Targeting Interacting Agents

Nikhil Vellodi, Joshua Weiss*

October 2, 2021 (click here for the latest version)

Abstract

We introduce a framework to study targeted policy interventions. Agents: 1) differ both in their likelihood of and payoff from interaction, 2) exert externalities through interaction, and 3) can alter their interaction rates at a cost. A planner selects a subset of agents to alter their interaction rates at no cost (a *selection policy*). We study the second- and first-best selection policies; that is, when costly interaction choices are either voluntary or mandatory. Our main results contrast these policies. Two features emerge as particularly salient: *non-monotonicity* – in the second-best, agents with intermediate payoffs are selected more – and *risk compensation*, which describes how endogenous responses to policy shape policy design. We apply our results to settings including vaccine allocation and information aggregation.

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

1 Introduction

We develop a framework to study targeted policy interventions within groups of interacting, heterogeneous agents. Agents differ both in their likelihood of interacting with other agents (*exposure* type) and their payoff from doing so (*payoff* type), and exert externalities through interaction. Finally, agents may alter their interaction rates at a cost, and a benevolent planner can select a subset of agents to alter their interaction rates at zero cost.

*Vellodi: Paris School of Economics, Weiss: Federal Reserve Bank of Richmond and the Institute for International Economic Studies (IIES), Stockholm University. We thank S. Nageeb Ali, Francis Bloch, Kfir Eliaz, Alessandro Lizzeri, Paula Onuchic, Ludovic Renou, Evan Sadler, Olivier Tercieux, Adrien Vigier and participants at various seminars for helpful comments. Vellodi acknowledges support from the EUR grant ANR-17-EURE-0001. The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

Our framework allows for both negative and positive externalities, and thus speaks to a variety of seemingly disparate economic applications. Consider the optimal allocation of vaccines among a population of agents that differ both in their exposure to others as well as their risk of adverse reaction to the illness. An authority can vaccinate some agents, while others can engage in costly self-isolation. Alternatively, consider a group of farmers deciding whether to adopt a new technology. If they adopt, farmers differ both in the amount of socially valuable information they generate as well as in their productivity gain. Farmers can pay to adopt, while an authority can select a subset to receive free product trials. Notice that these examples exhibit negative and positive externalities respectively, both accommodated by our framework. (Section 7 provides details, while Section 8 suggests further applications.)

We fully characterize optimal selection policies, with a particular focus on the features that derive from voluntary incentives. To do so, we study first- and second-best interventions, where the latter is constrained to take agents' choices, such as self-isolation and costly technology adoption in the examples above, as given. Any disparities between these two policies thus highlight the role played by voluntary incentives in shaping optimal policy. Such distinctions are often of important practical concern. For instance, in the case of vaccine allocation, priority lists are often designed in tandem with voluntary self-isolation decisions.

Two differences emerge as particularly salient, the interpretations of which rely crucially on whether interaction imposes negative or positive externalities, i.e., whether interaction choices form strategic substitutes or complements. First, the second-best is *non-monotone* – agents with intermediate payoffs from interaction are targeted most heavily. With strategic substitutes, such agents interact too much relative to the social optimum, as they do not internalize negative spillovers. Conversely, with strategic complements, such agents interact too little relative to the social optimum, as they do not internalize positive spillovers.

Next, the *shape* of the second-best policy – how selection depends on both payoff and exposure types – explicitly accounts for voluntary incentives. With strategic substitutes, the second-best distorts selection away from the *interactive* relative to the first-best, i.e., those likely to interact, toward the *vulnerable*, i.e., those with high losses from interaction; conversely, with complements, the second-best distorts selection *toward* the interactive. With substitutes, the planner would ideally target highly interactive agents to internalize negative spillovers. However, by doing so, they reduce the risk others face when interacting and thereby relax voluntary incentives for self-isolation; in the first-best, the planner can target these agents without concern for adverse effects on incentives (Proposition 2). To illustrate, consider a health authority designing vaccine priority lists. Absent a strictly enforced mandatory lockdown, vaccinating people with high contact rates might perversely result in those not vaccinated interacting more, as they face a reduced risk of contagion. As such, by vaccinating relatively isolated people, the authority maintains private incentives to self-isolate. With complements, the planner seeks to stimulate interaction in order to enhance other agents' incentives to interact. In the example of technology adoption, the planner has a stronger incentive to target highly informative farmers for free trials when adoption is voluntary – such farmers increase the value of adoption to others and thus stimulate uptake.

Thus, with substitutes, we uncover a form of *risk compensation*, a phenomenon that describes how mitigating the downsides to risky actions might not reduce the prevalence of adverse outcomes if outweighed by greater risk-taking behavior. With complements, risk compensation is reversed.

Taken together, non-monotonicity and risk compensation point to how the combination of agency and spillovers between interacting agents shapes prudent policy design.[1] While the non-monotonicity of the second-best is a novel result, risk compensation is a well-documented phenomenon, both theoretically and empirically. Previous work studies the concept from a positive viewpoint, specifically 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.[2] We offer three distinct contributions to this literature. First, we take a normative approach, and show how risk compensation emerges when policy design and policy responses shape each other in equilibrium, broadening the concept's reach and the settings to which it applies. Second, we connect risk compensation to our novel exposure premium concept, which crucially relies on heterogeneity in both interaction rates and payoffs. Finally, by studying a unified framework, we establish how risk compensation has a natural converse in settings with strategic complements.

[1] In this regard, our paper relates to the literature on mechanism design with allocative externalities, wherein the designer must account for the incentives of agents who are not allocated the good. In particular, see Jehiel and Moldovanu, 2001, Jehiel et al., 1996.

[2] 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.

While our baseline framework is general enough to fit a variety of settings, we generalize our analysis in Section 6 to heterogeneous interaction costs, as is often the case in reality, and provide natural conditions under which non-monotonicity and risk compensation are preserved.

The paper proceeds as follows. In Section 2, we introduce and discuss the model. In Sections 3 we discuss examples of selection policies. In Section 4 and 5, we characterize optimal selection policies in setting where interaction choices are strategic substitutes and complements, respectively. In Section 6, we consider an extension to the baseline model that allows for heterogeneous costs. Finally, in Section 7, we apply our findings to two leading applications, concluding in Section 8 with a brief discussion of further applications.

2 Baseline 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 with density function f. Each agent's likelihood of interaction is governed by the choices of both themselves and others. Formally, an agent with type (x, y) interacts with probability $\sigma(x, y)x\lambda$, where $\sigma : [0, 1]^2 \to [0, 1]$ is an interaction choice that scales the probability that type (x, y) interacts, and λ is the *interaction rate* given by

$$\lambda \equiv \alpha \int_0^1 \int_0^1 \sigma(x, y)xf(x, y)dxdy,$$

with $\alpha \in (0, 1]$ an exogenous parameter. We describe how $\sigma(\cdot, \cdot)$ is determined when we solve the model in Sections 4 and 5. The interaction decision generates a direct payoff $\sigma(x, y)xc$, where $c \in \mathbb{R}$ is exogenous. Type (x, y) receives payoff by conditional on interaction, where $b \in \mathbb{R}$ is exogenous, and otherwise they receive 0. Thus, we refer to x as an agent's *exposure type* and y as an agent's *payoff type*. If $c > 0$ and $b < 0$, agents pay a cost to avoid interaction and its associated losses, and thus interaction choices form strategic substitutes. Conversely, if $c < 0$ and $b > 0$, agents pay a cost to interact and enjoy its associated benefits, so that interaction choices form strategic complements.

Simple algebra confirms that an agent's ex-ante expected utility from interaction is $b\lambda\sigma(x, y)xy$, the magnitude of which is increasing in λ, x, y and crucially is determined by the product xy – which we refer to as the agent's *risk type* – rather than each component separately. Since the magnitude is increasing in λ, agents exert interaction spillovers on each other. In the strategic substitutes case, these spillovers are negative, and in the strategic complements case, these spillovers are positive.

Selection Policy – A policy maker can select an exogenous fraction $\beta \in (0,1)$ of agents who no longer face any directs costs associated with their interaction decisions. Specifically, in the strategic substitutes case, the policy maker selects agents to isolate at zero cost (they receive xc despite not interacting), and in the strategic complements case, the policy maker selects agents to interact at zero cost. Formally, a *selection policy* is a mapping $v : [0,1]^2 \to [0,1]$ where $v(x,y)$ is the probability that an agent with type (x,y) is selected and such that v is feasible:

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

Under policy v, in the strategic substitutes case, the interaction rate is

$$\lambda = \alpha \int_0^1 \int_0^1 \big(1 - v(x,y)\big) \sigma(x,y) x f(x,y) dx dy, \tag{2}$$

where now $\sigma(\cdot,\cdot)$ is conditional on not being selected, and a type (x,y) agent receives total payoff

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

In the strategic complements case, the interaction rate is

$$\lambda = \alpha \int_0^1 \int_0^1 \big(v(x,y) + (1 - v(x,y)) \sigma(x,y) \big) x f(x,y) dx dy,$$

and type (x,y) receives total payoff

$$u(x,y) = b\lambda xy - (1 - v(x,y)) \big[b\lambda (1 - \sigma(x,y)) xy + \sigma(x,y) x |c| \big]. \tag{4}$$

2.1 Model Discussion

We have cast our model of interaction at a purposefully abstract level, to allow interpretations that vary by application. One natural setting is when agents enter a pool of interacting agents with probability $\sigma(x,y)x$, and conditional on entering the pool, interact at a rate that scales with the pool's size, as in the applications in Section 7.1.

The separation of an agent's risk type into their likelihood of interaction and loss or benefit from interaction is crucial to better understand the distinct role each plays in shaping aggregate behavior – equations 3 and 4 captures cleanly the fundamental notion that an agent's payoff is a function of their rate of exposure, the magnitude of their payoff from interaction, and aggregate behavior. The marginal distribution $F(\cdot, y)$ models interaction in a tractable manner, in the spirit of random graphs. Specifically, $F(\cdot, y)$ resembles the *degree distribution* $G(\cdot)$ of a random graph, where $G(x)$ denotes the probability of drawing at most x nodes (partners).[3] A key difference to previous work is that we allow agents to scale their contact rate, thus giving rise to an endogenous group of interacting agents.[4] Farboodi et al. (2020) adopt a similar approach, a crucial point of contrast being that in their paper, an agent's interaction rate depends on neither the types nor decisions of other agents. Relatedly, our model could easily be generalized so that the interaction rate is any continuous and strictly increasing function of λ.

We can generalize the model so that the total direct payoff from the interaction decision for an agent with exposure type x who interacts with probability η is $c(x, \eta)$, where $c(x, \eta)$ is strictly increasing in η in the strategic substitutes case, and strictly decreasing in the strategic complements case. Our linearity assumption yields a characterization that makes clear the interaction of behavior with spillovers. One alternative would be $c(x, \eta) = c\eta$, in which the direct interaction payoff does not scale with x. In this case, the non-monotonicity would be in risk type, xy, rather than in payoff type, y, and risk compensation would be unchanged. Nonetheless, we interpret x as an agent's natural rate of encountering activities that might involve interaction, and that it is costly or beneficial to engage in each such encounter. In Section 6, we allow agents to draw idiosyncratic direct interaction payoffs, c.

Finally, in the strategic substitutes case, we model selection as isolating agents at zero cost. Alternatively, we can interpret selection as allowing an agent to interact without consequence, both to themselves and to others through the interaction rate λ. The appropriate interpretation differs by application, as we highlight in Section 7.

[3] As such, our paper relates to the literature on targeted interventions on networks. See Bloch (2015) for a survey. Galeotti et al. (2020) and Ballester et al. (2006) explicitly consider incentives and behavior within far richer networks of interactions, but abstract both from multi-dimensional heterogeneity as well as the comparison between first- and second-best policies.

[4] Recently, Golub and Sadler (2021) study games with both extensive margin (network formation) and intensive margin (investment) choices, studying equilibrium network structures whilst abstracting from policy design.

3 Policy Examples

Before proceeding to solve for optimal policies, it is instructive to define some natural policies, some of which form the building blocks for the subsequent analysis. Throughout this section, we focus on selection policies where $v(\cdot,\cdot)$ only takes on values in $\{0,1\}$, which we later prove holds at the optimal policies.

Definition 1. A policy v is *monotone on* $S \subset [0,1]^2$ if $v(x,y)=1$ and $(x,y)\in S$ imply that $v(x',y')=1$ for all $(x',y')\in S$ such that $x'\geqslant x$ and $y'\geqslant y$. 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 monotone policy v is a *y-policy* if $v(x,y)=1$ for some x implies that $v(x',y')=1$ for all $y'\geqslant y$, $x'\in[0,1]$. A monotone policy v is an *x-policy* if $v(x,y)=1$ for some y implies that $v(x',y')=1$ for all $x'\geqslant x$, $y'\in[0,1]$.

These 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] In contrast, y-policies are commonplace in practice, for instance in the allocation of vaccines to mitigate pandemics.[6]

Definition 3. A monotone policy v is *risk ranking* or an *xy-policy* if $v(x,y)=1$ implies $v(x',y')=1$ for all (x',y') such that $x'y'\geqslant xy$.

Risk ranking policies are characterized by an *iso-risk* threshold such that agents are selected if and only if their risk type, xy, is greater than the threshold. That is, agents are ranked according to their risk type, and selected in descending order.

Definition 4. A policy v is an *x-threshold policy* if there exists an *x-threshold function* $g:[0,1]\to[0,1]$ such that $v(x,y)=1$ if and only if $x\geqslant g(y)$. A policy v is a *y-threshold policy* if there exists a *y-threshold function* $g:[0,1]\to[0,1]$ such that $v(x,y)=1$ if and only if $y\geqslant g(x)$.[7]

[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.

[7] In what follows, we often refer to x- or y-threshold functions as just threshold functions when the difference is clear given the context. Furthermore, we refer to a threshold policy v using its threshold function g.

For example, a risk ranking policy is characterized by the x-threshold function $g(y) = A/y$, for some $A \geqslant 0$; risk, xy, is constant along the boundary.

To motivate our final policy example, notice that since risk xy represents the *private value* of selecting type (x, y), risk ranking policies ignore the spillovers agents exert through interaction. In contrast, policies that place a premium on selecting agents that tend to interact more aim to internalize these spillovers in their design. To formalize this property, we introduce the concept of the *exposure elasticity*, which measures precisely how much of a premium a policy places on an agent's exposure, x, relative to their vulnerability, y.

Definition 5. Let $\varepsilon(\cdot)$ denote the *exposure elasticity* of an x-threshold policy characterized by $g(\cdot)$:

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

To develop this concept further, simple algebra shows that $\varepsilon(y)$ can be written as the ratio of the slopes of the threshold function $g(y)$ and the threshold function $h(y) = A/y$ for the tangent xy-policy $(A = yg(y))$:

$$\varepsilon(y) = -\frac{dg(y)/dy}{dh(y)/dy}.$$

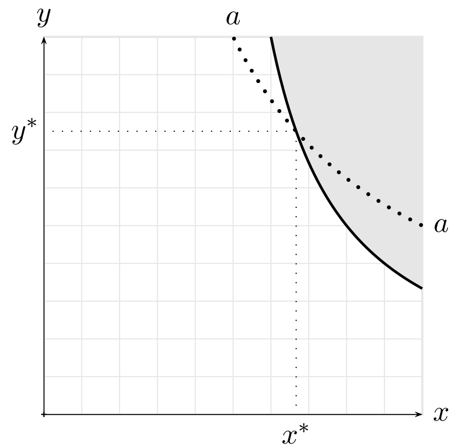
Thus, $\varepsilon(y)$ measures how flat the slope of the policy threshold is relative to risk ranking, and thus how much more willing it is to substitute x for y than simply ranking by risk would imply. In particular, a risk ranking policy has $\varepsilon(y) = -1$, an x-policy has $\varepsilon(y) = 0$ and a y-policy has $\varepsilon(y) = \infty$.

Definition 6. A policy v exhibits an *exposure premium on S* if it is characterized by a x-threshold function $g(\cdot)$ such that $\varepsilon(y) > -1$ for all $y \in S$. A monotone policy v exhibits an *exposure premium* if it exhibits an exposure premium on $[0, 1]$.[8]

In words, if a policy exhibits an exposure premium, then agents with higher x are selected with lower levels of risk, xy. Thus, such policies place a premium on an agent's exposure, x, relative to their payoff type, y, and $\varepsilon(y)$ quantifies this premium precisely. Figure 1 shows an exposure premium policy, demonstrating how the slope of the policy threshold is steeper than the corresponding risk ranking policy at one particular point (x^*, y^*), and thus $\varepsilon(y^*) > -1$.

[8] An equivalent definition exists in terms of y-threshold functions.

Figure 1: Exposure Premium Policy



Exposure premium policy. Solid line: policy threshold function $g(y)$. Dotted line: xy-policy threshold function $h(y)$ with iso-risk threshold $xy = a$. Grey shaded area: $v(x, y) = 1$.

4 Strategic Substitutes

In this section, we study two types of optimal policies under strategic substitutes, i.e., with $c > 0$ and $b < 0$ so that agents can incur a cost to avoid interaction and its associated losses. We refer to agents with high payoff types as *vulnerable.* We begin with the *second-best policy* in which the policy maker chooses selection, $v(\cdot, \cdot)$ to maximize utilitarian welfare, taking as given that interaction choices $\sigma(\cdot, \cdot)$ are individually optimal. We then proceed to study the *first-best policy* in which the policy maker jointly sets selection and interaction to maximize utilitarian welfare. In both cases, we define welfare to be the integral over agents' utilities:

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

where $u(\cdot, \cdot)$ is given by equation (3). We denote variables in the second-best with subscript s and in the first-best with subscript f.

4.1 Second-Best Policy

Equilibrium – Given a selection policy $v(\cdot,\cdot)$ and an interaction rate λ, a non-selected type (x, y) agent solves:

$$\max_{\sigma\in[0,1]} \{b\lambda\sigma xy + \sigma xc\}.$$

Thus, the individually optimal strategy is $\sigma^*(x, y) = \mathbb{I}_{y\leqslant y^*}$, where[9]

$$y^* \equiv c/(\lambda|b|). \tag{6}$$

Since $\sigma^*(x, y) \in \{0, 1\}$ for all (x, y), we say that type (x, y) agents *isolate* if $\sigma^*(x, y) = 0$ or *interact* if $\sigma^*(x, y) = 1$. An agent thus isolates if they are sufficiently vulnerable. Equilibrium is defined in the usual fashion for anonymous large games: given the aggregate distribution over actions, each agent best responds, while the distribution itself is consistent with individual optimization.

Definition 7. Fix a selection policy $v(\cdot,\cdot)$. An *equilibrium* is a pair $\{\sigma^*(\cdot,\cdot), \lambda\}$ such that equations 2 and 6 are satisfied.

Since externalities in this case are purely negative, the induced game is one of strategic substitutes and consequently admits a unique equilibrium.

Lemma 1. *For all selection policies $v(\cdot,\cdot)$, there exists a unique equilibrium.*

Selection Policy – We now turn to the main question of the paper: what is the optimal policy, taking incentives into account? Given equilibrium strategies, this problem reduces to choosing an interaction threshold for the non-selected y_s^*, an interaction rate λ_s, and a selection policy $v_s : [0, 1]^2 \to [0, 1]$ to maximize the Lagrangian:

$$\begin{aligned}\mathcal{L}_s =& \mathcal{W}_s + \gamma_{1,s}\left(\beta - \int_0^1\int_0^1 v_s(x, y)f(x, y)dxdy\right) + \gamma_{2,s}\left(y_s^* - \frac{c}{\lambda_s|b|}\right)\\ &+ \gamma_{3,s}\left(\lambda_s - \alpha\int_0^{y_s^*}\int_0^1 (1 - v_s(x, y))xf(x, y)dxdy\right),\end{aligned} \tag{7}$$

where the constraints are the supply constraint, the incentive compatibility constraint derived from

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

equation (6), and the equilibrium interaction rate constraint derived from equation (2).[10] The following proposition characterizes the second-best policy.

Proposition 1. *The second-best policy is characterized by an x-threshold function $g_s(\cdot)$ such that*

$$g_s(y) = \begin{cases} g_s^* & y \in (y_s^*, 1] \\ \min\left\{\frac{\gamma_{1,s}}{\lambda_s|b|y+\gamma_{3,s}\alpha}, 1\right\} & y \in [0, y_s^*], \end{cases}$$

where

$$g_s^* \equiv \min\left\{\frac{\gamma_{1,s}}{c}, 1\right\},$$

and $\gamma_{1,s}, \gamma_{3,s}$ are strictly positive.

The threshold function $g_s(\cdot)$ tells us the exposure type of the marginal selected agent for each payoff type. The benefit of selecting an agent with type $(g_s(y), y)$ on the threshold must equal the cost, which is precisely the Lagrange multiplier on the supply constraint, $\gamma_{1,s}$. Above y_s^*, agents isolate and the benefit of selecting one on the threshold is avoiding the cost of isolation, so that $\gamma_{1,s} = g_s(y)c$. Below y_s^*, agents interact and the benefit of selecting one on the threshold is comprised of both the private value from isolating the agent as well as the social value from reducing interaction spillovers. The former is $g_s(y)\lambda_s|b|y$ and the latter is $g_s(y)\gamma_{3,s}\alpha$ since $\gamma_{3,s}$ is the social benefit of reducing the equilibrium interaction rate, λ_s.

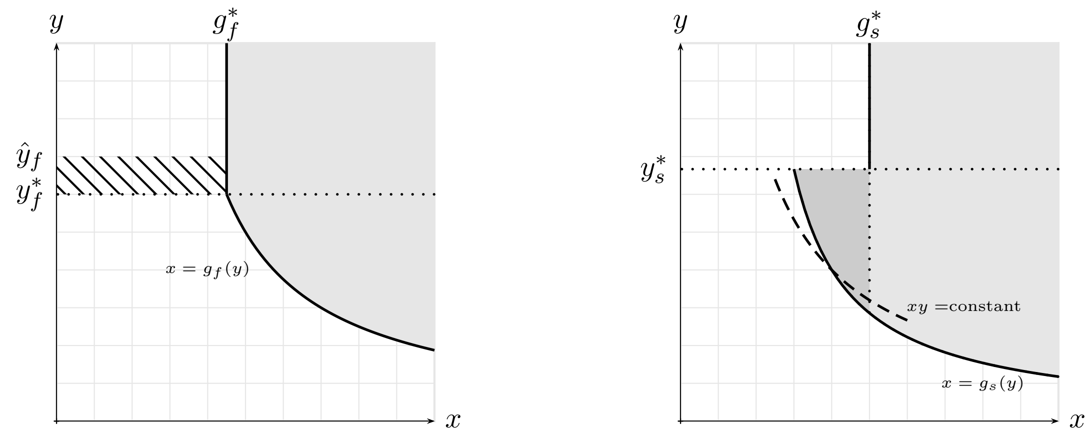
As a corollary to Proposition 1, we describe two striking features of the second-best policy.

Corollary 1. *The threshold function $g_s(\cdot)$ achieves its unique minimum at y_s^*, and thus the second-best is non-monotone if $y_s^* < 1$. Furthermore, the second-best exhibits an exposure premium on each of $[0, y_s^*]$ and $(y_s^*, 1]$.*

First, if $y_s^* < 1$, the second-best policy is *non-monotone*: the optimal policy aggressively selects agents with an intermediate payoff type (y just below y_s^*) relative to more or less vulnerable agents. While the most vulnerable agents efficiently self-isolate and the least vulnerable efficiently interact, agents with intermediate vulnerability interact when they should isolate. They are thus particularly

[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 distribution function F and its density f to be defined on all of $\mathbb{R}^2$, and equal 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: risk ranking policy intersecting optimal policy, demonstrating exposure premium. Shaded area: $v_s(x, y) = 1$. Dark shaded area: additional selection due to incentive distortions.

costly to society when non-selected. Comparing the two groups—the isolating and the interacting—agents over-interact relative to the social optimum, so the policy targets the interacting group more.

Second, the optimal policy exhibits an *exposure premium* within each of the two groups. Although the reasoning is different for the isolating and for the interacting, in each case the benefit of selecting an agent relative to others in their group depends more on their exposure type than on their payoff type. Among the interacting, if the planner only took into account an agent's private value of selection, then the optimal policy would be an xy-policy. Taking into account spillovers, the planner is willing to select highly exposed agents with relatively low risk levels. Among the isolating, the policy not only exhibits an exposure premium, but is an x-policy, i.e., it is invariant to y. Isolating behavior eliminates the importance of vulnerability altogether, while differences in exposure remain relevant as the cost of isolation scales with an agent's exposure.

Thus, while the non-monotonicity illustrates how to select agents *between* the two groups, the exposure premium illustrates how to select agents *within* the two groups. Figure 2 provides a graphical illustration of the second-best policy.

4.2 First-Best Policy

To further understand the distortions imposed by incentives, we solve the first-best problem, in which the planner ignores the incentive constraint for interaction decisions. This is best thought of as a policy that jointly prescribes selection as well as *mandatory isolation*. We allow the planner to make interaction decisions as a function of type (x, y) as well as whether an agent is selected. The planner's problem is then to choose a selection policy $v_f : [0,1]^2 \to [0,1]$, an interaction policy for the non-selected $\sigma_f : [0,1]^2 \to [0,1]$, and a interaction rate λ_f to maximize the Lagrangian:

$$\begin{aligned}\mathcal{L}_f = \mathcal{W}_f &+ \gamma_{1,f}\left(\beta - \int_0^1 \int_0^1 v_f(x,y) f(x,y) dx dy\right) \\ &+ \gamma_{3,f}\left(\lambda_f - \alpha \int_0^1 \int_0^1 (1 - v_f(x,y)) \sigma_f(x,y) x f(x,y) dx dy\right), \end{aligned} \tag{8}$$

where we call the Lagrange multipliers $\gamma_{1,f}$ and $\gamma_{3,f}$ to facilitate comparison with the second-best. Proposition 2 fully characterizes the first-best policy.

Proposition 2. *The first-best interaction policy is characterized by a threshold*

$$y_f^* = \frac{c - \gamma_{3,f}\alpha}{\lambda_f |b|}$$

such that

$$\sigma_f(x,y) = \begin{cases} 0 & y \in \left(y_f^*, 1\right] \\ 1 & y \in \left[0, y_f^*\right]. \end{cases}$$

The first-best selection policy is characterized by an x-threshold function $g_f(\cdot)$ such that

$$g_f(y) = \begin{cases} g_f^* & y \in (y_f^*, 1] \\ \min\left\{\frac{\gamma_{1,f}}{\lambda_f |b| y + \gamma_{3,f}\alpha}, 1\right\} & y \in [0, y_f^*], \end{cases}$$

where

$$g_f^* \equiv \frac{\gamma_{1,f}}{c},$$

and $\gamma_{1,f}, \gamma_{3,f}$ are strictly positive.

The following corollary is a direct analogue to Corollary 1 for the second-best. We make a thorough comparison between the two in Section 4.3.

Corollary 2. *The threshold function $g_f(\cdot)$ is continuous and weakly decreasing, and thus the first-best is monotone. Furthermore, it exhibits an exposure premium.*

Implementation – The first-best interaction policy can be simply implemented with a strictly positive, flat interaction tax. This implies that a planner who seeks to implement the first-best interaction policy can do so with transfers, without the need to observe agents' types. That the tax is positive reflects that there exists a measure of agents—the dashed area in the left panel in Figure 2—who would like to interact in the equilibrium induced by the first-best policy, but do not.

Corollary 3. *The planner can decentralize the optimal interaction policy with a flat interaction tax τ equal to $\gamma_{3,f}\alpha$ per interaction if $y^* < 1$ and 0 otherwise.*

4.3 Price of Anarchy

We highlight the precise role self-interested behavior plays in shaping optimal policy by comparing the second- and first-best policies. Two differences stand out as particularly striking, and constitute the main findings of our paper.

Non-monotonicity – A key qualitative difference between the first- and second-best policies is that the first-best policy is monotone while the second-best policy is not. The discontinuity at y_s^* in the second-best policy is driven by the fact that the threshold y_s^* is inefficiently high. In response, the planner reallocates selection toward agents who inefficiently over-interact (the dark shaded area in Figure 2). This is no longer the case in the first-best in which y_f^* is chosen efficiently.

Risk compensation – A more subtle difference relates to how the exposure premium is shaped by behavior. The exposure premium is present in both the first- and second-best policies, with the same intuition within each behavior group: among the interacting, exposure is more important than vulnerability because only the former generates spillovers, and among the isolating, exposure is more important because only exposure affects the cost of isolation. Nonetheless, the *size* of the exposure premium differs across the two policies. In the first-best, since y_f^* is efficient, an envelope argument implies that marginal changes in y_f^* do not impact welfare. In contrast, in the second-best, the exposure premium takes into account the response of y_s^* to changes in λ_s that result from selection policy.

To quantify this distinction more clearly, we compare the exposure elasticities (Definition 5) $\varepsilon_s(y)$ and $\varepsilon_f(y)$. Using the expressions in Propositions 1 and 2, we derive a simple yet intuitive expression for $\varepsilon_i(y)$ in each case ($i \in \{s, f\}$) for interacting agents: if $y < y_i^*$ and $g_i(y) < 1$, then

$$\varepsilon_i(y) = -\frac{\lambda_i |b| y}{\lambda_i |b| y + \gamma_{3,i}\alpha}.$$

The denominator reflects the marginal social benefit from selecting type $(g_i(y), y)$, which is the sum of both the marginal private benefit $\lambda_i|b|y$ and the marginal social benefit $\gamma_{3,i}\alpha$ from reducing spillovers. As discussed in Section 3, were spillovers ignored by setting $\gamma_{3,i}$ to 0, the policy would reduce to a risk ranking policy with $\varepsilon_i(y) = -1$. We therefore turn to the question: How do the benefits from reducing spillovers compare across the second- and first-best?

Proposition 3. *For $i \in \{s, f\}$, define $\tilde{y}_i$ to be the interaction-weighted average vulnerability among non-selected interacting agents:*

$$\tilde{y}_i = \frac{\int_0^{y_i^*} \int_0^{g_i(y)} yxf(x,y)dxdy}{\lambda_i/\alpha}.$$

The marginal benefit, $\gamma_{3,i}$, of reducing the interaction rate 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

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

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

In the first-best, the benefit of reducing the interaction rate is the direct effect of reducing the risk of interaction among those not selected and already interacting. In the second-best, this direct effect is mitigated by the response of agents in equilibrium: Ω_0 and Ω_1 capture the agents who choose to interact as the interaction rate falls and remain non-selected or become optimally selected, respectively. We immediately have the following corollary.

Corollary 4 (Risk compensation)**.** *For all $y < y_f^*$ such that $g_f(y) < 1$, $\varepsilon_f(y) = -y/\left(y + \tilde{y}_f\right)$. For all $y < y_s^*$ such that $g_s(y) < 1$, $\varepsilon_s(y) < -y/\left(y + \tilde{y}_s\right)$.*

Corollary 4 captures a novel form of "risk compensation" – it demonstrates how the optimal targeting of interventions between different agent characteristics depends on the response of agents to the interventions. In the second-best relative to the first-, the selection policy for the interacting is less affected by spillovers, so it is closer to a policy that ranks purely based on private risk.

Prohibitive Costs – To further isolate the role of incentives, we briefly consider the case in which the relative cost of isolation is sufficiently high such that $y^* = 1$ under any selection policy.[11] In this case, the analysis reduces to a model without the option to isolate, and thus second- and first-best policies are identical. The optimal policy is monotone and exhibits an exposure premium, much as in the right panel of Figure 1.

5 Strategic Complements

In this section, we perform a similar analysis but under strategic complements, i.e., with $c < 0$ and $b > 0$. In the previous section, agents with high y faced the greatest private loss from interaction, whereas now they face the greatest private benefit.

Most features of optimal policies, both in the first- and second-best, remain the same, as we show below. There are two important differences in the problem setups. First, the sign of the equilibrium interaction rate constraint in the Lagrangian, $\gamma_{3,i}$ for $i \in \{s, f\}$, is flipped. As such $\gamma_{3,i}$ is now the marginal *benefit* of increasing the interaction rate, rather than the marginal cost. Second, in the second-best problem, an equilibrium always exists, but there may be multiple equilibria. We focus on locally strong Nash equilibria:

Definition 8. An equilibrium is a *locally strong Nash equilibrium* if there exists an $\epsilon > 0$ such that no set of agents with measure less than ϵ can jointly change their interaction decisions and strictly improve the utility of each agent in the set except a subset of measure 0.

Intuitively, if an equilibrium is not locally strong Nash, then existence of the equilibrium relies heavily on the presence of infinitely many agents so that any one agent cannot have even a small effect on the interaction rate. If we further restricted to strong Nash equilibria so that no set of agents of any measure could profitably deviate, then given a selection policy, the only remaining equilibrium would be the one with the highest λ.

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

Proposition 4. *Let y_s^*, $g_s(\cdot)$, y_f^*, and $g_f(\cdot)$ be defined as before as functions of parameters and equilibrium objects in the equilibria induced by optimal policies, but replace c with $|c|$.*

1. *In the second-best, agents interact if and only if $y > y_s^*$. If the second-best selection policy induces a locally strong Nash equilibrium, then Proposition 1 and Corollary 1 hold.*

2. *In the first-best, Proposition 2 and Corollary 2 hold, except that agents interact rather than isolate if and only if $y > y_f^*$.*

Although the equilibrium interaction rate and the Lagrange multipliers vary between the strategic substitutes and complements cases—even holding parameters fixed—Proposition 4 states that the form of optimal selection policy is the same. However, the logic is reversed. Now, agents with intermediate y *inefficiently under-interact*, as they do not internalize the positive spillovers from doing so, and thus the planner targets them for selection more intensely. As such, an exposure premium emerges. Similarly, among the interacting, agents with higher x pay more to interact, and thus the planner targets them for selection to spare them the associated costs.

The crucial difference from strategic substitutes is that risk compensation is reversed, resulting in a stronger exposure premium in the second-best than in the first-best. Following the analysis in Section 4.3, we have the following proposition, analogous to Proposition 3 and Corollary 4.

Proposition 5. *The marginal benefit, $\gamma_{3,i}$, of increasing the interaction rate 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 are defined as in Proposition 3 and are positive with strict positivity if $y_s^* \leqslant 1$.

For all $y < y_f^*$ such that $g_f(y) < 1$, $\varepsilon_f(y) = -y/\left(y + \tilde{y}_f\right)$.

For all $y < y_s^*$ such that $g_s(y) < 1$, $\varepsilon_s(y) > -y/\left(y + \tilde{y}_s\right)$.

Intuitively, with strategic complements, increases in the interaction rate induce further increases through agent behavior. If agents are under-interacting, as in the second-best, then this force makes it optimal for selection policy to focus more on increasing interaction rather than on reducing interaction costs. This stands in contrast to the strategic substitutes case, where agent behavior mitigated the benefits of reducing the interaction rate relative to saving on isolation costs. In the extremes, in the strategic complements case, as Ω_0 approaches 1, optimal policy converges to an

$x-$policy everywhere that focuses exclusively on maximizing interaction ($\varepsilon_s(\cdot) = 0$), whereas in the strategic substitutes case, as Ω_0 goes to infinity, optimal policy converges to a risk ranking policy that ignores spillovers ($\varepsilon_s(\cdot) = -1$).

6 Heterogeneous Costs

We extend our analysis in the strategic substitutes case to allow for heterogeneous, unobservable, and independently drawn costs of isolation. (The insights for the strategic complements case are similar.) Specifically, the isolation cost c is now drawn from a distribution H supported on $[0, \infty]$. For simplicity, we assume H is continuously differentiable with density h. Each agent's cost draw is independent, and the total cost of isolation is linear in the probability of interaction, as before. Crucially, we do not allow the planner to condition selection based upon an agent's realized c.[12]

This exercise serves several purposes. First, in many applications, agents face differential costs of isolation. Second, this exercise is qualitatively similar to an extension in which the cost of isolation is strictly convex rather than linear, insofar as strict convexity would also smooth the discontinuity in behavior.

The following proposition characterizes the first- and second-best policies. In light of Corollary 3 and for a clear comparison to the second-best, we allow the first-best interaction policy to condition on private isolation costs.

Proposition 6. *Optimal interaction policies are characterized by isolation cost threshold functions*

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

such that, for $i \in \{s, f\}$, an agent with payoff type y interacts if and only if their cost draw satisfies $c > c_i^*(y)$. Define $\tilde{c}_i(y)$ to be the expected isolation cost conditional on y: $\tilde{c}_i(y) \equiv \int_0^{c_i^*(y)} ch(c)dc$.

The optimal selection policy is characterized by the x-threshold function

$$g_i(y) = \min\left\{\frac{\gamma_{1,i}}{(1 - H(c_i^*(y)))(\lambda_i |b| y + \gamma_{3,i}\alpha) + \tilde{c}_i(y)}, 1\right\}.$$

[12] 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.

The optimal selection policy retains features from the baseline model, with the intuition from the isolating and interacting groups blended together since an agent with type (x, y) may now interact or isolate depending on their cost draw. Specifically, the benefit of selecting any agent consists of two components: the private and social value from reducing interaction—which are scaled by the ex-ante likelihood of interaction, $1 - H\left(c_i^*(y)\right)$— and the value from saving on isolation costs, which is the expected cost paid by those with $c \leqslant c_i^*(y)$. The following corollary describes the key features of the first- and second-best selection policies.

Corollary 5.

1. *g_f and g_s are continuous and exhibit exposure premiums.*

2. *$g_f(\cdot)$ is strictly decreasing.*

3. *$g_s'(y) \leqslant 0$ if and only if*

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

Unlike in the baseline model, the second-best policy is continuous since now the distribution of isolation costs is continuous. Nonetheless, it may still be non-monotone. Indeed, this more general formulation shows clearly the two opposing forces at play that potentially generate the non-monotonicity and are at work in our baseline setup. On the one hand, an increase in vulnerability implies more private risk of a bad outcome, which increases the value of selection. On the other hand, agents interact less as vulnerability rises, imposing fewer externalities and thus decreasing the value of selection. 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 directly on the inverse hazard rate $\frac{1-H(\cdot)}{h(\cdot)}$. Figure 3 provides a graphical example in which the inverse hazard rate is strictly decreasing.

Regarding risk compensation, the following result forms an analogue to Proposition 3 and Corollary 4:

Proposition 7. *The marginal benefit, $\gamma_{3,i}$, of reducing the interaction rate is given by*

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

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

$$\varepsilon_f(y) = \frac{-\Big(1 - H\big(c_f^*(y)\big)\Big)\,\lambda_f |b| y}{\Big(1 - H\big(c_f^*(y)\big)\Big)\,\lambda_f |b| \left(y + \tilde{y}_f\right) + \tilde{c}_f(y)}.$$

In the second-best, if $g_s(y) < 1$, there exist a value $\bar{H}_s > 0$ such that

$$\varepsilon_s(y) < \frac{-\Big(1 - H\big(c_s^*(y)\big)\Big)\lambda_s |b| y}{\Big(1 - H\big(c_s^*(y)\big)\Big)\lambda_s |b| \left(y + \tilde{y}_s\right) + \tilde{c}_s(y)}$$

if and only if

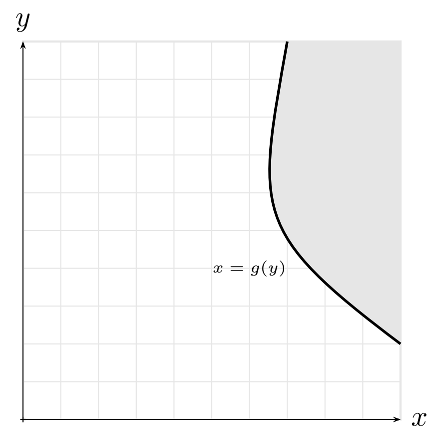
$$\frac{1 - H\big(c_s^*(y)\big)}{h\big(c_s^*(y)\big)} > \bar{H}_s.$$

Intuitively, risk compensation is as before, but another force is present that muddies the comparison between the first- and second-best. In the first-best, as in the baseline model, the exposure premium depends on the marginally selected agent's private risk of a bad outcome relative to the total cost associated with that agent, which also includes their spillover and expected isolation costs. In the second-best, if the inverse hazard rate is sufficiently large, then risk compensation is the dominant force in comparing to the first-best, i.e., the exposure premium is smaller because as the interaction rate falls, all non-selected agents inefficiently interact more (the second term in the denominator of $\gamma_{3,s}$). On the other hand, informed by the discussion on non-monotonicity above, if the inverse hazard rate is small, then another force dominates and pushes toward a larger exposure premium in the second-best: as vulnerability rises, agent behavior becomes more efficient and the value of selection falls.

7 Economic Applications

Owing to its generality, our model sheds new light on a plethora of economic applications that share important common features. In this section, we detail two leading examples.

Figure 3: Second-Best Policy with Heterogeneous Costs



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

7.1 Optimal Vaccination and Self-Isolation

A leading application for our analysis is optimal vaccine allocation. Agents differ both in their natural rate of interaction, x, as well as their severity of symptomatic response to the virus, y. Conditional on interacting, an agent contracts the virus with a likelihood that is increasing in the number of other agents interacting. We can interpret the parameter α as the contagiousness of the virus. Agents can exert costly effort to isolate, while a planner can select a subset of agents to vaccinate. Whether vaccinated or in isolation, agents do not contract the virus or contribute to viral contagion.

Which agents should be vaccinated? Our main findings demonstrate that the answer crucially relies on whether lockdowns are mandatory or voluntary.[13] In the latter case, non-monotonicity implies that agents with *intermediate* vulnerability to the disease should be selected for vaccination more than others, as they do not internalize the risk of contagion they impose on others through interaction. On risk compensation, when the planner cannot enforce mandatory isolation, they fear that the reduction in infections due to vaccination will relax voluntary incentives to self-isolate.[14]

[13] This pertains to *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.

[14] Such fears received widespread coverage in the case of COVID-19. See https://tinyurl.com/z7mh5bdx.

As such, the second-best policy targets vulnerability more than exposure, relative to the first-best. In practice, many authorities target vulnerability before exposure.[15] Our analysis provides a novel rationale for doing so – by vaccinating vulnerable agents who do not interact heavily, the rate of contagion is relatively unaffected within the pool of interacting agents, and thus incentives to voluntarily self-isolate are preserved.[16]

7.2 Information Aggregation

Next, we apply our results to a classical setting exhibiting strategic complementarities – information spillovers. Consider a large group of agents facing uncertainty regarding an unknown state $\theta \in \mathbb{R}$, who share a common prior $\theta \sim \mathcal{N}(\mu_0, \tau_0^{-1})$. Agents can access and share additional information regarding θ through social interaction. Specifically, if a type (x, y) agent's belief over θ has precision τ, they receive a payoff of $y\tau$.[17] Moreover, they join a *social group* with probability $\sigma(x, y)x$, which generates a direct payoff of $\sigma(x, y)xc$, where $c < 0$ is an exogenous parameter, so that it is costly to join. Within the social group, agents receive a public signal $X \sim \mathcal{N}(\lambda\theta, (\lambda b)^{-1})$, where λ depends on the measure of other agents in the social group and $b > 0$ is an exogenous parameter.[18] The Projection Theorem for normal distributions implies that type (x, y) solves the following problem:

$$\max_{\sigma \in [0,1]} \left\{(1 - \sigma x)y\tau_0 + \sigma xy(\tau_0 + \lambda b) - \sigma x|c|\right\} = y\tau_0 + \max_{\sigma \in [0,1]} \left\{\sigma xy\lambda b - \sigma x|c|\right\},$$

as in Section 5 but with a baseline payoff if $\tau_0 > 0$.

For an important application, consider voters that can devote time and attention to be active on social media to both contribute to and follow collective discussion, ultimately informing their vote. Voters differ in the importance they place on political outcomes, y, and in how well they engage with social media, x. Who should a social media platform target to provide incentives to

[15] 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.

[16] In the context of COVID-19, Akbarpour et al. (2021) also present a model with allocative externalities, focusing on how pricing mechanisms help screen agents and improve both efficiency and equity.

[17] This preference for precision could derive from an auxiliary choice m that results in a terminal payoff $y(m-\theta)^{-2}$. See Herskovic and Ramos (2020) for such an example in the context of information diffusion within an endogenously-formed network.

[18] To micro-found this learning story, consider a finite-agent model in which $n \in \mathbb{N}$ agents each receive an i.i.d signal $x \sim \mathcal{N}(\frac{\theta}{n}, \frac{\sigma^2}{n})$. The agents in the social group collectively share their signals, so that if μ agents enter the group and each shares with probability α, they each effectively receive a signal $X \sim \mathcal{N}(\frac{\lambda\theta}{n}, \frac{\lambda\sigma^2}{n})$, where $\lambda = \alpha\mu$. As $n \to \infty$, each agent's own signal becomes uninformative, while the group signal becomes $X \sim \mathcal{N}(\lambda\theta, \lambda\sigma^2)$.

be political engaged?[19]

Alternatively, consider a setting in which agents need not interact to learn from others, but abstaining from interaction precludes an agent from benefiting from information, and they instead receive a payoff of zero. Imagine farmers who must decide whether to pay to adopt a new, unknown production technology. If they adopt, then they generate publicly observable information about the technology's type, θ. Farmers are heterogeneous in the size of their plot, x, and in their productivity, y. The former scales their cost of adoption, xc, the information they generate, and their profits from using the new technology, $xy(\tau_0 + \lambda b)$, where λb is the precision of the signal generated by other adopting farmers. A farmer's productivity only scales their profits. In this case, type (x, y) solves the problem:

$$\max_{\sigma \in [0,1]} \left\{ \sigma xy(\tau_0 + \lambda b) - \sigma x|c| \right\},$$

as in Section 5 if $\tau_0 = 0$. Which farmers should a benevolent authority target for a free trial?[20]

In both these settings, non-monotonicity stems from agents with moderate incentives inefficiently opting out, forcing the second-best policy to aggressively target these agents for subsidized participation. Proposition 5 on risk compensation implies that targeting more informative agents becomes even more valuable when voluntary incentives matter.

8 Conclusion

In this paper, 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. A planner designs a *selection policy*, which alters the interaction patterns of a targeted subset of agents at zero cost. Our main results contrast optimal selection policies in two regimes: the first-best, wherein the planner can additionally control agents' costly effort decisions, and the second-best, wherein these decisions are made by agents and taken as given by the planner. We document two striking differences. First,

[19] Bond et al. (2012) and Jones et al. (2017) document empirical evidence showing how Facebook positively distorted voter turnout through targeted advertising during the 2010 US Congressional Elections and 2012 US Presidential Elections. Their results confirm that both network centrality and political activism were important determinants of the intervention's effectiveness.

[20] Beaman et al. (forthcoming) provides empirical evidence for the value of targeting highly-connected agents in this setting.

the second-best is *non-monotone*, targeting agent with intermediate interaction payoffs. Second, a novel form of *risk compensation* emerges, wherein the shape of the selection policy is designed with the endogenous response of agents in mind.

While we detailed two leading settings to which our framework and results port well, we believe many more applications can be studied using our model. For instance, consider firms competing on an online platform. Firms differ in both their quality, y, as well as their brand recognition, $1 - x$. Firms either are either monopolists or enter a competitive pool; greater brand recognition entails greater monopoly power. Firms can also invest in intangible capital (e.g., advertising) that alters their likelihood of remaining monopolists. Finally, the platform can select a subset of firms to steer consumers toward, thus increasing their search prominence and market power.[21] Which firms should the platform target in order to maximize industry profits? Finally, our framework can also be used to study optimal congestion management in traffic networks comprised of users that differ in both their likelihood of usage x and their outside options y. We leave a detailed exploration of these and other applications for future work.

References

Akbarpour, M., E. Budish, P. Dworczak, and S.D. Kominers, "An Economic Framework for Vaccine Prioritization," *working paper*, 2021.

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

Armstrong, M., J. Vickers, and J. Zhou, "Prominence and Consumer Search," *The RAND Journal of Economics*, 2009, *40* (2), 209–233.

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

Beaman, L., A. Benyishay, J. Magruder, and A.M. Mobarak, "Can Network Theory-based Targeting Increase Technology Adoption?," *American Economic Review*, forthcoming.

Bloch, F., *Targeting and pricing in social networks* 2015.

[21] Armstrong et al. (2009) model the practice of steering as providing firms with market power by increasing their search prominence, while Haan and Moraga-González (2011) allow for firms to increase their prominence through costly advertising.

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.

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.

Golub, B. and E. Sadler, "Games on Endogenous Networks," *working paper*, 2021.

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).

Haan, M.A. and J.L. Moraga-González, "Advertising for Attention in a Consumer Search Model," *The Economic Journal*, 2011, *121* (552), 552–579.

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.

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

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.

A Proofs

A.1 Lemma 1

Recall that the threshold is $y^* = c/(\lambda|b|)$, which is decreasing in λ. It follows that the right-hand side of equation 2 is decreasing in λ since the integrand is always positive and as λ increases, y^* decreases, and the integral is over a smaller set. The right-hand side is also continuous by the Fundamental Theorem of Calculus. Furthermore, it is strictly positive when $\lambda = 0$ and is always strictly less than 1. Since the left-hand side is λ, the result follows from the Intermediate Value Theorem.

A.2 Proposition 1

First, we have the following lemma that shows that the second-best planner always strictly prefers greater supply for selection.

Lemma A.1. *The Lagrange multiplier on the supply constraint, $\gamma_{1,s}$, is strictly positive.*

Proof. If there is a strictly positive measure of isolating agents, then they can be selected for a strictly positive private benefit. As such, the marginal benefit, $\gamma_{1,s}$, must be strictly positive. If there is not a strictly positive measure of isolating agents, then since $\beta < 1$, there is a strictly positive measure of non-selected interacting agents. Selecting them cannot lead to an increase in interaction since all non-selected agents are already interacting. As such, the non-selected interacting agents can be selected for a strictly positive private benefit with spillovers that yield a strictly positive public benefit. □

We now prove the following variation on the first part of Proposition 1. The difference is that if the Lagrange multiplier on the equilibrium interaction rate constraint, $\gamma_{3,s}$, is negative, then the expression for $g_s(y)$ in Proposition 1 may be negative, in which case it is optimal not to select any agents with payoff type y even though for all such agents, $x > g_s(y)$. To complete the proof of Proposition 1, all that remains is to show that $\gamma_{3,s} > 0$.

Proposition A.1. *The second-best policy is characterized by an x-threshold function $g_s(\cdot)$ such that*

$$g_s(y) = \begin{cases} g_s^* & y \in (y_s^*, 1] \\ min\left\{\frac{\gamma_{1,s}}{\lambda_s|b|y+\gamma_{3,s}\alpha}, 1\right\} & y \in (y_0, y_s^*] \\ 1 & y \in [0, y_0], \end{cases}$$

where

$$g_s^* \equiv min\left\{\frac{\gamma_{1,s}}{c}, 1\right\}$$

and

$$y_0 \equiv -\frac{\gamma_{3,s}\alpha}{\lambda_s|b|}.$$

Proof. First, consider type (x, y) with $y > y_s^*$. The first derivative of the planner's Lagrangian in equation (7) with respect to $v_s(x, y)$ is

$$(cx - \gamma_{1,s})f(x, y),$$

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 planner optimally sets $v_s(x, y) = 1$ if $x > \gamma_{1,s}/c$ and optimally sets $v_s(x, y) = 0$ if $x < \gamma_{1,s}/c$. A measure 0 of agents have precisely $x = \gamma_{1,s}/c$, so if $x = \gamma_{1,s}/c$, then any choice of $v_s(x, y)$ is optimal.

Next, consider type (x, y) with $y \in (y_0, y_s^*]$. The first derivative of the planner's Lagrangian with respect to $v_s(x, y)$ is

$$(\lambda_s|b|yx - \gamma_{1,s} + \gamma_{3,s}\alpha x)f(x, y),$$

where the first term is the benefit that the agent is no longer vulnerable, the second term is the selection supply cost, and the third term captures the externality of lowering the equilibrium interaction rate by selecting an agent who was interacting when non-selected. Again, the second derivative is 0. Since $y > y_0$, we know that $\lambda_s|b|y + \gamma_{3,s}\alpha > 0$, which implies that the first derivative is strictly increasing in x. It follows that the planner optimally sets $v_s(x, y) = 1$ if $x > \gamma_{1,s}/(\lambda_s|b|y + \gamma_{3,s}\alpha)$ and optimally sets $v_s(x, y) = 0$ if $x < \gamma_{1,s}/(\lambda_s|b|y + \gamma_{3,s}\alpha)$. A measure 0 of agents have precisely $x = \gamma_{1,s}/(\lambda_s|b|y + \gamma_{3,s}\alpha)$, so any choice of $v_s(x, y)$ for such an x is optimal.

Finally, consider type (x, y) with $y \leqslant y_0$. The first derivative of the planner's Lagrangian with respect to $v(x, y)$ is

$$(\lambda_s|b|yx - \gamma_{1,s} + \gamma_{3,s}\alpha x)f(x, y),$$

as in the previous case with $y \in (y_0, y_s^*]$. Again, the second derivative is 0. Since $y \leqslant y_0$, we know that $\lambda_s|b|y + \gamma_{3,s}\alpha \leqslant 0$. Since $x \geqslant 0$ and $\gamma_{1,s} > 0$, it follows that the first derivative is always strictly negative. Thus, the planner optimally set $v_s(x, y) = 0$. □

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

Lemma A.2. *The Lagrange multiplier on the equilibrium interaction rate constraint, $\gamma_{3,s}$, is strictly positive.*

Proof. From Proposition A.1, we can write the second-best planner's Lagrangian as

$$\mathcal{L}_s = \mathcal{W}_s + \gamma_{1,s}\left(\beta - \int_0^1 \int_{g_s(y)}^1 f(x,y)dxdy\right) + \gamma_{2,s}\left(y_s^* - \frac{c}{\lambda_s|b|}\right) + \gamma_{3,s}\left(\lambda_s - \alpha \int_0^{y_s^*} \int_0^{g_s(y)} xf(x,y)dxdy\right),$$

where welfare is

$$\mathcal{W}_s = \int_0^1 \int_0^1 xcf(x,y)dxdy - \int_0^{y_s^*} \int_0^{g_s(y)} \lambda_s|b|yxf(x,y)dxdy - \int_{y_s^*}^1 \int_0^{g_s^*} cxf(x,y)dxdy.$$

Note that the first term in welfare is not relevant for the planner's problem. We will need to use the First Order Conditions (FOC) for y_s^* and λ_s, so we begin with the derivatives of the Lagrangian with respect to each:

$$\frac{\partial \mathcal{L}_s}{\partial y_s^*} = \int_0^{g_s^*} cxf\left(x, y_s^*\right)dx - \int_0^{g_s(y_s^*)} (\lambda_s|b|y_s^* + \gamma_{3,s}\alpha)xf\left(x, y_s^*\right)dx - \int_{g_s(y_s^*)}^{g_s^*} \gamma_{1,s} f\left(x, y_s^*\right)dx + \gamma_{2,s},$$

where the first term is the reduction in isolation costs, the second term captures the increase in interaction, the third term is the cost (or benefit, depending on whether $g_s\left(y_s^*\right) < \gamma_{1,s}/c$) of the change in the quantity of agents selected implied by a discontinuity in $g_s(\cdot)$ at y_s^*, and the fourth term is the effect on the incentive compatibility constraint for the interaction threshold; and

$$\frac{\partial \mathcal{L}_s}{\partial \lambda_s} = -\int_0^{y_s^*} \int_0^{g_s(y)} |b|yxf(x,y)dxdy + \gamma_{2,s}\frac{c}{\lambda_s^2|b|} + \gamma_{3,s},$$

where the first term is the direct cost of increasing the interaction rate on those who are interacting, the second term is the effect of the equilibrium interaction rate on the individually optimal interaction threshold, and the third term is the effect on the equilibrium interaction rate constraint.

The necessary FOC for each of λ_s and y_s^* is that the derivative of the Lagrangian with respect to each is 0. We can see that λ_s cannot be at its lower bound of 0 because if it were, then all agents would optimally interact, which would imply that $\lambda_s > 0$. We can see that y_s^* cannot equal its lower bound of 0 by looking at the incentive compatibility constraint for y_s^* since λ_s is finite. Neither λ_s nor y_s^* has an upper bound.

Plugging in the incentive compatibility constraint, $\lambda_s|b|y_s^* = c$, into the FOC for y_s^* yields

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

It must be that $\lambda_s|b|y_s^* \leqslant \gamma_{1,s}/g_s(y_s^*) - \gamma_{3,s}\alpha$. To see this, suppose $\lambda_s|b|y_s^* \neq \gamma_{1,s}/g_s(y_s^*) - \gamma_{3,s}\alpha$. It follows from Proposition A.1 that for all $y \leqslant y_s^*$, $g_s(y) = 1$. To satisfy the planner's supply constraint, it must therefore be that $g_s^* < 1$. Moreover, from Proposition A.1, we can see that $\lambda_s|b|y_s^* + \gamma_{3,s}\alpha < \gamma_{1,s}$, which implies that $\lambda_s|b|y_s^* < \gamma_{1,s}/g_s(y_s^*) - \gamma_{3,s}\alpha$.

Now, it follows from the FOC for y_s^* that

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

Plugging into the FOC for λ_s yields

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

The first term in the numerator is strictly positive because $y_s^* > 0$, for all y, $g_s(y) > 0$, and the joint distribution of (x, y) has full support on $[0,1]^2$. The second term in the numerator is weakly positive: if $g_s^* > g_s(y_s^*)$, then for all $x \in (g_s(y_s^*), g_s^*]$, $x/g_s(y_s^*) > 1$, and if $g_s^* < g_s(y_s^*)$, then weak positivity follows from flipping the limits of the integral, multiplying the integrand by negative one, and noting that for all $x \in [g_s^*, g_s(y_s^*))$, $x/g_s(y_s^*) < 1$. It follows that $\gamma_{3,s} > 0$. □

To complete the proof of Proposition 1, note that since $\gamma_{3,s} > 0$, we have that $y_0 < 0$ in Proposition A.1.

A.3 Proof of Corollary 1

For the first part of the corollary, we can see from Proposition 1 that $g_s(\cdot)$ is strictly decreasing on $[0, y_s^*]$ and constant on $(y_s^*, 1]$. Moreover, plugging in the incentive compatibility constraint, $\lambda_s|b|y_s^* = c$, we can see that $g_s(y_s^*) < g_s^*$ since $\gamma_{3,s} > 0$ (Lemma A.2).

For the second part, we can see that for all $y \in (y_s^*, 1]$, the exposure elasticity is $\varepsilon_s(y) = 0 > -1$, and for all $y \in [0, y_s^*]$, $\varepsilon_s(y) = -\lambda_s|b|y/(\lambda_s|b|y + \gamma_{3,s}\alpha)$, which is strictly greater than -1 since $\gamma_{3,s} > 0$. As such, on each interval, the second-best policy exhibits an exposure premium.

A.4 Proof of Proposition 2

We begin with the following lemma that the Lagrange multipliers are strictly positive.

Lemma A.3. *The Lagrange multipliers in the first-best, $\gamma_{1,f}$ and $\gamma_{3,f}$, are strictly positive.*

Proof. First, consider the Lagrange multiplier on the supply constraint, $\gamma_{1,f}$. Since the planner controls interaction, they can select any agent without affecting other agents' behavior. As such, as long as there are non-selected agents, whether they are interacting or isolating, the planner can strictly improve their utility, and therefore welfare overall, by selecting them. It follows that $\gamma_{1,f} > 0$.

Next, consider the Lagrange multiplier on the equilibrium interaction rate constraint, $\gamma_{3,f}$. The derivative of the Lagrangian with respect to λ_f is

$$\frac{\partial \mathcal{L}_f}{\partial \lambda_f} = -\int_0^1 \int_0^1 (1 - v_f(x,y))\sigma_f(x,y) bxy f(x,y) dx dy + \gamma_{3,f}.$$

Since $\lambda_f > 0$, the FOC for λ_f is that the above derivative is equal to 0. To see that $\lambda_f > 0$, note that since $\beta < 1$ implies a positive measure of non-selected agents, $\lambda_f = 0$ implies a positive measure of non-selected isolating agents. On the margin, if $\lambda_f = 0$, these agents can be made to interact to improve welfare since they would no longer have to pay the isolation cost and would face no risk of interaction. It follows that $\lambda_f = 0$ cannot be optimal. Hence, the FOC for λ_f implies that

$$\gamma_{3,f} = \int_0^1 \int_0^1 (1 - v_f(x,y))\sigma_f(x,y) |b| xy f(x,y) dx dy,$$

which is strictly greater than 0 since $\lambda_f > 0$. □

We proceed by proving the first part of Proposition 2 concerning the optimal interaction policy, which we restate here.

Proposition A.2. *The first-best interaction policy is characterized by a threshold*

$$y_f^* = \frac{c - \gamma_{3,f}\alpha}{\lambda_f |b|}$$

such that

$$\sigma_f(x,y) = \begin{cases} 0 & y \in \left(y_f^*, 1\right] \\ 1 & y \in \left[0, y_f^*\right]. \end{cases}$$

Proof. The derivative of the Lagrangian with respect to the probability that a non-selected agent with type (x,y) interacts is

$$\frac{\partial \mathcal{L}_f}{\partial \sigma_f(x,y)} = (cx - \lambda_f |b| yx - \gamma_{3,f}\alpha x)(1 - v_f(x,y)) f(x,y),$$

where the first term in the first parentheses is the benefit from reduced isolation costs, the second term is the direct cost of increased interaction on the agent in question, and the third term is the cost of increased interaction due to the spillover effect on the equilibrium interaction rate. The second derivative is zero. If $v_f(x,y) = 1$, then all agents with type (x,y) are selected and the planner can optimally set any value $\sigma_f(x,y) \in [0,1]$. Suppose $v_f(x,y) < 1$. The first derivative is strictly decreasing in y since, as we saw in the proof of Lemma A.3, $\lambda_f > 0$. Moreover, the first derivative is linear in x. From the definition of y_f^* in the proposition, it follows that the planner optimally sets $\sigma_f(x, y) = 0$ if $y > y_f^*$ and optimally sets $\sigma_f(x, y) = 1$ if $y < y_f^*$. Since a measure 0 of agents have $y = y_f^*$, any interaction policy is optimal among those agents. $\square$

To complete the proof of Proposition 2, we can use the same arguments as in the proof of Proposition A.1 (part of the proof of Proposition 1).

A.5 Proof of Corollary 2

To see that $g_f(\cdot)$ is continuous, first note that it is continuous on $\left[0, y_f^*\right)$ and on $\left(y_f^*, 1\right]$. Moreover, $g_f(\cdot)$ is continuous at y_f^* because

$$g_f\left(y_f^*\right) = \frac{\gamma_{1,f}}{\lambda_f|b|\frac{c-\gamma_{3,f}\alpha}{\lambda_f|b|} + \gamma_{3,f}\alpha} = \frac{\gamma_{1,f}}{c} = g_f^*.$$

To see that $g_f(\cdot)$ is weakly decreasing, note that on $\left(y_f^*, 1\right]$ it is constant, on $\left[0, y_f^*\right)$ its derivative is

$$g_f'(y) = -g_f(y)\frac{\lambda_f|b|}{\lambda_f|b|y + \gamma_{3,f}\alpha},$$

which is negative, and $g_f(\cdot)$ is continuous at y_f^*. Finally, on $\left(y_f^*, 1\right]$, the exposure elasticity is $\varepsilon_f(y) = 0 > -1$, and on $\left[0, y_f^*\right)$, the exposure elasticity is $\varepsilon_f(y) = -\lambda_f|b|y/(\lambda_f|b|y + \gamma_{3,f}\alpha)$, which is strictly greater than -1 since $\gamma_{3,f} > 0$. Since $g_f(\cdot)$ is continuous at y_f^*, it follows that the first-best policy exhibits an exposure premium.

A.6 Proof of Corollary 3

Suppose there is a flat interaction tax τ. Since selected agents isolate, the tax has no impact on them. Non-selected agents take the tax as given and choose their individual probability of interaction to maximize their individual utility. Then, an agent with type (x, y) interacts with probability $\sigma(x, y)x$ and gets utility

$$-(1 - \sigma(x, y))cx - \sigma(x, y)(\lambda|b|xy + \tau x).$$

It follows that an agent with type (x, y) interacts if and only if their vulnerability y is below a threshold y^* such that $c = \lambda|b|y^* + \tau$, i.e.,

$$y^* = \frac{c - \tau}{\lambda|b|}.$$

Agents with $y = y^*$ are indifferent between interacting and isolating, but there is a measure 0 of such agents. It follows from Proposition 2 that if $\lambda = \lambda_f$, then to get $y^* = y_f^*$, we must have $\tau = \gamma_{3,f}\alpha$. Moreover, from Lemma 1, we can see that if the planner chooses the flat interaction tax $\tau = \gamma_{3,f}\alpha$ and the first-best selection policy, then the unique equilibrium is the first-best in which the interaction rate is λ_f and agents interact if and only if $y \leqslant y_f^*$ (ignoring interaction decisions among the measure 0 of agents with precisely $y = y_f^*$).

Finally, if $y_f^* \geqslant 1$, then the tax is irrelevant because even with the tax $\tau = \gamma_{3,f}\alpha$, all agents optimally interact. As such, the planner can decentralize the optimal interaction policy with a tax of 0 (or any tax sufficiently small so that all agents interact when $\lambda = \lambda_f$).

A.7 Proof of Proposition 3

Using the expression for $\tilde{y}_f$, the expression for $\gamma_{3,f}$ follows immediately from the proof of Lemma A.3.

Using the First Order Condition for y_s^* along with the incentive compatibility condition, $\lambda_s|b|y_s^* = c$, to plug in for $\gamma_{2,s}$ in the First Order Condition for λ_s (the derivatives are given in the proof of Lemma A.2) yields

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

Multiplying both sides by α, plugging in the expression for $\tilde{y}_s$, and plugging in $y_s^* = c/(\lambda_s|b|)$ yields the expression for $\gamma_{3,s}\alpha$ in the proposition.

The strict positivity of Ω_0 when $y_s^* \leqslant 1$ follows from the strict positivity of the threshold function, which must hold since $\gamma_{1,s} > 0$, and the fact that the distribution of (x,y) has full support on $[0,1]^2$. For Ω_1, it is sufficient to note from Proposition 1 that $\gamma_{1,s}/c \geqslant g_s^*$ and that if $y_s^* \leqslant 1$, then $g_s^* > g_s\left(y_s^*\right)$.

A.8 Proof of Propositions 4 and 5

We extend the proofs for the second-best problem from the strategic substitutes case to the strategic complements case. The argument for the first-best problem is nearly identical and we omit it for brevity.

We can write the second-best Lagrangian in the strategic complements case as

$$\begin{aligned}\mathcal{L}_s =& \mathcal{W}_s + \gamma_{1,s}\left(\beta - \int_0^1\int_0^1 v_s(x,y)f(x,y)dxdy\right) + \gamma_{2,s}\left(y_s^* - \frac{|c|}{\lambda_s b}\right)\\ &+ \gamma_{3,s}\left(\alpha\int_0^1\int_0^1 xf(x,y)dxdy - \alpha\int_0^{y_s^*}\int_0^1(1 - v_s(x,y))xf(x,y)dxdy - \lambda_s\right),\end{aligned}$$

where welfare is

$$\begin{aligned}\mathcal{W}_s =& \int_0^1\int_0^1 \lambda_s byxf(x,y)dxdy - \int_0^{y_s^*}\int_0^1(1 - v_s(x,y))\lambda_s byxf(x,y)dxdy\\ &- \int_{y_s^*}^1\int_0^1(1 - v_s(x,y))|c|xf(x,y)dxdy.\end{aligned}$$

Written this way and comparing to the second-best Lagrangian in the strategic substitutes case (equation (7)), we can see that there are four differences: an additional positive term multiplying $\gamma_{3,s}$, the disappearance of the irrelevant first term in welfare, an additional positive term in welfare, and a negative sign in front of λ_s. The only impacts of these differences are on the value of λ_s and on the derivative of the Lagrangian with respect to λ_s. As such, the only proof from the strategic substitutes case that is affected is the proof that $\gamma_{3,s} > 0$.

Before showing that $\gamma_{3,s} > 0$, we prove the following lemma.

Lemma A.4. *In a locally strong Nash equilibrium,*

$$\frac{|c|}{\lambda_s^2 b}\alpha\int_0^{g_s(y_s^*)} xf\left(x, y_s^*\right)dx < 1.$$

Proof. Suppose, for some $\delta > 0$, non-selected agents with payoff type $y \in [y_s^* - \delta, y_s^*]$ interact instead of isolating, holding fixed the selection policy $v_s(\cdot,\cdot)$ and the behavior of other agents. The interaction rate increases from λ_s to $\lambda_s + \lambda'$, where

$$\lambda' = \alpha\int_{y_s^*-\delta}^{y_s^*}\int_0^{g_s(y)} xf(x,y)dxdy.$$

An agent with type (x,y) who changed their behavior gets utility

$$\lambda' xby + \lambda_s xby - |c|x$$

instead of 0. Since the change in utility is linear in x and strictly increasing in y (for all but a measure 0 of agents with $x = 0$), it follows that all but a set of measure 0 of the agents who changed their behavior are strictly better off if agents with payoff type $y_s^* - \delta$ are weakly better off, i.e.,

$$\lambda'\left(y_s^* - \delta\right) \geqslant \lambda_s\delta,$$

where we use $\lambda_s b y_s^* = |c|$. Since $g_s(\cdot)$ is continuous below y_s^* and $f(\cdot,\cdot)$ is continuous, it follows that for sufficiently small δ, λ' is well approximated by

$$\alpha\delta\int_0^{g_s(y_s^*)} xf\left(x, y_s^*\right)dx.$$

Thus, plugging the limit into our previous inequality and dividing both sides by δ, we can see that if

$$\alpha\left(\int_0^{g_s(y_s^*)} xf\left(x, y_s^*\right)dx\right)y_s^* \geqslant \lambda_s,$$

then there is an arbitrarily small coalition that can deviate and strictly improve the welfare of all its members except a set of measure 0. Dividing both sides by λ_s and plugging in $y_s^* = |c|/(\lambda_s b)$ yields the desired result. $\square$

To see that we still have $\gamma_{3,s} > 0$, take the derivative of the Lagrangian with respect to λ_s, using that the optimal selection policy will be a threshold function $g_s(\cdot)$ that is constant at g_s^* above y_s^*:

$$\frac{\partial \mathcal{L}_s}{\partial \lambda_s} = \int_0^1 \int_0^1 byxf(x,y)dxdy - \int_0^{y_s^*} \int_0^{g_s(y)} byxf(x,y)dxdy + \gamma_{2,s}\frac{|c|}{\lambda_s^2 b} - \gamma_{3,s}.$$

Plugging in the First Order Condition for y_s^*, which is the same as in the strategic substitutes case, the First Order Condition for λ_s (setting the derivative equal to 0) implies that

$$\gamma_{3,s} = \frac{\int_0^1 \int_0^1 byxf(x,y)dxdy - \int_0^{y_s^*} \int_0^{g_s(y)} byxf(x,y)dxdy + \frac{|c|}{\lambda_s^2 b}\int_{g_s(y_s^*)}^{g_s^*} (\gamma_{1,s} - |c|x)\, f(x, y_s^*)\, dx}{1 - \frac{|c|}{\lambda_s^2 b}\alpha \int_0^{g_s(y_s^*)} xf(x, y_s^*)\, dx}.$$

The third term in the numerator is positive: if $g_s^* > g_s(y_s^*)$, then since $\gamma_{1,s}/|c| \geqslant g_s^*$, it follows that for all $x \in [g_s(y_s^*), g_s^*)$, $\gamma_{1,s} > |c|x$; if $g_s^* < g_s(y_s^*)$, then flip the endpoints of the integral, multiply the integrand by -1, and note that since $g_s^* < 1$, it must be that $\gamma_1/|c| = g_s^*$, which implies that for all $x \in (g_s^*, g_s(y_s^*)]$, $|c|x > \gamma_{1,s}$. Therefore, since the denominator is strictly positive in a locally strong Nash equilibrium by Lemma A.4, it follows that $\gamma_{3,s} > 0$. Moreover, multiplying $\gamma_{3,s}$ by α and plugging in the definitions of $\tilde{y}_s$ and y_s^* shows that $\gamma_{3,s}\alpha$ is as given in Proposition 5.

A.9 Proofs of Proposition 6, Corollary 5, and Proposition 7

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 cost of risking interaction, i.e., $cx > \lambda_s|b|xy$. 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 individually optimal interaction policy, the second-best planner maximizes the Lagrangian

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

where welfare is

$$\mathcal{W}_s = \int_0^1 \int_0^1 \int_0^\infty cxh(c)f(x,y)dcdxdy$$
$$- \int_0^1 \int_0^1 (1 - v_s(x,y)) \left[(1 - H(c_s^*(y))) \lambda_s |b| xy + \int_0^{c_s^*(y)} cxh(c)dc \right] f(x,y)dxdy,$$

and where we call the second Lagrange multiplier $\gamma_{3,s}$ to facilitate comparison with the baseline model. The same argument as for Lemma A.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:

$$\frac{\partial \mathcal{L}_s}{\partial \lambda_s} = - \int_0^1 \int_0^1 (1 - v_s(x,y)) (1 - H(c_s^*(y))) |b| xyf(x,y)dxdy$$
$$+ \gamma_{3,s} \left(1 + \alpha \int_0^1 \int_0^1 (1 - v_s(x,y)) h(c_s^*(y)) |b| yxf(x,y)dxdy \right),$$

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

$$\gamma_{3,s} = \frac{\int_0^1 \int_0^1 (1 - v_s(x,y)) (1 - H(c_s^*(y))) |b| xyf(x,y)dxdy}{1 + \alpha \int_0^1 \int_0^1 (1 - v_s(x,y)) h(c_s^*(y)) |b| yxf(x,y)dxdy},$$

which is strictly greater than 0.

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

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

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 $c_s^*(y) = \lambda_s |b| y$ and the definition of $\tilde{c}_s(y)$, the exposure elasticity is

$$\varepsilon_s(y) = \frac{-(1 - H(c_s^*(y))) \lambda_s |b| y}{(1 - H(c_s^*(y))) (\lambda_s |b| y + \gamma_{3,s} \alpha) + \tilde{c}_s(y)} \left[1 - \frac{h(c_s^*(y))}{1 - H(c_s^*(y))} \gamma_{3,s} \alpha \right].$$

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. It follows that the exposure elasticity is strictly greater than -1, and so the second-best selection policy exhibits an exposure premium. Moreover,

$$\varepsilon_s(y) < \frac{-(1 - H(c_s^*(y))) \lambda_s |b| y}{(1 - H(c_s^*(y))) \lambda_s |b| (y + \tilde{y}_s) + \tilde{c}_s(y)}$$

if and only if

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

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 in Section 4.2, but welfare is now

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

where $\tilde{\sigma}_f(x,y,c)$ is the probability that a non-selected agent with type (x,y) and cost draw c interacts 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 $\tilde{\sigma}_f(x,y,c)$:

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

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 A.3 for the baseline model shows that $\gamma_{1,f} > 0$. To see that $\gamma_{3,f} > 0$, take the derivative of the Lagrangian with respect to λ_f:

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

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

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

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 $c_f^*(y) = \lambda_f|b|y + \gamma_{3,f}\alpha$ and the definition of $\tilde{c}_f(y)$, the exposure elasticity is

$$\varepsilon_f(y) = \frac{-\left(1-H\left(c_f^*(y)\right)\right)\lambda_f|b|y}{\left(1-H\left(c_f^*(y)\right)\right)\lambda_f|b|\left(y+\tilde{y}_f\right)+\tilde{c}_f(y)},$$

which is strictly greater than -1, and so the first-best selection policy exhibits an exposure premium.