# Free and perfectly safe but only partially effective vaccines can harm everyone*

Eduard Talamàs & Rakesh Vohra†

**Abstract**

Risk compensation can undermine the ability of *partially-effective* vaccines to curb infectious-disease epidemics: Vaccinated agents may optimally choose to engage in more risky interactions and, as a result, may increase everyone's infection probability. We show how—in contrast to the prediction of standard models—things can be worse than that: Free and perfectly safe but only partially effective vaccines can harm *everyone*, and hence fail to satisfy—in a strong sense—the fundamental principle of "first, do no harm." Our main departure from standard economic epidemiological models is that we allow agents to strategically choose their partners, which we show creates strategic complementarities in risky interactions. As a result, the introduction of a partially-effective vaccine can lead to a much denser interaction structure—whose negative externalities overwhelm the beneficial direct effects of this intervention.

**Keywords:** Network formation, risk compensation, welfare, strategic complementarities, externalities.

***JEL***: C72, D85, I18.

*Date Printed:* February 28, 2019

†Department of Economics & Department of Electrical and Systems Engineering, University of Pennsylvania. We thank Simon Board, Craig Garthwaite, Edward Glaeser, Benjamin Golub, George Mailath and Pau Milán—as well as various audiences—for helpful comments. This work was supported by the Warren Center for Network & Data Sciences, and the Rockefeller Foundation (#2017PRE301). All errors are our own.

# 1 Introduction

Infectious-disease epidemics like HIV are a major source of human suffering. According to the World Health Organization, about 35 million people have died from HIV, and roughly the same number are currently living with this virus.[1] The development of effective vaccines is crucial for preventing infectious-disease epidemics. Developing an HIV vaccine, for instance, is a high priority for the US National Institutes of Health. Anthony S. Fauci, the Director of the National Institute of Allergy and Infectious Diseases, recently observed[2]

> The development and delivery of a preventive HIV vaccine that is safe and *at least moderately effective* would help bring about a durable end to the HIV/AIDS pandemic. We are committed to pursuing multiple vaccine development strategies to achieve this goal.

In this paper, we show that a free and perfectly safe but only partially-effective vaccine can harm *everyone*, and hence fail to satisfy—in a strong sense—the fundamental principle of "first, do no harm." This implies that—as suggested by Dr. Fauci in the quote above—developing and delivering a vaccine whose effectiveness is below a certain threshold might not be a good idea, and it underscores the importance of developing models that help practitioners estimate this threshold in particular applications.

A partially-effective vaccine has two opposing effects on welfare. On the one hand, it allows agents to have more risky interactions, making them better off. On the other hand, it can increase the probability that agents become infected (as a consequence of the increase in risky interactions), making them worse off. We show that—in contrast to the prediction of standard models—the second effect can uniformly dominate the first.

Our key departure from the standard economic epidemiological models (e.g., Kremer 1996 and Fenichel et al. 2011) is that we allow agents to strategically choose their partners—instead of only allowing each agent to choose her *number* of partners, and then having matches occur at random. This departure is crucial because it uncovers the existence—even in low-risk settings—of strategic complementarities in consensual risky interactions.

Let us illustrate the intuition behind the existence of these strategic complementarities with the simplest example. Suppose that there are two pairs of agents having risky interactions to begin with: Ann and Bob are one pair, and Chloé and Dane the other (Network

[1] See "Global Health Observatory (GHO) data" here.

[2] See the article "NIH and partners launch HIV vaccine efficacy study" here. Emphasis added.

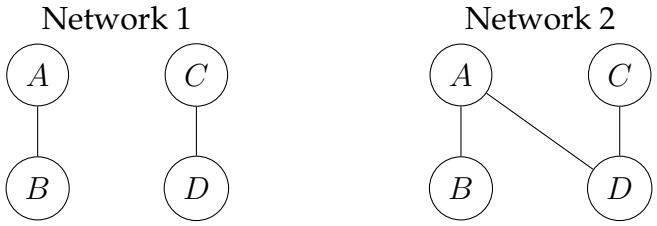



Figure 1: Two illustrative interaction networks.

1 in Figure 1). Each individual has a fixed probability of contracting a given virus independently of her interactions, and an infected individual transmits the virus in any given interaction with probability $p$. Infection and transmission are independent across agents and interactions, respectively. To build intuition, consider first the extreme case in which each interaction transmits the virus with probability one—that is, $p = 1$. An interaction between Chloé and Bob is risky for each of them, since under some states of the world only one of them is infected, and hence an interaction between them would infect the other. We claim that an interaction between Ann and Dane (that is, a switch from Network 1 to Network 2 in Figure 1) increases Chloé and Bob's incentives to interact—conditional on their initial interactions being sufficiently valuable that neither of them wants to become isolated instead. Indeed, in Network 2, the states of the world where one catches the virus are the same as the states of the world where the other one catches it, so their interaction is risk free. A similar intuition holds for *all* positive values of the transmission probability $p$: Ann and Dane's interaction decreases the probability that only one of Chloé and Bob is infected, hence increasing their incentives to interact.

The introduction of a partially-effective vaccine has obvious direct effects: Fixing the network of social interactions, it reduces everyone's probability of becoming infected, and hence makes everyone better off. But the introduction of a partially-effective vaccine can also have subtle indirect effects: By reducing the (ceteris-paribus) cost of each risky interaction, it can destabilize the existing interaction structure. We show how—because of the strategic complementarities in risky interactions illustrated above—the best stable interaction structure after this intervention can be much denser than before such intervention, and—as a consequence of the negative externalities of risky interactions—be worse for everyone. More generally, our analysis highlights that relatively high infection transmission probabilities can play a beneficial role by preventing deviations from the efficient social structure. As a result, the beneficial effects of partially-effective vaccines—in terms of decreased infection probability given any social structure—must be traded off against the welfare effects of the change

in social structure that they unleash.

In the context of a general game with externalities, Hoy and Polborn (2015) show that the combination of strategic complementarities and negative externalities implies that a safety technology improvement can be welfare reducing. From this perspective, the contribution of this paper is to (i) show how strategic complementarities in consensual risky interactions arise naturally when the agents strategically choose their partners, and (ii) illustrate how this can imply that a free and perfectly-safe but only partially-effective vaccine can reduce *everyone's* welfare.

Many social scientists have long realized that social networks play a central role in epidemiological processes (see for example Jacquez et al. 1988, Barnard 1993 and Friedman et al. 2006). Standard economic epidemiological models, however, abstract away from the structure of social interactions, so they are unable to capture the mechanism that we illustrate in this paper. Indeed, a free and perfectly safe but only partially effective vaccine necessarily makes everyone better off in these models. The logic is simple; Kremer (1996, page 555) explains it as follows:[3]

> Adoption of an imperfectly effective vaccine could not cause the number of partners to increase so much that [the per-interaction probability of infection] increased, because people would not be willing to have more partners if the probability of infection from an additional partner increased.

Hence, in these canonical models, everyone is better off after the adoption of a free and perfectly safe but only partially-effective vaccine. Indeed, since such a vaccine decreases the per-interaction probability of infection, everyone can choose the same amount of interaction as she was choosing before its introduction, and in this way obtain the same benefits from her interactions with a reduced probability of infection. In order to clarify this point, and to highlight the role that the strategic choice of partners plays in the mechanism that we illustrate, in Appendix B we describe a random-matching version of our model, and we show that, as long as the infection risk is extreme, strategic complementarities in risky interactions do not arise in this model. As a result, a partially-effective vaccine *increases* everyone's welfare in this random-matching version of our model.

[3] In this quote, we have substituted the symbol $\beta Y$ with its corresponding words: The per-interaction probability of infection. The sentence that follows the one in this quote is: "However, the combined costs of the increased prevalence, plus the expense and side effects of the vaccine, could outweigh the benefits of a reduced risk of infection per partner and so introduction of an imperfect vaccine could make everybody worse off." In this paper we show that an imperfect vaccine can reduce everyone's welfare *even if it is free and has no side effects.*

The remainder of this paper is organized as follows. In section 2, we introduce the model that we use to illustrate our argument and, in section 3, we discuss how strategic complementarities in risky interactions naturally arise in this model. In section 4, we characterize the set of stable networks and, in section 5, we show how a free and perfectly safe but only partially-effective vaccine can harm everyone. We further discuss the contribution of this paper in the context of the related literature in section 6, and we conclude in section 7. In Appendix A, we derive the infection probabilities that we use to prove some of the statements in the main body of the paper. Finally, in Appendix B, we describe a random-matching version of our model, and we show how, in this case, (i) no strategic complementarities arise, and hence (ii) a free and perfectly safe but only partially-effective vaccine necessarily makes everyone better off.

# 2 Epidemiological model with strategic choice of partners

There are $n \geq 2$ men and $n$ women, and four stages, listed below. In order to illustrate the mechanism as simply as possible, we focus on the case in which $n$ is even, and in which agents value only (up to two) partners of the opposite sex.

Stage 1: **Network Formation.** Each agent announces which partners he or she wants to have. An edge between two agents is formed if and only if both of them have announced that they want to partner with the other.

Stage 2: **Infection.** Each agent becomes exogenously infected with probability $q > 0$. Infection is independent across agents.

Stage 3: **Contagion.** Each edge becomes *live* with probability $p > 0$. Each agent connected via a path of live edges to an infected agent becomes infected. Edges become live independently of each other.

Stage 4: **Utility Realized.** The utility of each agent is the benefit that he or she derives from his or her partners ($0$ if no opposite-sex partners, $s_1$ if one opposite-sex partner, and $s_1 + s_2$ if two opposite-sex partners, with $s_2 \leq s_1$) minus the cost of infection ($c$ if infected, and $0$ otherwise).

Stage 1 is the only stage in which the agents take actions. We focus on situations in which having a risky interaction involves mutual consent. To capture this idea, we assume that the outcome in stage 1 is a *stable network*—in the sense that no agent can profitably drop any

subset of his or her edges, and no pair of agents can both benefit by adding an edge between them (while possibly removing some of their existing edges).

This model is similar to the one in Blume et al. (2011): The main difference is that Blume et al. (2011) assume that infected agents do not benefit from their edges, whereas we assume that infected agents benefit from their edges but pay a cost $c$ when they become infected. More importantly, their objective is different: Whereas we focus on the effects of partially-effective vaccines—which we think of as reductions in the probabilities $q$ and $p$—they focus on characterizing the structural differences between optimal and (pairwise) stable networks.

Notation 2.1 reviews standard notation that we shall use throughout the text.

**Notation 2.1.** Given a network $G = (N, E)$ with node set $N$ and edge set $E$, and given $N' \subseteq N$ and $E' \subseteq E$, we say that $(N', E')$ is a *subnetwork of* $G$. We say that a subnetwork $G' = (N', E')$ of $G$ is a *component of* $G$ if $E$ contains no edges between $N'$ and $N \backslash N'$, and, for any two nodes $i$ and $j$ in $N'$, there is a *walk* (sequence of edges) in $E'$ from $i$ to $j$.

# 3 Strategic complementarities in risky interactions

We start by showing how strategic complementarities in risky interactions naturally arise in this model. For simplicity, we focus on the case in which the value $s_1$ of having one partner is high enough so that no network with an isolated agent is stable.

Figure 2 depicts all the possible networks (up to isomorphism) of four non-isolated agents that can emerge in the network formation stage. Let $\mu_i$ denote the probability that agent $i$ becomes infected (exogenously—i.e., in stage 2—or endogenously—i.e., in stage 3); for simplicity we denote by $\mu_I$ and $\mu_X$ the infection probability of any given agent in the symmetric networks $I$ and $X$, respectively. In Appendix A we derive the infection probability of the agent in each relevant network position.

Proposition 3.1 uses Definition 3.1 to formalize the idea that risky interactions are strategic complements.[4]

**Definition 3.1.** Given a network $G$, the *risk of the edge $ij$ for agent $i$* is the difference in $i$'s infection probability in $G \cup ij$ and $i$'s infection probability in $G$. When the risk of edge $ij$ is the same for agents $i$ and $j$, we refer to it simply by the *risk of the edge $ij$*.

**Proposition 3.1.** *The risk of the edge $N_2N_3$ in Network $N$ is smaller than the risk of the edge $I_2I_3$ in Network $I$.*

[4]For brevity, we denote the edge between nodes $i$ and $j$ by $ij$.

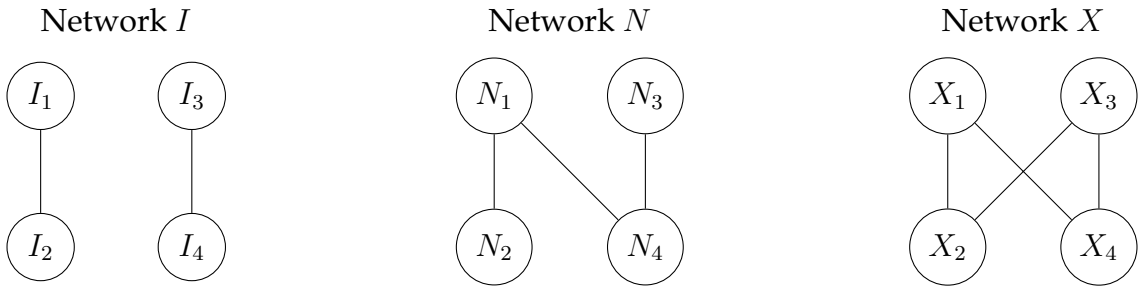



Figure 2: The networks $I$, $N$ and $X$.

*Proof.* The risk of edge $N_2N_3$ is $\mu_X - \mu_{N_2}$, and the risk of edge $I_2I_3$ is $\mu_{N_1} - \mu_I$. Using the expressions derived in Appendix A, it is easily verified that $\mu_X - \mu_{N_2} < \mu_{N_1} - \mu_I$ for all values of $p$ and $q$. □

Figure 4 depicts the risk of edge $N_2N_3$ and $I_2I_3$ as a function of the transmission probability $p$ when the exogenous infection probability is $q = \frac{1}{4}$; the picture looks similar for all $q \in (0, 1)$. The risk of edge $N_2N_3$ is increasing for low values of $p$ and decreasing for high values of $p$. Intuitively, the risk of edge $N_2N_3$ is highest when the transmission probability is high enough so that this edge has a substantial probability of transmitting an infection but low enough so that there is a substantial probability that only one of $N_2$ and $N_3$ are infected.

# 4 Stable networks

Proposition 4.1 shows that only the *pair-complete* and the *cross-complete networks* (Definition 4.1) can be stable. Every pair-complete network is isomorphic to $n/2$ copies of *Network $I$*, and every cross complete network is isomorphic to $n/2$ copies of *Network $X$*.

**Definition 4.1.** We say that a network is *pair complete* if each agent is part of exactly one edge (that is, each agent has exactly one partner). We say that a network is *cross complete* if each agent is in a component of the network that is isomorphic to *Network $X$* (see Figure 2).

Proposition 4.1 shows that pair-complete networks are stable for intermediate values of the transmission probability $p$, which is intuitive: When the transmission probability $p$ is small enough, *Network $I$* is not stable because agents have incentives to form the diagonal links. In contrast, when the transmission probability $p$ is high enough, *Network $I$* is not stable because agents have incentives to remove their one link. Proposition 4.1 also shows that

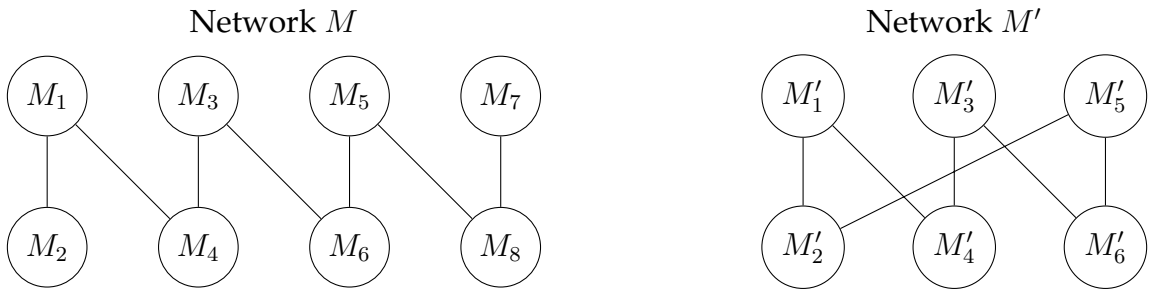



Figure 3: The networks $M$ and $M'$.

cross-complete networks are stable with the only potential exception of a range of intermediate values of transmission probability, which is also intuitive: The benefit from removing the edge $X_2X_3$ for its adjacent vertices is highest for intermediate values of $p$, when it is most likely that exactly one of them is infected in network $N$.

**Proposition 4.1.** *Only* pair-complete *and* cross-complete networks *are ever stable. Moreover:*

- *There exist $p^\star \leq p^{\star\star}$ such that any* pair-complete network *is stable if and only if $p$ is in $[p^\star, p^{\star\star}]$.*
- *There exist $p^\star < \underline{p} \leq \overline{p}$ such that any* cross-complete network *is stable if and only if $p$ is not in $(\underline{p}, \overline{p})$.*

*Proof.* A pair-complete network is stable if and only if (i) no agent wants to remove her existing edge (that is, the cost $s_1$ of removing this edge is greater than the associated benefit $c\mu_I$) and (ii) no two agents have incentives to partner up (that is, the cost $c(\mu_{N_1} - \mu_I)$ of an extra edge is greater than its benefit $s_2$). Using Equation 1 in Appendix A, it is easily verified that there exists $p^{\star\star}$ such that condition (i) holds for all small $p < p^{\star\star}$. Using Equation 3 in Appendix A, it is easily verified that there exists $p^\star < p^{\star\star}$ such that condition (ii) holds for all $p > p^\star$.

A cross-complete network is stable if and only (i) the benefit $c(\mu_X - \mu_{N_2})$ from deleting an edge is smaller than its cost $s_2$ and (ii) no two agents have incentives to remove one of their edges and create an edge between them; that is, $\mu_{M_3} \geq \mu_X$ (see Figure 3). Using Equation 4 in Appendix A, it is easily verified that condition (i) is satisfied for all $p$ except possibly those in an intermediate range $(\underline{p}, \overline{p})$. The fact that $p^\star < \underline{p}$ follows from Proposition 3.1. Condition (ii) follows from the fact that combining Equation 2 and Equation 6 in Appendix A gives $\mu_{M_3} > \mu_X$.

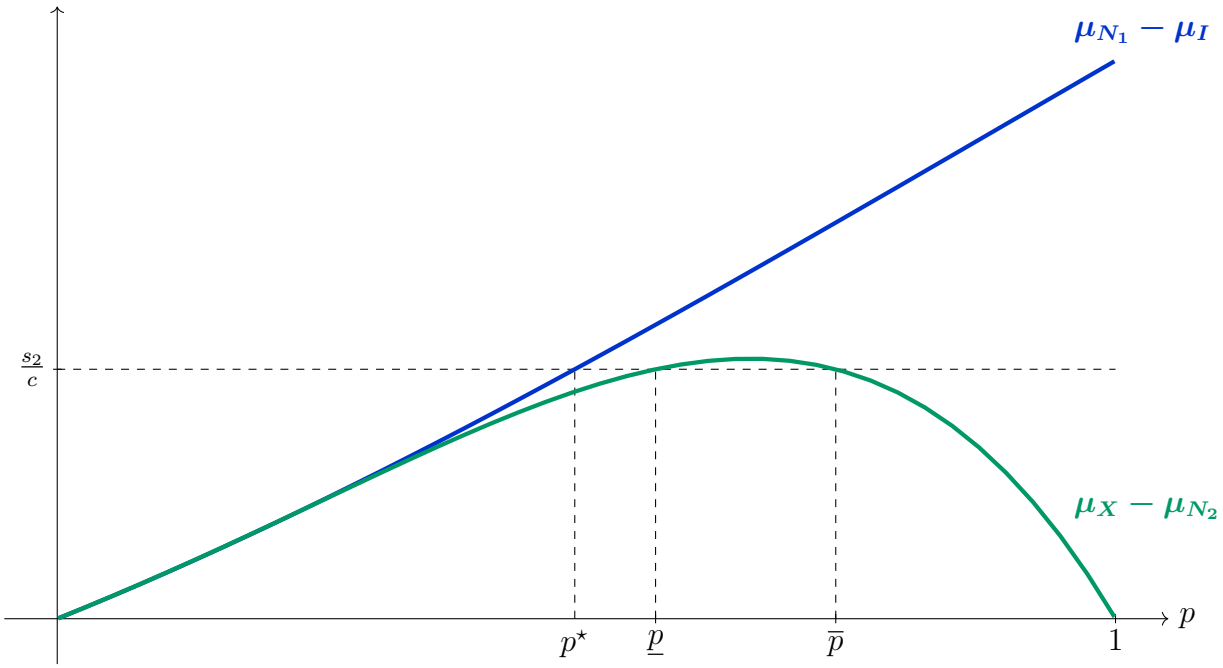



Figure 4: Illustration of Proposition 4.1 and when $q = \frac{1}{4}$ and $\frac{s_2}{c} = .11$. Assuming $s_1$ is large enough, a pair-complete network is stable if and only if $\frac{s_2}{c}$ is below $\mu_{N_1} - \mu_I$. A cross-complete network is stable if and only if $\frac{s_2}{c}$ is above $\mu_X - \mu_{N_2}$.

It only remains to be shown that a network that is neither pair complete nor cross complete is not stable. The fact that $\mu_{M_3} > \mu_X$ implies that no subnetwork of a stable network is isomorphic to network $M$. Note that no agent is part of more than two edges in any stable network, so a component of such a network contains a node that is part of two edges if and only if all of the nodes in this component are part of two edges (otherwise, this component contains at least two agents that are part of only one edge, who can profitably deviate by removing their one edge and matching to each other). Hence, we only have to show that no subnetwork of a stable network is isomorphic to network $M'$ (see Figure 3). This follows from the fact that $\mu_X < \mu_{M'}$, where $\mu_{M'}$ denotes the infection probability of each agent in $M'$.[5] This last inequality is intuitive: The probability that each of the agents $M'_1, M'_2, M'_3$ and $M'_4$ becomes infected goes down when the three edges $M'_3M'_6$, $M'_6M'_5$, and $M'_5M'_2$ are replaced by the edge $M'_3M'_2$, since this leaves each of them with the same number of edges but decreases the independent sources of infection. □

Figure 4 illustrates the determinants of the cutoff $p^\star$ when the infection probability is $q = \frac{1}{4}$ and $\frac{s_2}{c} = .11$. For simplicity, in Figure 4, we don't show the determinants of the cutoff $p^{\star\star}$;

[5]Indeed, $\mu_X < \mu_{M'}$ implies that $M'_2$ and $M'_3$ strictly benefit from partnering while dropping their partnerships with $M'_5$ and $M'_6$, respectively.

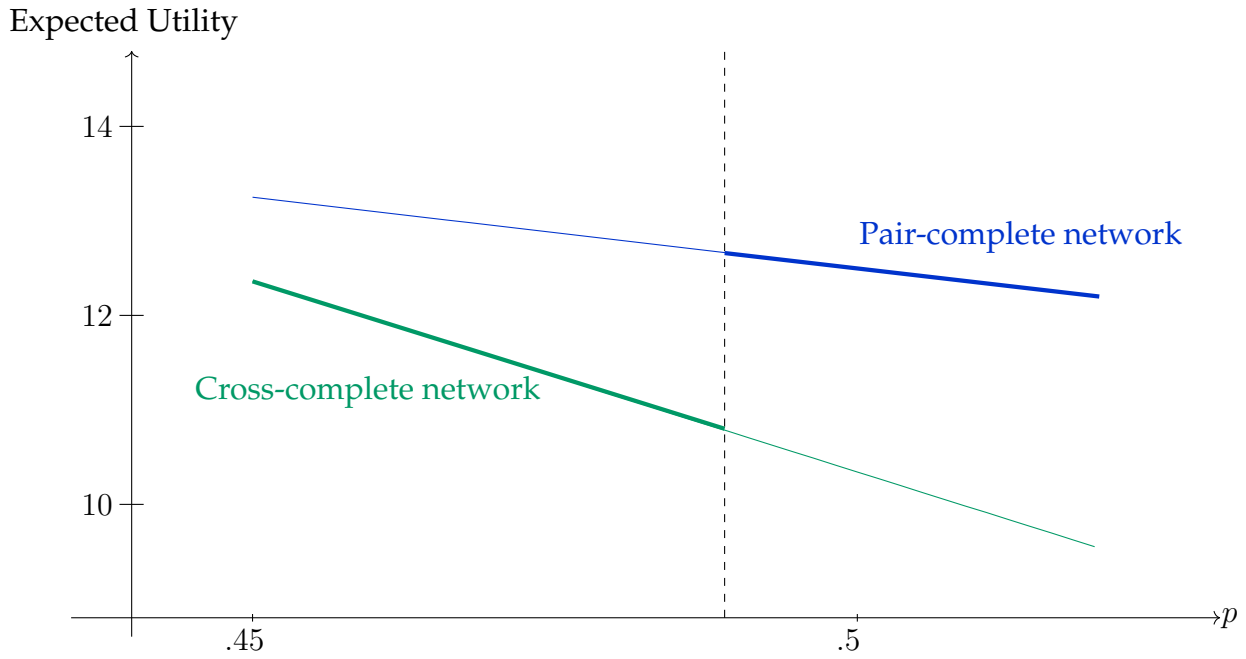



Figure 5: Illustration of Theorem 5.1 when $s_1 = 40, c = 80, \frac{s_2}{c} = .11$ and $q = 1/4$. Each agent's expected utility under the most efficient stable network structure is in bold.

this cutoff is 1 if $s_1$ large enough. Figure 4 also illustrates the determinants of the cutoff values $\underline{p}$ and $\overline{p}$ when the infection probability is $q = \frac{1}{4}$ and $\frac{s_2}{c} = .11$. For large enough values of $\frac{s_2}{c}$, every cross-complete network is stable for all transmission probabilities $p$.

# 5 Partially-effective vaccines can harm everyone

Pair-complete and cross-complete networks—the only two (non-trivial) potentially stable types of networks—are fully symmetric. Hence, in every stable network, all agents' expected utilities are the same. Therefore, the total welfare in each network scales with the expected utility of a single agent in this network.

From Proposition 4.1, we have that there always exists a nonempty region $(p^\star, \underline{p})$ of values of the transmission probability $p$ in which (i) both pair-complete and cross-complete networks are stable and (ii) a reduction in $p$ leads to only pair-complete networks being stable. Theorem 5.1 follows directly from this observation and the fact that, for all values of $p$ close enough to $p^\star$, everyone's expected utility in a pair-complete network is strictly greater than in a cross-complete network. To see this last fact, note that, when $p = p^\star$, agent 1's expected utility in network $I$ is the same as that in network $N$, and hence, each agent's expected utility in a pair-complete network is strictly greater than that in a cross-complete

network.

**Theorem 5.1.** *There exists $\Delta > 0$ such that each agent's expected utility in the most efficient stable network when $p \in (p^\star, p^\star + \Delta)$ is strictly greater than when $p \in (p^\star - \Delta, p^\star)$.*

Figure 5 illustrates Theorem 5.1 for a particular utility function ($s_1 = 40$, $c = 80$ and $\frac{s_2}{c} = .11$) and exogenous infection probability $q = \frac{1}{4}$. In this case, $p^\star$ is approximately $.48$.

**Remark 5.2.** *Theorem 5.1 implies that there is always a non-empty range of transmission probabilities above $p^\star$ such that—assuming that we start with a transmission probability in this range and from an efficient stable network (a pair-complete network)—there is a non-empty range of reductions in the transmission probability $p$ that necessarily harm everyone. A similar statement holds for interventions that either reduce $q$, or that reduce both $p$ and $q$.*

# 6 Relation to the existing literature

The well-known phenomenon of *risk compensation* is an important element of the mechanism that we describe in this paper. Observed at least as early as the Victorian era (see for example Adams 1879), it was popularized by Peltzman (1975), who controversially suggested that automobile safety regulations would not diminish automobile-related deaths.

In the context of HIV, the evidence on risk compensation is mixed.[6] For example, on the one hand, Eaton and Kalichman (2007) (see also Chan et al. 2015, Delavande and Kohler 2015, and Blumenthal and Haubrich 2017) review the empirical literature on risk compensation in HIV prevention and conclude that "risk compensation is evident in response to prevention technologies that are used in advance of HIV exposure and at minimal personal cost." On the other hand, Marcus et al. (2013) argue that there is no evidence of risk compensation in a recent trial of Daily Oral HIV Preexposure Prophylaxis (iPrEx).

In this paper we show how strategic complementarities in consensual risky interactions naturally arise when agents strategically choose their partners, and how this—combined with risk compensation—can imply that a free and perfectly safe but only partially effective vaccine can make *everyone worse off*. This implies that there exists a non-trivial welfare trade-off when considering whether to distribute partially-effective vaccines: While such an intervention increases the welfare associated with any given interaction structure, it can disrupt the existing structure in favor of a more inefficient one.

[6] See Philipson and Posner (1993) for an interesting in-depth economic perspective on the HIV epidemic.

The mechanism that we illustrate in this paper is distinct from the one described in Kremer (1996), which relies on heterogeneities in agents' preferences, and can be summarized as follows: If low-activity people reduce their activity by a higher proportion than high-activity people in response to an *increase in the prevalence of a disease*, the composition of the pool of available partners worsens after such a change, which creates positive feedbacks. In contrast with our mechanism, however, the feedback effects in Kremer (1996) only make partially-effective vaccines more desirable. For example, for those with sufficiently many partners, the introduction of a partially-effective vaccine will actually increase the marginal risk of infection from an additional partner, reducing their optimal number of partners, and hence making the pool of available partners safer for everyone. These feedback effects are absent in our analysis because—in order to illustrate our mechanism as simply as possible—we focus on the case of homogeneous preferences.

This paper complements the growing body of theoretical literature that studies the effects of different interventions on epidemiological processes (e.g., Galeotti and Rogers 2013, Chen and Toxvaerd 2014, Rowthorn and Toxvaerd 2015, Goyal and Vigier 2015 and Goyal, Jabbari, Kearns, Khanna, and Morgenstern 2016). The main difference between this paper and most of this literature is that we focus on the *welfare effects* of such interventions—rather than the effects on *infection rates*. In a similar vein, Toxvaerd (2017) uses a dynamic version of a standard economic epidemiological model to argue that partially-effective vaccines can have negative welfare consequences in the transition between steady states. In contrast, we show—using a different model that allows agents to strategically choose whom to interact with—that the conclusion that a free and perfectly safe but only partially-effective vaccine necessarily makes everyone better off in steady state is an artifact of the anonymous-mixing assumption of the standard models.

# 7 Conclusion

The capacity of infectious-disease epidemics to disrupt societies is comparable to that of wars and natural disasters. For this reason, considerable resources are expended to manage and ameliorate the effects of such epidemics. Because of risk compensation, however, the effects of different potential interventions are subtle. As a consequence, before deciding whether and how to intervene, we might wish to ensure that our interventions at least do no harm.

In this paper we show that an intervention that consists of distributing a free and perfectly

safe but only partially effective vaccine can fail this fundamental principle of "first, do no harm" in a strong sense, since it can actually harm *everyone*. A key force in the mechanism that we uncover is that consensual risky interactions can feature strategic complementarities, even in low-risk environments. These strategic complementarities can generate feedback effects that can lead to large changes in the structure of risky interactions after a relatively small intervention—overwhelming the direct effects of this intervention.

The result of this paper underscores the importance of taking into account the network of social interactions in theoretical and empirical epidemiological studies in order to understand the potential effects that different interventions have on social structure—and hence on behavior and welfare. Measuring the relevant interaction structure—and how it changes with different interventions—can be crucial for understanding which social groups are more likely to feature strategic complementarities in risky interactions, and hence which parts of a society are more vulnerable to the potentially-negative welfare effects of different interventions.

# A Appendix: Infection probabilities

Lemma A.1 describes the probability that an agent becomes infected (exogenously—i.e., in stage 2—or endogenously—i.e., in stage 3) conditional on her network position.

**Lemma A.1.** *The probability that any given agent in network $I$ is infected is*

$$\mu_I = qp + (1 - qp)q. \tag{1}$$

*the probability that any given agent in network $X$ is infected is*

$$\begin{aligned} \mu_X \;=\;& (1-p)^2\mu_I \\ &+\; p(1-p)\left[q(2-q) + (1-q)^2 qp(2-qp)\right] \\ &+\; p(1-p)\left[q + (1-q)p(pq + (1-pq)q(2-q))\right] \\ &+\; p^2\left[q(2-q) + (1-q(2-q))q(2-q)p(2-p)\right], \end{aligned} \tag{2}$$

*the probability that $N_1$ is infected is*

$$\mu_{N_1} = \mu_I + (1-q)(1-pq)\mu_I p, \tag{3}$$

*the probability that $N_2$ is infected is*

$$\mu_{N_2} = q + (1-q)pq + (1-q)^2 p^2 \mu_I, \tag{4}$$

*the probability that $M_3$ is infected is*

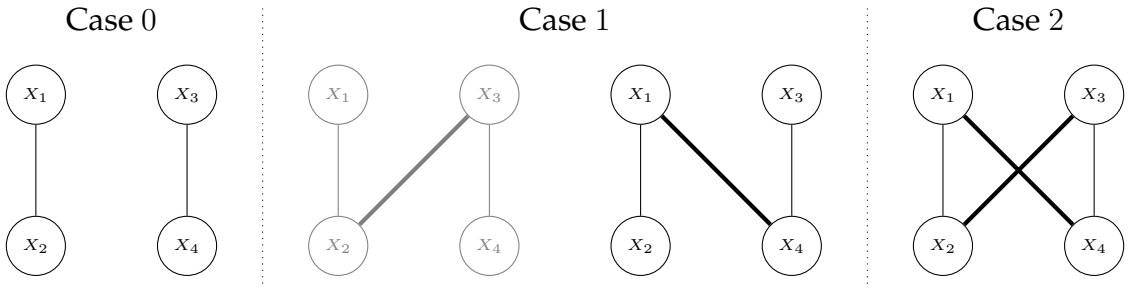
$$\mu_{M_3} = \mu_{N_2} + (1 - \mu_{N_2})\mu_I p, \tag{5}$$

*and, finally, the probability that $M_3$ is infected is*

$$\mu_{M_3} = \mu_{N_2} + (1 - \mu_{N_2})\mu_I p. \tag{6}$$

*Proof.* To see Equation 1, consider for concreteness the probability that $I_1$ is infected. The probability that $I_2$ infects $I_1$ is $qp$ and, conditional on not being infected by $I_2$, $I_1$ is infected with probability $q$.

To derive Equation 2, consider the three exhaustive and mutually exclusive cases depicted below, where thick edges correspond to live edges. We say that $i$ *is infected from* $j$ if $j$ is exogenously infected and there is an live path from $i$ to $j$.



***Case 0:*** *None of the edges $X_1X_4$ and $X_2X_3$ are live.* This happens with probability $(1-p)^2$. The probability that any given agent is infected is $\mu_I$.

***Case 1:*** *Exactly one of the edges $X_1X_4$ and $X_2X_3$ is live.* This happens with probability $2p(1-p)$. Assume without loss of generality that $X_1X_4$ is live (and hence $X_2X_3$ is not live). The probability that node $X_1$ is infected is $q(2-q)+(1-q(2-q))(qp+(1-qp)qp)$ or

$$q(2-q)+(1-q)^2qp(2-qp) \tag{7}$$

To see this, note that the probability that $X_1$ is infected exogenously or from $X_4$ is $1-(1-q)^2 = q(2-q)$, and the probability that $X_1$ is infected from $X_2$ or $X_3$ is $qp+(1-qp)qp$.

The probability that node $X_2$ is infected is

$$q+(1-q)p(pq+(1-pq)q(2-q)) \tag{8}$$

To see this, note that the probability that $X_2$ is exogenously infected is $q$. Conditional on this not happening, the probability that $X_2$ is infected is $p$ times the probability that $X_1$ is infected from $X_3$, or $X_1$ or $X_4$, which is $pq+(1-pq)q(2-q)$.

We conclude that each agent's expected probability of infection in this case is the average of expressions (7) and (8).

***Case 2:*** *Both edges $X_1X_4$ and $X_2X_3$ are live.* This happens with probability $p^2$. The probability that $X_1$ is infected is

$$q(2-q)+(1-q(2-q))q(2-q)p(2-p).$$

To see this, note that the probability that $X_1$ is infected exogenously or from $X_4$ is $1-(1-q)^2 = q(2-q)$, and the probability that $X_1$ is infected from $X_3$ or $X_4$ is the probability $q(2-q)$ that either of them is infected times the probability $p(2-p)$ that at least one of the edges $X_1X_2$ and $X_3X_4$ is live.

To see Equation 3, note that $\mu_{N_1}-\mu_I = (1-q)(1-pq)\mu_I p$, since the probability that $N_1$ is infected from either $N_3$ or $N_4$ and is not infected from either $N_1$ or $N_2$ is the probability $1-q$ that $N_1$ is not infected from $N_1$ times the probability $1-qp$ that $N_1$ is not infected from $N_2$ times the probability $\mu_I$ that $N_4$ is infected either exogenously or from $N_3$ times the probability $p$ that the edge $N_1N_4$ is live.

To see Equation 4, note that the probability that $N_2$ is infected is the probability $q$ that she becomes exogenously infected plus the probability $(1-q)p$ that she does not become infected and $N_1N_2$ is live times the probability $q+(1-q)p\mu_I$ that $N_1$ is infected exogenously, from $N_3$ or from $N_4$. That is, $\mu_{N_2} = q+(1-q)p\left[q+(1-q)p\mu_I\right]$, which is equivalent to Equation 4.

To see Equation 6, note that the probability that $M_3$ is infected is the probability $\mu_{N_2}$ that she becomes infected exogenously or from either $M_1, M_2$ or $M_4$ plus the complement probability times the probability that the edge $M_3M_6$ is live times the probability that $M_6$ is infected exogenously or from $M_5, M_7$ or $M_8$. □

# B Epidemiological model with random matching

In this section we describe a model similar to the one described in section 2 that features random matching instead of strategic choice of partners. We then discuss how both the strategic complementarities in risky interactions and the negative welfare effects of partially-effective vaccines naturally vanish in this case. There are $n \geq 2$ agents, and four stages, listed below. For simplicity, we focus on the case in which $n$ is even.

Stage 1: **Network Formation.** Each agent announces how many partners he or she wants to have. Edges are then formed as follows: One pair of agents is selected uniformly at random, and if both of them have less edges than the number that they have announced, an edge is formed between them. This process continues until everyone but at most one agent has less edges than the number that she has announced.

Stage 2: **Infection.** Each agent becomes exogenously infected with probability $q$. Infection is independent across agents.

Stage 3: **Contagion.** Each edge becomes *live* with probability $p$. Each agent connected via a directed path of live edges to an infected agent becomes infected. Edges become live independently of each other.

Stage 4: **Utility Realized.** The utility of each agent is the benefit that she derives from his partners ($0$ if no partners, $s_1$ if one partner, and $s_1 + s_2$ if two partners) less the cost of infection ($c$ if infected, and $0$ otherwise).

When everyone announces that she or she wants to have one partner, the outcome of the random network formation is necessarily a pair-complete network (Definition 4.1). Hence, it follows from our analysis in section 4 that this is an equilibrium if and only if $p \in (p^\star, p^{\star\star})$.

When the number of agents $n$ is large, and each agent is part of at most two edges, the process of random matching implies that the event that an agent $i$ transmits the infection to an agent $j$ is approximately independent from the event that an agent $k \neq i$ transmits the

infection to agent $j$. For simplicity, in what follows, we shall assume that $n$ is large enough, so that these events can be safely treated as being independent (alternatively, we can assume, as is standard in the literature using epidemiological models with random matching, that there is a continuum of agents, so that these approximations are exact). Also, we focus on the natural case in which the equilibrium *per-interaction* infection probability is below $1/2$ (this is guaranteed, for example, by $c \geq 2s_1$).[7]

When an agent decides to have more partners, the probability of becoming infected from each interaction increases. Hence, everyone's incentives to have additional partners diminish after this deviation. In other words, *with random matching, there are no strategic complementarities in risky interactions*. As a consequence, as we now discuss, *partially-effective vaccines cannot harm anyone in this random-matching version of our model*.

Suppose for contradiction that a partially-effective vaccine reduces someone's expected utility. Denote by $\mu$ and $\mu'$ the average equilibrium per-interaction infection probability before and after introducing the vaccine, respectively. We have that

$$\mu' > \mu. \tag{9}$$

Indeed, otherwise, each agent can have an expected utility at least as high after the intervention as she had before the intervention—by doing exactly as she was doing before.

Equation 9 implies that some agents must be choosing strictly more partners after the intervention. Moreover, the vaccine cannot reduce anyone's expected utility in equilibrium if either everyone is choosing zero partners or everyone is choosing two partners before the intervention, so we have that

$$s_1 \geq \underbrace{\mu pc}_{\text{expected infection cost of first interaction}} \quad \text{and } s_2 \leq \underbrace{(1-\mu p)\mu pc}_{\text{expected infection cost of second interaction}}$$

Hence, the following two cases are exhaustive. On the one hand, if no one chooses to have zero partners before the intervention, Equation 9 implies that some agents choose to have two partners after the intervention, so $s_2 \geq (1-\mu' p)\mu' pc > (1-\mu p)\mu pc$, a contradiction.[8] On the other hand, if some (but not all) agents choose to have zero partners before the intervention, we have that $\mu pc = s_1$, so that $\mu' pc > s_1$ (i.e., no one chooses to have any partners after the intervention), also a contradiction.

[7] This rules out situations in which agents are *fatalistic*, in the sense that they are so likely to become infected by their first interaction that they might as well choose additional interactions (e.g., Kremer 1996). Note that the mechanism that we illustrate in this paper does not rely on such extreme scenarios.

[8] This inequality follows from the assumption that the equilibrium per-interaction probability $\mu' p$ is below $1/2$.

# References


ADAMS, C. F. (1879): "Notes on railroads accidents," *G.P. Putnam Sons*.

BARNARD, M. A. (1993): "Needle sharing in context: patterns of sharing among men and women injectors and HIV risks," *Addiction*, 88, 805–812.

BLUME, L., D. EASLEY, J. KLEINBERG, R. KLEINBERG, AND É. TARDOS (2011): "Network formation in the presence of contagious risk," in *Proceedings of the 12th ACM conference on Electronic commerce*, ACM, 1–10.

BLUMENTHAL, J. AND R. HAUBRICH (2017): "Risk compensation in PrEP: An old debate emerges yet again," *The Virtual Mentor*, 909–915.

CHAN, T. Y., B. H. HAMILTON, AND N. W. PAPAGEORGE (2015): "Health, risky behaviour and the value of medical innovation for infectious disease," *The Review of Economic Studies*, 83, 1465–1510.

CHEN, F. AND F. TOXVAERD (2014): "The economics of vaccination," *Journal of Theoretical Biology*, 363, 105–117.

DELAVANDE, A. AND H.-P. KOHLER (2015): "HIV/AIDS-related expectations and risky sexual behaviour in Malawi," *The Review of Economic Studies*, 83, 118–164.

EATON, L. A. AND S. C. KALICHMAN (2007): "Risk compensation in HIV prevention: Implications for vaccines, microbicides, and other biomedical HIV prevention technologies," *Current HIV/AIDS*, 165–172.

FENICHEL, E. P., C. CASTILLO-CHAVEZ, M. G. CEDDIA, G. CHOWELL, P. A. G. PARRA, G. J. HICKLING, G. HOLLOWAY, R. HORAN, B. MORIN, C. PERRINGS, ET AL. (2011): "Adaptive human behavior in epidemiological models," *Proceedings of the National Academy of Sciences*, 108, 6306–6311.

FRIEDMAN, S. R., R. CURTIS, A. NEAIGUS, B. JOSE, AND D. C. DES JARLAIS (2006): *Social networks, drug injectors' lives, and HIV/AIDS*, Springer Science & Business Media.

GALEOTTI, A. AND B. W. ROGERS (2013): "Strategic immunization and group structure," *American Economic Journal: Microeconomics*, 5, 1–32.

GOYAL, S., S. JABBARI, M. KEARNS, S. KHANNA, AND J. MORGENSTERN (2016): "Strategic network formation with attack and immunization," in *International Conference on Web and Internet Economics*, Springer, 429–443.

GOYAL, S. AND A. VIGIER (2015): "Interaction, protection and epidemics," *Journal of Public Economics*, 125, 64–69.

HOY, M. AND M. K. POLBORN (2015): "The value of technology improvements in games with externalities: A fresh look at offsetting behavior," *Journal of Public Economics*, 131, 12–20.

JACQUEZ, J. A., C. P. SIMON, J. KOOPMAN, L. SATTENSPIEL, AND T. PERRY (1988): "Modeling and analyzing HIV transmission: the effect of contact patterns," *Mathematical Biosciences*, 92, 119–199.

KREMER, M. (1996): "Integrating behavioral choice into epidemiological models of AIDS," *The Quarterly Journal of Economics*, 111, 549–573.

MARCUS, J., D. GLIDDEN, K. MAYER, A. LIU, S. BUCHBINDER, K. AMICO, ET AL. (2013): "No evidence of sexual risk compensation in the iPrEx trial of daily oral HIV preexposure prophylaxis," *PLoS ONE*, 8(12).

PELTZMAN, S. (1975): "The effects of automobile safety regulation," *Journal of Political Economy*, 83, 677–725.

PHILIPSON, T. AND R. POSNER (1993): *Private choices and public health: An economic interpretation of the AIDS epidemic*, Harvard University Press.

ROWTHORN, B. R. AND F. TOXVAERD (2015): "The optimal control of infectious diseases via prevention and treatment," *Mimeo*.

TOXVAERD, F. (2017): "Rational disinhibition and externalities in prevention," *Mimeo*.