}_4(\eta, \beta, p'), \quad \text{for all } p' \geq p, \beta, \eta.$$

This lemma directly follows from the definition of the sets \mathcal{A}_1 and \mathcal{A}_4 and the fact that the infection probabilities are increasing in η , β , and p . The boundary between sets \mathcal{A}_2 and \mathcal{A}_3 (for large enough n) also increases as we increase either η , β , or p . The sets \mathcal{A}_2 and \mathcal{A}_3 , however, can either shrink or expand.

4 Optimal Testing Policy

We now discuss the design of optimal testing policy (α_L, α_H) to maximize social welfare. Throughout we assume that there is a limited testing capacity represented by θn where $\theta \in [0, 1]$ and n is the number of individuals. We refer to θ as *testing capacity*. Social welfare is

$$W(\mathbf{x}, \alpha_L, \alpha_H) = \sum_{i=1}^n u_i(\mathbf{x}, \alpha_L, \alpha_H).$$

The game among the planner and the agents is a two-stage game with the following timing:

1. The social planner chooses the testing policy (α_L, α_H) .
2. Given the testing policy, the unique symmetric equilibrium from Proposition 2 is played.

With this timing and notation, the social planner's problem becomes

$$\begin{aligned} \max_{(\alpha_L, \alpha_H) \in [0, 1]^2} & W(\mathbf{x}^e, \alpha_L, \alpha_H) \\ \text{s.t. } & \mathbf{x}^e \text{ is the unique (symmetric) equilibrium,} \\ & \alpha_H |\mathcal{H}| + \alpha_L |\mathcal{L}| \leq \theta n. \end{aligned}$$

Our main result of this section, stated next, characterizes the optimal policy.

Theorem 2. *There exist $M \in \mathbb{N}_+$, $\bar{c} \in \mathbb{R}_+$, and $\theta^{(1)} \geq \theta^{(2)} \geq \theta^{(3)}$ such that for $c \leq \bar{c}$ and $n \geq M$ we have:*

- (a) *If $\theta > \theta^{(1)}$, then the optimal testing policy is to test all individuals with probability θ and in the corresponding equilibrium all agents are fully active, i.e., $x_i^e = 1$ for all $i \in \mathcal{V}$.*
- (b) *If $\theta^{(2)} < \theta < \theta^{(1)}$, then the optimal testing policy is to test only high-value individuals with probability $\frac{\theta^{(2)}}{\tau_H}$ and in the corresponding equilibrium high-value agents are fully active and low-value agents are inactive, i.e., $x_i^e = 1$ for all $i \in \mathcal{H}$ and $x_j^e = 0$ for all $j \in \mathcal{L}$.*

Equilibrium with optimal policy: (high-value, low-value)	(0, 0)	(1, 0)	(1, 0)	(1, 1)
Testing capacity in the optimal policy	Underutilized	Fully-utilized	Underutilized	Fully-utilized

θ

Figure 4: Depending on θ , the fraction of society that can be tested, there are four possibilities for the optimal testing policy and the corresponding social activity profile in equilibrium.

- (c) If $\theta^{(3)} < \theta < \theta^{(2)}$, then the optimal testing policy is to test only high-value individuals with probability $\frac{\theta}{r_H}$ and in the corresponding equilibrium high-value agents are fully active and low-value agents are inactive, i.e., $x_i^e = 1$ for all $i \in \mathcal{H}$ and $x_j^e = 0$ for all $j \in \mathcal{L}$.
- (d) If $\theta < \theta^{(3)}$, then the optimal testing policy is to have zero tests and in the corresponding equilibrium all agents are inactive, i.e., $x_i^e = 0$ for all $i \in \mathcal{V}$.⁵

The most important result in Theorem 2 is that the optimal policy does not necessarily use all available tests. In particular, when there are enough tests that all agents can be fully active with the appropriate testing in isolation, the social planner is (obviously) happy to deploy all testing and allow all agents to be fully active. This is the case when $\theta > \theta^{(1)}$. However, when $\theta^{(2)} < \theta < \theta^{(1)}$, the social planner prefers not to use all available tests. The intuition for this result is related to the non-monotonicity of the comparative statics derived in Theorem 1: greater testing will encourage more social activity, in this case from low-value agents. The social planner, on the other hand, prefers zero activity from low-value agents so as to slow down the spread of the virus. Therefore, she opts for a policy that does not test low-value agents, discouraging their social activity and keeps the social network less dense. When $\theta < \theta^{(3)}$, the optimal policy is even more extreme. It does not test any agents. This is because just testing high-value agents would encourage sufficient social activity to lead to faster spread of the virus, which the social planner prefers to avoid.

One implication of our model, highlighted in Theorem 2, is that testing high-value agents ahead of low-value agents is optimal. In Section 5 we show that this is a consequence of individuals choosing a general social activity level. If, instead, they can target their social activity to low-value and high-value agents separately, then the social planner may want to test low-value agents and refrain from testing high-value agents, in order to discourage everybody from interacting with these “super spreaders” (because they are socially very active).

⁵The thresholds on θ relate to the ones found in Lemma 2 and in particular we have $\theta^{(1)} = r_H \alpha_H^{(1)}$, $\theta^{(2)} = \alpha_H^{(2)} r_H$, and $\theta^{(3)} = \alpha_H^{(3)} r_H$. These thresholds are such that $\frac{\theta^{(2)}}{r_H} \in [0, 1]$ and for $\theta \in [\theta^{(3)}, \theta^{(2)})$, we have $\frac{\theta}{r_H} \in [0, 1]$.

4.1 Optimal Uniform Testing Policy

It may be impossible for authorities to discriminate or identify high-value and low-value agents, in which case testing policy would have to be uniform, $\alpha_L = \alpha_H = \alpha$. With uniform testing policy, the social planner's problem becomes

$$\begin{aligned} \max_{\alpha \in [0,1]} W(\mathbf{x}^e, \alpha_L = \alpha, \alpha_H = \alpha) \\ \text{s.t. } \mathbf{x}^e \text{ is the unique (symmetric) equilibrium,} \\ \alpha \leq \theta. \end{aligned}$$

The following corollary readily follows from Proposition 2 and characterizes the optimal uniform policy.

Corollary 1. *Let $\theta^{(1)} \geq \theta^{(2)} \geq \theta^{(3)}$ be thresholds found in Theorem 2. There exist $\bar{c} \in \mathbb{R}_+$ and $M \in \mathbb{N}_+$ such that for $c \leq \bar{c}$ and $n \geq M$ we have:*

- (a) *If $\theta > \theta^{(1)}$, then the optimal testing policy is to test all individuals with probability θ and in the corresponding equilibrium both agent types have full social activities, i.e., $x_i^e = 1$ for all $i \in \mathcal{V}$.*
- (b) *If $\min\{\frac{\theta^{(2)}}{r_H}, \theta^{(1)}\} < \theta < \theta^{(1)}$, then the optimal testing policy is to test all individuals with probability $\min\{\frac{\theta^{(2)}}{r_H}, \theta^{(1)}\}$ and in the corresponding equilibrium high-value agents have full social activities while low-value agents have zero social activities, i.e., $x_i^e = 1$ for all $i \in \mathcal{H}$ and $x_j^e = 0$ for all $j \in \mathcal{L}$.*
- (c) *If $\min\{\frac{\theta^{(3)}}{r_H}, \theta^{(1)}\} < \theta < \min\{\frac{\theta^{(2)}}{r_H}, \theta^{(1)}\}$, then the optimal testing policy is to test all with probability θ and in the corresponding equilibrium high-value agents have full social activities while low-value agents have zero social activities, i.e., $x_i^e = 1$ for all $i \in \mathcal{H}$ and $x_j^e = 0$ for all $j \in \mathcal{L}$.*
- (d) *If $\theta < \min\{\frac{\theta^{(3)}}{r_H}, \theta^{(1)}\}$, then the optimal testing policy is to have zero tests in the corresponding equilibrium and all agents have zero social activities, i.e., $x_i^e = 0$ for all $i \in \mathcal{V}$.*

Comparing Corollary 1 with Theorem 2, we see that the top region, in which all agents are fully active, does not change. The second region shrinks (because $\frac{\theta^{(2)}}{r_H} > \theta^{(2)}$) and involves greater spread of the virus, because uniform policies are less effective at identifying and isolating the “super spreader” agents that are more likely to be infected and more likely to spread the virus (because of their greater social activity). The third region may expand, but in this region individuals continue to have the same infection probability as targeted policies. Finally, the fourth region expands, and in this region individuals have the same infection probability as under targeted policy. Overall, uniform policies make testing less effective, but do not change our qualitative conclusions.

4.2 Optimal Testing Policy with Mandatory Social Distancing

The non-monotonicity in our comparative statics and the unwillingness of the social planner to always use all testing capacity are related to the fact that greater testing reduces voluntary so-

cial distancing. This naturally suggests that testing should be combined with mandatory social distancing. The next proposition shows that when this is the case, the social planner would always like to use all available testing capacity and would then deploy mandatory social distancing measures to the adverse behavioral effects of testing.

Formally, we suppose the social planner, in addition to the testing policy (α_L, α_H) , can choose \bar{x}_L and \bar{x}_H which indicates the maximum social activity levels that low- and high-value agents can have. We call such a policy, denoted by $(\alpha_L, \alpha_H, \bar{x}_L, \bar{x}_H)$, a *testing policy with mandatory social distancing*. With this notation, the social planner's problem becomes

$$\begin{aligned} \max_{(\alpha_L, \alpha_H, \bar{x}_L, \bar{x}_H) \in [0,1]^4} & W(\mathbf{x}^e, \alpha_L, \alpha_H) \\ \text{s.t. } & \mathbf{x}^e \text{ is the unique (symmetric) equilibrium,} \\ & x_i \leq \bar{x}_H \text{ for } i \in \mathcal{H}, x_j \leq \bar{x}_L \text{ for } j \in \mathcal{L}, \\ & \alpha_H |\mathcal{H}| + \alpha_L |\mathcal{L}| \leq \theta n. \end{aligned}$$

For a given testing capacity θ , we denote the first best by $(\alpha_L^{\text{FB}}(\theta), \alpha_H^{\text{FB}}(\theta), x_l^{\text{FB}}(\theta), x_h^{\text{FB}}(\theta))$ which is the solution of

$$\begin{aligned} \max_{(\alpha_L, \alpha_H, x_l, x_h) \in [0,1]^4} & W((\mathbf{x}_{\mathcal{H}} = x_h \mathbf{1}, \mathbf{x}_{\mathcal{L}} = x_l \mathbf{1}), \alpha_L, \alpha_H) \\ \text{s.t. } & \alpha_H |\mathcal{H}| + \alpha_L |\mathcal{L}| \leq \theta n. \end{aligned}$$

Proposition 3. *For any testing capacity θ , a testing policy with mandatory social distancing with*

$$(\alpha_L, \alpha_H, \bar{x}_L, \bar{x}_H) = (\alpha_L^{\text{FB}}(\theta), \alpha_H^{\text{FB}}(\theta), x_l^{\text{FB}}(\theta), x_h^{\text{FB}}(\theta))$$

achieves the social welfare of the first best. Moreover, with this policy the social planner uses all the testings capacity.

5 Extensions

In this section, we consider two extensions. First, we show that all of our results extend to an environment with multiple types and non-monotonicities become more likely in this case. Second, we allow for social activity levels directed to different types of agents (for example, individuals choosing how much to interact with more active/popular agents and how much to interact with other agents). We show that with such targeted social activity behavior, optimal policy can try to discourage individuals from interacting with high-social value agents what are more likely to spread the virus.

5.1 Multiple Types

In our baseline model, we considered two types of individuals with different values from social activity. Here, we show that our main results and therefore the insights will carry over to a more general setting with m different values for individuals. In particular, we let $v_1 < v_2 < \dots < v_m$ denote the social activity value of different individual types and also let \mathcal{V}_k for $k = 1, \dots, m$ denote the set of individuals of type k (therefore, $\bigcup_{k=1}^m \mathcal{V}_k = \mathcal{V}$). We also denote the testing probability of type k individuals by α_k for $k = 1, \dots, m$ and let $\alpha = (\alpha_1, \dots, \alpha_m) \in [0, 1]^m$ denote the testing policy.

To characterize the equilibrium, we first introduce a few notations. Again, we let $\mathbf{1}$ and $\mathbf{0}$ denote the vectors of all 1 and all 0, respectively where their dimension will be clear from the context. We also define the following sets:

$$\begin{aligned}\mathcal{A}_1 &= \{\alpha \in [0, 1]^m : \mathbb{P}_k^{\text{inf}}(\mathbf{x} = \mathbf{1}, \alpha) \leq v_k + \frac{1}{n}, \text{ for } k = 1, \dots, m\} \\ \mathcal{A}_2 &= \{\alpha \in [0, 1]^m : \mathbb{P}_1^{\text{inf}}(\mathbf{x} = \mathbf{1}, \alpha) \geq v_1 + \frac{1}{n}, \\ &\quad \mathbb{P}_1^{\text{inf}}(\mathbf{x}_{\mathcal{V}_2} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, x_1 = 1, \mathbf{x}_{\mathcal{V}_1 \setminus \{1\}} = \mathbf{0}, \alpha) \leq v_1 + \frac{1}{n}, \\ &\quad \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_2} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \alpha) \leq v_k + \frac{1}{n} \text{ for } k = 2, \dots, m\}\end{aligned}$$

and similarly for $j = 2, \dots, m$ we let

$$\begin{aligned}\mathcal{A}_{2j-1} &= \{\alpha \in [0, 1]^m : \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_j} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \dots, \mathbf{x}_{\mathcal{V}_{j-1}} = \mathbf{0}, x_k = 1, \alpha) \geq v_k + \frac{1}{n}, \\ &\quad \text{for } k = 1, \dots, j-1, \\ &\quad \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_j} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \dots, \mathbf{x}_{\mathcal{V}_{j-1}} = \mathbf{0}, x_k = 1, \alpha) \leq v_k + \frac{1}{n}, \\ &\quad \text{for } k = j, \dots, m, \\ &\quad \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_k} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \dots, \mathbf{x}_{\mathcal{V}_{k-1}} = \mathbf{0}, x_k = 1, \alpha) \geq v_k + \frac{1}{n}, \\ &\quad \text{for } k = 1, \dots, j-1\} \text{ and} \\ \mathcal{A}_{2j} &= \{\alpha \in [0, 1]^m : \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_j} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \dots, \mathbf{x}_{\mathcal{V}_{j-1}} = \mathbf{0}, x_k = 1, \alpha) \geq v_k + \frac{1}{n}, \\ &\quad \text{for } k = 1, \dots, j, \\ &\quad \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_j} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \dots, \mathbf{x}_{\mathcal{V}_{j-1}} = \mathbf{0}, x_k = 1, \alpha) \leq v_k + \frac{1}{n}, \\ &\quad \text{for } k = j, \dots, m, \\ &\quad \mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{V}_k} = \mathbf{1}, \dots, \mathbf{x}_{\mathcal{V}_m} = \mathbf{1}, \mathbf{x}_{\mathcal{V}_1} = \mathbf{0}, \dots, \mathbf{x}_{\mathcal{V}_{k-1}} = \mathbf{0}, x_k = 1, \alpha) \geq v_k + \frac{1}{n}, \\ &\quad \text{for } k = 1, \dots, j-1\}.\end{aligned}$$

The following proposition is the analogue of Proposition 2 and characterizes equilibrium in this

setting.

Proposition 4. *There exists $M \in \mathbb{N}_+$ and functions $\gamma_k : [0, 1]^m \rightarrow [0, 1]$ for $k = 1, \dots, m$ such that for $n \geq M$, depending on the testing policy α there are $2m$ possibilities for the equilibrium. For $k = 1, \dots, m$, we have:*

- *If $\alpha \in \mathcal{A}_{2k-1}$, there exists a unique symmetric equilibrium where $x_i^e = 0$ for all $i \in \bigcup_{j=1}^{k-1} \mathcal{V}_j$ and $x_i^e = 1$ for all $i \in \bigcup_{j=k}^m \mathcal{V}_j$.*
- *If $\alpha \in \mathcal{A}_{2k}$, there exists a unique symmetric equilibrium where $x_i^e = 0$ for all $i \in \bigcup_{j=1}^{k-1} \mathcal{V}_j$, $x_i^e = 1$ for all $i \in \bigcup_{j=k+1}^m \mathcal{V}_j$, and a mixed-strategy μ_k^e for all $i \in \mathcal{V}_k$ that puts probability $\gamma_k(\alpha)$ on 1 and probability $1 - \gamma_k(\alpha)$ on 0.*

This proposition shows that for any k if $\alpha \in \mathcal{A}_{2k-1}$, then individual types $1, \dots, k-1$ have zero social activities and individual types k, \dots, m have full social activity. If $\alpha \in \mathcal{A}_{2k}$, then individual types $1, \dots, k-1$ still have zero social activities and individual types $k+1, \dots, m$ still have full social activity. Individual type k , however, plays according to a mixed-strategy.

Using Proposition 4, we next show how the infection probability of agents in equilibrium changes as the testing policy α varies, which is the analogue of Theorem 1 and establishes the counterfactual non-monotonicity of infection probability in testing policy.

Theorem 3. *There exists $M \in \mathbb{N}_+$ such that for $n \geq M$, in the unique symmetric equilibrium for $k = 1, \dots, m$ we have:*

- *Higher α in the interior of \mathcal{A}_{2k-1} decreases the infection probability of agent types $j = k, \dots, m$ and does not change the infection probability of agent types $j = 1, \dots, k-1$.*
- *Higher α in the interior of \mathcal{A}_{2k} increases the infection probability of agent type k and does not change the infection probability of agent types $j = 1, \dots, k-1, k+1, \dots, m$.*

Moreover, infection probabilities are continuous in α at the boundaries of the above sets.

Similar to Theorem 1, this theorem states that if one of the individuals is playing a mixed-strategy (i.e., for the sets with even indices), increasing the testing probabilities, adversely, increases the infection probability of individuals. Figure 5 depicts the infection probabilities for $m = 3$ for a uniform policy that tests each individual type with probability α , illustrating that the infection probability of all three types feature non-monotonicity as we increase α .

Finally, note that by having more types of individuals there are more regions in which increasing the testing policy, adversely, increases the infection probability of individuals. More precisely, suppose we have a society with $\mathcal{V} = \{1, \dots, n\}$ individuals and m individual types with values

$$v_1 < \dots < v_m.$$

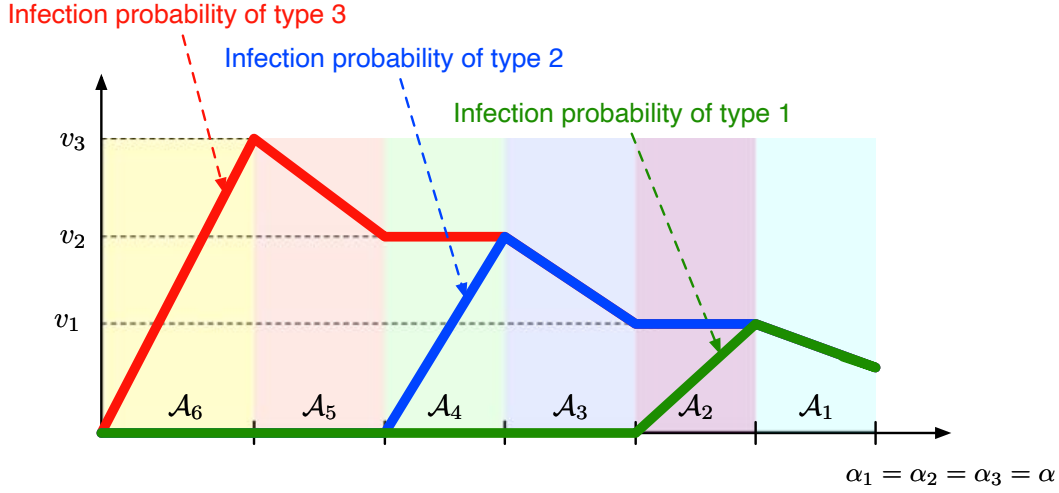


Figure 5: Illustration of Theorem 3 for $m = 3$ types with $v_3 \geq v_2 \geq v_1$. The schematic view of the infection probability of types 1, 2, and 3 agents, in the limit as $n \rightarrow \infty$, for uniform policy (i.e., $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$) as a function of α .

As shown in Theorem 3 in regions \mathcal{A}_{2k} for $k = 1, \dots, m$ infection probability of individuals increase as the testing policy increases. Now suppose we add some individuals with value v_0 which is smaller than v_1 . We let \mathcal{V}_0 denote the set of individuals with this new value and consider a society with $\mathcal{V} \cup \mathcal{V}_0$ many people whose values are

$$v'_1 = v_0 < v'_2 = v_1 < \dots < v'_{m+1} = v_{m+1}.$$

Since individuals with value $v'_1 = v_0$ are the last ones who start having non-zero social activities (because they have the lowest value), for this new society we have

$$\mathcal{A}'_{2k} = \mathcal{A}_{2(k-1)}, \mathcal{A}'_{2k-1} = \mathcal{A}_{2(k-1)-1}, \quad \text{for } k = 2, \dots, m+1,$$

and the infection probability of all individuals with types with values v'_k are equal to the infection probability of type $k-1$ individuals in the smaller society for $k = 2, \dots, m+1$. We will have two new sets: Set \mathcal{A}'_2 in which as we increase the testing policy the infection probability of all previous individual types remains equal to $v'_1 = v_0$ and the infection probability of individuals with value $v'_1 = v_0$ increases from $\frac{1}{n}$ and at the boundary between \mathcal{A}'_2 and \mathcal{A}'_1 , in the limit, becomes $v'_1 = v_0$. Set \mathcal{A}'_1 in which the infection probability of all individuals, in the limit, are equal and decreases as we increase the testing policy. Therefore, in the new society there will be more regions for which increasing the testing policy, counter-factually, increases the infection probability.

5.2 Targeted Social Activities

Here, we show that if agents choose two levels of activities targeted to low-value and high-value agents, then the optimal policy may involve testing low-value agents with a higher probability than high-value agents. In particular, we let the social activity of each agent i be a pair $x_i = (x_i^L, x_i^H)$, where x_i^L denotes agent i 's social activity directed to low-value agents and x_i^H denotes her social activity directed to high-value agents. In the network of contacts, the probability of an edge between agents i and j is:

$$\mathbb{P}[\mathbf{E}_{ij} = 1] = \begin{cases} \eta x_i^H x_j^H, & i, j \in \mathcal{H}, \\ \eta x_i^H x_j^L, & i \in \mathcal{H}, j \in \mathcal{L}, \\ \eta x_i^L x_j^H, & i \in \mathcal{L}, j \in \mathcal{H}, \\ \eta x_i^L x_j^L, & i, j \in \mathcal{L}. \end{cases}$$

The utility of agent i is similar to our baseline model and is given by

$$u_i(x_i, \mathbf{x}_{-i}, \alpha_L, \alpha_H) = v_i(x_i^L + x_i^H) - \mathbb{P}_i^{\text{inf}}(x_i, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c (\alpha_L \mathbf{1}\{i \in \mathcal{L}\} + \alpha_H \mathbf{1}\{i \in \mathcal{H}\}),$$

where $v_i \in \{v_L, v_H\}$, $c > 0$ is a (possibly small) cost of testing for an agent, and the infection probability is determined in the same way as our baseline model.

The next proposition provides conditions under which it is optimal to test only low-value agents.

Proposition 5. *If $r_H > \max\{\frac{1}{1+\beta}, \frac{v_H}{\beta}, 1-v_L\}$ and $\theta < r_H - \frac{v_H}{\beta}$, then there exist \bar{c} and $M \in \mathbb{N}_+$ such that for $n \geq M$ and $c \leq \bar{c}$ the optimal policy is to only test low-value agents and in the corresponding symmetric unique equilibrium all low-value agents play $x_j = (1, 0)$ and all high-value agents play a mixed-strategy between $(1, 0)$ and $(1, 1)$.*

Proposition 5 shows that for sufficiently small testing capacity θ and sufficiently large population fraction of high-value agents r_H , it is optimal to only test low-value agents and in the corresponding equilibrium low-value agents will not interact with high-value agents.

6 Concluding Remarks

This paper studied the effects of testing on social activity and voluntary social distancing in the context of an epidemic. Social activity levels determine the (endogenous) social network over which contacts take place and an infection spreads. Testing enables authorities to identify and isolate infected individuals who spread the virus, and has been identified by the recent literature on COVID-19 and policymakers as a key tool for combating epidemics. Our analysis, however, shows that the impact of testing on the spread of an epidemic may be more complex because, knowing that tests will lead to the isolation of infected individuals, agents can increase their social

activity levels and refrain from voluntary social distancing. As a result, our analysis showed that the effects of testing on the spread of the infection can be non-monotonic—greater testing can lead to greater infection probabilities.

Our analysis also characterized the optimal testing policies. The same forces that lead to non-monotonic comparative statics also imply that a benevolent social planner may prefer to leave her testing capacity partially or fully unused—because increasing testing can make the spread of the virus more likely. This implies that testing should often be combined with mandatory social distancing measures (which ensure that the adverse behavioral effects of testing can be countered by preventing excessively high social activity levels).

Our paper is part of a growing literature on the interaction between economic incentives and epidemiological dynamics. Two conceptual contributions of our approach are to conceptualize the problem of endogenous behavior as one of social network formation and to use the percolation model rather than the SIR dynamic model. Both of these contributions can be useful beyond the confines of our specific question, but the robustness of our conclusions to relaxing both assumptions and adopting different modeling strategies need to be investigated. Other interesting areas for research include the analysis of optimal testing and tracing when tests lead to type I and type II errors and policy is constrained by privacy considerations and non-obedience. Another interesting avenue is to enrich the setup to incorporate more heterogeneity and richer economic, social and epidemiological interactions so as to enable quantitative policy analysis.

A Appendix

This appendix includes the omitted proofs from the text and additional results.

A.1 Proofs

Proof of Lemma 1

First note that for any agent i we have

$$\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) = \sum_{s=1}^n \frac{1}{n} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = s),$$

where $\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = s)$ is the infection probability of agent i conditional on s being the source of infection. We next establish that $\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = s)$ is increasing in \mathbf{x} and concave in x_i . To this end, we generalize this claim and then use induction on the number of nodes to prove it. In particular, we prove that for any set of nodes $S \subseteq \mathcal{V}$ that are infected at time 0, the probability of infection reaching node i is increasing in \mathbf{x} and concave in x_i . We denote this probability by

$$\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S).$$

We also use an auxiliary infection probability in this lemma. In particular for any set of nodes S , we let

$$\tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S)$$

denote the probability of infection reaching node i in one round (i.e., only through the nodes in S).

Claim 1: For any set S and $i \in \mathcal{V}$, the probability $\tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S)$ is increasing in \mathbf{x} and concave in x_i .

Proof of Claim 1: We prove this claim by induction on the size of S . If $|S| = 1$ and contains only node j (i.e., $S = \{j\}$), then this probability is $\eta x_i x_j (\alpha_j \beta p + (1 - \alpha_j) \beta)$ which is increasing in \mathbf{x} and concave in x_i . Now suppose $|S| > 1$ and let $j \in S$. We can write

$$\begin{aligned} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \\ = \eta x_i x_j (\beta(1 - \alpha_j) + \beta p \alpha_j) + (1 - \eta x_i x_j (\beta(1 - \alpha_j) + \beta p \alpha_j)) \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\}). \end{aligned}$$

Taking derivative of this expression results in

$$\begin{aligned} \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \\ = \eta x_j (\beta(1 - \alpha_j) + \beta p \alpha_j) - \eta x_j (\beta(1 - \alpha_j) + \beta p \alpha_j) \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\}) \\ + (1 - \eta x_i x_j (\beta(1 - \alpha_j) + \beta p \alpha_j)) \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\}) \\ = \eta x_j (\beta(1 - \alpha_j) + \beta p \alpha_j) \left(1 - \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\})\right) \\ + (1 - \eta x_i x_j (\beta(1 - \alpha_j) + \beta p \alpha_j)) \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\}) \stackrel{(a)}{\geq} 0, \end{aligned}$$

where (a) follows from the induction hypothesis. This establishes monotonicity in x_i . Similarly, we can establish monotonicity in x_j . Finally, note that monotonicity in $\mathbf{x}_{-\{i,j\}}$ follows by induction hypothesis which says $\tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\})$ is increasing in $\mathbf{x}_{-\{i,j\}}$.

We next prove concavity in x_i . Taking a second order derivative of the above expression results in

$$\begin{aligned} \frac{\partial^2}{\partial^2 x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \\ = -\eta x_j (\beta(1 - \alpha_j) + \beta p \alpha_j) \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\}) \\ - \eta x_j (\beta(1 - \alpha_j) + \beta p \alpha_j) \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_H | \text{source} = S \setminus \{j\}) \\ + (1 - \eta x_i x_j (\beta(1 - \alpha_j) + \beta p \alpha_j)) \frac{\partial^2}{\partial^2 x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\}) \stackrel{(a)}{\leq} 0, \end{aligned}$$

where (a) follows from the monotonicity of $\tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S \setminus \{j\})$ and the induction hypothesis. ■

We now proceed with the proof of Lemma 1 by induction on the number of nodes in the network. The lemma evidently holds for $n = 2$. We can write

$$\begin{aligned} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) &= \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) + \left(1 - \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S)\right) \\ &\quad \times \sum_{S^{\text{new}} \subseteq \mathcal{V} \setminus S} \mathbb{P}(S^{\text{new}} | S) \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}). \end{aligned}$$

Taking derivative of this expression leads to

$$\begin{aligned} \frac{\partial}{\partial x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) &= \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \\ &\quad - \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}) \\ &\quad + \left(1 - \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S)\right) \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \frac{\partial}{\partial x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}) \\ &= \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \left(1 - \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}})\right) \\ &\quad + \left(1 - \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S)\right) \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \frac{\partial}{\partial x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}) \stackrel{(a)}{\geq} 0 \end{aligned}$$

where (a) follows from Claim 1 and the induction hypothesis. This establishes monotonicity in x_i . Similarly, the derivative with respect to x_j for $j \in S$ is non-negative, showing monotonicity in x_j for $j \in S$. Finally, the monotonicity in $\mathbf{x}_{-(S \cup \{i\})}$ follows by induction hypothesis.

We next take the second order derivative and show it is non-positive. We can write

$$\begin{aligned} \frac{\partial^2}{\partial^2 x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) &= \frac{\partial^2}{\partial^2 x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \left(1 - \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}})\right) \\ &\quad - \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \frac{\partial}{\partial x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}) \\ &\quad - \frac{\partial}{\partial x_i} \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S) \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \frac{\partial}{\partial x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}) \\ &\quad + \left(1 - \tilde{\mathbb{P}}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S)\right) \sum_{S^{\text{new}}} \mathbb{P}(S^{\text{new}} | S) \frac{\partial^2}{\partial^2 x_i} \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H | \text{source} = S^{\text{new}}) \stackrel{(a)}{\leq} 0, \end{aligned}$$

where (a) follows from Claim 1 and the induction hypothesis. This completes the proof ■

Proof of Proposition 1

We let $\{U_{i,t}\}_{i \in \{1, \dots, n\}, t \geq 0}$ be a sequence of i.i.d. random variables with uniform distribution over $[0, 1]$. We use a coupling argument together with stochastic dominance to relate the stochastic processes unfolding with (α_L, α_H) to the stochastic process unfolding with (α'_L, α'_H) . Consider the process with (α_L, α_H) and suppose at round t agent i is infected and consider the testing probability of this agent. The probability of agent i getting tested is α_i (where $i \in \{L, H\}$) and the probability of agent i not getting tested is $1 - \alpha_i$. Alternatively, if we let $Z_{i,t} \in \{0, 1\}$ denote whether agents i gets tested at time t or not then we can write this random variable by using the sequence of sequence of i.i.d. random variables $\{U_{i,t}\}_{i \in \{1, \dots, n\}, t \geq 0}$ as

$$Z_{i,t} = \begin{cases} 1, & \text{if } U_{i,t} \leq \alpha_i \\ 0, & \text{if } U_{i,t} > \alpha_i. \end{cases}$$

We now define a coupled independent cascade process by using the same draws of the sequence of i.i.d. random variables $\{U_{i,t}\}_{i \in \{1, \dots, n\}, t \geq 0}$ as

$$Z'_{i,t} = \begin{cases} 1, & \text{if } U_{i,t} \leq \alpha'_i \\ 0, & \text{if } U_{i,t} > \alpha'_i. \end{cases}$$

Given $\alpha'_i \geq \alpha_i$ for $i = 1, 2$ we have $\mathbb{P}(Z'_{i,t} \geq Z_{i,t}, \text{ for all } i = 1, \dots, n, t \geq 0) = 1$. We also let $Y_{i,j,t}$ be the probability of infection reaching from agent i to agent j which can be formulated by using a sequence of i.i.d. random variables $\{U_{i,j,t}\}_{i,j \in \{1, \dots, n\}, t \geq 0}$ as

$$Y_{i,j,t} = \begin{cases} 1, & \text{if } U_{i,j,t} \leq \beta p, Z_{i,t} = 1 \\ 0, & \text{if } U_{i,j,t} > \beta p, Z_{i,t} = 1 \\ 1, & \text{if } U_{i,j,t} \leq \beta, Z_{i,t} = 0 \\ 0, & \text{if } U_{i,j,t} > \beta, Z_{i,t} = 0. \end{cases}$$

We again define a coupled independent cascade process by using the same draws of the sequence of sequence of i.i.d. random variables $\{U_{i,j,t}\}_{i,j \in \{1, \dots, n\}, t \geq 0}$ as

$$Y'_{i,j,t} = \begin{cases} 1, & \text{if } U_{i,j,t} \leq \beta p, Z'_{i,t} = 1 \\ 0, & \text{if } U_{i,j,t} > \beta p, Z'_{i,t} = 1 \\ 1, & \text{if } U_{i,j,t} \leq \beta, Z'_{i,t} = 0 \\ 0, & \text{if } U_{i,j,t} > \beta, Z'_{i,t} = 0. \end{cases}$$

Given $\mathbb{P}(Z'_{i,t} \geq Z_{i,t}, \text{ for all } i = 1, \dots, n, t \geq 0) = 1$, we obtain $\mathbb{P}(Y'_{i,j,t} \leq Y_{i,j,t}, \text{ for all } i, j = 1, \dots, n, t \geq 0) = 1$. Note that node i gets infected if and only if $\sum_{t \geq 0} \sum_{j \neq i} Y_{i,j,t} \geq 1$. Therefore, the

probability of node i getting infected with the original process is smaller than the auxiliary process. Finally, note that the probability of getting infected in the auxiliary process is $\mathbb{P}_i^{\text{pinf}}(G, \alpha'_L, \alpha'_H)$, completing the proof. ■

Proof of Proposition 2

We first prove that the best response decision of each individual is either 1 or 0 (i.e., the best response social activity of each agent is at the boundaries of $[0, 1]$).

Lemma 4. *Given an individual $i \in \mathcal{V}$ and action profile \mathbf{x}_{-i} , let x_i^e denote the best response social activity of individual i , i.e.,*

$$x_i^e \in \arg \max_{x \in [0,1]} u_i(x, \mathbf{x}_{-i}, \alpha_L, \alpha_H).$$

Then, we have $x_i^e \in \{0, 1\}$.

Proof of Lemma 4: The utility of agent i is $v_i x_i - \mathbb{P}_i^{\text{pinf}}(x_i, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c\alpha_i$. Using Part (b) of Lemma 1 this utility function is convex in x_i and therefore its maximum is either at 0 or at 1. ■ Lemma 4 also shows that in any mixed-strategy equilibrium each agent i must be mixing between playing 1 and playing 0. In what follows we let $x = \gamma$ for some $\gamma \in [0, 1]$ to denote a mixed strategy that puts probability γ on 1 and puts probability $1 - \gamma$ on 0. We also let α_i denote the testing probability of agent i which is equal to α_L if $i \in \mathcal{L}$ and is equal to α_H if $i \in \mathcal{H}$.

We now proceed with the proof of proposition.

Proof of part (a): We first establish that $x_i = 1$ for all $i \in \mathcal{V}$ is an equilibrium and then show that it is the unique symmetric equilibrium. Consider $i \in \mathcal{H}$. The utility of i with action profile $\mathbf{x} = \mathbf{1}$ is

$$v_H - \mathbb{P}_i^{\text{pinf}}(x_i = 1, x_{-i} = 1, \alpha_H, \alpha_L) - c\alpha_H.$$

If agent i deviates and plays $x_i = 0$ (using Lemma 4, this is the only candidate for a profitable deviation), then her utility becomes

$$-\mathbb{P}_i^{\text{pinf}}(x_i = 0, x_{-i} = 1, \alpha_H, \alpha_L) - c\alpha_H = \frac{-1}{n} - c\alpha_H,$$

where we used the fact that if $x_i = 0$, then the only way for agent i to get infected is be the source of infection. Therefore, given $(\alpha_H, \alpha_L) \in \mathcal{A}_1$ we have $u_i(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) > u_i(x_i = 0, \mathbf{x}_{-i}, \alpha_L, \alpha_H)$. Similarly, any $j \in \mathcal{L}$ does not have a profitable deviation.

We next prove that this is the unique symmetric equilibrium. First note that for a high-value agent $i \in \mathcal{H}$ no matter what the strategy of other agents are, the dominant strategy is to play $x_i = 1$. This is because for any \mathbf{x}_{-i} we can write

$$\begin{aligned} & u_i(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) \\ &= v_H - \mathbb{P}_i^{\text{pinf}}(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c\alpha_H \end{aligned}$$

$$\begin{aligned}
&\stackrel{(a)}{\geq} v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{-i} = \mathbf{1}, \alpha_L, \alpha_H) - c\alpha_H \\
&\stackrel{(b)}{>} -\frac{1}{n} - c\alpha_H = u_i(x_i = 0, \mathbf{x}_{-i}, \alpha_L, \alpha_H)
\end{aligned}$$

where (a) follows from Part (a) of Lemma 1 and (b) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_1$ and $v_H > v_L$. Therefore, there are two other candidates for a symmetric equilibrium: (i) the action profile $x_i = 1$ for all $i \in \mathcal{H}$ and $x_j = 0$ for all $j \in \mathcal{L}$ and (ii) $x_i = 1$ for all $i \in \mathcal{H}$ and $x_j = \gamma_L$ for all $j \in \mathcal{L}$ and some $\gamma_L \in (0, 1)$. We next show that none of these can be an equilibrium because an agent $j \in \mathcal{L}$ has a profitable deviation to 1:

$$\begin{aligned}
&u_j(x_j = 1, \mathbf{x}_{L \setminus \{j\}}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L) \\
&= v_L - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{L \setminus \{j\}}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L) - c\alpha_L \\
&\stackrel{(a)}{>} v_L - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{L \setminus \{j\}} = \mathbf{1}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L) - c\alpha_L \\
&\stackrel{(b)}{\geq} -\frac{1}{n} - c\alpha_L = u_j(x_j = 0, \mathbf{x}_{L \setminus \{j\}} = \mathbf{0}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L),
\end{aligned}$$

where (a) follows from Part (a) of Lemma 1 and the fact that we have either $\mathbf{x}_{L \setminus \{j\}} = \mathbf{1}$ or $\mathbf{x}_{L \setminus \{j\}} = \gamma_L \mathbf{1}$ which are both below $\mathbf{x}_{L \setminus \{j\}} = \mathbf{1}$ and (b) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_1$. This completes the proof of part (a).

Proof of part (b): Before proceeding with the proof of this part, note that as we decrease (α_L, α_H) , one of the constraints of \mathcal{A}_1 will be violated because the infection probability increases (by using Proposition 1). For any $\epsilon > 0$, there exists $M \in \mathbb{N}_+$ such that for $n \geq M$, we have

$$|\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_H, \alpha_L) - \mathbb{P}_h^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_H, \alpha_L)| < \epsilon,$$

In what follows we let $\epsilon < \frac{v_H - v_L}{2}$. Therefore, the constraint $\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_H, \alpha_L) \leq v_L + \frac{1}{n}$ will be violated first, resulting in (α_L, α_H) that belongs to the set \mathcal{A}_2 .

We now proceed with the proof of part (b). Consider any symmetric mixed strategy for agents $i \in \mathcal{H}$. Using Lemma 4, this mixed strategy must have only two atoms $\{0, 1\}$. We let γ_H (and similarly γ_L) denote the probability of being 1 for high-value agents (and similarly for low-value agents). With the abuse of notation whenever we write $x_i = \gamma_H$ this means expectation over x_i which is 1 with probability γ_H and 0 with probability $1 - \gamma_H$.

We define $\gamma_L : [0, 1]^2 \rightarrow [0, 1]$ such that for any pair (α_L, α_H) , we have

$$\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \gamma_L(\alpha_L, \alpha_H), \alpha_L, \alpha_H) = \frac{1}{n} + v_L.$$

Note that for any $(\alpha_L, \alpha_H) \in \mathcal{A}_2$ there exists $\gamma_L(\alpha_L, \alpha_H)$ in $[0, 1]$ that satisfies the above equality. This is because by using Part (a) of Lemma 1 the function $f : [0, 1] \rightarrow [0, 1]$ where

$$f(y) = \mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = y\mathbf{1}, \alpha_L, \alpha_H)$$

is increasing in y . For $y = 0$, we have

$$f(0) = \mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) \stackrel{(a)}{\leq} \frac{1}{n} + v_L$$

where (a) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_2$. For $y = 1$ we have

$$f(1) = \mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}, \alpha_L, \alpha_H) \stackrel{(a)}{\geq} \frac{1}{n} + v_L,$$

where (a) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_2$. Using, mean-value Theorem shows that $\gamma_L(\alpha_L, \alpha_H) \in [0, 1]$ exists.

We first prove that $x_i = 1$ for all $i \in \mathcal{H}$ and $x_j = \gamma_L(\alpha_L, \alpha_H)$ for all $j \in \mathcal{L}$ is an equilibrium. For $i \in \mathcal{H}$, we have

$$\begin{aligned} u_i(x_i = 1, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \alpha_H, \alpha_L) \\ &= v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \alpha_H, \alpha_L) - c\alpha_H \\ &\stackrel{(a)}{\geq} v_H - \mathbb{P}_i^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \alpha_H, \alpha_L) - \epsilon - c\alpha_H \\ &\stackrel{(b)}{=} v_H - (v_L + \frac{1}{n}) - \epsilon - c\alpha_H \stackrel{(c)}{\geq} -\frac{1}{n} - c\alpha_H \\ &= u_i(x_i = 0, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \alpha_H, \alpha_L), \end{aligned}$$

where (a) follows from $n \geq M$, (b) follows from the definition of $\gamma_L(\alpha_L, \alpha_H)$, and (c) follows from $\epsilon < \frac{v_H - v_L}{2}$. This shows that high-value agents do not have a profitable deviation. For $j \in \mathcal{L}$, we have

$$\begin{aligned} u_j(x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L) \\ &= v_L - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L) - c\alpha_L \\ &\stackrel{(a)}{=} -\frac{1}{n} - c\alpha_L = u_j(x_j = 0, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_H, \alpha_L), \end{aligned}$$

where (a) follows from the definition of $\gamma_L(\alpha_L, \alpha_H)$. This proves that each low-value agent is indifferent between playing 0 and 1 given the action profile of others in this equilibrium.

We next prove that there exists no other symmetric equilibrium by listing all possibilities:

1. $(x_l = 0, x_h = 0)$: This is not an equilibrium because any agent i can deviate and receive $1 - \frac{1}{n} - c\alpha_i$ instead of $-\frac{1}{n} - c\alpha_i$.
2. $(x_l = 0, x_h = 1)$: This is not an equilibrium because any agent $j \in \mathcal{L}$ has a profitable deviation:

$$\begin{aligned} u_j(x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{0}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H) \\ &= v_L - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{0}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H) - c\alpha_L \end{aligned}$$

$$\stackrel{(a)}{\geq} \frac{-1}{n} - c\alpha_L = u_j(x_j = 0, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{0}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H),$$

where (a) follows from $(\alpha_H, \alpha_L) \in \mathcal{A}_2$.

3. $(x_l = 0, x_h = \gamma_H)$: This is not an equilibrium because any agent $j \in \mathcal{L}$ can deviate with a similar reason to case 2.
4. $(x_l = 1, x_h = 0)$: This is not an equilibrium because any agent $i \in \mathcal{H}$ has a profitable deviation for a similar reason as :

$$\begin{aligned} & u_i(x_i = 1, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{0}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) \\ &= v_H - \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{0}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) - c\alpha_H \\ &\stackrel{(a)}{\geq} v_H - \mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{H}} = \mathbf{0}, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}, \alpha_L, \alpha_H) - \epsilon - c\alpha_H \\ &\stackrel{(b)}{\geq} v_H - \epsilon - v_L - \frac{1}{n} - c\alpha_H \\ &\stackrel{(c)}{\geq} \frac{-1}{n} - c\alpha_H = u_i(x_i = 0, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{0}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H), \end{aligned}$$

where (a) follows from n being large, (b) follows from $x_l = 1$ being equilibrium, and (c) follows from $v_H > v_L$ and $\epsilon < v_H - v_L$.

5. $(x_l = 1, x_h = 1)$: This is not an equilibrium because any agent $j \in \mathcal{L}$ has a profitable deviation:

$$\begin{aligned} & u_j(x_j = 0, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{1}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H) \\ &= \frac{-1}{n} - c\alpha_L \\ &\stackrel{(a)}{\geq} v_L - \mathbb{P}_l^{\text{inf}}(x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{1}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H) - c\alpha_L \\ &= u_j(x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{1}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H), \end{aligned}$$

where (a) follows from $(\alpha_H, \alpha_L) \in \mathcal{A}_2$.

6. $(x_l = 1, x_h = \gamma_H)$: This is not an equilibrium because any agent $j \in \mathcal{L}$ has a profitable deviation for a similar reason as case 5.

Proof of part (c): By further decreasing the pair (α_L, α_H) , using Proposition [1](#) the infection probabilities increase and therefore one of the following constraints

$$\begin{aligned} \mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) &\leq v_L + \frac{1}{n}, \\ \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) &\leq v_H + \frac{1}{n} \end{aligned}$$

will be violated first. For $n \geq M$, the first constraint that will be violated is $\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) \leq v_L + \frac{1}{n}$, resulting in the set \mathcal{A}_3 .

In region \mathcal{A}_3 , we next list all the candidate symmetric equilibria and conclude that $x_i = 1$ for $i \in \mathcal{H}$ and $x_j = 0$ for all $j \in \mathcal{L}$ is the only symmetric equilibrium.

1. $(x_l = \gamma_L, x_h = 1)$: This is not equilibrium because if this was equilibrium we would have been in region \mathcal{A}_2 more precisely this is not equilibrium because $\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}$.
2. $(x_l = \gamma_L, x_h = \gamma_H)$: This cannot be equilibrium for large enough n . This is because if it is equilibrium then we must have

$$\mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}\gamma_L, \mathbf{x}_{\mathcal{H}} = \mathbf{1}\gamma_H, \alpha_L, \alpha_H) = v_L + \frac{1}{n}$$

and

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}\gamma_H, \mathbf{x}_{\mathcal{L}} = \mathbf{1}\gamma_L, \alpha_L, \alpha_H) = v_H + \frac{1}{n}.$$

for large enough n the difference between the left-hand of the above equations becomes smaller than $v_H - v_L$ which is a contradiction.

3. $(x_l = \gamma_L, x_h = 0)$: This cannot be equilibrium for large enough n . This is because if it is equilibrium then we must have

$$\mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}\gamma_L, \mathbf{x}_{\mathcal{H}} = \mathbf{0}, \alpha_L, \alpha_H) = v_L + \frac{1}{n}$$

and

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{0}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}\gamma_L, \alpha_L, \alpha_H) \geq v_H + \frac{1}{n}.$$

for large enough n the difference between the left-hand of the above equations becomes smaller than $v_H - v_L$ which again is a contradiction.

4. $(x_l = 0, x_h = 1)$: This is a symmetric equilibrium because a low-value agent has no profitable deviation:

$$\begin{aligned} & u_l(x_l = 0, \mathbf{x}_{\mathcal{L} \setminus \{l\}}, \mathbf{x}_{\mathcal{H}}, \alpha_L, \alpha_H) \\ &= -\frac{1}{n} - c\alpha_L \stackrel{(a)}{\geq} v_L - \mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}}, \mathbf{x}_{\mathcal{H}}, \alpha_L, \alpha_H) - c\alpha_L \\ &= u_l(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}}, \mathbf{x}_{\mathcal{H}}, \alpha_L, \alpha_H), \end{aligned}$$

where (a) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_3$. Also, a high-value agent has no profitable deviation:

$$\begin{aligned} & u_h(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \\ &= v_H - \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) - c\alpha_H \\ &\stackrel{(a)}{\geq} \frac{-1}{n} - c\alpha_H = u_h(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H), \end{aligned}$$

where (a) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_3$

5. $(x_l = 0, x_h = \gamma_H)$: This cannot be equilibrium for large enough n . This is because if it is equilibrium then we must have

$$\mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{0}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}\gamma_H, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}$$

and

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}\gamma_H, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) = v_H + \frac{1}{n}.$$

for large enough n the difference between the left-hand of the above equations becomes smaller than $v_H - v_L$ which again is a contradiction.

6. $(x_l = 0, x_h = 0)$: This cannot be equilibrium because any agent i can deviate and increase its utility from $\frac{-1}{n} - c\alpha_i$ to $v_i - \frac{1}{n} - c\alpha_i$.
7. $(x_l = 1, x_h = 1)$: This cannot be equilibrium because if it was equilibrium we would have been in region \mathcal{A}_1 .
8. $(x_l = 1, x_h = \gamma_H)$: This cannot be equilibrium because if it is equilibrium we must have

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}\gamma_H, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) = \frac{1}{n} + v_H$$

and

$$\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}\gamma_H, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) \leq \frac{1}{n} + v_L.$$

For large enough n , the difference between the left-hand side of the above inequalities is below ϵ which is a contradiction.

9. $(x_l = 1, x_h = 0)$: This cannot be equilibrium because if it is equilibrium we must have

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}\gamma_H, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) \geq \frac{1}{n} + v_H$$

and

$$\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}_{\gamma_H}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) \leq \frac{1}{n} + v_L.$$

Again, for large enough n , the difference between the left-hand side of the above inequalities is below ϵ which is a contradiction.

Proof of part (d): By further decreasing the pair (α_L, α_H) , the infection probabilities increase and therefore the constraint $\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \leq v_H + \frac{1}{n}$ will be violated and we get to region \mathcal{A}_4 .

We define $\gamma_H : [0, 1]^2 \rightarrow [0, 1]$ such that for any pair (α_L, α_H) , we have

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \gamma_H(\alpha_L, \alpha_H) \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) = \frac{1}{n} + v_H.$$

Note that for any $(\alpha_L, \alpha_H) \in \mathcal{A}_4$ there exists $\gamma_H(\alpha_L, \alpha_H)$ in $[0, 1]$ that satisfies the above equality. This is because by using Part (a) of Lemma 1 the function $f : [0, 1] \rightarrow [0, 1]$ where $f(y) = \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = y \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H)$ is increasing in y . For $y = 0$, we have

$$f(0) = \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{0}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) = \frac{1}{n} \stackrel{(a)}{\leq} \frac{1}{n} + v_H.$$

For $y = 1$ we have

$$f(1) = \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \stackrel{(a)}{\geq} \frac{1}{n} + v_H,$$

where (a) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_4$. Using, mean-value Theorem shows that $\gamma_H(\alpha_L, \alpha_H) \in [0, 1]$ exist.

We first establish that $x_j = 0$ for all $j \in \mathcal{L}$ and $x_i = \gamma_H(\alpha_L, \alpha_H)$ for all $i \in \mathcal{H}$ is an equilibrium. For $j \in \mathcal{L}$, we have

$$\begin{aligned} & u_j(x_j = 1, \mathbf{x}_{\mathcal{H}} = \mathbf{1}_{\gamma_H(\alpha_L, \alpha_H)}, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{0}, \alpha_H, \alpha_L) \\ &= v_L - \mathbb{P}_j^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}_{\gamma_H(\alpha_L, \alpha_H)}, x_j = 1, \mathbf{x}_{\mathcal{L} \setminus \{j\}} = \mathbf{0}, \alpha_H, \alpha_L) - c\alpha_L \\ &\stackrel{(a)}{\leq} v_L - \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}_{\gamma_H(\alpha_L, \alpha_H)}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_H, \alpha_L) + \epsilon - c\alpha_L \\ &\stackrel{(b)}{=} v_L - \frac{1}{n} - v_H + \epsilon - c\alpha_L \\ &\stackrel{(c)}{\leq} \frac{-1}{n} - c\alpha_L = u_j(x_j = 0, \mathbf{x}_{\mathcal{H}} = \mathbf{1}_{\gamma_H(\alpha_L, \alpha_H)}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_H, \alpha_L), \end{aligned}$$

where (a) follows from $n \geq M$, (b) follows from the definition of $\gamma_H(\alpha_L, \alpha_H)$, and (c) follows from $\epsilon < \frac{v_H - v_L}{2}$. This shows that low-value agents do not have a profitable deviation. For $i \in \mathcal{H}$, we

have

$$\begin{aligned}
u_i(x_i = 1, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1} \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_H, \alpha_L) \\
= v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1} \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_H, \alpha_L) - c\alpha_H \\
\stackrel{(a)}{=} \frac{-1}{n} - c\alpha_H = u_i(x_i = 0, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1} \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_H, \alpha_L),
\end{aligned}$$

showing that high-value agents are indifferent between playing 1 and 0 with this activity profile, where (a) follows from the definition of $\gamma_H(\alpha_L, \alpha_H)$. Similar to the proof of previous parts, listing all possible symmetric equilibria shows this is the unique symmetric equilibrium. ■

Proof of Lemma 2

We first define a few notations that we use in the proof. Consider a society of n individuals and suppose that all agents are fully active, i.e., $\mathbf{x} = \mathbf{1}$. For any (α_L, α_H) we let

$$Q(n, \alpha_L, \alpha_H) = \mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H)$$

be the infection probability of any agent when the testing probability of high-value agents is α_H and the testing probability of low-value agents is α_L . In the next lemma we establish that for large enough n the quantities $Q(n, \alpha_L, \alpha_H)$ and $Q(n, r_H\alpha_H + r_L\alpha_L, r_H\alpha_H + r_L\alpha_L)$ are close to each other:

Claim 1: We have

$$\lim_{n \rightarrow \infty} |Q(n, \alpha_L, \alpha_H) - Q(n, r_H\alpha_H + r_L\alpha_L, r_H\alpha_H + r_L\alpha_L)| = 0.$$

Proof of Claim 1: We define another quantity $\tilde{Q}(n, \alpha)$ that denotes the infection probability of any agent in a society of n agents where $n\alpha$ agents are tested (we drop the ceiling and assume this number is integer, to simplify the notation). To prove Claim 1, it suffices to show

$$\lim_{n \rightarrow \infty} |Q(n, \alpha_L, \alpha_H) - \tilde{Q}(n, \alpha = r_L\alpha_L + r_H\alpha_H)| = 0, \quad (3)$$

showing that only the expected number of tests determines the limit of the infection probability. We next establish (3). We let S_n denote the number of tested agents. Note that the distribution of S_n is Binomial(n, α). We also let $f(k, n)$ denote the infection probability when k agents are being tested and the rest of them are not tested. With these notations, the statement of (3) becomes

$$\lim_{n \rightarrow \infty} |f(\alpha n, n) - \mathbb{E}_{S_n \sim \text{Binomial}(n, \alpha)} [f(S_n, n)]| = 0.$$

The above inequality holds because S_n concentrates around αn . We next formally prove this. We can write

$$\mathbb{E}_{S_n \sim \text{Binomial}(n, \alpha)} [f(S_n, n)] \stackrel{(a)}{\geq} (1 - e^{-2n\epsilon^2}) f((\alpha + \epsilon)n, n)$$

$$\stackrel{(b)}{\geq} (1 - e^{-2n\epsilon^2}) \left(f(\alpha n, n) - \epsilon - \frac{\epsilon}{1 - \alpha} \right) \quad (4)$$

where (a) follows from Chernoff-Hoeffding inequality $\mathbb{P}(S_n \leq (\alpha + \epsilon)n) \geq 1 - e^{-2n\epsilon^2}$ and (b) follows from the following argument. We let $S = \{1, \dots, \alpha n\}$ be the set of nodes that are being tested in finding $f(n\alpha, n)$ and $\tilde{S} = \{1, \dots, (\alpha + \epsilon)n\}$ be the set of nodes that are being tested in finding $f(n(\alpha + \epsilon), n)$. With probability ϵ , the source of infection is in set $\tilde{S} \setminus S$. In this case we bound $f(\alpha n, n)$ by ϵ . With probability $1 - \epsilon$ the source is not in the set $\tilde{S} \setminus S$. Without loss of generality we assume it is agent 0 and find the relationship between $f(n\alpha, n)$ and $f(n(\alpha + \epsilon), n)$ given the event that the source is agent 0. We let $h(k, n)$ be the infection probability of agent n when we do not test agents $n, \dots, n - k$ and test the rest of the agents. Note that $h(k, n)$ is increasing in k . With this new notation, we have

$$\begin{aligned} f(n(\alpha + \epsilon), n) &= h((1 - (\alpha + \epsilon))n, n), \\ f(n\alpha, n) &= h(n(1 - \alpha), n). \end{aligned}$$

We can write a telescopic summations

$$h(n(1 - \alpha), n) = h(0, n) + \sum_{k=1}^{n(1-\alpha)} (h(k, n) - h(k - 1, n)), \quad (5)$$

and

$$h(n(1 - (\alpha + \epsilon)), n) = h(0, n) + \sum_{k=1}^{n(1-(\alpha+\epsilon))} (h(k, n) - h(k - 1, n)), \quad (6)$$

In Claim 2, proved next, we establish that $(h(k, n) - h(k - 1, n))$ is decreasing in k . Together with (5) and (6), this implies that

$$\begin{aligned} f(n\alpha, n) &= h(n(1 - \alpha), n) = h(n(1 - (\alpha + \epsilon)), n) + \sum_{k=n(1-(\alpha+\epsilon))+1}^{n(1-\alpha)} (h(k, n) - h(k - 1, n)) \\ &\stackrel{(a)}{\leq} h(n(1 - (\alpha + \epsilon)), n) + \frac{\epsilon}{1 - \alpha} = f(n(\alpha + \epsilon), n) + \frac{\epsilon}{1 - \alpha}, \end{aligned}$$

where (a) follows from the fact that $h(n(1 - \alpha), n)$ given in (5) is less than or equal to 1 and $(h(k, n) - h(k - 1, n))$ is decreasing in k . Therefore, combining the two events for the source of infection, we obtain

$$f(\alpha n, n) \leq \epsilon + f((\alpha + \epsilon)n, n) + \frac{\epsilon}{1 - \alpha},$$

which is used in part (b) of (4).

Similarly, we get

$$\mathbb{E}_{S_n \sim \text{Binomial}(n, \alpha)} [f(S_n, n)] \leq (1 - e^{-2n\epsilon^2}) f((\alpha - \epsilon)n, n)$$

$$\leq (1 - e^{-2n\epsilon^2}) \left(f(\alpha n, n) + \epsilon + \frac{\epsilon}{1 - \alpha} \right). \quad (7)$$

Combining the inequalities (4) and (7) for $\epsilon = \frac{\log n}{\sqrt{n}}$ shows that

$$\lim_{n \rightarrow \infty} |f(\alpha n, n) - \mathbb{E}_{S_n \sim \text{Binomial}(n, \alpha)} [f(S_n, n)]| = 0,$$

completing the proof of Claim 1. ■

Claim 2: For the function $h(k, n)$ defined above, we have

$$h(k + 1, n) - h(k, n) \leq (h(k, n) - h(k - 1, n)), \quad \text{for all } k, n. \quad (8)$$

Proof of Claim 2: Let us fix whether we test all agents in $\mathcal{V} \setminus \{i, j\}$. This claim is equivalent to

$$\begin{aligned} & \mathbb{P}(0 \rightsquigarrow n | i \text{ not tested, } j \text{ not tested}) - \mathbb{P}(0 \rightsquigarrow n | i \text{ tested, } j \text{ not tested}) \\ & \geq \mathbb{P}(0 \rightsquigarrow n | i \text{ not tested, } j \text{ tested}) - \mathbb{P}(0 \rightsquigarrow n | i \text{ tested, } j \text{ tested}), \end{aligned}$$

where $s \rightsquigarrow n$ denotes the event that infection reaches agent n starting from agent 0. The difference between the left-hand side and right-hand side of the above inequality is non-zero only when the infection reaches node n through a path that uses agents i and j which we denote by $0 \rightsquigarrow i \rightsquigarrow j \rightsquigarrow n$. We also denote the probability of this event excluding the transmissions from i and j , by $\mathbb{P}(0 \rightsquigarrow i \rightsquigarrow j \rightsquigarrow n)$. In this case, the left-hand side becomes $\mathbb{P}(0 \rightsquigarrow i \rightsquigarrow j \rightsquigarrow n) (\beta - \beta p)$, which is larger than the right-hand side that is $\mathbb{P}(0 \rightsquigarrow i \rightsquigarrow j \rightsquigarrow n) (\beta - \beta p)(\beta - \beta p)$. This completes the proof of Claim 2. ■

We now proceed with the proof of Lemma 2. We recall that set \mathcal{A}_1 must satisfy

$$\begin{aligned} \mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) &\leq v_L + \frac{1}{n}, \\ \mathbb{P}_h^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) &\leq v_H + \frac{1}{n}. \end{aligned}$$

For large enough n the difference between the left-hand side of the above inequalities become smaller than $v_H - v_L$, showing that the constraints of the set \mathcal{A}_1 will be satisfied if and only if we have

$$\mathbb{P}_l^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{1}, \alpha_L, \alpha_H) \leq v_L + \frac{1}{n}.$$

Using Claim 1, as $n \rightarrow \infty$, the above inequality becomes equivalent to

$$Q(\alpha_H r_H + \alpha_L r_L, \alpha_H r_H + \alpha_L r_L) \leq v_L,$$

where $Q(\alpha) = \lim_{n \rightarrow \infty} Q(n, \alpha)$. Noting that $Q(\alpha)$ is decreasing in α (which follows from Proposition 2) establishes the existence of a threshold which we denote by $r_H \alpha^{(1)}$ such that $(\alpha_L, \alpha_H) \in \mathcal{A}_1^*$

if and only if we have $\alpha_H r_H + \alpha_L r_L \geq r_H \alpha_H^{(1)}$ ⁶ The constraint $\mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{1}, \alpha_L, \alpha_H) \geq v_L + \frac{1}{n}$ in the other sets also becomes $\alpha_H r_H + \alpha_L r_L \leq r_H \alpha_H^{(1)}$ as $n \rightarrow \infty$. The two other constraints of set \mathcal{A}_2 are

$$\begin{aligned}\mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, x_l = 1, \mathbf{x}_{L \setminus \{l\}} = \mathbf{0}, \alpha_L, \alpha_H) &\leq v_L + \frac{1}{n}, \\ \mathbb{P}_h^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{0}, \alpha_L, \alpha_H) &\leq v_H + \frac{1}{n}.\end{aligned}$$

The left-hand side of these constraints in the limit converge to each other and therefore, using $v_H > v_L$, these two constraints can be written as $\lim_{n \rightarrow \infty} \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \mathbf{0}, \alpha_L, \alpha_H) = Q(\alpha_H) \leq v_L$. This shows the existence of a threshold $\alpha_H^{(2)}$ such that these two constraints of set \mathcal{A}_2 in the limit as $n \rightarrow \infty$ hold if and only if we have $\alpha_H \geq \alpha_H^{(2)}$. A similar argument shows the existence of a threshold $\alpha_H^{(3)}$ that defines the sets \mathcal{A}_3^* and \mathcal{A}_4^* . ■

Proof of Theorem 1

Proof of part(a): Using Proposition 2, the unique symmetric equilibrium for $(\alpha_L, \alpha_H) \in \mathcal{A}_1$ is $x_i = 1$ for all $i \in \mathcal{V}$ and therefore the infection probability of agent i becomes

$$\mathbb{P}_i^{\text{inf}}(\mathbf{x} = \mathbf{1}, \alpha_L, \alpha_H),$$

which is decreasing in (α_L, α_H) as shown in Proposition 1.

Proof of part (b): Using Proposition 2, the unique symmetric equilibrium for $(\alpha_L, \alpha_H) \in \mathcal{A}_2$ is $x_i = 1$ for all $i \in \mathcal{H}$ and mixed action for all $j \in \mathcal{L}$ where $x_j = 1$ with probability $\gamma_L(\alpha_L, \alpha_H)$ and $x_j = 0$ with probability $1 - \gamma_L(\alpha_L, \alpha_H)$. Here, $\gamma_L(\alpha_L, \alpha_H)$ is such that low-value agent are indifferent between playing $x = 1$ and $x = 0$ which implies

$$\mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{L \setminus \{l\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_H = \mathbf{1}, \alpha_L, \alpha_H) = \frac{1}{n} + v_L. \quad (9)$$

The infection probability of low value agents is

$$\begin{aligned}\gamma_L(\alpha_L, \alpha_H) \mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{L \setminus \{l\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_H = \mathbf{1}, \alpha_L, \alpha_H) \\ + (1 - \gamma_L(\alpha_L, \alpha_H)) \mathbb{P}_l^{\text{inf}}(x_l = 0, \mathbf{x}_{L \setminus \{l\}} = \mathbf{1} \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_H = \mathbf{1}, \alpha_L, \alpha_H) \\ = \gamma_L(\alpha_L, \alpha_H) \left(\frac{1}{n} + v_L \right) + (1 - \gamma_L(\alpha_L, \alpha_H)) \frac{1}{n} = \gamma_L(\alpha_L, \alpha_H) v_L - \frac{1}{n}.\end{aligned}$$

We next prove that the $\gamma_L(\alpha_L, \alpha_H)$ which is the solution of (9) is increasing in (α_L, α_H) . We prove this by assuming the contrary and reaching a contradiction. In particular, suppose that $(\alpha'_L, \alpha'_H) >$

⁶The multiplier r_H in this threshold is included to have a better geometric interpretation of the limiting set \mathcal{A}_1^* . In particular, with this notation $\alpha_H^{(1)}$ becomes the intersection of the boundary between \mathcal{A}_1^* and \mathcal{A}_2^* with the α_H -axis.

(α_L, α_H) and $\gamma_L(\alpha'_L, \alpha'_H) < \gamma_L(\alpha_L, \alpha_H)$. We can write

$$\begin{aligned} v_L + \frac{1}{n} &\stackrel{(a)}{=} \mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}, \gamma_L(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H) \\ &\stackrel{(b)}{>} \mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}, \gamma_L(\alpha'_L, \alpha'_H), \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha_L, \alpha_H) \\ &\stackrel{(c)}{>} \mathbb{P}_l^{\text{inf}}(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \mathbf{1}, \gamma_L(\alpha'_L, \alpha'_H), \mathbf{x}_{\mathcal{H}} = \mathbf{1}, \alpha'_L, \alpha'_H) \stackrel{(d)}{=} v_L + \frac{1}{n}, \end{aligned}$$

which is a contradiction. In the above derivation (a) and (d) follow by invoking (9), (b) follow from the assumption that $\gamma_L(\alpha'_L, \alpha'_H) < \gamma_L(\alpha_L, \alpha_H)$ and part (a) of Lemma 1 and (c) follows from $(\alpha'_L, \alpha'_H) > (\alpha_L, \alpha_H)$ and Proposition 1. The infection probability of high-value agents is given in (9) which remains equal to $\frac{1}{n} + v_L$.

Proof of part (c): Using Proposition 2, the unique symmetric equilibrium for $(\alpha_L, \alpha_H) \in \mathcal{A}_3$ is $x_j = 0$ for all $j \in \mathcal{L}$ and $x_i = 1$ for all $i \in \mathcal{H}$ and therefore the infection probability of an agent k becomes

$$\mathbb{P}_k^{\text{inf}}(\mathbf{x}_{\mathcal{H}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H),$$

which is decreasing in (α_L, α_H) as shown in Proposition 1.

Proof of part (d): Using Proposition 2, the unique symmetric equilibrium for $(\alpha_L, \alpha_H) \in \mathcal{A}_4$ is $x_j = 0$ for all $j \in \mathcal{L}$ and mixed action for all $i \in \mathcal{H}$ where $x_i = 1$ with probability $\gamma_H(\alpha_L, \alpha_H)$ and $x_i = 0$ with probability $1 - \gamma_H(\alpha_L, \alpha_H)$. Here, $\gamma_H(\alpha_L, \alpha_H)$ is such that high-value agents are indifferent between playing $x = 1$ and $x = 0$ which implies

$$\mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) = \frac{1}{n} + v_H. \quad (10)$$

The infection probability of high value agents is

$$\begin{aligned} &\gamma_H(\alpha_L, \alpha_H) \mathbb{P}_l^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{l\}} = \mathbf{1}, \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \\ &+ (1 - \gamma_H(\alpha_L, \alpha_H)) \mathbb{P}_h^{\text{inf}}(x_h = 0, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \\ &= \gamma_H(\alpha_L, \alpha_H) \left(\frac{1}{n} + v_H \right) + (1 - \gamma_H(\alpha_L, \alpha_H)) \frac{1}{n} = \gamma_H(\alpha_L, \alpha_H) v_H - \frac{1}{n}. \end{aligned}$$

We next prove that the $\gamma_H(\alpha_L, \alpha_H)$ which is the solution of (10) is increasing in (α_L, α_H) . We establish this by assuming the contrary and reaching a contradiction. In particular, suppose that $(\alpha'_L, \alpha'_H) > (\alpha_L, \alpha_H)$ and $\gamma_H(\alpha'_L, \alpha'_H) < \gamma_H(\alpha_L, \alpha_H)$. We can write

$$\begin{aligned} v_H + \frac{1}{n} &\stackrel{(a)}{=} \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \gamma_H(\alpha_L, \alpha_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \\ &\stackrel{(b)}{>} \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \gamma_H(\alpha'_L, \alpha'_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) \\ &\stackrel{(c)}{>} \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}, \gamma_H(\alpha'_L, \alpha'_H), \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha'_L, \alpha'_H) \stackrel{(d)}{=} v_H + \frac{1}{n}, \end{aligned}$$

which is a contradiction. In the above derivation (a) and (d) follow by invoking (10), (b) follow from the assumption that $\gamma_H(\alpha'_L, \alpha'_H) < \gamma_H(\alpha_L, \alpha_H)$ and part (a) of Lemma 1, and (c) follows from $(\alpha'_L, \alpha'_H) > (\alpha_L, \alpha_H)$ and Proposition 1. The infection probability of low-value agents remains $\frac{1}{n}$, completing the proof. ■

Proof of Lemma 3

All parts follows from the fact that the infection probability is increasing in η , β , and p . The fact that the infection probability is increasing in η follows from a similar argument to the proof of Part (a) of Lemma 1. The fact that the infection probability is increasing in β and p follows from a similar argument to the proof of Proposition 1. ■

Proof of Theorem 2

We first prove the theorem for the limit when $n \rightarrow \infty$ with $\theta^{(1)} = r_H \alpha_H^{(1)}$, $\theta^{(2)} = \alpha_H^{(2)} r_H$, and $\theta^{(3)} = \alpha_H^{(3)} r_H$ where $\alpha_H^{(1)}$, $\alpha_H^{(2)}$, and $\alpha_H^{(3)}$ are the thresholds found in Lemma 2 and then use Lemma 2 to show it holds for sufficiently large n . Similar to the proof of Lemma 2 we let $Q(\alpha)$ be the limiting infection probability of agents in a society of n individuals when each agent is tested with probability α .

Proof of Part (a): For $\theta \geq \theta^{(1)}$, the number of tests is such that the equilibrium can be in region \mathcal{A}_1^* with normalized social welfare equal to

$$v_H r_H + v_L r_L - Q(\alpha) - c\alpha,$$

where α is the expected number of tests. Since $\pi(\alpha, \beta, p)$ is decreasing in α for

$$c < \max_{x \in [0,1]} \left\{ \frac{d}{d\alpha} Q(\alpha) |_{\alpha=x} \right\},$$

it is optimal to let $\alpha = \theta$.

Proof of Part (b): For $\theta^{(2)} \leq \theta \leq \theta^{(1)}$, the number of tests is such that the is in region \mathcal{A}_2^* with normalized social welfare equal to

$$\begin{aligned} & v_H r_H + v_L r_L \gamma_L - r_L \gamma_L \mathbb{P}_l^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \gamma_L \mathbf{1}, \alpha_L, \alpha_H) \\ & - r_H \mathbb{P}_h^{\text{inf}}(\mathbf{x}_H = \mathbf{1}, \mathbf{x}_L = \gamma_L \mathbf{1}, \mathbf{x}_H = \mathbf{1}, \alpha_L, \alpha_H) - c(\alpha_H r_H + \alpha_L r_L) \\ & = r_H(v_H - v_L) - c(\alpha_H r_H + \alpha_L r_L). \end{aligned}$$

Therefore, for any point (α_L, α_H) in this region the social welfare is equal to $r_H(v_H - v_L) - c(\alpha_H r_H + \alpha_L r_L)$. Therefore, the optimal testing policy (α_L, α_H) has the minimum number of tests which brings the social activity of low-value agents to 0 (i.e., the boundary between \mathcal{A}_2^* and \mathcal{A}_3^*) and we test high-value agents with probability $\frac{\theta^{(2)}}{r_H}$ so that on average we are testing the minimum fraction of agents that belongs to $[\theta^{(2)}, \theta^{(1)}]$.

Proof of Part (c): For $\theta^{(3)} \leq \theta \leq \theta^{(2)}$, the number of tests is such that the equilibrium can be in region \mathcal{A}_3^* with normalized social welfare equal to

$$v_H r_H - r_H Q\left(\frac{\theta}{r_H}\right) - c(\alpha_H r_H + \alpha_L r_L),$$

where we test each high-value agent with probability $\frac{\theta}{r_H}$ which makes the average number of tests equal to α . Since $Q\left(\frac{\theta}{r_H}\right)$ is decreasing in $\frac{\theta}{r_H}$ and $c < \max_{x \in [0,1]} \left\{ \frac{d}{d\alpha} Q(\alpha) |_{\alpha=x} \right\}$, it is optimal to let $\alpha = \theta$.

Proof of Part (d): For $\theta \leq \theta^{(3)}$, the number of tests is such that the equilibrium can only be in region \mathcal{A}_4^* with normalized social welfare equal to

$$\begin{aligned} v_H r_H \gamma_H - r_H \gamma_H \mathbb{P}_h^{\text{inf}}(x_h = 1, \mathbf{x}_{\mathcal{H} \setminus \{h\}} = \mathbf{1}_{\gamma_H}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) - c(\alpha_H r_H + \alpha_L r_L) \\ = -c(\alpha_H r_H + \alpha_L r_L). \end{aligned}$$

Therefore, for any point (α_L, α_H) in this region the social welfare is equal to $-c(\alpha_H r_H + \alpha_L r_L)$ and the optimal testing policy is 0. Finally, note that the theorem holds for large enough n with strict inequalities by invoking Lemma 2. ■

Proof of Corollary 1

The following lemma readily follow from Proposition 2 and characterizes equilibrium with uniform policies.

Lemma 5. *With uniform policy $\alpha_L = \alpha_H = \alpha$, there exists $M \in \mathbb{N}_+$, functions $\gamma_L : [0, 1] \rightarrow [0, 1]$ and $\gamma_H : [0, 1] \rightarrow [0, 1]$ and $\alpha^{(1)} \geq \alpha^{(2)} \geq \alpha^{(3)}$ such that for $n \geq M$, depending on the testing policy α there are four possibilities for the equilibrium:*

- (a) *For $\alpha > \alpha^{(1)}$, there exists a unique symmetric equilibrium given by $x_i^e = 1$ for all $i \in \mathcal{V}$.*
- (b) *For $\alpha^{(2)} < \alpha < \alpha^{(1)}$, there exists a unique symmetric equilibrium given by pure-strategy $x_i^e = 1$ for all $i \in \mathcal{H}$ and mixed-strategy μ_j^e for all $j \in \mathcal{L}$ where μ_j^e puts $\gamma_L(\alpha)$ probability on 1 and $1 - \gamma_L(\alpha)$ probability on 0.*
- (c) *For $\alpha^{(3)} < \alpha < \alpha^{(2)}$, there exists a unique symmetric equilibrium given by $x_i^e = 1$ for all $i \in \mathcal{H}$ and $x_j^e = 0$ for all $j \in \mathcal{L}$.*
- (d) *For $\alpha < \alpha^{(3)}$, there exists a unique symmetric equilibrium given by pure-strategy $x_j^e = 0$ for all $j \in \mathcal{L}$ and mixed-strategy μ_i^e for all $i \in \mathcal{H}$ where μ_i^e puts $\gamma_H(\alpha)$ probability on 1 and $1 - \gamma_H(\alpha)$ probability on 0.*

Corollary follows from this lemma and Theorem 2. ■

Proof of Proposition 3

Consider the testing policy with mandatory social distancing, the social planner uses all the testing capacity. This is because

$$(\alpha_L, \alpha_H, \bar{x}_L, \bar{x}_H) = (\alpha_L^{\text{FB}}(\theta), \alpha_H^{\text{FB}}(\theta), x_L^{\text{FB}}(\theta), x_H^{\text{FB}}(\theta)).$$

We prove that each agent $i \in \mathcal{H}$ plays $x_i = x_H^{\text{FB}}$ in equilibrium. Similarly, we can establish that each agent $j \in \mathcal{L}$ plays $x_j = x_L^{\text{FB}}$ in equilibrium. Since the utility function of each agent $i \in \mathcal{H}$ is concave in x_i , the equilibrium action of agent i (similar to Lemma 4), is either $x_i = 0$ or $x_i = x_H^{\text{FB}}$. We next show that playing x_H^{FB} is preferable for agent i :

$$\begin{aligned} u_i(x_i = x_H^{\text{FB}}, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{1} x_H^{\text{FB}}, \mathbf{x}_{\mathcal{L}} = \mathbf{1} x_L^{\text{FB}}, \alpha_L, \alpha_H) \\ \stackrel{(a)}{\geq} u_i(x_i = 0, \mathbf{x}_{\mathcal{H} \setminus \{i\}} = \mathbf{0}, \mathbf{x}_{\mathcal{L}} = \mathbf{1} x_L^{\text{FB}}, \alpha_L, \alpha_H) \\ \stackrel{(b)}{\geq} u_i(x_i = 0, \mathbf{x}_{\mathcal{H}} = \mathbf{1} x_H^{\text{FB}}, \mathbf{x}_{\mathcal{L}} = \mathbf{1} x_L^{\text{FB}}, \alpha_L, \alpha_H), \end{aligned}$$

where (a) follows from the fact that social welfare with $x_i = x_H^{\text{FB}}$ for $i \in \mathcal{H}$ and $x_j = x_L^{\text{FB}}$ for $j \in \mathcal{L}$ is larger than social welfare with $x_i = 0$ for $i \in \mathcal{H}$ and $x_j = x_L^{\text{FB}}$ for $j \in \mathcal{L}$ and (b) follows from the fact that the infection probability is increasing in the social activities (i.e., Part (a) of Lemma 1).

We next prove that with this testing policy with mandatory social distancing, the social planner uses all the testing capacity. This is because the infection probability decreases by increasing the testing probabilities (by using Proposition 1) and therefore the first best solution improves by using all the testing capacity. ■

Proof of Proposition 4

The proof is similar to the proof of Proposition 2. ■

Proof of Theorem 3

The proof is similar to the proof of Theorem 1. ■

Proof of Proposition 5

First, note that a similar argument to Lemma 4 shows that the best response of each agent is to play one of the four actions $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$.

Claim 1: Under the assumptions of the proposition, high-value agents will have full social activities toward low-value agents, i.e., $x_i^L = 1$ for all $i \in \mathcal{H}$.

The maximum increase in the infection probability of a high-value agent from fully connecting with low-value agents is r_L . This is because this is the probability with which the infection hits low-value agents. Given $r_H \geq \frac{v_H}{\beta}$, we have $v_H \geq r_L$, showing that high-value agents will fully

connect to low-value agents.

Claim 2: Under the assumptions of the proposition, low-value agents will have zero social activities toward high-value agents, i.e., $x_j^H = 0$ for all $j \in \mathcal{L}$.

The minimum infection probability of a low-value j agent if she connects to high-value agents is $r_H \beta \left(1 - \frac{\theta}{r_H}\right)$. This is because with probability r_H the infection hits one of the agents in \mathcal{H} and we do not test each of those with minimum probability $1 - \frac{\theta}{r_H}$ in which case the infection reaches agent j . Using the assumptions of the proposition, we have $r_H \beta \left(1 - \frac{\theta}{r_H}\right) \geq v_H \geq v_L$, showing that low-value agents will not connect to high-value agents.

Claim 3: Under the assumptions of the proposition, low-value agents will have full social activities toward low-value agents, i.e., $x_j^L = 1$ for all $j \in \mathcal{L}$.

This is because the maximum infection probability of a low-value agent j if she connects to other low-value agents is $r_L(1 - \frac{1}{n})$ which is below v_L given the assumption $r_H > 1 - v_L$ and for large enough n .

Claim 4: Under the assumptions of the proposition, high-value agents will have a mixed-strategy activity toward high-value agents.

They will not have zero activity, because in that case one of them has a profitable deviation to connect to others. They will not have full social activity because in that case their infection probability is at least $r_H(1 - \frac{\theta}{r_H})$ which is above v_H . We let γ_H denote the mixed strategy of high-value agents toward other high-value agents. This γ_H satisfies

$$\begin{aligned} v_H - \mathbb{P}_h(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, x_h = (1, 0), \mathbf{x}_{\mathcal{H} \setminus \{h\}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H) \\ = 2v_H - \mathbb{P}_h(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, x_h = (1, 1), \mathbf{x}_{\mathcal{H} \setminus \{h\}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H). \end{aligned} \quad (11)$$

The normalized social welfare in this unique symmetric equilibrium is

$$\begin{aligned} & r_H(\gamma_H v_H + v_H) + r_L v_L - r_L \mathbb{P}_l(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, \mathbf{x}_{\mathcal{H}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H) \\ & - r_H \gamma_H \mathbb{P}_h(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, x_h = (1, 1), \mathbf{x}_{\mathcal{H} \setminus \{h\}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H) \\ & - r_H(1 - \gamma_H) \mathbb{P}_h(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, x_h = (1, 0), \mathbf{x}_{\mathcal{H} \setminus \{h\}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H) - c(r_H \alpha_H + r_L \alpha_L) \\ & \stackrel{(a)}{=} r_H v_H + r_L v_L - r_L \mathbb{P}_l(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, \mathbf{x}_{\mathcal{H}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H) \\ & - r_H \mathbb{P}_h(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, x_h = (1, 0), \mathbf{x}_{\mathcal{H} \setminus \{h\}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H) - c(r_H \alpha_H + r_L \alpha_L), \end{aligned}$$

where (a) follows from (11). By decreasing α_H , the term $\mathbb{P}_h(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, x_h = (1, 0), \mathbf{x}_{\mathcal{H} \setminus \{h\}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H)$ remains equal to $\frac{1}{n}$, the term $\mathbb{P}_l(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, \mathbf{x}_{\mathcal{H}} = (1, \gamma_H)\mathbf{1}, \alpha_L, \alpha_H)$ does not change as well because low-value agents are not connected to high-value agents. Therefore, in the optimal policy we have $\alpha_H = 0$. Finally, by letting

$$c \leq \max_x \left\{ \frac{d}{d\alpha} \mathbb{P}_l(\mathbf{x}_{\mathcal{L}} = (1, 0)\mathbf{1}, \mathbf{x}_{\mathcal{H}} = (1, 0)\mathbf{1}, \alpha_L = \alpha, \alpha_H = 0) |_{\alpha=x} \right\},$$

the optimal testing policy will have maximum α_L . Given the constraint on testing capacity, the optimal testing policy becomes

$$(\alpha_L, \alpha_H) = \left(\min \left\{ 1, \frac{\theta}{r_L} \right\}, 0 \right),$$

completing the proof. ■

A.2 Additional Results

Characterization of Asymmetric Equilibrium

Theorem 4. *For any $\epsilon > 0$, there exists $M \in \mathbb{N}_+$ such that for $n \geq M$, depending on $\alpha = (\alpha_L, \alpha_H) \in \mathcal{A}_1$ there are four possibilities for the equilibrium:*

- (a) *For $\alpha \in \mathcal{A}_1$, there exists a unique (symmetric or asymmetric) pure-strategy equilibrium given by $x_i^e = 1$, for all $i \in \mathcal{V}$.*
- (b) *For $\alpha \in \mathcal{A}_2$, there exists a unique (symmetric or asymmetric) pure-strategy equilibrium given by $x_i^e = 1$, for all $i \in \mathcal{H}$, $x_j^e = 1$ for $\lceil \tilde{\gamma}_L(\alpha_L, \alpha_H) |\mathcal{L}| \rceil$ of low-value agents and $x_j^e = 0$ for the rest of them. Moreover, the difference between the expected number of infected individuals in this pure-strategy equilibrium and the the expected number of infected individuals in the unique symmetric mixed equilibrium is at most ϵn .*
- (c) *For $\alpha \in \mathcal{A}_3$, there exists a unique (symmetric or asymmetric) pure-strategy equilibrium given by $x_i^e = 1$, for all $i \in \mathcal{H}$ and $x_j^e = 0$, for all $j \in \mathcal{L}$.*
- (d) *For $\alpha \in \mathcal{A}_4$, there exists a unique (symmetric or asymmetric) pure-strategy equilibrium given by $x_j^e = 1$, for all $j \in \mathcal{L}$, $x_i^e = 1$ for $\lceil \tilde{\gamma}_H(\alpha_L, \alpha_H) |\mathcal{L}| \rceil$ of high-value agents and $x_i^e = 0$ for the rest of them. Moreover, the difference between the expected number of infected individuals in this pure-strategy equilibrium and the the expected number of infected individuals in the unique symmetric mixed equilibrium is at most ϵn .*

Before proceeding with the proof of this theorem, we highlight that this theorem implies that the social welfare of the unique symmetric equilibrium we studied in the main text is close to the unique pure-strategy equilibrium. This establishes that all the results regarding the non-monotonicity of infection probability in testing policy as well as the optimal testing hold true for asymmetric equilibrium as well.

Proof of Theorem 4: Throughout this proof we use Lemma 4, showing that in any pure-strategy equilibrium the social activity of agents is either 1 or 0.

Proof of Part (a): Part (a) follows from the fact that for any action profile of other agents \mathbf{x}_{-i} , it is optimal for agent i to have $x_i = 1$. This is because

$$u_i(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) = v_i - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c(\alpha_L \mathbf{1}\{i \in \mathcal{L}\} + \alpha_H \mathbf{1}\{i \in \mathcal{H}\})$$

$$\begin{aligned}
&\stackrel{(a)}{\geq} v_i - \left(v_L + \frac{1}{n} \right) - c(\alpha_L \mathbf{1}\{i \in \mathcal{L}\} + \alpha_H \mathbf{1}\{i \in \mathcal{H}\}) \\
&\geq -\frac{1}{n} - c(\alpha_L \mathbf{1}\{i \in \mathcal{L}\} + \alpha_H \mathbf{1}\{i \in \mathcal{H}\}) = u_i(x_i = 0, \mathbf{x}_{-i}, \alpha_L, \alpha_H),
\end{aligned}$$

where (a) follow from the definition of set \mathcal{A}_1 .

In the rest of the proof we use the following two lemmas.

Lemma 6. *There exists $M \in \mathbb{N}_+$ such that for $n \geq M$, in any pure-strategy equilibrium we cannot have $x_j = 1$ for some $j \in \mathcal{L}$ and $x_i = 0$ for some $i \in \mathcal{H}$.*

Proof of Lemma 6: To arrive at a contradiction suppose the contrary holds, i.e., $x_j = 1$ for some $j \in \mathcal{L}$ and $x_i = 0$ for some $i \in \mathcal{H}$. We prove that agent i has a profitable deviation to $x_i = 1$. This is because we can write

$$\begin{aligned}
u_i(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) &= v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c\alpha_H \\
&\stackrel{(a)}{\geq} v_H - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{-j}, \alpha_L, \alpha_H) - \epsilon - c\alpha_H \\
&\stackrel{(b)}{\geq} v_H - \left(v_L + \frac{1}{n} \right) - \epsilon - c\alpha_H \\
&\stackrel{(c)}{>} \frac{-1}{n} - c\alpha_H = u_i(x_i = 0, \mathbf{x}_{-i}, \alpha_L, \alpha_H)
\end{aligned}$$

where (a) follows by having sufficiently large n which guarantees that the impact of one agent in the infection probability of is smaller than ϵ , (b) follows from the fact that agent j is playing $x_j = 1$ in equilibrium, and (c) follows by choosing $\epsilon < v_H - v_L$. This completes the proof of Lemma 6. ■
 Lemma 6 shows that high-value agents are always the first ones who connect with others.

Lemma 7. *Consider a society of n individuals and let \mathbf{x} be an action profile in which γ fraction of high-value agents are playing 1 and the rest of them are playing 0. For any $\epsilon > 0$, there exists $M \in \mathbb{N}_+$ such that for $n \geq M$ we have*

$$|\mathbb{P}_j^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) - \mathbb{P}_j^{\text{inf}}(\mathbf{x} = \mathbf{1}\gamma, \alpha_L, \alpha_H)| \leq \epsilon \text{ for all } j \text{ such that } x_j = 1.$$

Proof of Lemma 7: The proof of this lemma is similar to the proof of Lemma 2 hence omitted. ■
 We now proceed with the proof of Theorem 4.

Proof of Part (b): Using Lemma 6 in any equilibrium all high-value agents must have $x_i = 1$. This is because otherwise a high-value agent $i \in \mathcal{H}$ with $x_i = 0$ has a profitable deviation:

$$\begin{aligned}
u_i(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) &= v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c\alpha_H \\
&\stackrel{(a)}{\geq} v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{\mathcal{H} \setminus \{i\}}, \mathbf{x}_{\mathcal{L}} = 0, \alpha_L, \alpha_H) - c\alpha_H \\
&\stackrel{(b)}{\geq} \frac{-1}{n} - c\alpha_H \geq u_i(x_i = 0, \mathbf{x}_{-i}, \alpha_L, \alpha_H),
\end{aligned}$$

where (a) follows from Lemma 6 and (b) follow from $(\alpha_L, \alpha_H) \in \mathcal{A}_2$. Now suppose a fraction γ of low-value agents also play 1 and the rest of them play 0. We next establish that we must have $\gamma = \gamma_L(\alpha_L, \alpha_H)$. Letting $j, j' \in \mathcal{L}$ be such that $x_j = 1$ and $x_{j'} = 0$, this holds because we have

$$\begin{aligned}\mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{-j}, \alpha_L, \alpha_H) &\leq v_L + \frac{1}{n} \text{ and} \\ \mathbb{P}_{j'}^{\text{inf}}(x_{j'} = 1, \mathbf{x}_{-j'}, \alpha_L, \alpha_H) &\geq v_L + \frac{1}{n}.\end{aligned}$$

For large enough n the difference between the left-hand sides of the above inequalities is less than ϵ . This shows that γ is such that with $\mathbf{x}_{\mathcal{H}} = \mathbf{1}$ and $\lceil \gamma |\mathcal{L}| \rceil$ of low-values playing 1 the infection probability of a low-value agent who plays 1 is $\frac{\epsilon}{2}$ away from $v_L + \frac{1}{n}$. Finally, note that using Lemma 7 this infection probability is $\frac{\epsilon}{2}$ away from $\mathbb{P}_l(x_l = 1, \mathbf{x}_{\mathcal{L} \setminus \{l\}} = \gamma \mathbf{1}, \mathbf{x}_{\mathcal{H}} = \mathbf{1}) = v_L + \frac{1}{n}$. Therefore, γ is such that

$$\begin{aligned}|\mathbb{P}_i^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) - \left(v_L + \frac{1}{n}\right)| &\leq \epsilon, \quad \text{for all } i \in \mathcal{H}, \\ |\mathbb{P}_j^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) - \left(v_L + \frac{1}{n}\right)| &\leq \epsilon, \quad \text{for all } j \in \mathcal{L}, x_j = 1,\end{aligned}$$

and

$$\mathbb{P}_j^{\text{inf}}(\mathbf{x}, \alpha_L, \alpha_H) = \frac{1}{n}, \quad \text{for all } j \in \mathcal{L}, x_j = 0.$$

Therefore, the difference between the expected number of infected individuals in the pure-strategy equilibrium and the unique symmetric mixed-strategy equilibrium is at most ϵn .

Proof of Part (c): Again, using Lemma 6, in equilibrium either all high-value agents together with some of low-value agents are playing 1 or a subset of high-value agents are playing 1 or only high value agents play 1. The former cannot be equilibrium because a low value agent $j \in \mathcal{L}$ has a profitable deviation:

$$\begin{aligned}u_j(x_j = 0, \mathbf{x}_{-j}, \alpha_L, \alpha_H) &= \frac{-1}{n} - c\alpha_L \\ &\stackrel{(a)}{\geq} v_L - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{\mathcal{H} \setminus \{j\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) - c\alpha_L \\ &\stackrel{(b)}{\geq} v_L - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{-j}, \alpha_L, \alpha_H) - c\alpha_L,\end{aligned}$$

where (a) follow from $(\alpha_L, \alpha_H) \in \mathcal{A}_3$ and (b) follows from the assumption that in the equilibrium we started from all high-value agent play 1. Therefore, none of the low-value agents will play 1 in any equilibrium. Moreover, all the high-value agents must play 1 because otherwise if $i \in \mathcal{H}$ is playing 0, then she has a profitable deviation:

$$\begin{aligned}u_i(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) &= v_H - \mathbb{P}_i^{\text{inf}}(x_i = 1, \mathbf{x}_{-i}, \alpha_L, \alpha_H) - c\alpha_H \\ &\stackrel{(a)}{\geq} v_H - \mathbb{P}_j^{\text{inf}}(x_j = 1, \mathbf{x}_{\mathcal{H} \setminus \{j\}} = \mathbf{1}, \mathbf{x}_{\mathcal{L}} = \mathbf{0}, \alpha_L, \alpha_H) - c\alpha_H\end{aligned}$$

$$\stackrel{(b)}{\geq} -\frac{1}{n} - c\alpha_H = u_i(x_i = 0, \mathbf{x}_{-i}, \alpha_L, \alpha_H),$$

where (a) follows from $(\alpha_L, \alpha_H) \in \mathcal{A}_3$ and (b) follows from the assumption that in the equilibrium we started from a subset of high-value agents play 1. This shows that the unique equilibrium is to have all high-value agents play 1 and all low-value agents play 0.

Proof of Part (d): The proof of this part follows from a similar argument to part (b). ■

References

- D. Acemoglu, A. Malekian, and A. Ozdaglar. Network security and contagion. *Journal of Economic Theory*, 166:536–585, 2016.
- D. Acemoglu, A. Makhdoumi, A. Malekian, and A. Ozdaglar. Privacy-constrained network formation. *Games and Economic Behavior*, 105:255–275, 2017a.
- D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi. Microeconomic origins of macroeconomic tail risks. *American Economic Review*, 107(1):54–108, 2017b.
- D. Acemoglu, V. Chernozhukov, I. Werning, and M. D. Whinston. A multi-risk sir model with optimally targeted lockdown. Technical report, National Bureau of Economic Research, 2020.
- F. E. Alvarez, D. Argente, and F. Lippi. A simple planning problem for covid-19 lockdown. Technical report, National Bureau of Economic Research, 2020.
- A. Atkeson. What will be the economic impact of covid-19 in the us? rough estimates of disease scenarios. Technical report, National Bureau of Economic Research, 2020.
- V. Bala and S. Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.
- D. W. Berger, K. F. Herkenhoff, and S. Mongey. An seir infectious disease model with testing and conditional quarantine. Technical report, National Bureau of Economic Research, 2020.
- J. R. Birge, O. Candogan, and Y. Feng. Controlling epidemic spread: Reducing economic losses with targeted closures. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2020-57), 2020.
- L. Blume, D. Easley, J. Kleinberg, R. Kleinberg, and E. Tardos. Network formation in the presence of contagious risk. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 1–10. ACM, (2011).
- B. Bollobás and B. Béla. *Random graphs*. Number 73. Cambridge university press, 2001.
- L. Brotherhood, P. Kircher, C. Santos, and M. Tertilt. An economic model of the covid-19 epidemic: The importance of testing and age-specific policies. 2020.

- G. Brumfiel. To end the coronavirus crisis we need widespread testing, experts say. *NPR*, 2020. URL <https://www.npr.org/sections/health-shots/2020/03/24/820157519/to-end-the-coronavirus-crisis-we-need-widespread-testing-experts-say>.
- S. BruSoe-Lin and R. Hecht. To suppress covid-19, we need to test those with no symptoms. *The New York Times*, 2020. URL <https://www.nytimes.com/2020/04/26/opinion/coronavirus-test-asymptomatic.html>.
- R. W. Crandall and J. D. Graham. Automobile safety regulation and offsetting behavior: Some new empirical estimates. *The American Economic Review*, 74(2):328–331, 1984.
- S. Currarini, M. O. Jackson, and P. Pin. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045, 2009.
- K. Drakopoulos and R. S. Randhawa. Demand control of information products: Why perfect tests may not be worth waiting for. *Available at SSRN*, 2020.
- K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis. When is a network epidemic hard to eliminate? *Mathematics of Operations Research*, 42(1):1–14, 2017.
- M. S. Eichenbaum, S. Rebelo, and M. Trabandt. The macroeconomics of epidemics. Technical report, National Bureau of Economic Research, 2020.
- M. Elliott, B. Golub, and M. O. Jackson. Financial networks and contagion. *The American economic review*, 104(10):3115–3153, 2014.
- J. Ely, A. Galeotti, and J. Steiner. Optimal test allocation. Technical report, Mimeo, 2020.
- M. Farboodi, G. Jarosch, and R. Shimer. Internal and external effects of social distancing in a pandemic. Technical report, National Bureau of Economic Research, 2020.
- A. Glover, J. Heathcote, D. Krueger, and J.-V. Ríos-Rull. Health versus wealth: On the distributional effects of controlling a pandemic. Technical report, National Bureau of Economic Research, 2020.
- S. Goyal. *Sustainable communication networks*. Econometric Institute, Erasmus University Rotterdam, 1993.
- M. Jackson. Social and economic networks. 2008. *Princeton: Princeton University Press. Princeton*, 2008.
- M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.
- C. J. Jones, T. Philippon, and V. Venkateswaran. Optimal mitigation policies in a pandemic: Social distancing and working from home. Technical report, National Bureau of Economic Research, 2020.

- E. H. Kaplan. Containing 2019-ncov (wuhan) coronavirus. *Health care management science*, pages 1–4, 2020.
- M. Kasy, A. Teytelboym, et al. Adaptive targeted infectious disease testing. *Oxford Review of Economic Policy*, 2020.
- T. E. Keeler. Highway safety, economic behavior, and driving environment. *The American Economic Review*, 84(3):684–693, 1994.
- D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. *Theory OF Computing*, 11(4):105–147, 2015.
- T. Kruse and P. Strack. Optimal control of an epidemic through social distancing. 2020.
- K. Y. Leung, F. Ball, D. Sirl, and T. Britton. Individual preventive social distancing during an epidemic may have negative population-level outcomes. *Journal of The Royal Society Interface*, 15(145):20180296, 2018.
- B. Lindgren and C. Stuart. The effects of traffic safety regulation in sweden. *Journal of Political Economy*, 88(2):412–427, 1980.
- W. Maloney and T. Taskin. Determinants of social distancing and economic activity during covid-19: A global view, 2020.
- S. Morris. Contagion. *The Review of Economic Studies*, 67(1):57–78, 2000.
- M. E. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical review E*, 64(2):026118, 2001.
- S. Peltzman. The effects of automobile safety regulation. *Journal of political Economy*, 83(4):677–725, 1975.
- F. Piguillem and L. Shi. Optimal covid-19 quarantine and testing policies. 2020.
- T. Searchinger, A. LaMantia, and G. Douglas. Mass testing is the only sustainable solution to the coronavirus. *Washington Post*, 2020. URL <https://www.washingtonpost.com/opinions/2020/03/23/mass-testing-is-only-sustainable-solution/>.
- E. Tardos and T. Wexler. Network formation games and the potential function method. *Algorithmic Game Theory*, pages 487–516, 2007.
- F. Toxvaerd. Equilibrium social distancing. 2020.
- F. Vega-Redondo. *Complex social networks*. Number 44. Cambridge University Press, 2007.