Targeting Interacting Agents

Nikhil Vellodi, Joshua Weiss*

August 16, 2023 (click here for the latest version)

Abstract

We introduce a framework to study targeted policy interventions. Agents differ both in their likelihood of and payoff from interaction, exert externalities through interaction, and can alter their interaction rates at a cost. A policy maker selects a subset of agents to alter their interaction rates at no cost (a *selection policy*). We fully characterize both the second- and first-best selection policies; that is, when costly interaction choices are either voluntary or mandatory. Our main result introduces the concept of *normative risk compensation*, which describes how the second-best selection policy internalizes spillovers to a lesser (greater) extent than the first-best when externalities are negative (positive), in order to manipulate voluntary incentives. We apply our results to various settings including vaccine allocation and information aggregation.

JEL Classification: H4, D62. **Keywords**: Risk compensation, allocative externalities, targeted interventions.

1 Introduction

In many settings, authorities design targeted interventions for diverse groups of interacting agents. Consider viral contagion within a population, wherein people differ both in their rate of contact with others as well as the severity of their symptoms conditional on contracting the virus. They can take costly measures to avoid potentially harmful interactions, while a policy maker allocates a limited supply of vaccines. Alternatively, consider a group of agents who can spend time and effort to engage on a social media platform. They differ in how connected they are within the platform,

*Vellodi: Paris School of Economics, Weiss: Institute for International Economic Studies (IIES), Stockholm University. We thank S. Nageeb Ali, Roland Bénabou, Francis Bloch, Piotr Dworczak, Kfir Eliaz, Ben Golub, Dan Gottlieb, Alessandro Lizzeri, Paula Onuchic, Ludovic Renou, and participants at various seminars for helpful comments. Vellodi acknowledges support from the EUR grant ANR-17-EURE-0001, as well as the Simons Institute for the Theory of Computing at UC Berkeley, where part of this research was performed.

as well as how much they value interaction. The platform can select who to "promote", for instance in the allocation of "blue check" marks on Twitter.

In each case, a designer must decide who to target. During the COVID-19 pandemic, many authorities first vaccinated those most at risk conditional on an infection.[1] On the other hand, many advocated targeting those most at risk of contracting and spreading an infection, so-called "super-spreaders" to achieve "herd immunity", an approach echoed by the literature on targeting central nodes in networks.[2] The trade-off underlying the debate between these two approaches is between prioritizing direct benefits to those vaccinated versus indirect benefits to those not vaccinated. A crucial determinant of the value of these indirect benefits is how unvaccinated people respond. For example, a person's willingness to self-isolate or wear a mask depends on the risks they face, and therefore on the vaccination policy and implied equilibrium choices of others.

This connection between targeted interventions and voluntary behavior is our central focus. We develop a framework with three central ingredients. First, agents differ along two separate dimensions — their *exposure* type indexes their likelihood of interacting and the externality they exert on other interacting agents, and their *payoff* type indexes their payoff from interacting. Second, agents can incur costs to adjust their likelihood of interacting. Third, a benevolent policy maker can select a subset of agents who no longer need to pay these adjustment costs. Our analysis is general and speaks both to situations in which an agent's interaction imposes a negative externality (Sections 4 and 6.1) or a positive externality (Sections 5 and 6.2) on other interacting agents, as with viral contagion and social media engagement, respectively.

We develop the concept of an *exposure elasticity* to characterize the extent to which policy prioritizes based on exposure type rather than payoff type, and thus indirect rather than direct benefits. Our primary interest lies in showing precisely how this prioritization should take into account its effects on voluntary incentives.

Our main findings compare the exposure elasticity in the first-best and the second-best policies — respectively, when voluntary incentive constraints are relaxed or respected — and thus uncover the role of incentives. In general, optimal policy places more weight on exposure type than payoff type because they both contribute to an agent's private benefits from being selected by the policy,

[1] See for instance the European Union's guidelines: https://tinyurl.com/y8n2ayrh.

[2] See Pastor-Satorras and Vespignani (2002), Shaw and Schwartz (2008), Gross et al. (2006), and Epstein et al. (2008). In contrast to most countries, Indonesia explicitly aimed to achieve herd immunity: see https://tinyurl.com/y8vvpfkm and https://tinyurl.com/yybj498g.

but only the former determines the spillover benefits to others. In the case of negative externalities, the second-best policy has a *lower* exposure elasticity than the first-best (Theorem 1), and so does not prioritize exposure so heavily. Conversely, with positive externalities, the second-best policy has a *greater* exposure elasticity than the first-best (Theorem 2). Intuitively, vaccinating super-spreaders reduces the risk of contagion, which encourages the unvaccinated to stop self-isolating, ultimately undermining the aim of reducing contagion. Conversely, promoting highly-connected users of a social media platform increases the interaction incentives of unpromoted users, which furthers the goal of increasing interaction.

An important corollary is that it is relatively efficient to allocate vaccines with a market clearing price, i.e., to those with the highest private benefit, if behavior is voluntary rather than mandatory because in that case, super-spreaders do not internalize that their vaccination reduces risk for others, but also do not internalize that their vaccination inefficiently discourages self-isolation. On the other hand, it is particularly *in*efficient to allocate Twitter blue check marks with a price if behavior is voluntary because in that case, highly-connected users do not internalize that their promotion provides benefits to other users *and* encourages more participation on the platform.

Our findings speak to the literature on *risk compensation*, which argues that mitigating the downsides to risky actions might not reduce the prevalence of adverse outcomes if outweighed by greater risk-taking behavior. Risk compensation has been well-documented, both theoretically and empirically.[3] Previous work studies the concept from a positive viewpoint, typically via a comparative static on the efficacy of a certain intervention or technology; a more effective technology might lead to overcompensating risk-taking by agents, thus counteracting its desirability.

We offer three distinct contributions to this literature. First, we take a normative approach, and show precisely how risk compensating behavior affects the *design* of a targeted intervention. As such, we term the phenomenon *normative risk compensation*. Second, we connect risk compensation to our novel exposure elasticity concept, which crucially relies on hitherto unexplored heterogeneity in both interaction rates and payoffs. Finally, we establish how normative risk compensation has a natural converse in settings with positive externalities in which agents' choices are strategic complements, broadening the scope of applications.

[3] Geoffard and Philipson, 1996, Greenwood et al., 2019, and Kremer, 1996 study risk compensating behavior in public health management, Ehrlich and Becker, 1972 in a moral hazard setting, and Talamás and Vohra, 2020 and Hoy and Polborn, 2015 more generally.

Beyond normative risk compensation, we derive additional qualitative insights into optimal policies. We show that the second-best policy is *non-monotone*—agents with intermediate payoffs from interaction are selected most heavily—whereas the first-best is monotone (Corollaries 2 and 5). With negative externalities and strategic substitutes, such agents interact too much relative to the social optimum, and are particularly costly to society when not selected. With positive externalitites and strategic complements, such agents interact too little. Finally, in Appendix A, we extend our results to a more general model with heterogeneous interaction adjustment costs.

Our paper relates to the broad literature on targeted interventions within groups of heterogeneous agents. For example, Pastor-Satorras and Vespignani (2002), Shaw and Schwartz (2008), Gross et al. (2006), and Epstein et al. (2008) study targeted interventions on random / adaptive networks. A key difference is that they abstract from the effects of equilibrium behavior on optimal policy. Alternatively, Jehiel and Moldovanu (2001), Jehiel et al. (1996), Ostrizek and Sartori (2022), and Akbarpour et al. (2022) study mechanism design problems with allocative externalities, wherein which agent receives the good affects the utilities of the agents who do not receive it, and thus their incentives to report their type truthfully. In the context of COVID-19, Akbarpour et al. (2022) show that requiring agents to pay for vaccines helps screen agents and improve both efficiency and equity. These papers focus on eliciting private information from agents, whereas we suppose the policy maker has perfect information, and study how agents respond to the effects of policy on the aggregate environment. Finally, our analysis bears some similarity to problems of optimal taxation with externalities (Akcigit et al., 2022; Piketty et al., 2014).

Beyond the question we seek to address, the tractability of our framework makes it a useful input into future applied work, as suggested by the applications we explore in Section 6. From a modeling perspective, we draw inspiration from random search, network models, and aggregative games (see Section 2.4 for a detailed discussion). Galeotti et al. (2020) and Ballester et al. (2006) explicitly consider incentives and behavior in network models, but abstract from endogenous participation as well as an explicit comparison between first- and second-best policies. Demange (2016) introduces a "threat index" that measures the marginal spillover exerted by firms in a financial network using a multiplier approach similar to ours (Proposition 2). Unlike in her setting, we can fully characterize these spillover effects by employing a tractable aggregative game approach. Finally, in the search tradition, the model in Farboodi et al. (2020) is similar to ours, except that an agent's interaction rate depends on neither the types nor decisions of other agents.

2 Model

A unit measure of agents are indexed by two-dimensional type $(x, y) \in [0, 1]^2$ drawn from a continuously differentiable distribution F with full support and density f. We call x an agent's *exposure type*, and y an agent's *payoff type*. Each agent faces the possibility of interacting with others. This likelihood is governed by their choices, the choices of others, and the intervention of a policy maker.

More precisely, agents make *interaction choices* $\sigma : [0, 1]^2 \to [0, 1]$ where $\sigma(x, y)$ is the probability that a type (x, y) agent chooses to interact, and the policy maker designs a *selection policy* $v : [0, 1]^2 \to [0, 1]$ where $v(x, y)$ is the probability that a type (x, y) agent is selected. Up to a fraction $\beta \in (0, 1)$ of agents can be selected:

$$\int_0^1 \int_0^1 v(x, y) f(x, y)\, dxdy \leqslant \beta. \tag{1}$$

Type (x, y) receives an *interaction payoff* with probability equal to the product of their *interaction probability* $\Sigma(\sigma(x, y) \mid x, v(x, y))$ and the *aggregate interaction probability* λ, where the former is a function (described in detail below) of their interaction choice, exposure type, and whether they are selected, and the latter is

$$\lambda \equiv \alpha \int_0^1 \int_0^1 \Sigma(\sigma(x, y) \mid x, v(x, y)) f(x, y)\, dxdy, \tag{2}$$

where $\alpha \in (0, 1]$ is an exogenous parameter. To provide further details, it is convenient to define two broad cases that ultimately form the backbone of the analysis.

2.1 Negative Externalities / Strategic Substitutes

In this case, interaction payoffs are negative, each agent can pay to avoid interaction, and selecting an agent allows them to do so at zero cost. Type (x, y)'s interaction probability is

$$\Sigma(\sigma(x, y) \mid x, v(x, y)) = x(1 - v(x, y))\sigma(x, y).$$

If they are not selected ($v(x,y) = 0$), then they receive a negative interaction payoff $-by$ with probability $x\lambda$ if they choose to interact ($\sigma(x,y) = 1$), and a negative payoff (the isolation cost) $-xc$ if they choose to isolate ($\sigma(x,y) = 0$), where $b, c > 0$ are exogenous parameters. If selected ($v(x,y) = 1$), they receive 0 regardless of their interaction choice. Their ex-ante utility is

$$u(x,y) = (1 - v(x,y))\big[\sigma(x,y) \cdot (-x\lambda by) + (1 - \sigma(x,y)) \cdot (-xc)\big]. \tag{3}$$

Crucially, selecting an agent both protects them from losses and prevents them from exerting negative externalities on others: if $v(x,y) = 1$, then $\Sigma(\sigma(x,y) \mid x, v(x,y)) = 0$, so type (x,y) does not contribute to λ.

2.2 Positive Externalities / Strategic Complements

Here, interaction payoffs are positive, each agent can pay to interact, and selecting an agent allows them to do so at zero cost. Type (x,y)'s interaction probability is

$$\Sigma(\sigma(x,y) \mid x, v(x,y)) = x[(1 - v(x,y))\sigma(x,y) + v(x,y)].$$

If they are not selected ($v(x,y) = 0$), then they must incur the negative payoff $-xc$ if they choose to interact ($\sigma(x,y) = 1$), in which case they receive a positive interaction payoff by with probability $x\lambda$, where $b, c > 0$. If they are not selected and choose to isolate ($\sigma(x,y) = 0$), they receive a payoff 0. If they are selected ($v(x,y) = 1$), then they receive the positive interaction payoff by with probability $x\lambda$ regardless of their interaction choice. Their ex-ante utility is

$$u(x,y) = (1 - v(x,y))\sigma(x,y) \cdot (x\lambda by - xc) + v(x,y) \cdot (x\lambda by). \tag{4}$$

Now, selecting an agent saves them from paying to interact, and imposes positive externalities on others: if $v(x,y) = 1$, then $\Sigma(\sigma(x,y) \mid x, v(x,y)) = 1$, so type (x,y) contributes to λ.

2.3 Welfare

In all cases, the policy maker's objective function is utilitarian welfare:

$$\mathcal{W} = \int_0^1 \int_0^1 u(x,y) f(x,y)\, dxdy, \tag{5}$$

where u is defined in equations (3) and (4).

2.4 Model Discussion

We cast our model at an abstract level to allow interpretations that vary by application, as we highlight in Section 6. For instance, in the case of negative spillovers / substitutes, selection can be viewed either as isolating agents at zero cost, or as allowing an agent to interact without consequence to herself or others. The probability of an interaction payoff for non-selected agents, $x\sigma(x,y)\lambda$, can be unpacked in various ways. One natural setting is where agent (x,y) receives an opportunity to enter an interaction pool with probability x, accepts this opportunity with probability $\sigma(x,y)$, and conditional on entering the pool, interacts with a probability that scales with the pool's size, λ, as in Section 6.1. Alternatively, as in Section 6.2, x can index the strength of payoff externality exerted by type (x,y) on those who interact, rather than the contribution of type (x,y) to the quantity of aggregate interaction.

An agent's payoff is a function of idiosyncratic variables—their type, behavior, and whether they are selected—and a single aggregate statistic λ. The model is thus an aggregative game.[4] Nonetheless, the marginal distribution $F(\cdot,y)$ models interaction in the spirit of random graphs; $F(\cdot,y)$ resembles the *degree distribution* $G(\cdot)$ of a random graph, where $G(x)$ is the probability of drawing at most x nodes (partners). We can easily generalize the model so that the probability an agent receives an interaction payoff is any continuous and strictly increasing function of λ. The parameter α allows for some flexibility on this front, and can be interpreted as a matching friction.

We can generalize the model so that the direct cost from interaction choice σ for an agent with exposure type x is $c(x,\sigma)$, which is weakly increasing in x and strictly increasing in σ. In our baseline model, $c(x,\sigma) = x\sigma c$, which implies that interaction choices do not depend on exposure types, and so yields an easy to understand characterization of optimal selection policies. Our main

[4]See Corchón 2021 for a survey and definition.

results hold as long as $c(x, \sigma)$ is linear in σ so that agents choose $\sigma \in \{0, 1\}$. Intuitively, optimal selection policies for interacting agents in the case of negative externalities / strategic substitutes and for non-interacting agents in the case of positive externalities / strategic complements only depend on the cost function $c(\cdot, \cdot)$ insofar as it determines the response of agents to changes in the aggregate interaction probability. This response depends quantitatively on the particular choice of cost function, but the sign of its effect on welfare, which is key for our results, does not. Finally, in Appendix A, we consider the case in which agents draw idiosyncratic costs c.

3 Policy Examples

Before solving for optimal policies, we define some natural policies, which form the building blocks for the subsequent analysis. We focus on *pure* selection policies where $v(x, y) \in \{0, 1\}$ for all $(x, y) \in [0, 1]^2$, which we later prove include the optimal policies. We define properties of policies that hold on all of $[0, 1]^2$, but later, we discuss cases in which these properties hold only on a subset.

Definition 1. A policy v is *monotone* if $v(x, y) = 1$ implies that $v(x', y') = 1$ for all (x', y') such that $x' \geqslant x$ and $y' \geqslant y$.

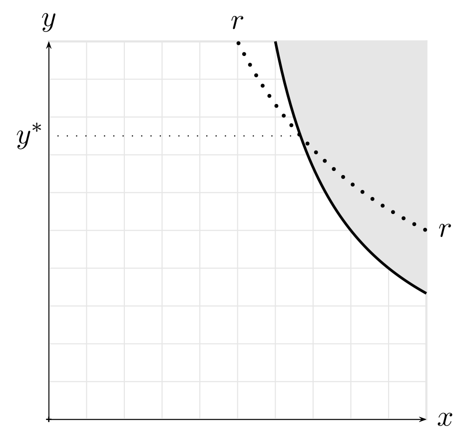
Definition 2. A policy v is an *x-policy* if $v(x, y) = 1$ implies that $v(x', y') = 1$ for all $x' \geqslant x$, $y' \in [0, 1]$. A policy v is a *y-policy* if $v(x, y) = 1$ implies that $v(x', y') = 1$ for all $y' \geqslant y$, $x' \in [0, 1]$. A policy v is an *xy-policy* if $v(x, y) = 1$ implies $v(x', y') = 1$ for all (x', y') such that $x'y' \geqslant xy$.

Monotone policies prioritize agents with higher exposure and payoff types, whereas x-, y-, and xy- policies give priority to a particular dimension of agents' types. x-policies select the most interactive or connected agents and arise naturally in models that study targeted interventions in networks, and where impacting interaction is the sole objective.[5] By contrast, y-policies are commonplace in practice, for instance in the allocation of vaccines to mitigate pandemics.[6] Finally, xy-policies select agents with the highest expected interaction payoff conditional on choosing to interact, which depends only on x and y through xy.

[5]See Pastor-Satorras and Vespignani, 2002, Shaw and Schwartz, 2008, Gross et al., 2006, and Epstein et al., 2008 and Akbarpour et al., 2020.

[6]In the context of the COVID-19 pandemic, the European Union guidelines state that when vaccine supply is low, priority should be given to the most vulnerable. See https://tinyurl.com/y8n2ayrh.

Figure 1: Exposure Premium Policy



Example of a threshold policy that exhibits an exposure premium. Solid line: policy threshold function $g(y)$. Dotted line: xy-policy threshold function $h(y) = r/y$. Grey shaded area: $v(x, y) = 1$. The exposure elasticity of $g(\cdot)$ at y^* is negative 1 times the ratio of the slope of $g(\cdot)$ to the slope of $h(\cdot)$ at y^*.

Definition 3. A policy $v(\cdot, \cdot)$ is a *threshold policy* if there exists an *threshold function* $g : [0, 1] \to [0, 1]$ such that $v(x, y) = 1$ if and only if $x \geqslant g(y)$.[7]

For example, an x-policy is characterized by a constant threshold function. An xy-policy is characterized by the threshold function $g(y) = r/y$, for some $r \geqslant 0$; for each y, the policy selects any agent with $x \geqslant r/y$, and so with $xy \geqslant r$. We often refer to a threshold policy $v(\cdot, \cdot)$ by its threshold function $g(\cdot)$. Figure 1 displays an example of a threshold policy.

Much of our analysis will focus on optimal selection policy for agents who would choose to interact in the negative externality case, or choose not to interact in the positive externality case, i.e., make the choice that is worse for others. Thus, since xy represents the *private value* of selecting such an agent with type (x, y), xy-policies ignore the spillovers agents exert through interaction, which scale only with their exposure type x. We introduce the *exposure elasticity* to capture the extent to which a policy places a premium on agents' exposure types and thus on spillovers:

Definition 4. Let

$$\varepsilon(y) \equiv \frac{y}{g(y)} \frac{dg(y)}{dy}$$

[7]We can similarly define a threshold policy in terms of a threshold for y as a function of x, which we can use in the following definitions. However, optimal policies can only be characterized by thresholds for x as functions of y.

denote the *exposure elasticity* of a threshold policy characterized by $g(\cdot)$.

To develop the exposure elasticity concept further, suppose there is a threshold policy with differentiable threshold function $g(\cdot)$ and for a given y^*, consider the alternative xy-policy with threshold function $h(\cdot)$ that intersects $g(\cdot)$ at y^*, i.e., $h(y) = r/y$, where $r = g(y^*)y^*$. Simple algebra shows that $-\varepsilon(y^*)$ is the ratio of the slope of $g(\cdot)$ to the slope of $h(\cdot)$ at y^*:

$$\varepsilon(y^*) = -g'(y^*)/h'(y^*).$$

Thus, the exposure elasticity $\varepsilon(y)$ measures how flat the slope of the policy threshold function is relative to an xy-policy, and thus how much more willing it is to substitute x for y. In particular, an xy-policy has $\varepsilon(\cdot) = -1$, an x-policy has $\varepsilon(\cdot) = 0$, and a y-policy has $\varepsilon(\cdot) = -\infty$.

Definition 5. A policy $v(\cdot, \cdot)$ exhibits an *exposure premium* if it is characterized by a threshold function $g(\cdot)$ such that $\varepsilon(y) > -1$ for all $y \in [0, 1]$.

If a policy exhibits an exposure premium, then agents with higher exposure types x are selected with lower levels of xy. Such a policy places a premium on an agent's exposure type x relative to their payoff type y, and $\varepsilon(y)$ quantifies this premium precisely. Figure 1 displays an exposure premium policy, and shows that the slope of the policy threshold is steeper than the corresponding xy-policy at one particular point $(g(y^*), y^*)$, which implies that $\varepsilon(y^*) > -1$.

4 Optimal Policies I: Negative Externalities / Substitutes

In this section, we study optimal policy in the case in which interaction payoffs are negative, and so an agent's choice to interact imposes negative externalities on others who do the same. As such, interaction choices are strategic substitutes. We consider two types of optimal policies: the *second-best policy* in which the policy maker chooses selection, $v(\cdot, \cdot)$, taking as given that interaction choices $\sigma(\cdot, \cdot)$ are individually optimal, and the *first-best policy* in which the policy maker jointly sets selection and interaction, so that $\sigma(\cdot, \cdot)$ need not be individually optimal. We denote variables in the second-best with subscript s and in the first-best with subscript f.

Before presenting the policy maker's problem, we state the following lemma that we can restrict attention to interaction choices characterized by an *interaction threshold*. If an agent's payoff type y

is below the threshold, then they interact with maximal probability, and we say they are *interacting*. If it is above the threshold, then they never interact, and we say they are *isolating*.

Lemma 1. *Fix a selection policy $v(\cdot,\cdot)$. Let $\sigma_s(\cdot,\cdot)$ and $\sigma_f(\cdot,\cdot)$ be the individually and socially optimal interaction choices, respectively. For $i \in \{s, f\}$, there exists an interaction threshold y_i^* such that $\sigma_i(x, y) = \mathbb{I}_{y \leqslant y_i^*}$, where $\mathbb{I}$ is the indicator function.*[8] *Under individually optimal behavior, there is a unique equilibrium: the interaction threshold is $y_s^* = c/(\lambda b) > 0$, where $\lambda > 0$ is the equilibrium aggregate interaction probability given by equation (2) as a function of $v(\cdot,\cdot)$ and y_s^*.*[9]

Intuitively, the private cost of isolating, xc, scales with an agent's exposure type x, and the private cost of interacting, $x\lambda by$, scales with an agent's exposure type x and payoff type y. Thus, for individually optimal interaction choices, x is irrelevant, and an agent isolates if y is sufficiently high. Socially optimal interaction choices also follow a payoff type threshold rule because the only difference is that the social cost of interacting includes an additional effect that scales with the agent's x but not their y: interacting increases the aggregate interaction probability, which increases the likelihood other interacting agents get negative payoffs. Finally, the equilibrium under individually optimal behavior is unique because interaction choices are strategic substitutes.

The policy maker's problem is to choose an interaction threshold y_i^*, a selection policy $v_i(\cdot,\cdot)$, and an aggregate interaction probability λ_i to maximize the Lagrangian:

$$\begin{aligned}\mathcal{L}_i =& \mathcal{W}_i + \gamma_{1,i}\left(\beta - \int_0^1 \int_0^1 v_i(x,y) f(x,y)\, dxdy\right) + \gamma_{2,i}\left(y_i^* - \frac{c}{\lambda_i b}\right) \\ &+ \gamma_{3,i}\left(\lambda_i - \alpha \int_0^{y_i^*} \int_0^1 (1 - v_i(x,y)) x f(x,y)\, dxdy\right),\end{aligned} \tag{6}$$

where $i \in \{s, f\}$ denotes the second- or the first-best, the constraints are the supply constraint (equation (1)), the incentive compatibility constraint derived from Lemma 1, and the equilibrium

[8] A measure 0 of agents have $y = y_i^*$, so their behavior is not relevant for aggregate outcomes or optimal policy.

[9] We use the concept of a Wardrop equilibrium ((Wardrop, 1952)), the natural analogue of Nash equilibrium for anonymous large games: given the aggregate distribution over actions, each agent best responds, while the distribution itself is consistent with individual optimization.

aggregate interaction probability constraint derived from equation (2)[10], and welfare is

$$\mathcal{W}_i = -\int_0^{y_i^*}\int_0^1 (1 - v_i(x, y))x\lambda_i b y f(x, y)\, dx dy - \int_{y_i^*}^1\int_0^1 (1 - v_i(x, y))xcf(x, y)\, dx dy. \tag{7}$$

The only difference between the second- and the first-best problems is that the incentive compatibility constraint need not hold in the first-best, and so we set $\gamma_{2,f} = 0$. The following proposition characterizes optimal policy.

Proposition 1. *Optimal selection policy for $i \in \{s, f\}$ denoting the second- or first-best is characterized by a threshold function $g_i(\cdot)$ such that*

$$g_i(y) = \begin{cases} g_i^* & y \in (y_i^*, 1] \\ \min\left\{\frac{\gamma_{1,i}}{\lambda_i b y + \gamma_{3,i}\alpha}, 1\right\} & y \in [0, y_i^*], \end{cases}$$

where

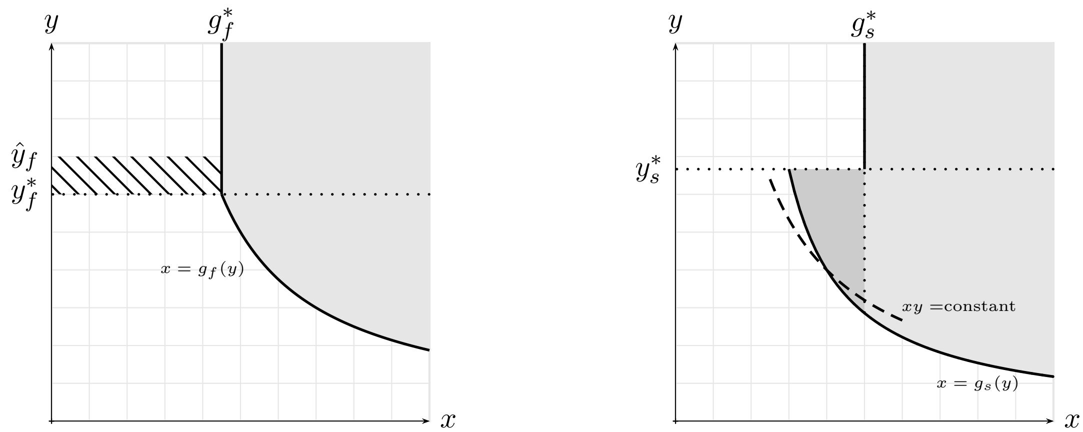
$$g_i^* \equiv \min\left\{\frac{\gamma_{1,i}}{c}, 1\right\},$$

and $\gamma_{1,i}, \gamma_{3,i}$ are strictly positive. The first-best interaction threshold is $y_f^* = (c - \gamma_{3,f}\alpha)/(\lambda_f b)$, where y_f^* and λ_f are strictly positive.

Figure 2 illustrates the first- and second-best policies. The threshold function $g_i(\cdot)$ tells us the exposure type of the marginal selected agent for each payoff type: an agent with payoff type y is selected if and only if their exposure type x is greater than $g_i(y)$. The benefit of selecting an agent with type $(g_i(y), y)$ on the threshold must equal the cost, which is the Lagrange multiplier on the supply constraint, $\gamma_{1,i}$. Above y_i^*, non-selected agents isolate, so the benefit of selecting type (x, y) is avoiding the cost of isolation, xc. Below y_i^*, non-selected agents interact, so the benefit of selecting type (x, y) is the private value from isolating them as well as the spillover value from reducing the aggregate interaction probability λ_i, and therefore the likelihood of a negative payoff for other interacting agents. The former is $x\lambda_i b y$ and the latter is $x\gamma_{3,i}\alpha$ because $\gamma_{3,i}$ is the social benefit of reducing the aggregate interaction probability.

[10]We allow y_s^* to be greater than 1 so that the incentive compatibility constraint always holds exactly. For all equations to be well-defined, we can extend the density f to be defined on all of $\mathbb{R}^2$, and equal to 0 outside of $[0, 1]^2$.

Figure 2: Optimal Policies



Left panel: first-best policy. Solid line: $g_f(y) = x$. Shaded area: $v_f(x, y) = 1$. $\hat{y}_f$: individually optimal interaction threshold. Right panel: second-best policy. Solid line: $g_s(y) = x$. Dashed line: xy-policy intersecting the optimal policy, demonstrating the exposure premium. Shaded area: $v_s(x, y) = 1$. Dark shaded area: additional selection due to inefficient interaction choices.

4.1 Normative Risk Compensation

We now use our purpose-built exposure elasticity measure $\varepsilon_i(\cdot)$ (Definition 4) to characterize the extent to which optimal selection policy for interacting agents prioritizes exposure types x over payoff types y. Moreover, we demonstrate the dependence of $\varepsilon_i(\cdot)$ on voluntary interaction behavior.

Using the expression for the threshold function in Proposition 1, if $y < y_i^*$ and $g_i(y) < 1$, then the exposure elasticity is the private benefit of selecting type (x, y) relative to the social benefit:

$$\varepsilon_i(y) = -\frac{x\lambda_i by}{x\lambda_i by + x\gamma_{3,i}\alpha}. \tag{8}$$

As discussed after Proposition 1, the social benefit in the denominator consists of the private benefit, which scales with xy, and the spillover benefit, which scales only with x. If the spillover benefit is 0, then the elasticity is -1, so the policy is an xy-policy. As the spillover benefit increases, the elasticity shifts toward 0, so the policy exhibits an exposure premium (Definition 5)—it prioritizes exposure type x as well as private benefits xy, and thus favors exposure type x over payoff type y.

It is useful to first consider a baseline that eliminates the role of voluntary behavior: the optimal selection policy taking as given interaction choices $\sigma(\cdot, \cdot)$. In this case, if agents with payoff type y

interact and are selected with $x > g(y)$, where $g(y) < 1$, then the exposure elasticity is

$$E(y, \tilde{y}) = -y/(y + \tilde{y}), \tag{9}$$

where $\tilde{y}$ is the interaction-weighted average payoff type:

$$\tilde{y} \equiv \mathbb{E}(y \mid \sigma(x,y) = 1, v(x,y) = 0) = \frac{\int_0^1 \int_0^1 yx(1 - v(x,y))\sigma(x,y)f(x,y)\,dxdy}{\int_0^1 \int_0^1 x(1 - v(x,y))\sigma(x,y)f(x,y)\,dxdy}. \tag{10}$$

The private benefit is reducing negative interaction payoffs for the selected agent, which depends on their payoff type y, and the spillover benefit is reducing negative interaction payoffs for others, which depends on their average payoff type $\tilde{y}$. Thus, the lower is the selected agent's payoff type relative to the average of those it would interact with, the less valuable are private benefits relative to spillover benefits, and so the closer the exposure elasticity is to 0 rather than -1.

Returning to optimal policy allowing equilibrium variables to respond, we have our main result:

Theorem 1. *Let $\tilde{y}_s$ and $\tilde{y}_f$ satisfy (10) in the second- and first-best optimal policy, respectively. If $y < y_f^*$ and $g_f(y) < 1$, then $\varepsilon_f(y) = E(y, \tilde{y}_f) > -1$.
If $y < y_s^*$ and $g_s(y) < 1$, then $-1 < \varepsilon_s(y) \leqslant E(y, \tilde{y}_s)$, and the inequality is strict if $y_s^* \leqslant 1$.*

Theorem 1 contains two results. First, both the first- and second-best policies exhibit exposure premia among interacting agents, i.e., $\varepsilon_i(\cdot) > -1$. Thus, as in the baseline above, optimal policy prioritizes exposure type x over payoff type y in that it selects high x agents at lower levels of xy.

Second, the second-best exhibits a weaker exposure elasticity than the first-best, for a given average payoff type $\tilde{y}$ among those interacting, i.e., if behavior is voluntary, then exposure does not receive such high priority. Since the first-best optimally sets both interaction choices and selection policy, an envelope argument implies that the marginal spillover benefit of selection takes interaction choices as given. Thus, the first-best selection policy exhibits an exposure elasticity of $E(y, \tilde{y}_f)$, as in the baseline above. On the other hand, in the second-best, the spillover benefit is mitigated by the equilibrium response of agents to changes in the risk they face; selecting some inadvertently encourages the inefficient interaction of others.

Taking the two results together, the second-best policy sits between one that considers only private benefits and one that accounts for spillovers but takes agent behavior as given.

Normative risk compensation is the effect of the anticipated response to policy on the second-best policy, which drives the second result in Theorem 1. To quantify this effect and understand it more precisely, we compute the marginal benefit $\gamma_{3,i}$ of decreasing the aggregate interaction probability λ, which determines the spillover benefit of selection and, from (8), the exposure elasticity.

Proposition 2. *For $i \in \{s, f\}$ denoting the second- or first-best, the marginal benefit $\gamma_{3,i}$ of reducing the aggregate interaction probability is given by*

$$\gamma_{3,f}\alpha = \lambda_f b\tilde{y}_f \qquad\qquad \gamma_{3,s}\alpha = \frac{\lambda_s b\tilde{y}_s - \Omega_1}{1+\Omega_0},$$

where

$$\Omega_0 \equiv \frac{y_s^*\alpha}{\lambda_s}\int_0^{g_s(y_s^*)} xf\left(x, y_s^*\right)dx \qquad\qquad \Omega_1 \equiv \frac{y_s^*\alpha}{\lambda_s}\int_{g_s(y_s^*)}^{g_s^*}(\gamma_{1,s} - xc)f\left(x, y_s^*\right)dx,$$

are weakly positive and strictly so if and only if $y_s^* \leqslant 1$.

Ω_0 and Ω_1 capture the offsetting effect of equilibrium behavior on the marginal spillover benefit when behavior is voluntary. As the aggregate interaction probability falls, they count agents who were non-selected and isolating but choose to interact and, respectively, remain non-selected or become optimally selected. Crucially, this response lowers welfare because agents do not internalize the negative externalities of interacting. If Ω_0 and Ω_1 are larger, then spillovers are weaker relative to private benefits, so priority in the second-best policy shifts from exposure type to payoff type.

4.2 Further Results

Optimal Policy for Isolating Agents – An immediate corollary of Proposition 1 is that optimal selection policy exhibits an exposure premium among agents who would isolate if not selected.

Corollary 1. *For $i \in \{s, f\}$, if $y_i^* < 1$, then for all $y > y_i^*$, $\varepsilon_i(y) = 0$.*

Among the isolating, optimal selection policy not only exhibits an exposure premium, but is an x-policy, i.e., is invariant to payoff type y. The benefit of selecting such an agent is only in avoiding the isolation cost, which scales with their exposure type, but not their payoff type. Together with Theorem 1, it follows that optimal selection policies exhibit exposure premia both for interacting and isolating agents. Moreover, the second-best policy is once again between the policy that considers only private values and the policy that takes agent behavior as given, which are equal here.

Non-monotonicity – Another immediate corollary of Proposition 1 is that if some non-selected agents isolate in the second-best, i.e., $y_s^* < 1$ (guaranteed if $c/(\alpha b)$ is sufficiently small), then the selection policy is *non-monotone*, whereas the first-best selection policy is always monotone.

Corollary 2. *If $y_s^* < 1$, then the second-best threshold function $g_s(\cdot)$ achieves its unique minimum at y_s^*. The first-best threshold function $g_f(\cdot)$ is continuous and decreasing.*

The second-best policy aggressively selects agents with intermediate payoff types (y at or just below y_s^*—the dark shaded area in the right panel of Figure 2). Whereas those with the highest payoff types efficiently isolate and those with the lowest efficiently interact, intermediate agents interact though they should isolate. They are thus particularly costly to society when non-selected. Put another way, comparing the isolating and the interacting, optimal policy focuses on the latter because agents over-interact relative to the social optimum. In the first-best, there is no reason to target in particular either the interacting or the isolating because interaction choices are efficient.

Implementation – The first-best interaction policy can be simply implemented with a strictly positive, flat interaction tax.

Corollary 3. *The policy maker can decentralize the optimal interaction policy in the first-best with a flat interaction tax equal to $\gamma_{3,f}\alpha$ per interaction if $y_f^* < 1$, and 0 otherwise.*

The policy maker can implement the first-best interaction policy without observing agents' types. That the tax is positive reflects that there exist agents—the dashed area in the left panel of Figure 2—who would like to interact in the equilibrium induced by the first-best policy, but do not.

Prohibitive Costs – To highlight the role of behavior, consider the case in which the relative cost of isolation is sufficiently high so that all agents interact ($y^* = 1$) under any selection policy.[11] The model reduces to one without the option to isolate, and thus second- and first-best policies are identical. The optimal policy is monotone and exhibits an exposure premium, as in Figure 1.

[11] For example, if $c/(\alpha b) > 1$, then even an agent facing maximal loss upon interacting would choose to do so. Alternatively, if β is sufficiently large, then $\lambda < c/b$ for any policy that satisfies the supply constraint.

5 Optimal Policies II: Positive Externalities / Complements

In this section, we perform a similar analysis but in the case (Section 2.2) in which interaction payoffs are positive, and so an agent's choice to interact imposes positive externalities on others who do the same. As such, interaction choices are strategic complements.

There are two important differences in the setup from Section 4. First, the sign of the equilibrium aggregate interaction probability constraint in the Lagrangian is flipped: $\gamma_{3,i}$ is now the marginal benefit of *increasing* λ, rather than decreasing it. Second, in the second-best problem, an equilibrium always exists, but there may be multiple equilibria. We follow the literature on partial implementation and henceforth allow the policy maker to select their preferred equilibrium.[12]

The proposition below, analogous to Lemma 1 and Proposition 1, characterizes optimal policy.

Proposition 3. *For $i \in \{s, f\}$, let y_i^* and $g_i(\cdot)$ be defined as before (in Lemma 1 and Proposition 1) as functions of parameters and equilibrium objects in the equilibria induced by optimal policies.*

1. *In the first-best, agents interact if and only if $y > y_f^*$ ($\sigma_f(x, y) = \mathbb{I}_{y>y_f^*}$), the selection policy is characterized by the threshold function $g_f(\cdot)$, and $\gamma_{1,f}$ and $\gamma_{3,f}$ are strictly positive.*

2. *In the second-best, agents interact if and only if $y > y_s^*$, the selection policy is characterized by the threshold function $g_s(\cdot)$, and $\gamma_{1,s}$ and $\gamma_{3,s}$ are strictly positive.*

The form of optimal policy is the same as in Proposition 1, except that agents now interact if their payoff type is sufficiently *high*, rather than low. The intuition for optimal behavior and policy is the same but flipped. For a type (x, y) agent's interaction choice, the private cost of interacting is now xc, the private cost of isolating is $x\lambda_i by$, and the social cost of isolating now includes the effect on the aggregate interaction probability. When selecting an interacting type (x, y), the benefit is now that they continue to interact but without paying the cost of doing so, xc. When selecting an isolating type (x, y), the benefit is now the private value of allowing them to interact, $x\lambda_i by$, and the spillover value of increasing the aggregate interaction probability, $x\gamma_{3,i}\alpha$.

The crucial difference from Section 4 is that normative risk compensation is reversed, as we see in the analogue to Theorem 1, which now concerns the exposure elasticity for *isolating* agents.

[12] Alternatively, the preferred equilibrium is the unique strong Nash equilibrium of the induced game (Aumann, 1959). This follows directly by noting that a higher λ is Pareto-improving. Moreover, the results hold as long as agents do not select an equilibrium in which an arbitrarily small measure can profitably deviate.

Theorem 2. *Let $E(\cdot,\cdot)$ be defined by (9), and let $\tilde{y}_s$ and $\tilde{y}_f$ be the interaction-weighted average payoff type (defined in (10)) in the second- and first-best optimal policy, respectively.*
If $y < y_f^*$ and $g_f(y) < 1$, then $\varepsilon_f(y) = E\left(y, \tilde{y}_f\right) > -1$.
If $y < y_s^*$ and $g_s(y) < 1$, then $E\left(y, \tilde{y}_s\right) \leqslant \varepsilon_s(y) < 0$, and the inequality is strict if $y_s^* < 1$.

Theorem 2 contains two results, like Theorem 1. The first is that both the second- and first-best selection policies exhibit exposure premia among isolating agents, i.e., $\varepsilon_i(\cdot) > -1$.

The second result is that the second-best now exhibits a *stronger exposure elasticity* than the first-best, for a given average payoff type $\tilde{y}$ among those interacting, i.e., if behavior is voluntary, then exposure receives *higher* priority. As before, an envelope argument implies that the first-best policy has an exposure elasticity of $E(y, \tilde{y})$, as in the baseline that takes interaction choices as given. On the other hand, in the second-best, the marginal spillover benefit is now amplified by the equilibrium response of agents; selecting some encourages the efficient interaction of others.

The second-best policy no longer sits between one that considers only private benefits and one that accounts for spillovers but takes agent behavior as given. Instead, it prioritizes exposure even more than what either policy would suggest.

To quantify normative risk compensation—the effect of the anticipated response to policy on the second-best policy—we again compute the marginal benefit $\gamma_{3,i}$ of increasing the aggregate interaction probability λ, which drives the spillover benefit of selection and so the exposure elasticity.

Proposition 4. *For $i \in \{s, f\}$ denoting the second- or first-best, the marginal benefit, $\gamma_{3,i}$, of increasing the aggregate interaction probability is given by*

$$\gamma_{3,f}\alpha = \lambda_f b \tilde{y}_f \qquad\qquad \gamma_{3,s}\alpha = \frac{\lambda_s b \tilde{y}_s + \Omega_1}{1 - \Omega_0},$$

where Ω_0 and Ω_1, defined in Proposition 2, are weakly positive and strictly so if and only if $y_s^* \leqslant 1$.

Ω_0 and Ω_1 now capture the *amplifying* rather than offsetting effect of equilibrium behavior on the marginal spillover benefit when behavior is voluntary. As the aggregate interaction probability rises, they count agents who would have chosen to isolate but now interact and were, respectively, non-selected or selected. This response raises welfare because in this case, agents do not internalize the positive externalities of interacting. If Ω_0 and Ω_1 are larger, then spillovers are stronger relative to private benefits, so priority in the second-best policy shifts from payoff type to exposure type.

5.1 Further Results

We briefly state the results analogous to Corollaries 1, 2, and 3.

Selection Policy for Interacting Agents – Optimal selection policy exhibits an exposure premium among agents who would interact if not selected.

Corollary 4. *For $i \in \{s, f\}$, if $y_i^* < 1$, then for all $y > y_i^*$, $\varepsilon_i(y) = 0$.*

Among the interacting, optimal selection policy not only exhibits an exposure premium, but is an x-policy, i.e., is invariant to payoff type y. Together with Theorem 2, it follows that optimal selection policies exhibit exposure premia both for interacting and isolating agents.

Monotonicity – If some non-selected agents interact in the second-best, i.e., $y_s^* < 1$, then the selection policy is *non-monotone*, whereas the first-best selection policy is always monotone.

Corollary 5. *If $y_s^* < 1$ then the second-best threshold function $g_s(\cdot)$ achieves its unique minimum at y_s^*. The first-best threshold function $g_f(\cdot)$ is continuous and decreasing.*

The intuition is the same as with negative externalities / strategic substitutes, but now comparing the isolating and the interacting, optimal policy under voluntary behavior focuses on the *isolating* because agents under-interact relative to the social optimum.

Implementation – The first-best interaction policy can be simply implemented with a strictly positive, flat interaction *subsidy.*

Corollary 6. *The policy maker can decentralize the optimal interaction policy in the first-best with a flat interaction subsidy equal to $\gamma_{3,f}\alpha$ per interaction if $y_f^* < 1$, and 0 otherwise.*

6 Economic Applications

6.1 Optimal Vaccination and Self-Isolation

A leading application for our analysis is optimal vaccine allocation, which maps to our results from Section 4. People differ both in their natural rate of interaction, x, as well as their severity of symptomatic response to the virus, y. Specifically, type (x, y) receives an interaction opportunity with probability x. Upon receiving the opportunity, they can pay a cost c to isolate and avoid the interaction. If not, they contract the virus with likelihood λ, the product of the number of other people interacting and the contagiousness of the virus, α.[13] Upon contracting the virus, they have an adverse reaction with probability y, which imposes cost b. A benevolent policy maker can select a subset of people to vaccinate, who can interact without any risk of contracting the virus or infecting others. We can also view the policy maker as spending a limited supply of resources to advertise or dispense vaccines.

Given a limited vaccine supply, whom should the policy maker target? Our findings demonstrate that the answer depends on whether lockdowns are mandatory (first-best) or voluntary (second-best), and in the latter case, on the responsiveness of behavior to the riskiness of the environment.[14]

Whether isolation is mandatory or voluntary, among those not sufficiently vulnerable to isolate if unvaccinated (y below y^*), the policy maker should weigh exposure x more heavily than vulnerability y to take into account the spillover risk highly exposed people impose on others when they interact. If isolation is voluntary or not strictly enforced, then the policy maker should weigh exposure less than if isolation were mandatory or than what holding fixed the behavior of the unvaccinated would imply is optimal.

More specifically, normative risk compensation (Theorem 1), implies that if the policy maker cannot enforce mandatory isolation, then they fear that the reduction in risk to the unvaccinated due to vaccination will relax voluntary incentives to isolate, and thus encourage some to inefficiently begin to interact. Such fears received widespread coverage in the case of COVID-19.[15] In practice, many authorities targeted vulnerability before exposure.[16] Our analysis provides a novel rationale for doing so: by vaccinating vulnerable people who do not interact heavily, the rate of contagion is relatively unaffected for those who interact, thus preserving incentives to voluntarily isolate.

Among those sufficiently vulnerable to isolate, the policy maker should target people who would like to interact frequently and therefore face particularly high isolation costs. If isolation is voluntary, then between the two groups—the isolating and the interacting—the policy maker

[13]We interpret this expression as capturing the steady state of a dynamic disease model in which conditional on leaving isolation, the likelihood of meeting a contagious person is increasing in the rate at which other people interact.

[14]One contributor is *lockdown fatigue*, a well-documented phenomenon whereby people fail to follow mandatory lockdown rules. See `https://tinyurl.com/y43jpxne` in the case of COVID-19.

[15]See `https://tinyurl.com/z7mh5bdx`.

[16]The WHO official guidelines suggest giving vaccine priority to health-workers as well as those above a certain age, as a proxy for vulnerability. See `https://tinyurl.com/nanndxx5`.

should target the latter more aggressively: non-monotonicity (Corollary 2) implies that optimal policy aggressively targets people with *intermediate* vulnerability—the most vulnerable that do not isolate—because they would isolate if they internalized the risk their interaction imposes on others. If isolation is mandatory, then the policy maker should not have any such preference for targeting the interacting over the isolating.

Finally, if the policy maker simply sets a price and sells the vaccine supply, then vaccines are allocated to people with the highest private benefit xy. Though this policy fails to internalize any spillovers, it is less costly if isolation is voluntary rather than mandatory. In that case, private benefits are closer to social benefits because spillover benefits are mitigated by risk compensation.

6.2 Information Aggregation

We next apply our results to a setting with information aggregation, which exhibits positive externalities and strategic complementarities and thus utilizes our results from Section 5. Users of a social media platform choose how much to engage with each other. They bear a cost to do so, subsequently both contributing to and benefiting from collective discussion to learn about an issue of common interest. Users differ both in their connectedness on the platform, x, as well as their interest in learning about the issue, y. A policy maker that works for the platform can select a limited subset of users to "promote" who no longer need to incur costs to engage. One interpretation is that promoted users are made more salient on the platform, and do not need to work to get others to engage with them and convey information. An example is the verified blue check mark on Twitter. Prior to 2022, Twitter allocated blue check marks to selected users for free, and a check significantly promoted a user's visibility (Hearn, 2017). After a change in ownership in 2022, the blue check became a paid service, available to all.

Formally, users are interested in an uncertain state $\theta \in \mathbb{R}$ over which they each hold a common prior $\theta \sim G(\cdot)$. A type (x, y) user gets the opportunity to interact with probability x, which is free if they are promoted, but costs $c > 0$ if not. If they interact, they receive a signal $s \sim \pi(\cdot \mid \theta, \lambda)$ that induces a posterior distribution $\hat{G}(\cdot \mid s, \lambda)$, where $\lambda \equiv \mu$ is the mass of other users interacting.[17] The signal structure is such that the precision of the posterior distribution is increasing in λ. Finally, we suppose that each user has a preference for precision, specifically that their final payoff is λby. We

[17] We set $\alpha = 1$ for simplicity. Otherwise, we can interpret α as scaling signal precision.

can micro-found this setup further by supposing that users receive private signals, that interacting users share their signals with each other, and that users must take an action the value of which depends on estimating θ.[18]

Given that the policy maker can only promote a limited number of users, whom should they target? For those not sufficiently interested in the issue at hand to engage if not promoted (y below y^*), the policy maker should weigh connectedness x more heavily than interest in the issue at hand y to take into account the spillover benefits highly connected users provide to others. Moreover, if interaction is voluntary, then due to normative risk compensation (Theorem 2), the policy maker should target connectedness *even more so* than under mandated engagement, or taking the interaction choices of unpromoted users as given. Promoting the highly connected encourages some users to engage who would otherwise inefficiently choose not to. Among those sufficiently interested to engage, the policy maker should promote highly connected users that otherwise spend a lot of effort interacting with others. If interaction is voluntary, then non-monotonicity (Corollary 5) implies that between the two groups—those putting in effort to interact and those not willing to do so if not promoted—the policy maker should prioritize the latter because they do not internalize the benefits their interactions create for others.

In stark contrast to the vaccination application, if interaction is voluntary then selling a fixed amount of promotion, i.e., allocating it to users with the highest private benefit, is particularly inefficient because spillover benefits are amplified by risk compensation: highly connected users do not internalize that their promotion benefits other interacting users *and* encourages others to interact. Applying our result to Twitter implies that decentralizing access to the blue check mark through prices inefficiently targets users who are not so well connected—and thus do not provide much value to others—but place a high private value on interacting with and learning from others.

Regarding implementation of the first-best interaction policy, suppose the policy maker can effectively subsidize interaction, for instance through promotional advertising. Corollary 6 tells us that the policy maker should subsidize non-interacting agents that place the highest value on information (y just below y^*). Moreover, to push each such agent to interact, the policy maker should advertise to an agent in proportion to their connectedness, i.e., advertising expenditures on a type (x, y) agent scale with x. Even though more connected agents who would not otherwise

[18]See Herskovic and Ramos 2020 for a related example in the context of endogenous network formation.

interact face the highest private cost from putting in the effort to do so, their interaction generates the most information and therefore the highest benefit to others. Indeed, Bond et al. (2012) and Jones et al. (2017) document empirical evidence showing that Facebook positively distorted voter turnout through targeted advertising during the 2010 US Congressional Elections and 2012 US Presidential Elections. Their results confirm that network centrality was an important determinant of the intervention's effectiveness, specifically by encouraging others to participate more actively.

7 Conclusion

We introduced a novel framework to study targeted interventions within groups of interacting, heterogeneous agents. Agents differ in their exposure to others, as well as their payoffs from interaction, and can exert costly effort to change their interaction rates. We study two cases, in which an agent's interaction imposes negative externalities on other interacting agents, and so interaction choices form strategic substitutes, and in which an agent's interaction imposes positive externalities on other interacting agents, and so interaction choices form strategic complements. A benevolent policy maker designs a *selection policy*, which alters the interaction patterns of a targeted subset of agents at zero cost, subject to a capacity constraint.

Our main insight, which we term *normative risk compensation*, demonstrates how the voluntary interaction decisions of non-selected agents shape optimal selection policy. The policy maker always takes into account spillovers and thus bases selection more on an agent's exposure to others than on their payoff from interaction. However, if interaction entails negative externalities, then relative to what holding fixed the behavior of the non-selected would suggest is optimal, the policy maker shifts selection away from highly exposed agents toward agents with particularly high negative payoffs from interaction, and thus shifts priority from spillover benefits to the non-selected toward private benefits to the selected. The opposite holds under positive externalities. In each case, the inefficiency of voluntary agent behavior implies that one goal of selection policy should be to manipulate non-selected agents' values of interacting, and thus their behavior.

We detailed two leading settings to which our framework and results port well, and believe many more applications can be studied using our model, or slight variations of it. For instance, we can apply our results from the negative externalities / strategic substitutes case to a setting in which firms can pay to differentiate their products and thus escape competition, and a policy maker can select a subset of firms to promote for free. Alternatively, in a congested traffic network, agents can incur costs to avoid participating, and a policy maker can select a subset of agents that are allowed to use an alternative uncongested route. Finally, we can apply our results from the positive externalities / strategic complements case to a setting in which firms can pay to adopt a technology that generates private profits as well as information for other adopters, and a policy maker can select a subset of firms to use the technology for free.

References

Akbarpour, M., E. Budish, P. Dworczak, and S.D. Kominers, "An Economic Framework for Vaccine Prioritization," *conditionally accepted, Quarterly Journal of Economics*, 2022.

__ , S. Maladi, and A. Saberi, "Just a Few Seeds More: Value of Network Information for Diffusion," *Working Paper*, 2020.

Akcigit, U., D. Hanley, and S. Stantcheva, "Optimal taxation and R&D policies," *Econometrica*, 2022, *90* (2), 645–684.

Aumann, R., *Acceptable Points in General Cooperative n-Person Games. In Contributions to the Theory of Games*, Vol. 4, Contributions to the Theory of Games, 1959.

Ballester, C., A. Calvó-Armengol, and Y. Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 2006, *74* (5), 1403–1417.

Bond, R.M., C.J. Fariss, J.J. Jones, A.D.I. Kramer, C. Marlow, J.E. Settle, and J.H. Fowler, "A 61-million-person experiment in social influence and political mobilization," *Nature*, 2012, *489*, 295–298.

Corchón, L.C., "Aggregative games," *SERIEs*, 2021, *12*, 49–71.

Demange, G., "Contagion in financial networks: a threat index," *Management Science*, 2016, *64* (2), 955–970.

Ehrlich, I. and G.S. Becker, "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy*, 1972, *80* (4), 623–648.

Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond, "Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations," *PLoS ONE*, 2008, *3* (12).

Farboodi, M., G. Jarosch, and R. Shimer, "The Emergence of Market Structure," *Working Paper*, 2020.

Galeotti, A., B. Golub, and S. Goyal, "Targetting Interventions in Networks," *Econometrica*, 2020, *88* (6), 2445–2471.

Geoffard, P.Y. and T. Philipson, "Rational Epidemics and Their Public Control," *International Economic Review*, 1996, *37* (3), 603–624.

Greenwood, J., P. Kircher, C. Santos, and M. Tertilt, "An Equilibrium Model of the African HIV/AIDS Epidemic," *Econometrica*, 2019, *87* (4), 1081–1113.

Gross, T., C. J. D. D'Lima, and B. Blasius, "Epidemic Dynamics on an Adaptive Network," *Phys Rev Lett*, 2006, *96* (20).

Hearn, A., "Verified: Self-presentation, identity management, and selfhood in the age of big data," *Popular Communication*, 2017, *15* (2), 62–77.

Herskovic, B. and J. Ramos, "Acquiring Information through Peers," *American Economic Review*, 2020, *110* (7), 2128–52.

Hoy, M. and M.K. Polborn, "The value of technology improvements in games with externalities: a fresh look at offsetting behavior," *Journal of Public Economics*, 2015, *131*, 12–20.

Jehiel, P. and B. Moldovanu, "Efficient Design with Interdependent Valuations," *Econometrica*, 2001, *69* (5), 1237–1259.

_ , _ , and E. Stacchetti, "How (not) To Sell Nuclear Weapons," *American Economic Review*, 1996, *86* (4), 814–829.

Jones, J.J., R.M. Bond, E.Bakshy, D.Eckles, and J.H. Fowler, "Social influence and political mobilization: Further evidence from a randomized experiment in the 2012 U.S. presidential election," *PLoS ONE*, 2017, *12* (4).

Kremer, M., "Integrating Behavioral Choice into Epidemiological Models of AIDS," *Quarterly Journal of Economics*, 1996, *111*, 549–573.

Ostrizek, F. and E. Sartori, "Screening While Controlling an Externality," *working paper*, 2022.

Pastor-Satorras, R. and A. Vespignani, "Immunization of Complex Networks," *Phys Rev E*, 2002, *65* (3).

Piketty, T., E. Saez, and S. Stantcheva, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 2014, *6* (1), 230–271.

Shaw, L. B. and I. B. Schwartz, "Fluctuating Epidemics on Adaptive Networks," *Phys Rev E*, 2008, *77* (6).

Talamás, E. and R. Vohra, "Free and Perfectly Safe but Only Partially Effective Vaccines Can Harm Everyone," *Games and Economic Behaviour*, 2020, *122*, 277–289.

Wardrop, J.G., "Some theoretical aspects of road traffic research," *Proceedings of the Institute of Civil Engineers*, 1952, *1* (2), 325–378.

A Heterogeneous Costs

We extend our analysis in the negative externalities / strategic substitutes case to allow for heterogeneous, unobservable, and independently drawn costs of isolation. (The insights for the positive externalities case are similar.) Specifically, the isolation cost c is now drawn independently for each agent from a continuously differentiable distribution H with support $[0,\infty]$, and density h. Crucially, the policy maker cannot condition selection on an agent's realized c.[19] The first-best policy maker can condition interaction policy on c, which facilitates a clear comparison with the second-best, and is feasible if the policy maker implements with a flat interaction tax, as in Corollary 3.

This exercise serves several purposes. First, in many applications, agents face different isolation costs. Second, it sheds light on the more general forces that drive the results in the baseline model. Finally, it is qualitatively similar to an extension in which the total isolation cost is strictly convex rather than linear in $1-\sigma(x,y)$, which would also smooth the discontinuity in behavior at y^*.

The following proposition characterizes the first- and second-best policies.

Proposition A.1. *Optimal interaction policies are characterized by isolation cost threshold functions*

$$c_s^*(y) = \lambda_s by \qquad\qquad c_f^*(y) = \lambda_f by + \gamma_{3,f}\alpha,$$

such that, for $i \in \{s, f\}$, an agent with payoff type y interacts if and only if their cost draw satisfies $c > c_i^*(y)$.

Optimal selection policies are characterized by the threshold function

$$g_i(y) = \min\left\{\frac{\gamma_{1,i}}{\left(1-H\left(c_i^*(y)\right)\right)\left(\lambda_i by + \gamma_{3,i}\alpha\right) + \int_0^{c_i^*(y)} ch(c)dc}, 1\right\}.$$

Optimal selection policies are qualitatively as before, but the intuition for isolating and interacting agents are blended because type (x, y) now interacts or isolates depending on their isolation cost draw. Specifically, the benefit of selecting an agent consists of two components: the benefit of reducing interaction, which scales with the likelihood of interaction $1 - H\left(c_i^*(y)\right)$, and of avoiding expected isolation costs $\int_0^{c_i^*(y)} ch(c)dc$.

On normative risk compensation, the following theorem is analogous to Proposition 2 and Theorem 1:

Theorem A.1. *The marginal benefit, $\gamma_{3,i}$, of reducing the aggregate interaction probability λ is given by*

$$\gamma_{3,f}\alpha = \lambda_f b\tilde{y}_f \qquad\qquad \gamma_{3,s}\alpha = \frac{\lambda_s b\tilde{y}_s}{1 + \alpha\int_0^1\int_0^{g_s(y)} h\left(c_s^*(y)\right) xbyf(x,y)\,dxdy},$$

[19] Were each agent's cost observable, it would be irrelevant for interacting agents and would play the same role as x for isolating agents, scaling their cost of isolation.

where $\tilde{y}_i$ *is the interaction-weighted average payoff type:*

$$\tilde{y} = \frac{\int_0^1 \int_0^1 yx(1 - v(x,y))H(c^*(y))f(x,y)\,dxdy}{\int_0^1 \int_0^1 x(1 - v(x,y))H(c^*(y))f(x,y)\,dxdy}.$$

In the first-best, if $g_f(y) < 1$*, then the exposure elasticity is*

$$\varepsilon_f(y) = \frac{-\left(1 - H\left(c_f^*(y)\right)\right)\lambda_f by}{\left(1 - H\left(c_f^*(y)\right)\right)\lambda_f b\,(y + \tilde{y}_f) + \int_0^{c_f^*(y)} ch(c)dc} > -1.$$

In the second-best, if $g_s(y) < 1$*, then* $\varepsilon_s(y) > -1$ *and there exists a* $\bar{h}_s > 0$ *such that*

$$\varepsilon_s(y) < \frac{-\left(1 - H\left(c_s^*(y)\right)\right)\lambda_s by}{\left(1 - H\left(c_s^*(y)\right)\right)\lambda_s b\,(y + \tilde{y}_s) + \int_0^{c_s^*(y)} ch(c)dc}$$

if and only if

$$\frac{h\left(c_s^*(y)\right)}{1 - H\left(c_s^*(y)\right)} < \bar{h}_s.$$

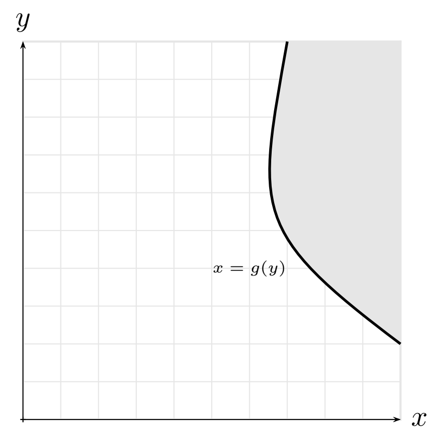
As in the baseline, the first- and second-best policies exhibit exposure premia, i.e., $\varepsilon_i(\cdot) > -1$. Thus, they prioritize exposure type x over payoff type y in that they select high x agents at lower levels of xy.

Due to normative risk compensation, as long as the hazard rate $\frac{h(c_s^*(y))}{1-H(c_s^*(y))}$ is sufficiently low, the second-best exhibits a weaker exposure elasticity than the first-best, for a given interaction-weighted average payoff type $\tilde{y}$ and expected isolation cost $\int_0^{c^*(y)} ch(c)dc$. The denominator in the expression for $\gamma_{3,s}$ reflects that as the aggregate interaction probability falls, for each payoff type y, agents on the interaction cost threshold $c_s^*(y)$ inefficiently shift from isolating to interacting.

The continuous distribution of isolation costs on which selection policy cannot condition complicates the exposure elasticity expressions, and the comparison between the first- and second-best. In the first-best, the exposure elasticity is the component of the private benefit of selection that scales with y—the expected interaction payoff but not the expected isolation cost—relative to the social benefit, rather than the entire private benefit relative to the social benefit. In the second-best, *going against normative risk compensation*, the policy maker shifts priority from payoff type to exposure type because agents with high payoff types interact less often, and agents over-interact relative to the social optimum. The final result in the theorem follows because the strength of this force is increasing in the density $h\left(c_s^*(y)\right)$, and the strength of normative risk compensation is increasing in the probability that a payoff type y agent chooses to interact, $1 - H\left(c_s^*(y)\right)$.

This last effect is the same force that drives non-monotonicity of the second-best selection policy in the baseline model (Corollary 2), except now it is active for all payoff types, rather than just at the interaction

Figure 3: Second-Best Policy with Heterogeneous Costs



Solid line: policy threshold function. Shaded area: $v(x, y) = 1$.

threshold y_s^*. The following corollary characterizes when it generates non-monotonicity in the general model.

Corollary A.1.

1. *$g_f(\cdot)$ is continuous and strictly decreasing.*

2. *$g_s(\cdot)$ is continuous. $g_s'(y) \leqslant 0$ if and only if*

$$\frac{h\left(c_s^*(y)\right)}{1 - H\left(c_s^*(y)\right)} \leqslant \frac{1}{\gamma_{3,s}\alpha}.$$

The second-best policy is continuous unlike in the baseline model because the distribution of isolation costs is continuous. Nonetheless, it may still be non-monotone. The private benefit of selecting an agent is strictly increasing in their payoff type as long as they interact with some probability. In the baseline model, this is the only force for interacting agents. Now however, the spillover benefit of selecting an agent is strictly decreasing in their payoff type because the probability of interaction is decreasing in payoff type. The first effect scales with $1 - H\left(c_s^*(y)\right)$, and the second scales with $h\left(c_s^*(y)\right)$, so which force dominates depends on the hazard rate $\frac{h(\cdot)}{1-H(\cdot)}$. Figure 3 illustrates an example in which the hazard rate is strictly increasing.

A.1 Proofs of Proposition A.1, Theorem A.1, and Corollary A.1

Second-best – For the individually optimal interaction policy, note that a non-selected agent with type (x, y) and cost draw c interacts if and only if the benefit of forgoing the isolation cost exceeds the expected negative interaction payoff, i.e., $xc > x\lambda_s by$. It follows that an agent with type (x, y) interacts with probability $1 - H(c_s^*(y))$, and pays an expected isolation cost of $\int_0^{c_s^*(y)} cxh(c)dc$, where $c_s^*(\cdot)$ is as defined in the proposition.

Plugging in, the second-best policy maker maximizes the Lagrangian

$$\begin{aligned}\mathcal{L}_s =& \mathcal{W}_s + \gamma_{1,s}\left(\beta - \int_0^1 \int_0^1 v_s(x,y) f(x,y)\, dxdy\right) \\ &+ \gamma_{3,s}\left(\lambda_s - \alpha \int_0^1 \int_0^1 (1 - v_s(x,y))\left(1 - H\left(c_s^*(y)\right)\right) x f(x,y)\, dxdy\right),\end{aligned}$$

where welfare is

$$\begin{aligned}\mathcal{W}_s =& \int_0^1 \int_0^1 \int_0^\infty xch(c) f(x,y) dcdxdy \\ &- \int_0^1 \int_0^1 (1 - v_s(x,y))\left[\left(1 - H\left(c_s^*(y)\right)\right) x\lambda_s by + \int_0^{c_s^*(y)} xch(c)dc\right] f(x,y)\, dxdy,\end{aligned}$$

and where we call the second Lagrange multiplier $\gamma_{3,s}$ to facilitate comparison with the baseline model. The same argument as for Lemma B.1 shows that $\gamma_{1,s} > 0$. To see that $\gamma_{3,s} > 0$, take the derivative of the Lagrangian with respect to λ_s:

$$\begin{aligned}\frac{\partial \mathcal{L}_s}{\partial \lambda_s} =& - \int_0^1 \int_0^1 (1 - v_s(x,y))\left(1 - H\left(c_s^*(y)\right)\right) xbyf(x,y)\, dxdy \\ &+ \gamma_{3,s}\left(1 + \alpha \int_0^1 \int_0^1 (1 - v_s(x,y)) h\left(c_s^*(y)\right) xbyf(x,y)\, dxdy\right),\end{aligned}$$

where the effects on welfare through changes in $c_s^*(\cdot)$, holding fixed λ_s, net to 0. The First Order Condition for λ_s, setting the derivative to 0, then implies that

$$\gamma_{3,s} = \frac{\int_0^1 \int_0^1 (1 - v_s(x,y))\left(1 - H\left(c_s^*(y)\right)\right) xbyf(x,y)\, dxdy}{1 + \alpha \int_0^1 \int_0^1 (1 - v_s(x,y)) h\left(c_s^*(y)\right) xbyf(x,y)\, dxdy},$$

which is strictly greater than 0.

To find the optimal selection policy, take the derivative of the Lagrangian with respect to $v_s(x, y)$:

$$\frac{\partial \mathcal{L}_s}{\partial v_s(x,y)} = \left[\left(1 - H\left(c_s^*(y)\right)\right) x\lambda_s by + \int_0^{c_s^*(y)} xch(c)dc - \gamma_{1,s} + \gamma_{3,s}\alpha\left(1 - H\left(c_s^*(y)\right)\right) x\right] f(x,y).$$

The derivative is strictly increasing in x and the second derivative with respect to $v_s(x, y)$ is 0. It follows that the optimal policy is of the form in the proposition.

The remaining results for the second-best follow from taking the derivative of $g_s(\cdot)$. In particular, if

$g_s(y) < 1$, then plugging in $c_s^*(y) = \lambda_s by$, the exposure elasticity is

$$\varepsilon_s(y) = \frac{-\left(1 - H\left(c_s^*(y)\right)\right)\lambda_s by}{\left(1 - H\left(c_s^*(y)\right)\right)\left(\lambda_s by + \gamma_{3,s}\alpha\right) + \int_0^{c_s^*(y)} ch(c)dc}\left[1 - \frac{h\left(c_s^*(y)\right)}{1 - H\left(c_s^*(y)\right)}\gamma_{3,s}\alpha\right].$$

Since $\gamma_{3,s} > 0$, it follows that the first fraction outside the brackets is negative but strictly greater than -1, and that the difference inside the brackets is strictly less than 1. Thus, the exposure elasticity is strictly greater than -1. Moreover,

$$\varepsilon_s(y) < \frac{-\left(1 - H\left(c_s^*(y)\right)\right)\lambda_s by}{\left(1 - H\left(c_s^*(y)\right)\right)\lambda_s b\left(y + \tilde{y}_s\right) + \int_0^{c_s^*(y)} ch(c)dc}$$

if and only if

$$\frac{h\left(c_s^*(y)\right)}{1 - H\left(c_s^*(y)\right)} < \frac{\left(1 - H\left(c_s^*(y)\right)\right)\left(\frac{\lambda_s b\tilde{y}_s}{\gamma_{3,s}\alpha} - 1\right)}{\left(1 - H\left(c_s^*(y)\right)\right)\left(\lambda_s by + \lambda_s b\tilde{y}_s\right) + \int_0^{c_s^*(y)} ch(c)dc},$$

which is strictly positive since $\gamma_{3,s}\alpha < \lambda_s b\tilde{y}_s$.

First-best – The Lagrangian is as in the baseline model, but welfare is now

$$\begin{aligned}\mathcal{W}_f = &\int_0^1 \int_0^1 \int_0^\infty xch(c)f(x,y)dcdxdy \\ &- \int_0^1 \int_0^1 \left(1 - v_f(x,y)\right)\left[\sigma_f(x,y)x\lambda_f by + \int_0^\infty \left(1 - \tilde{\sigma}_f(x,y,c)\right)xch(c)dc\right] f(x,y)\,dxdy,\end{aligned}$$

where $\tilde{\sigma}_f(x,y,c)$ is the probability that a non-selected agent with type (x,y) and cost draw c chooses to interact, and $\sigma_f(x,y)$ is the probability that a non-selected agent with type (x,y) interacts, integrating over all possible cost draws: $\sigma_f(x,y) = \int_0^\infty \tilde{\sigma}_f(x,y,c)h(c)dc$. To find the optimal interaction policy, take the derivative of the Lagrangian with respect to $\tilde{\sigma}_f(x,y,c)$:

$$\frac{\partial \mathcal{L}_f}{\partial \tilde{\sigma}_f(x,y,c)} = -\left(1 - v_f(x,y)\right)\left[x\lambda_f by - xc\right]h(c)f(x,y) - \gamma_{3,f}\alpha(1 - v_f(x,y))xh(c)f(x,y).$$

The second derivative with respect to $\tilde{\sigma}_f(x,y,c)$ is 0 and the first derivative is strictly increasing in c for non-selected agents. It follows that an agent with type (x,y) interacts with probability $1 - H\left(c_f^*(y)\right)$, and pays an expected isolation cost of $\int_0^{c_f^*(y)} cxh(c)dc$, where $c_f^*(\cdot)$ is as defined in the proposition.

The same argument from Lemma B.1 shows that $\gamma_{1,f} > 0$. To see that $\gamma_{3,f} > 0$, take the derivative of the Lagrangian with respect to λ_f:

$$\frac{\partial \mathcal{L}_f}{\partial \lambda_f} = -\int_0^1 \int_0^1 \left(1 - v_f(x,y)\right)\left(1 - H\left(c_f^*(y)\right)\right)xbyf(x,y)\,dxdy + \gamma_{3,f},$$

where the effects on the Lagrangian through changes in $c_f^*(\cdot)$ net to 0. The First Order Condition for λ_f, setting the derivative to 0, then implies that

$$\gamma_{3,f} = \int_0^1 \int_0^1 (1 - v_f(x,y)) \left(1 - H\left(c_f^*(y)\right)\right) xbyf(x,y)\, dxdy,$$

which is strictly greater than 0.

The remainder of the proof follows the same structure as for the second-best. In particular, if $g_f(y) < 1$, then plugging in $c_f^*(y) = \lambda_f by + \gamma_{3,f}\alpha$ yields the exposure elasticity.

B Proofs

B.1 Proof of Lemma 1

First, consider the individually optimal interaction choice of a type (x, y) agent, taking as given the aggregate interaction probability λ. The benefit of isolating is avoiding the risk of interaction, which generates expected payoff $x\lambda by$, and the cost is xc. Thus, if $y < c/(\lambda b)$, then the agent maximizes the probability of interaction, i.e., sets $\sigma(x, y) = 1$, and if $y > c/(\lambda b)$, then the agent sets $\sigma(x, y) = 0$. The agent is indifferent if $y = c/(\lambda b)$, and there is a measure 0 of such agents, so their decision is not significant.

Next, take as given a selection policy v. The equilibrium under individually optimal behavior is unique because if y_s^* increases, then more agents interact, which implies a higher λ by its definition in equation (2), and so a lower y_s^*. In the unique equilibrium, $y_s^* > 0$ and $\lambda > 0$ because in either case, $\lambda = 0$, which implies that all agents optimally interact, and so $\lambda > 0$ since there is insufficient supply to select all agents ($\beta < 1$).

Finally, consider the socially optimal interaction policy. Suppose the policy does not follow a threshold rule for a positive measure of agents, i.e., there exists an $\epsilon > 0$ and a y' such that there is a positive measure of agents with $y < y'$ and $\sigma(x, y) < 1 - \epsilon$, and a positive measure of agents with $y > y'$ and $\sigma(x, y) > \epsilon$. We show that the planner can change the interaction policy to strictly improve welfare. The policy maker can reduce $\sigma(x, y)$ for a positive measure of agents with $y > y'$ and increase $\sigma(x, y)$ for a positive measure of agents with $y < y'$ such that the aggregate interaction probability λ is held fixed. Since λ is held fixed, each agent's expected payoff given $\sigma(x, y)$ is held fixed. Thus, the effect of the policy change on welfare is the integral across the affected agents of the changes in their individual payoffs from changing their probabilities of interaction. This integral is the same as the effect on the aggregate interaction probability, except the integrand for each agents is multiplied by $(b\lambda y - c)/\alpha$. Thus, since the effect on the aggregate interaction probability is 0 and since the agents with an increase in $\sigma(x, y)$ have a lower y than the agents with a decrease in $\sigma(x, y)$, it follows that the effect on welfare of the change is strictly positive.

B.2 Proof of Proposition 1

First, we prove the last part of the proposition concerning the first-best interaction threshold, i.e., that $y_f^* = (c - \gamma_{3,f}\alpha)/(\lambda_f b)$, where y_f^* and λ_f are strictly positive.

Proof. The first derivative of the policy maker's Lagrangian in equation (6) with respect to y_f^* is

$$\int_0^1 (xc - x\lambda_f b y_f^*)(1 - v_f(x, y_f^*))f(x, y_f^*)dx - \gamma_{3,f}\alpha \int_0^1 (1 - v_f(x, y_f^*))xf(x, y_f^*)dx$$
$$= (c - \lambda_f b y_f^* - \gamma_{3,f}\alpha) \int_0^1 (1 - v_f(x, y_f^*))xf(x, y_f^*)dx,$$

where the first integral on the first line is the private effect of switching non-selected agents with $y = y_f^*$ from isolating to interacting, and the second integral is the spillover effect that results from increasing the equilibrium aggregate interaction probability λ_f. Recall that this expression holds even for $y_f^* \geqslant 1$ if we extend f to be defined on all of $\mathbb{R}^2$, and equal to 0 outside of $[0, 1]^2$.

First, $y_f^* > 0$ and $\lambda_f > 0$. If $y_f^* = 0$ or $\lambda_f = 0$, then $\lambda_f = 0$, which implies that switching some agents from isolating to interacting yields them strictly positive benefits (since $\lambda_f = 0$) and generates no externalities (since no other agents are interacting). Non-selected isolating agents must exist if $\lambda_f = 0$ because there is insufficient supply to select all agents, i.e., $\beta < 1$. More formally, we will see in the proof of Lemma B.3 that if $\lambda_f = 0$, then $\gamma_{3,f} = 0$, which implies that the derivative above is strictly positive.

Now, suppose $c - \lambda_f b y_f^* - \gamma_{3,f}\alpha < 0$. If there exists an $\epsilon > 0$ such that the measure of non-selected agents with $y \in [y_f^* - \epsilon, y_f^*]$ is 0, then the policy maker can decrease y_f^* without any effect on welfare. If there does not exist such an $\epsilon > 0$, then the policy maker can strictly improve welfare by decreasing y_f^*.

Finally, suppose $c - \lambda_f b y_f^* - \gamma_{3,f}\alpha > 0$. If there exists an $\epsilon > 0$ such that the measure of non-selected agents with $y \in [y_f^*, y_f^* + \epsilon]$ is 0, then the policy maker can increase y_f^* without any effect on welfare. If there does not exist such an $\epsilon > 0$, then the policy maker can strictly improve welfare by increasing y_f^*.

Thus, $\lambda_f b y_f^* = c - \gamma_{3,f}\alpha$, which must be strictly positive. □

Next, the following lemma states that the policy maker always strictly prefers greater supply for selection.

Lemma B.1. *For $i \in \{s, f\}$, the Lagrange multiplier on the supply constraint, $\gamma_{1,i}$, is strictly positive.*

Proof. We show that if the policy maker can select additional agents, then welfare strictly increases. It is sufficient to show that holding fixed the rest of the selection policy, the policy maker can strictly increase welfare. In the case of the first-best, we can also hold fixed the interaction policy. If there is a strictly positive measure of isolating agents, then they can be selected for a strictly positive private benefit without any effect on individually optimal interaction decisions. As such, welfare strictly increases. If there is not a strictly positive measure of isolating agents, then since supply is insufficient to select all agents, i.e., $\beta < 1$, there is a strictly positive measure of non-selected interacting agents. Holding fixed interaction decisions, selecting some of these agents gives them a strictly positive private benefit, and generates spillovers to the interacting agents who remain non-selected that yield a strictly positive public benefit. Even in the second-best with individually optimal interaction decisions, this increase in selection cannot lead to an increase in the aggregate interaction probability because all non-selected agents are already interacting. If non-selected agents respond by isolating, then their private value must go up, and there are further public benefits to the non-selected that continue to interact. $\square$

We now prove a variation on the first part of Proposition 1. The function we define here is only different from the one defined in the proposition if the Lagrange multiplier on the equilibrium aggregate interaction probability constraint, $\gamma_{3,i}$, is negative. We first show that the function takes the form defined here without knowing whether $\gamma_{3,i} > 0$, and then complete the proof of Proposition 1 by showing that $\gamma_{3,i} > 0$.

Lemma B.2. *Optimal selection policy for $i \in \{s, f\}$ is characterized by a threshold function $g_i(\cdot)$ such that*

$$g_i(y) = \begin{cases} g_i^* & y \in (y_i^*, 1] \\ min\left\{\frac{\gamma_{1,i}}{\lambda_i by + \gamma_{3,i}\alpha}, 1\right\} & y \in (y_{0,i}, y_i^*] \\ 1 & y \in [0, y_0], \end{cases}$$

where

$$g_i^* \equiv min\left\{\frac{\gamma_{1,i}}{c}, 1\right\}$$

and

$$y_{0,i} \equiv -\frac{\gamma_{3,i}\alpha}{\lambda_i b}.$$

Proof. First, consider type (x, y) with $y > y_i^*$. The first derivative of the policy maker's Lagrangian in equation (6) with respect to $v_i(x, y)$ is

$$(xc - \gamma_{1,i})f(x, y),$$

where the first term is the benefit that the agent no longer has to pay the isolation cost, and the second term is the selection supply cost. The second derivative is 0. It follows that the policy maker optimally sets $v_i(x, y) = 1$ if $x > \gamma_{1,i}/c$ and optimally sets $v_i(x, y) = 0$ if $x < \gamma_{1,i}/c$. A measure 0 of agents have precisely $x = \gamma_{1,i}/c$, so in that case, any choice of $v_i(x, y)$ is optimal.

Next, consider type (x, y) with $y \in (y_{0,i}, y_i^*]$. The first derivative of the Lagrangian with respect to $v_i(x, y)$ is

$$(x\lambda_i by - \gamma_{1,i} + \gamma_{3,i}\alpha x)f(x, y),$$

where the first term is the private benefit to the agent, the second term is the selection supply cost, and the third term captures the externality of lowering the equilibrium aggregate interaction probability. Again, the second derivative is 0. Since $y > y_{0,i}$, we know that $\lambda_i by + \gamma_{3,i}\alpha > 0$, which implies that the first derivative is strictly increasing in x. It follows that the policy maker optimally sets $v_i(x, y) = 1$ if $x > \gamma_{1,i}/(\lambda_i by + \gamma_{3,i}\alpha)$ and optimally sets $v_i(x, y) = 0$ if the opposite holds. A measure 0 of agents have precisely $x = \gamma_{1,i}/(\lambda_i by + \gamma_{3,i}\alpha)$, so any choice of $v_i(x, y)$ for such an x is optimal.

Finally, consider type (x, y) with $y \leqslant y_{0,i}$. The first derivative of the planner's Lagrangian with respect to $v_i(x, y)$ is the same as in the previous case with $y \in (y_{0,i}, y_i^*]$, and again the second derivative is 0. Since $y \leqslant y_{0,i}$, we know that $\lambda_i by + \gamma_{3,i}\alpha \leqslant 0$. Since $x \geqslant 0$ and $\gamma_{1,i} > 0$, it follows that the first derivative is strictly negative. Thus, the policy maker optimally sets $v_i(x, y) = 0$. □

We can now complete the proof of Proposition 1 with the following lemma.

Lemma B.3. *For $i \in \{s, f\}$, the Lagrange multiplier on the equilibrium aggregate interaction rate constraint, $\gamma_{3,i}$, is strictly positive.*

Proof. From Lemma B.2, we can write the policy maker's Lagrangian as

$$\begin{aligned}\mathcal{L}_i =& \mathcal{W}_i + \gamma_{1,i}\left(\beta - \int_0^1 \int_{g_i(y)}^1 f(x,y)\,dxdy\right) + \gamma_{2,i}\left(y_i^* - \frac{c}{\lambda_i b}\right)\\ &+ \gamma_{3,i}\left(\lambda_i - \alpha\int_0^{y_i^*}\int_0^{g_i(y)} xf(x,y)\,dxdy\right),\end{aligned}$$

and welfare as

$$\mathcal{W}_i = -\int_0^{y_i^*}\int_0^{g_i(y)} x\lambda_i byf(x,y)\,dxdy - \int_{y_i^*}^1\int_0^{g_i^*} xcf(x,y)\,dxdy.$$

We will need to use the First Order Conditions (FOC) for y_i^* and λ_i, so we begin with the derivative of the Lagrangian with respect to each:

$$\frac{\partial\mathcal{L}_i}{\partial y_i^*} = \int_0^{g_i^*} xcf\left(x, y_i^*\right)dx - \int_0^{g_i(y_i^*)} (\lambda_i by_i^* + \gamma_{3,i}\alpha)xf\left(x, y_i^*\right)dx - \int_{g_i(y_i^*)}^{g_i^*} \gamma_{1,i}f\left(x, y_i^*\right)dx + \gamma_{2,i},$$

where the first term is the reduction in isolation costs, the second term is the increase in costs associated with interaction, the third term is the cost (or benefit, depending on whether $g_i\left(y_i^*\right) < g_i^*$) of the change in the quantity of agents selected implied by a discontinuity in $g_i(\cdot)$ at y_i^*, and the fourth term is the Lagrange multiplier on the incentive compatibility constraint for the interaction threshold (which is potentially non-zero only if $i = s$); and

$$\frac{\partial\mathcal{L}_i}{\partial\lambda_i} = -\int_0^{y_i^*}\int_0^{g_i(y)} xbyf(x,y)\,dxdy + \gamma_{2,i}\frac{c}{\lambda_i^2 b} + \gamma_{3,i},$$

where the first term is the cost of increased risk for interacting agents, the second term is the effect on the incentive compatibility constraint for the individually optimal interaction threshold (for $i = s$), and the third term is the Lagrange multiplier on the equilibrium aggregate interaction probability constraint.

The necessary FOC for each of λ_i and y_i^* is that the derivative of the Lagrangian with respect to each is 0. We saw in Lemma 1 that λ_s and y_s^* are strictly positive, and in the proof at the beginning of this section that λ_f and y_f^* are strictly positive. Moreover, neither λ_i nor y_i^* has an upper bound.

For the first-best, $\gamma_{2,f} = 0$, so the FOC for λ_f yields

$$\gamma_{3,f} = \int_0^{y_f^*} \int_0^{g_f(y)} xbyf(x,y)\, dxdy, \tag{B.1}$$

which is strictly positive because $y_f^* > 0$ and $g_f(\cdot) > 0$. For the second-best, plugging in the incentive compatibility constraint, $\lambda_s b y_s^* = c$, into the FOC for y_s^* yields

$$\gamma_{2,s} = \gamma_{3,s}\alpha \int_0^{g_s^*} xf\left(x, y_s^*\right) dx - \int_{g_s\left(y_s^*\right)}^{g_s^*} \left(\left(\lambda_s b y_s^* + \gamma_{3,s}\alpha\right) x - \gamma_{1,s}\right) f\left(x, y_s^*\right) dx. \tag{B.2}$$

It must be that $\lambda_s b y_s^* + \gamma_{3,s}\alpha = \gamma_{1,s}/g_s\left(y_s^*\right)$ or $\lambda_s b y_s^* + \gamma_{3,s}\alpha < \gamma_{1,s}/g_s\left(y_s^*\right)$ and $g_s(y_s^*) = 1$ because if $\lambda_s b y_s^* + \gamma_{3,s}\alpha \neq \gamma_{1,s}/g_s\left(y_s^*\right)$, then the first term in the minimum in the definition of $g_s(y)$ in Lemma B.2 must be strictly greater than 1 at y_s^*. Hence for all x, $\left(\lambda_s b y_s^* + \gamma_{3,s}\alpha\right) x \leqslant \gamma_{1,s} x/g_s\left(y_s^*\right)$, and if $g_s^* > g_s\left(y_s^*\right)$, then the inequality is an equality. Thus whether $g_s^* > g_s\left(y_s^*\right)$ or the opposite holds, it follows that

$$\gamma_{2,s} \leqslant \gamma_{3,s}\alpha \int_0^{g_s^*} xf\left(x, y_s^*\right) dx - \gamma_{1,s} \int_{g_s\left(y_s^*\right)}^{g_s^*} \left(\frac{x}{g_s\left(y_s^*\right)} - 1\right) f\left(x, y_s^*\right) dx.$$

Plugging into the FOC for λ_s yields

$$\gamma_{3,s} \geqslant \frac{\int_0^{y_s^*} \int_0^{g_s(y)} xbyf(x,y)\, dxdy + \frac{c}{\lambda_s^2 b}\gamma_{1,s} \int_{g_s\left(y_s^*\right)}^{g_s^*} \left(\frac{x}{g_s\left(y_s^*\right)} - 1\right) f\left(x, y_s^*\right) dx}{\frac{c}{\lambda_s^2 b}\alpha \int_0^{g_s^*} xf\left(x, y_s^*\right) dx + 1}. \tag{B.3}$$

The first term in the numerator is strictly positive because $y_s^* > 0$ and $g_s(\cdot) > 0$. The second term in the numerator is positive: if $g_s^* > g_s\left(y_s^*\right)$, then for all $x \in \left(g_s\left(y_s^*\right), g_s^*\right]$, $x/g_s\left(y_s^*\right) > 1$, and if $g_s^* < g_s\left(y_s^*\right)$, then flip the limits of the integral, multiply the integrand by negative one, and note that for all $x \in \left[g_s^*, g_s\left(y_s^*\right)\right)$, $x/g_s\left(y_s^*\right) < 1$. It follows that $\gamma_{3,s} > 0$. □

B.3 Proof of Corollaries 1, 2, and 3

It is convenient to prove the three corollaries before proving Theorem 1 and Proposition 2. Corollary 1 follows immediately from Definition 4 of an exposure elasticity because $g_i(y)$ is constant for all $y > y_i^*$.

For Corollary 2, first note that $g_i(\cdot)$ is continuous and strictly decreasing on $[0, y_i^*)$ and constant on $[y_i^*, 1]$. In the second-best, $\gamma_{1,s}/\left(\lambda_s b y_s^* + \gamma_{3,s}\alpha\right) < \gamma_{1,s}/c$ because individually optimal interaction implies that $\lambda_i b y_s^* = c$, and from Lemma B.3, we know that $\gamma_{3,s} > 0$. It therefore follows from the definition of $g_s(\cdot)$ in Proposition 1 that $g_s\left(y_s^*\right) < \min\left\{g_s^*, 1\right\}$ because $g_s(\cdot)$ is strictly decreasing on $[0, y_s^*)$ and a positive measure of agents are selected ($\beta > 0$). Thus, $g_s(\cdot)$ achieves its unique minimum at y_s^*. The result for the first-best holds because plugging the first-best interaction threshold into the definition of $g_f\left(y_f^*\right)$ yields $g_f\left(y_f^*\right) = \gamma_{1,f}/c$, which must therefore be strictly less than 1 and equal to g_f^*.

For Corollary 3, let T be the tax per interaction, which implies that in the first-best, the private cost of isolating for a type (x, y) agent is xc and the private benefit is $x(\lambda_f by + T)$. An agent therefore finds it optimal to interact if $y \geqslant (c - T)/(\lambda_f b)$, and finds it optimal to isolate if $y < (c - T)/(\lambda_f b)$. Thus, if $y_f^* < 1$, then the individually optimal interaction policy aligns with the socially optimal interaction policy if and only if $T = \gamma_{3,f}\alpha$. If $y_f^* \geqslant 1$, then $T = 0$ works as well because $\gamma_{3,f} > 0$ from Lemma B.3, which implies that $c/(\lambda_f b) > 1$, i.e., the individually optimal interaction policy for all agents is to interact.

B.4 Proof of Theorem 1 and Proposition 2

We first prove Proposition 2, and then Theorem 1. The expression for $\gamma_{3,f}$ follows immediately from equation (B.1) in the proof of Lemma B.3. The expression for $\gamma_{3,s}$ follows from inequality (B.3) in the proof of Lemma B.3 because the proof shows the inequality holds with equality as long as $g_s\left(y_s^*\right) < 1$, which is the case from the proof of Corollary 2. If $y_s^* > 1$, then $\Omega_0 = \Omega_1 = 0$ because $f\left(x, y_s^*\right) = 0$. If $y_s^* \leqslant 1$, then $\Omega_0 > 0$ because $g_s\left(y_s^*\right) > 0$, and $\Omega_1 > 0$ because $g_s\left(y_s^*\right) < g_s^*$ by Corollary 2 and $g_s^* \leqslant \gamma_{1,s}/c$.

Now, for Theorem 1, plugging in the expression for $\gamma_{3,f}\alpha$ into (8) for $\varepsilon_i(y)$ yields the expression in the theorem, which is strictly greater than -1 because $\gamma_{3,f} > 0$. For the second-best, (8) shows that $\varepsilon_s(y) > -1$ because $\gamma_{3,s} > 0$. Plugging in the expression for $\gamma_{3,s}\alpha$ yields the other inequality because $\Omega_0 \geqslant 0$ and $\Omega_1 \geqslant 0$. The inequality is strict if $y_s^* \leqslant 1$ because in that case $\Omega_1 > 0$.

B.5 Proof of Proposition 3

First, the same argument as in the proof of Lemma 1 (except with the cost and benefit of isolation flipped) shows that the individually optimal interaction threshold is y_s^*. In this case, $y_s^* > 0$ because as y goes to 0, the benefit of interaction goes to 0, but the cost of interaction remains at xc. Moreover, the same argument as in the beginning of the proof of Proposition 1 (except with the cost and benefit of isolation flipped) shows that in the first-best selection policy, agents interact if and only if $y > y_f^*$. Note that we may now have that the definition of y_f^* is strictly negative in which case all agents interact.

Next, the Lagrange multiplier on the supply constraint, $\gamma_{1,i}$, is strictly positive. If there is a strictly positive measure of non-selected interacting agents, then they can be selected for a strictly positive benefit without any effects on other agents or other agents' individually optimal interaction choices. Thus, welfare strictly increases with more supply. If there is not a strictly positive measure of non-selected interacting agents, then since $\beta < 1$, there must be a strictly positive measure of non-selected isolating agents. Holding fixed interaction choices, selecting some of these agents gives them a strictly positive private benefit, and generates spillovers to selected interacting agents that yield a positive benefit. Moreover, allowing interaction choices to respond in the second-best, the equilibrium aggregate interaction probability can only increase because all non-selected agents are isolating. In that case, the agents who change their behavior are better off because their value of isolating remains the same, and the increase in the equilibrium aggregate interaction probability generates spillovers to interacting agents that yield a positive benefit. Thus, welfare strictly increases with more supply.

Now, we can write the Lagrangian as

$$\begin{aligned}\mathcal{L}_i =& \mathcal{W}_i + \gamma_{1,i}\left(\beta - \int_0^1 \int_0^1 v_i(x,y) f(x,y)\, dxdy\right) + \gamma_{2,i}\left(y_i^* - \frac{c}{\lambda_i b}\right) \\ &+ \gamma_{3,i}\left(\alpha \int_0^1 \int_0^1 x f(x,y)\, dxdy - \alpha \int_0^{y_i^*} \int_0^1 (1 - v_i(x,y)) x f(x,y)\, dxdy - \lambda_i\right),\end{aligned}$$

and welfare as

$$\begin{aligned}\mathcal{W}_i =& \int_0^1 \int_0^1 x\lambda_i b y f(x,y)\, dxdy - \int_0^{y_i^*} \int_0^1 (1 - v_i(x,y)) x\lambda_i b y f(x,y)\, dxdy \\ &- \int_{y_i^*}^1 \int_0^1 (1 - v_i(x,y)) x c f(x,y)\, dxdy,\end{aligned}$$

where $\gamma_{2,f} = 0$, i.e., the incentive compatibility constraint for the individually optimal interaction threshold need not hold in the first-best. Comparing the expressions for the Lagrangian and welfare to equations (6) and (7) for the Lagrangian and welfare, respectively, in the negative externalities / strategic substitutes case, we can see that there are three differences: an additional positive term multiplying $\gamma_{3,i}$, an additional positive term in welfare, and a negative sign in front of λ_i in the equilibrium aggregate interaction probability constraint. The changes don't affect the derivative of the Lagrangian with respect to $v_i(x,y)$, so the proof of Lemma B.2 holds, adapted for the case positive externalities / strategic complements case. Thus, all that remains to show Proposition 3 is that $\gamma_{3,i} > 0$.

To see that $\gamma_{3,i} > 0$, first take the derivative of the Lagrangian with respect to λ_i:

$$\frac{\partial \mathcal{L}_i}{\partial \lambda_i} = \int_0^1 \int_0^1 xbyf(x,y)\, dxdy - \int_0^{y_i^*} \int_0^{g_i(y)} xbyf(x,y)\, dxdy + \gamma_{2,i}\frac{c}{\lambda_i^2 b} - \gamma_{3,i}. \tag{B.4}$$

We must have $\lambda_i > 0$ because selected agents interact and there is a strictly positive measure of selected agents, i.e., $\beta > 0$. It follows that the derivative equals 0. In the first-best, $\gamma_{2,f} = 0$ and so $\gamma_{3,f} > 0$ because the second integral in the derivative equals the first if and only if all agents are isolating and non-selected, which cannot be the case given a strictly positive measure of selection supply, i.e., $\beta > 0$.

To see that $\gamma_{3,s} > 0$, we first prove the following lemma, which states that if strategic complementarities are sufficiently strong, then there is a better individually optimal equilibrium conditional on the selection policy, or there is a better selection policy.

Lemma B.4. *Under the optimal selection policy $v(\cdot,\cdot)$, in the policy maker's preferred individually optimal equilibrium*

$$\frac{c}{\lambda^2 b}\alpha \int_0^1 (1 - v(x, y^*))xf(x, y^*)dx < 1.$$

Proof. First, we show that given any selection policy, in the policy maker's preferred equilibrium, the inequality in the lemma holds weakly. The policy maker's preferred equilibrium—the one with the highest welfare—is the one with the lowest interaction threshold y^* because each agent receives the maximum of their value of isolating and interacting, and comparing two equilibria, the value of isolating is the same, but the value of interacting is higher in the one with a lower y^*.

Define a function $Y_1 : \mathbb{R}_+ \to \mathbb{R}_{++}$ so that if non-selected agents with payoff type greater than y interact, then based on the resulting equilibrium aggregate interaction probability λ, agents with payoff type greater than $Y_1(y) > 0$ prefer to interact, i.e., $Y_1(y) = c/(\lambda b)$. There is an equilibrium with interaction threshold $y^* > 0$ if and only if $Y_1(y^*) = y^*$.

We want to show that at the lowest y^* with $Y_1(y^*) = y^*$, the inequality in the lemma holds. The inequality trivially holds if $y^* > 1$ because then $f(\cdot, y^*) = 0$, so suppose that is not the case. Note that $Y_1(\cdot)$ is continuous and always has a left derivative. It follows that at the lowest equilibrium y^*, the left derivative of $Y_1(\cdot)$ is weakly less than 1; otherwise, there exists a $y < y^*$ such that $Y(y) < y$, which implies an equilibrium less than y^* by the Intermediate Value Theorem because $Y_1(0) > 0$. Thus, inequality in the lemma holds weakly because the left derivative of $Y_1(\cdot)$ at y is

$$\frac{c}{\lambda^2 b}\alpha \int_0^1 (1 - v(x, y))xf(x, y)dx.$$

Next, to complete the proof of the lemma, we show that if the inequality in the lemma holds with equality, then there is a selection policy such that in the policy maker's preferred equilibrium, welfare is strictly higher. In particular, the policy maker can shift selection near y^* away, some to agents with higher y to save on interaction costs, and some to agents with lower y so that if the interaction threshold falls below y^*, the equilibrium aggregate interaction probability is higher and the fall is self-fulfilling.

As an intermediate step, the selection policy is optimal only if (up to a measure 0 of agents), it is an

x-policy below y^* with a strictly positive threshold function, and does not select agents above y^*. If it is not an x-policy below y^*, then there exists a $y_1 < y^*$ and $x_1 < x_2$ such that there are strictly positive measures of selected agents with $y < y_1$ and $x < x_1$, and of non-selected agents with $y < y_1$ and $x > x_2$. Consider unselecting a measure δ of agents in the first group, and selecting agents in the second group instead, which changes the interaction threshold in the policy maker's preferred equilibrium. Since the inequality in the lemma holds with equality in the policy maker's preferred equilibrium under the initial selection policy, it follows that as δ goes to 0, the decrease in the interaction threshold in the policy maker's preferred equilibrium from the change becomes arbitrarily large relative to δ. Thus, the change improves welfare because the increase in the value of interaction is arbitrarily large relative to the change in value from possibly selecting agents with a lower average y. A similar argument shows that the optimal selection policy cannot select agents above y^*. Finally, the threshold above which the x-policy selects agents must be strictly positive because otherwise the left-hand side of the inequality in the lemma is 0.

Now, it follows that there exists an $M > 0$ such that for sufficiently small $\delta > 0$, there is more than a measure $M\delta$ of selected agents between $y^* - \delta$ and y^*. Moreover, there exists an $x_1 > 0$ and $\psi > 0$ such that for sufficiently small δ, the policy maker can unselect agents with $y \in (y^* - \delta, y^*)$, shift a share $1 - \psi$ of that selection to agents with $y > y^*$, and shift the remaining share ψ of that selection to agents with $y < y^* - \delta$, so that all selected agents have $x > x_1$, and so that the integral of x across selected agents with $y > y^* - \delta$ is unchanged. Define a function $Y_2 : \mathbb{R}_+ \to \mathbb{R}_{++}$ so that if the policy maker uses δ to make the specified change, and non-selected agents with payoff type greater than $y^* - \delta$ interact, then based on the resulting equilibrium aggregate interaction probability λ, agents with payoff type greater than $Y_2(\delta) > 0$ prefer to interact, i.e., $Y_2(\delta) = c/(\lambda b)$. If there exists a $\delta > 0$ so that $Y_2(\delta) \leqslant y^* - \delta$, then welfare must be higher after the specified change: among those who were already interacting with $y > y^* - \delta$ (by choice or by selection), total interaction costs are unchanged and the value of interaction is higher; among everyone else, some are the same if they were not interacting and continue not to, some are strictly better off if they switched from not interacting to interacting, and some are strictly better off if they are newly selected.

As δ converges to 0, $Y_2(\delta)$ converges to less than

$$y^* - \delta \frac{c}{\lambda^2 b} \alpha \left(\psi M x_1 + \int_0^1 (1 - v(x, y^*)) x f(x, y^*) dx \right),$$

where the terms in parentheses are the increase in the equilibrium aggregate interaction probability first due to the increase in selection below $y^* - \delta$, and second due to the increase in interaction net of the reduction in selection among agents with $y \in (y^* - \delta, y^*)$. Since the inequality in the lemma holds with equality, it follows that for sufficiently small $\delta > 0$, $Y_2(\delta) < y^* - \delta$, which competes the proof of the lemma. $\square$

Now, since $y_s^* > 0$ it follows that the derivative of the Lagrangian with respect to y_s^*, which is the same as in the negative externalities / strategic substitutes case, must be 0. Plugging equation (B.2) for $\gamma_{2,s}$ into the derivative of the Lagrangian with respect to λ_s, equation (B.4), set equal to 0 yields

$$\gamma_{3,s} = \frac{\int_0^1 \int_0^1 xbyf(x,y)\,dxdy - \int_0^{y_s^*} \int_0^{g_s(y)} xbyf(x,y)\,dxdy - \frac{c}{\lambda_s^2 b} \int_{g_s(y_s^*)}^{g_s^*} (xc - \gamma_{1,s}) f\left(x, y_s^*\right) dx}{1 - \frac{c}{\lambda_s^2 b} \alpha \int_0^{g_s(y_s^*)} xf\left(x, y_s^*\right) dx}, \tag{B.5}$$

where we use the individually optimal interaction threshold $\lambda_s b y_s^* = c$. The difference of the first two integrals in the numerator is strictly positive because otherwise $\lambda_s = 0$. The third term in the numerator (including the negative sign) is positive: if $g_s^* \geqslant g_s\left(y_s^*\right)$, then since $\gamma_{1,s}/c \geqslant g_s^*$, it follows that for all $x \in \left[g_s\left(y_s^*\right), g_s^*\right]$, $\gamma_{1,s} \geqslant xc$; if $g_s^* < g_s\left(y_s^*\right)$, then flip the endpoints of the integral, multiply the integrand by -1, and note that it must be that $g_s^* < 1$, which implies that $\gamma_{1,s}/c = g_s^*$, and so for all $x \in \left[g_s^*, g_s\left(y_s^*\right)\right]$, $\gamma_{1,s} \leqslant xc$. Thus, the numerator on the right-hand side of equation (B.5) is strictly positive. Finally, by Lemma B.4, the denominator on the right-hand side of equation (B.5) is strictly positive. Thus, $\gamma_{3,s} > 0$.

B.6 Proof of Theorem 2, Proposition 4, and Corollaries 4, 5, and 6

Given Proposition 3, the proofs follow the same arguments as the proofs of Theorem 1, Proposition 2, and Corollaries 1, 2, and 3.