Information Disclosure about Booster Efficacy in a Non-Stationary Environment

Sohil Shah, Saurabh Amin, Patrick Jaillet

Abstract— **This paper investigates the dynamic disclosure of information in non-stationary environments. In particular, a planner iteratively discloses information about the efficacy of an immunizing booster shot that stochastically evolves over time amid the long-run spread of an infectious disease whose severity also varies over time. Each time period, a heterogeneous population of agents uses the disclosed information to determine whether they should obtain the booster shot, and then whether to remain isolated or active. The central planner's objective is to ensure that the active population remains above a minimum threshold each period. We characterize a Markov decision process over the state of beliefs and how signalling mechanisms act on them. We highlight the "greedy" disclosure rule which provides the least amount of information possible subject to the planner maximizing the likelihood of achieving the active population threshold in the current period. Our results demonstrate that the greedy disclosure rule becomes optimal in finite time. We show this for settings where the population's belief over the booster's efficacy becomes more pessimistic than the belief required in the long-run.**

I. INTRODUCTION

A. Motivation

Repetitively administering multiple vaccine doses or "boosters" is a potent tool in mitigating long-run disease spread in order to enable in-person activities. However, as diseases mutate, boosters may require dynamic updates and could become unsafe or less effective [1]. This makes it difficult to persuade individuals to make use of boosters, especially if the previous information they have collected suggests that boosters are flawed. The COVID-19 outbreak renewed interest in how information affects serial immunization and social activity [2], [3], This paper investigates the use of *strategic dynamic information disclosure* to shape public beliefs and maintain desired activity levels over time. Each time period, a central planner carefully designs stochastic "experiments" that depend on the true uncertain booster efficacy at that time. The experiments' outcomes (i.e. signals) are used by the population to update their existing beliefs over the true booster efficacy in that time period. Planners can thereby influence the population towards more desirable beliefs and consequently desirable outcomes. Examples of these experiments may include choosing to fund specific types of studies that measure vaccine effectiveness or establishing rules that censor particular data sources.

B. Our work

In this paper, we consider a model where, each period, a unit mass of strategic agents has the choice to be active or remain isolated amid disease spread. The active agents can choose to take a booster shot at a cost to receive an additional level of protection against the disease. The planner aims to ensure the active population size remains above a specified threshold. The disease's *infectiousness* changes over time according to finite Markov chain. Independent of this, the booster shot's efficacy also evolves according to a finite Markov chain. The planner can provide information about the booster's efficacy in each period, but disclosing precise information in one period limits the planner's ability to shape agents' beliefs in future periods. Each period in our model practically could correspond to the time length between the emergence of new disease variants or booster technologies. We investigate the optimal dynamic disclosure rule and identify which signalling mechanisms the planner should optimally choose in each period by solving a non-stationary Markov decision process. Our work contributes to the literature by providing new insights into the optimal design of dynamic information for heterogenous agents in non-stationary environments such as long-run disease transmission. Our technical contribution includes a new strategy to identify the optimal dynamic disclosure rule when the set of beliefs the planner seeks to induce is non-stationary.

C. Literature Review

Information disclosure is the key tool we motivate as an effective disease mitigation tool [4], [5]. Several works have demonstrated the value of information disclosure as a means to disrupt disease spread [6], [7]. The authors of [8] consider a model of hybrid workers that we adapt by making dynamic and by allowing agents to immunize with boosters.

Critically, in contrast to existing work on information disclosure to mitigate disease spread, our model discloses information about booster efficacy each period. Our work falls under the subdomain of *dynamic information disclosure* which requires planners to account for both the latent evolution in uncertainty and the implications of information disclosed in the present on future beliefs.

Several works provide insightful models into dynamic information disclosure [9], [10], [11]. Most notably, [12] solves for the optimal dynamic information disclosure rule for forward-looking agents that choose when to halt a stochastic process to match a switch in the uncertain binary state. By contrast, agents in our setting continuously make decisions and we consider Markovian dynamics on the uncertain parameter which has direct modelling applications for disease evolution. Their results suggest that planners can delay information transmission to entice the agent to wait longer. We

The authors are affiliated with the Laboratory for Information and Decision Systems, MIT {sshah95, amins, jaillet}@mit.edu

recover a similar insight about keeping agents less informed in our model as we show that planners defer information disclosure when the current period reward is maximized.

Most closely related to our work, [13] presents a model with a myopic receiver where the uncertain parameter also evolves according to a Markov chain. Our work specifically improves upon their results and this subdomain by considering multiple heterogeneous agents and non-stationary objectives. Our model has two sources of non-stationary uncertainty – booster efficacy and disease infectiousness. Since beliefs on both uncertain quantities impact agents' actions, the time-varying disease infectiousness creates a non-stationary environment that cannot be captured in their setting. Using distinct technical tools, our results analogously show the robustness and optimality of the "greedy" mechanism that *minimizes* information disclosure subject to maximizing the current period reward.

More practically, various works studying disease severity over time have shown that diseases mutate in favor of *continued survival* which leads diseases to become easier to transmit, but less deadly with time [14], [15]. Moreover, the dynamic mutation of disease has empirically resulted in changes to the efficacy of pharmaceuticals such as boosters and vaccines [16]. Our model captures these features by allowing the uncertain parameters and beliefs to drift latently according to the Markov chain. Furthermore, the utility of information disclosure has been practically motivated as [17] and [18] have shown that communicating information about the efficacy or safety of vaccines significantly influences the adaptation of boosters and vaccines.

D. Outline

We outline the technical contributions of our work. In Section II, we describe the model of our agents and the evolution of both the disease infectiousness and the uncertain booster efficacy. We formalize the signalling mechanisms that the planner uses to disclose information to the agents and describe the planner's utility model.

In Section III, we compute a cutoff for the private benefit from activity. At equilibrium, agents with values below the cutoff isolate and agents with values above the cutoff are active – either all with a booster or all without a booster (Prop. 1). This enables us to compute the isolated mass in terms of the beliefs over the effectiveness of the booster and the infectiousness in the current period. We show that the map between beliefs and the isolated mass and vice versa are monotone and continuous (Lemmas 1, 2, 3).

Next, in Section IV, we characterize the value functions mapping current beliefs to the total discounted sum of future rewards. We show that the functions are concave (Prop. 2) and that identifying the segments where the value function is linear is sufficient to determine the optimal disclosure rule (Lemma 4). We define the "greedy" disclosure rule which discloses the minimal amount of information subject to maximizing the current period reward. We show that, after finite time, this disclosure rule becomes optimal for perpetuity (Theorem 1 and Prop. 7).

Finally, in Section V, we provide a brief numerical comparison of the greedy disclosure rule against fully-informative and non-informative disclosure rules.

Due to space limitation, proofs are deferred to [19].

II. PROBLEM FORMULATION

A. Agent Model and Disease Evolution

We consider a dynamic model of agents' decision-making subject to information disclosure in the presence of an infectious disease. The time horizon is discrete, indexed by $t = 1, \dots, \infty$. In each period t, a unit-mass of myopic, risk-neutral individuals (agents) face a choice between remaining isolated (ℓ_R) or participating in a public activity with (ℓ_B) or without (ℓ_S) taking a period-specific booster shot that provides an *uncertain* level of added immunity to the disease for that period only. The public location and activity are general to a variety of practical settings such as working in office spaces or stadium attendance for larger events. The agents are supervised by a strategic, long-run *central planner* that aims to manage the public activity level over the entire time horizon. The myopicity of agents is justified by the renewal of the population as in the case of stadium attendance or the agents' having shorter-term objectives than that of the planner as in the case of work.

As is common in the literature, we model the infectiousness of the disease $\theta^t \in \{\theta_L, \theta_H\}$ using a two-state, irreducible, homogeneous Markov chain with a commonly-known transition matrix $\mathcal{M}_\Theta$. In practice, new disease variants periodically emerge with stochastically generated mutations. The corresponding dynamics of the disease's infectiousness have been well-approximated by Markov chains [20]. In an abuse of notation, denote the probability of transitioning from $\theta^t = \theta_j$ to $\theta^{t+1} = \theta_k$ by $(\mathcal{M}_\Theta)_{jk} := \mathbb{P}[\theta^{t+1} = \theta_k | \theta^t = \theta_j]$. The initial distribution $m_\Theta^\circ = \left[\mathbb{P}[\theta^1 = \theta_L], \quad \mathbb{P}[\theta^1 = \theta_H]\right]$ is common knowledge at the start of period $t = 1$. We denote $\zeta_\Theta := (\mathcal{M}_\Theta)_{HL}$, $\nu_\Theta := \frac{(\mathcal{M}_\Theta)_{LH}}{(\mathcal{M}_\Theta)_{HL}}$, and the stationary distribution of $\{\theta^t\}_{t\geq 1}$ by m_Θ. We assume neither the agents nor the planner can directly observe the initial value θ^1 or any future states θ^t, so the infectiousness is *uncertain* across time.

Agents receive a private benefit v from participating in the activity (ℓ_B or ℓ_S), which is drawn from a commonly-known distribution denoted by G supported on $[0, M]$ which we assume for convenience is continuous with bounded, non-zero density $0 < \underline{g} \leq \frac{dG}{dv} \leq \bar{g}$. Agents that participate in the activity incur a stochastic cost that depends on the infectiousness of the disease in that time period θ^t and the size of the isolated population y^t (i.e. active population is $1 - y^t$). We model the infectious risk as being independent of the fraction of the active population that is boosted, as boosted agents may be partially immunized but are still capable of disease transmission which poses a risk to others. This stochastic cost incurred without a booster is denoted by $\beta(y^t; \theta^t)$.

Agents that receive the booster have their cost subsequently dampened, in proportion to the efficacy of the booster in that period, denoted by γ^t. The booster administered in this period only provides the agent with protection in the current period, as immunity is not carried over across periods. Specifically, the cost incurred by agents who receive the booster is $(1-\gamma^t)\beta(y^t;\theta^t)$, however, they incur an additional cost κ to acquire the booster. The utility of agents who choose to remain isolated (ℓ_R) is zero. When indifferent between any two choices, we assume that agents choose ℓ_B over ℓ_S and ℓ_R, and choose ℓ_S over ℓ_R. Consequently, the utilities of agents can be expressed as follows.

$$u_v(\ell_R, y^t; \theta^t, \gamma^t) = 0 \tag{1}$$
$$u_v(\ell_S, y^t; \theta^t, \gamma^t) = v - \beta(y^t; \theta^t) \tag{2}$$
$$u_v(\ell_B, y^t; \theta^t, \gamma^t) = v - \kappa - (1-\gamma^t)\beta(y^t; \theta^t) \tag{3}$$

Motivated by a simple epidemiological model of community transmission (see [8]), we assume that this cost is linear in the true state θ^t and decreasing in the mass of agents choosing to remain isolated y^t, and can be expressed as:

$$\beta(y^t; \theta^t) := \theta^t c_1(y^t) + c_2(y^t), \tag{4}$$

where $c_1, c_2 : [0,1] \to \mathbb{R}$ are publicly known functions with following properties: *(i)* $c_1(1) = c_2(1) = 0$, $c_1(0) = C_1$, $c_2(0) = C_2$; *(ii)* c_1 is strictly decreasing and continuous; and *(iii)* c_2 is weakly decreasing and continuous.

We assume that the booster's current efficacy $\gamma^t \in \{0, E\}$ for some commonly-known scalar $0 < E < 1$ and that the efficacy evolves over time according to a Markov chain parameterized by a known stochastic transition matrix $\mathcal{M}_\Gamma$. We denote the initial distribution by m_Γ° and the stationary distribution by m_Γ. Moreover, denote $\zeta_\Gamma := (\mathcal{M}_\Gamma)_{E0}$ and $\nu_\Gamma = \frac{(\mathcal{M}_\Gamma)_{0E}}{(\mathcal{M}_\Gamma)_{E0}}$.

In our model, the two stochastic processes $\{\gamma^t\}_{t\geq 1}$ and $\{\theta^t\}_{t\geq 1}$ are independent, i.e. $\{\gamma^t\}_{t\geq 1} \perp \{\theta^t\}_{t\geq 1}$. We believe this to be a practical assumption when considering settings where the technology used to develop boosters itself changes across time periods or the genetic characteristics of the disease change significantly from period to period.

B. Signalling

The central planner is a strategic entity and each period implements a *signaling mechanism* to publicly disclose information *only* about the current efficacy of the booster γ^t through a mechanism $\pi^t = \langle \mathcal{I}^t, \{z_\gamma^t\}_{\gamma\in\{0,E\}}\rangle$. The set $\mathcal{I}^t$ is an alphabet of signals, and z_γ^t are probability distributions over $\mathcal{I}^t$. Denote the set of all such signalling mechanisms π^t by the set Π. The disclosure of information in period t occurs as follows. First, the planner commits to and discloses a signaling mechanism $\pi^t \in \Pi$. Next, the true state γ^t is realized using the Markovian dynamics from the previous state γ^{t-1} with both states not directly observed by the agents or planner. The corresponding probability distribution $z_{\gamma^t}^t$ is used to disclose a signal to all the agents; that is, $i^t \in \mathcal{I}^t$ is publicly signaled with probability $z_{\gamma^t}^t(i^t)$. Finally, agents use the received signal to update their belief over the state γ^t and make simultaneous choices about their choice of action in $\{\ell_R, \ell_B, \ell_S\}$.

We believe it practical for the planner to only convey information about the booster efficacy and not the infectiousness of the disease in any given period due to the comparative cost in estimating both quantities dynamically. The infectiousness of the disease is a parameter that both encapsulates the cost of illness and the likelihood of transmission – both of these quantities have high variance and are costly to estimate via experiment. However, randomized control trials and studies requiring limited samples can provide accurate estimates of the efficacy of pharmaceutical interventions like boosters [21].

Notationally, we distinguish between the public belief just before and after the information i^t is disclosed in period t. We express the agents' beliefs prior to the revelation of i^t by $\underline{p}^t := \mathbb{P}[\gamma^t = E \mid \{i^j\}_{j\leq t-1}]$ Then, on receiving signal $i^t \in \mathcal{I}^t$, the agents update their belief $\bar{p}^t := \mathbb{P}[\gamma^t = E \mid \{i^j\}_{j\leq t}]$ according to Bayes' rule:

$$\begin{aligned} \bar{p}^t &= \mathbb{P}[\gamma^t = E \mid \{i^j\}_{j\leq t}] \\ &= \frac{\underline{p}^t z_E^t(i^t)}{\underline{p}^t z_E^t(i^t) + (1-\underline{p}^t) z_0^t(i^t)} \end{aligned} \tag{5}$$

To keep quantities notationally consistent, let $p^{eq} := (m_\Gamma)_E$. Between periods, observe that the states evolve according to known *linear* Markovian dynamics, so in each period t:

$$\begin{aligned} \underline{p}^t &:= \phi_\Gamma(\bar{p}^{t-1}) \\ &= \bar{p}^{t-1}(\mathcal{M}_\Gamma)_{EE} + (1-\bar{p}^{t-1})(\mathcal{M}_\Gamma)_{0E} \\ &= p^{eq} + (1 - \nu_\Gamma\zeta_\Gamma - \zeta_\Gamma)(\bar{p}^{t-1} - p^{eq}) \end{aligned} \tag{6}$$

Analogously, we can express the beliefs over θ^t at any given time by $r^t = \mathbb{P}[\theta^t = \theta_H]$ with $r^{eq} := (m_\Theta)_H$. Since the manipulation of beliefs with respect to the disclosed information does not affect r^t, this is a known sequence over time with $r^t \to r^{eq}$. Precisely, independent of the chosen $\{\pi^t\}$:

$$\begin{aligned} r^t &:= \phi_\Theta(r^{t-1}) \\ &= r^{eq} + (1 - \nu_\Theta\zeta_\Theta - \zeta_\Theta)(r^{t-1} - r^{eq}) \end{aligned} \tag{7}$$

In this paper, we restrict to settings where $0 < \nu_\Theta\zeta_\Theta + \zeta_\Theta < 1$ and $0 < \nu_\Gamma\zeta_\Gamma + \zeta_\Gamma < 1$ to study scenarios where the Markov chains update beliefs in infectiousness and booster efficacy monotonically. Here, $r^1 > r^{eq}$ (resp. $r^1 < r^{eq}$) corresponds to settings when the infectiousness is believed to be progressively getting weaker (resp. stronger) on average over time. That is, r^t is either strictly decreasing or increasing in t.

The mechanism π^t establishes a set of posterior distributions for each signal with each posterior distribution incident with a particular probability. We thus can equivalently represent π^t by a set of tuples $\{(q_i^t, \mu_i^t)_{i\in\mathcal{I}^t}\}$ of the probability that each signal $i \in \mathcal{I}^t$ is realized (q_i^t) and the posterior belief on observing that signal (μ_i^t). Formally, for all $i \in \mathcal{I}^t$:

$$q_i^t := \underline{p}^t(0) z_0^t(i) + \underline{p}^t(E) z_E^t(i) \quad \text{[signal probability]} \tag{8}$$
$$\mu_i^t := \frac{\underline{p}^t(E) z_E^t(i)}{\underline{p}^t(0) z_0^t(i) + \underline{p}^t(E) z_E^t(i)} \quad \text{[posterior belief]} \tag{9}$$

The intuition of the representation is that the mechanism equivalently sets $\bar{p}^t = \mu_i^t$ with probability q_i^t. Since, μ_i^t establishes a probability over a binary set, appealing to the results of [22], all mechanisms in Π can be bijectively mapped to the set of all tuples $\{(q_i, \mu_i)_{i\in\mathcal{I}}\}$ such that $q_i, \mu_i \in [0,1]$, $\sum_{i\in\mathcal{I}} q_i\mu_i = \underline{p}^t$ and $\sum_{i\in\mathcal{I}} q_i = 1$. As is customary in the literature, we will refer to these as *splittings* of the prior belief $\underline{p}^t$ into posterior beliefs $\bar{p}^t = \mu_i$ with *weights* q_i [22].

C. Planner Objective

The planner's aim is to manage the mass of the active population each period over the entire time horizon. Precisely, the planner receives unit reward each period if and only if the active population mass (ℓ_B and ℓ_S) lies in a compact interval $\mathcal{Y} := [0, x]$ where $0 < x < 1$[1]. Practically, this can be considered a capacity floor on the active population that is externally mandated or estimated based on the needs of the public location and activity. If the isolated population mass lies outside this interval, the planner receives zero reward. The planner discounts future periods by a factor δ so their total utility is a discounted sum of all future rewards. This utility is a function of the masses y^t which represent the aggregate choices agents make in response to the induced beliefs $\bar{p}^t$ in each period. These are generated as a function of the planner's choices $\{\pi^t\}_{t\geq 1}$ and the randomness in the Markovian dynamics of $\{\gamma^t\}_{t\geq 1}$. The planner's utility is evaluated on expectation over both sources of randomness. For any initial belief r^1 and $\underline{p}^1$, the planner utility equals:

$$V_{\delta,\mathcal{Y}}(\{\pi^t\}_{t\geq 1}; r^1, \underline{p}^1) = \mathbb{E}\big[\sum_{t=1}^{\infty}(1-\delta)\delta^{t-1}\mathbb{I}\{y^t \in \mathcal{Y}\}\big]$$

To determine the planner's optimal signalling mechanism π^t in period t, we recursively define value functions for the forward utility starting from period t. We express the current period reward using $v_{\mathcal{Y}}(\bar{p}^t) := \mathbb{I}\{y^t \in \mathcal{Y}\}$ since the posterior belief in period t is used by agents to determine their actions and the isolated mass in aggregate. The posterior belief is a function of the planner's mechanism π^t since the distribution over the beliefs $\bar{p}^t$ that can be induced from $\underline{p}^t$ is specified by π^t as in (9). Thus, in each period t, given the current belief $\underline{p}^t$, the planner's design problem is to maximize value and determine the optimal signalling mechanism for the given belief. We refer to the optimal mapping of beliefs $\underline{p}^t$ to signalling mechanisms as the optimal *disclosure* rule $\pi_*^t \in \Pi$:[2]

$$V_{\delta,\mathcal{Y}}^t(\underline{p}^t) = \max_{\pi_t\in\Pi}\mathbb{E}[(1-\delta)v_{\mathcal{Y}}(\bar{p}^t) + \delta V_{\delta,\mathcal{Y}}^{t+1}(\underline{p}^{t+1})] \quad (10)$$

$$\pi_*^t(\underline{p}^t) \in \arg\max_{\pi_t\in\Pi}\mathbb{E}[(1-\delta)v_{\mathcal{Y}}(\bar{p}^t) + \delta V_{\delta,\mathcal{Y}}^{t+1}(\underline{p}^{t+1})] \quad (11)$$

Observe that the setting we have described is a non-stationary Markov decision process where the planner's state is represented by the tuple $(t, \underline{p}^t)$. The action set can be represented by the set of mechanisms Π, with rewards $v_{\mathcal{Y}}(\mu_i^t)$ with probability q_i^t as dictated by the chosen mechanism $\pi^t \in \Pi$, and transitions from states $(t, \underline{p}^t)$ to states $(t+1, \phi_\Gamma(\mu_i^t))$ with probability q_i^t.

III. EQUILIBRIUM CHARACTERIZATION

We use the solution concept of Bayes-Nash equilibrium to characterize the outcome of agents' choices in period t upon receiving signal i^t as generated by the signaling mechanism $\pi^t = \{(q_i^t, \mu_i^t)_{i\in\mathcal{I}^t}\}$. We represent the equilibrium mass of isolated agents in period t by $y_{\pi^t}^*(i^t)$ which results from all the *myopic* agents simultaneously making their choices under the posterior belief $\bar{p}^t = \mu_{i^t}^t$ corresponding to i^t. We can represent the expected infectious cost faced by agents choosing ℓ_S with the posterior belief r^t and an active population of $1-w$ by $\tilde{\beta}(w; r^t) = c_1(w)(r^t(\theta_H - \theta_L) + \theta_L) + c_2(w)$. The following result shows that we can exactly characterize the equilibrium in terms of these two quantities. At equilibrium, the agents that choose to be active (ℓ_S and ℓ_B) will be those that have the largest private benefits from G. Moreover, those that will choose to be active will either all choose to take booster (ℓ_B) or all elect to not do so (ℓ_S).

Proposition 1: For any signal $i^t \in \mathcal{I}^t$ realized by mechanism π^t setting $\bar{p}^t = \mu_i^t$, the equilibrium mass of isolated agents $y^*(i^t) = m(\bar{p}^t; r^t)$ is given by the unique solution $w^* \in [0,1]$ to the following equation:

$$G^{-1}(w) = \min\{\tilde{\beta}(w; r^t), \kappa + (1-\bar{p}^t E)\tilde{\beta}(w; r^t)\} \quad (12)$$

Furthermore, at equilibrium, agents with private value from being active v choose ℓ_B if $v \geq G^{-1}(m(\bar{p}^t; r^t))$ and $\tilde{\beta}(m(\bar{p}^t; r^t); r^t) \geq \kappa + (1-\bar{p}^t E)\tilde{\beta}(m(\bar{p}^t; r^t); r^t)$, choose ℓ_S if $v \geq G^{-1}(m(\bar{p}^t; r^t))$ and $\tilde{\beta}(m(\bar{p}^t; r^t); r^t) < \kappa + (1-\bar{p}^t E)\tilde{\beta}(m(\bar{p}^t; r^t); r^t)$, and choose ℓ_R otherwise. ■

Leveraging Proposition 1, we can further identify the functional relationship of how the parameters and beliefs governing the model affect the outcome. The following lemmas show that the isolated mass in equilibrium is well-behaved in response to the beliefs over the booster efficacy and the infectiousness of the disease.

Lemma 1: For any $r^t \in [0,1]$, the map $m(\bar{p}^t; r^t)$ is continuous, bounded and weakly decreasing in $\bar{p}^t$. Moreover, $m(\bar{p}^t; r^t)$ is Lipschitz continuous in $\bar{p}^t$.

Lemma 2: For any $\bar{p}^t \in [0,1]$, the map $m(\bar{p}^t; r^t)$ is continuous, bounded and weakly increasing in r^t. Moreover, $m(\bar{p}^t; r^t)$ is Lipschitz continuous in r^t.

These lemmas establish that the mapping between the belief on the θ^t and γ^t and the isolated mass are continuous, monotone mappings. We depict an example of such a function m in Figure 1. The continuity of m implies that we can invert from isolated agent masses y_t^* to beliefs. Namely, observe that we can define $\mathcal{W}(r^t) = \{p : m(p, r^t) \in \mathcal{Y}\}$. Since $\mathcal{Y} = [0, x]$ and m is weakly decreasing in p, either $m(1, r^t) > x$ and $\mathcal{W}(r^t) = \emptyset$, or there exists some $W(r^t) := \inf\{\mathcal{W}(r^t)\}$ in $[0,1]$ such that $\mathcal{W}(r^t) = [W(r^t), 1]$. For convenience, we let

[1] Note that our design approach can be extended to more general sets of desired isolated masses. We can, with mild adjustment, address unions of intervals but for the sake of exposition we restrict to this family of preferences.

[2] We can use max because $V_{\delta,\mathcal{Y}}^t$ is Lipschitz over $\underline{p}^t$, the expression in braces of (10) is upper hemi-continuous w.r.t. $\underline{p}^t$ in the weak-* topology on Π, and Π is compact in that topology. We omit the details as they are standard [13], [5].

$W(r^t) = \infty$ if $\mathcal{W}(r^t) = \emptyset$. The following lemma establishes that $W(\cdot)$ also is weakly increasing.

Lemma 3: $W(r)$ is weakly increasing in r and locally Lipschitz continuous over the interior domain $dom(W) := \{r : 0 < W(r) < 1\}$.

The previous lemma shows that the relationship between the beliefs about the infectiousness of a disease and the set of beliefs we need to induce on the booster efficacy for the planner to achieve reward is monotone.

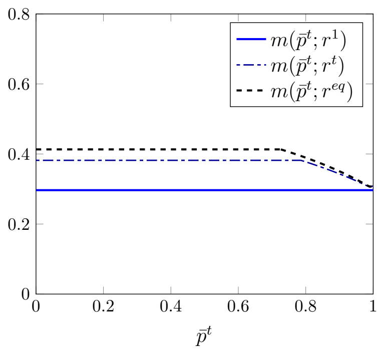

Fig. 1: $m(\bar{p}^t; r)$ for different values of r^t. $G = Unif[0, 10], E = \frac{99}{100}, \theta_L = 4, \theta_H = 16, r^1 = \frac{1}{6}, r^t = \frac{1}{2}, r^{eq} = \frac{2}{3}$, $\beta(y^t; \theta^t) = \theta^t (y^t)^2$.

Moreover, Lemma 3 ensures that as the severity of the disease increases, our ability to induce beliefs on the booster to achieve a desirable outcome becomes more limited since $\mathcal{W}(r)$ shrinks as r increases. This is evident in Figure 1, as the domain of $\bar{p}^t$ that map to any isolated mass $m(\bar{p}^t; r^t)$ in $[0, x]$ contracts as r^t increases.

The local Lipschitz continuity of $W(r)$ guarantees that the set of beliefs in period t, $\mathcal{W}(r^t) = [W(r^t), 1]$ that we need to induce in the booster effectiveness $\bar{p}^t$ to achieve an isolated mass $y^*_{\pi^t}(i^t)$ in $\mathcal{Y}$ converges over time. If $\mathcal{W}(r^t)$ becomes $\emptyset$ eventually, we will not have this closure and hence the planner can never induce beliefs that map to the required active population in the long-run. Likewise, if $W(r^t)$ converges to 0 or 1, trivial mechanisms become necessary in the long-run and W will not be locally Lipschitz at these limit points. Hence, in this paper we will ignore these edge cases where $W(r^{eq}) = 0, 1$, or ∞. With this assumption, we can show that W preserves the convergence we know we will observe in the belief r^t.

Assumption 1: $0 < W(r^{eq}) < 1$

Corollary 1: Under Assumption 1, $\lim_{t\to\infty} W(r^t) \to W(r^{eq})$.

IV. OPTIMAL DISCLOSURE RULE

A. Structure of Value Functions

In order to characterize the optimal mechanism, we first identify structural properties of the value functions specified in (10) that are maximized when using π^t_*. In particular, the shape of the value function $V^t_{\delta,\mathcal{Y}}$ provides insight into the structure of π^t_*

We first show the following proposition which establishes the concavity of the value function of the planner with respect to the population's belief over the booster's efficacy.

Proposition 2: For all $t \geq 1$, $V^t_{\delta,\mathcal{Y}}(\underline{p}^t)$ is concave in $\underline{p}^t$.

This result establishes a critical trade-off between the planner disclosing information (i.e. splitting across multiple posterior beliefs $\bar{p}^t$) to enhance their current stage reward $v_{\mathcal{Y}}(\bar{p}^t)$ and the cost of this information disclosure on the future value functions $V^{t+1}_{\delta,\mathcal{Y}}$. It suggests that if the planner is already maximizing their current stage reward with the population's current belief $\underline{p}^t$, there is no need to divulge additional information since it could lead to worse continuation payoffs in future periods due to the concavity of future value functions. To this end, we motivate the non-informative mechanism, $\pi_{NI}(\underline{p}^t) = \{(1, \underline{p}^t)\}$, which conveys no information and thus does not affect the belief.

We consequently show in the following proposition that $\pi^t_*(\underline{p}^t) = \pi_{NI}(\underline{p}^t)$ whenever $v_{\mathcal{Y}}(\underline{p}^t) = 1$. The non-informative mechanism $\pi_{NI}(\underline{p}^t)$ can be constructed by choosing $\mathcal{I}^t = 1$ as no information is conveyed on seeing the only possible signal and on employing this mechanism in period t, $\bar{p}^t = \underline{p}^t$ with probability one. Similarly, if the set of beliefs the planner seeks to induce with the signalling mechanism $\mathcal{W}(r^t) = \emptyset$ (i.e. $W(r^t) = \infty$), then regardless of the initial belief $\underline{p}^t \in [0, 1]$, the non-informative mechanism maximizes the stage reward as no reward is possible. Therefore, $\pi^t_*(\underline{p}^t) = \pi_{NI}(\underline{p}^t)$ as the non-informative mechanism is optimal.

Proposition 3: For all $t \geq 1$, if $\underline{p}^t \geq W(r^t)$ or if $W(r^t) = \infty$, then $\pi^t_*(\underline{p}^t) = \pi_{NI}(\underline{p}^t)$.

Next, we further extend this insight by highlighting that the current stage reward $v_{\mathcal{Y}}(\bar{p}^t)$ is binary and consequently necessitates only two signals at optimality. Namely, the optimal signalling mechanism should generate two posterior beliefs by using only two signals $\mathcal{I} = \{a, b\}$ – one that will yield an outcome that maximizes the current period reward (i.e. $\mu^t_a \geq W(r^t)$), and one that yields an outcome that does not achieve a current period reward ($\mu^t_b < W(r^t)$). Any additional information revelation is suboptimal by Proposition 4. We standardize the posteriors so that for all t, $\mathcal{I}^t = \{a, b\}$ and $\mu^t_b < W(r^t), \mu^t_a \geq W(r^t)$.

Proposition 4: For all $t \geq 1$, if $\underline{p}^t < W(r^t)$, $\pi^t_*(\underline{p}^t)$ is based on $\mathcal{I}^t = \{a, b\}$ with $\mu^t_b \leq \underline{p}^t < W(r^t)$ and $\mu^t_a \geq W(r^t)$.

Since the optimal disclosure rule specifies that a mechanism need split into only two posteriors when $\underline{p}^t < W(r^t)$, we need only specify the two posterior means μ^t_b and μ^t_a in any period t for any given $\underline{p}^t$. Observe that q^t_b and q^t_a are implicitly defined as a result since $q^t_b + q^t_a = 1$ and $q^t_b \mu^t_b + q^t_a \mu^t_a = \underline{p}^t$.

We next analyze the recursive relationship between value functions. We define the *posterior value function* in period t to be $\bar{V}^t_{\delta,\mathcal{Y}}(\bar{p}^t) := (1-\delta)v_{\mathcal{Y}}(\bar{p}^t) + \delta V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\bar{p}^t))$. Observe that these functions represent the value-to-go once the signal i^t has already been realized. By the known Markovian update between $\underline{p}^{t+1} = \phi_\Gamma(\bar{p}^t)$:

$$\begin{aligned} V^t_{\delta,\mathcal{Y}}(\underline{p}^t) &= \max_{\pi_t \in \Pi} \mathbb{E}[(1-\delta)v_{\mathcal{Y}}(\bar{p}^t) + \delta V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\bar{p}^t))] \\ &= \max_{\pi_t \in \Pi} \mathbb{E}[\bar{V}^t_{\delta,\mathcal{Y}}(\bar{p}^t)] \end{aligned} \tag{13}$$

Thus, each point on the value function $V^t_{\delta,\mathcal{Y}}(\underline{p}^t)$ is equal to largest possible linear average of points on the posterior value function. Hence, $V^t_{\delta,\mathcal{Y}}(\underline{p}^t)$ must be the *concave envelope* of the *posterior value function* $\bar{V}^t_{\delta,\mathcal{Y}}$. Following [23], we denote $k(p)$ as the concave envelope of $V^t_{\delta,\mathcal{Y}}(p)$ if $k(p) \geq V^t_{\delta,\mathcal{Y}}(p)$ for all $p \in [0,1]$ satisfying the condition that $\max\ \{\ \bar{V}^t_{\delta,\mathcal{Y}}(p)\ -\ k(p)\ ,\ k''(p)\ \} = 0$ for all p.

Proposition 5: $V^t_{\delta,\mathcal{Y}} = conc(\bar{V}^t_{\delta,\mathcal{Y}})$ where $conc(\cdot)$ denotes the concave envelope.

The previous proposition provides us with another formulation to compute $V^t_{\delta,\mathcal{Y}}$. We observe that this additional formulation allows us to characterize the optimal disclosure rule based on the shape of the value function $V^t_{\delta,\mathcal{Y}}$, as on intervals where $V^t_{\delta,\mathcal{Y}}$ is linear, the optimal disclosure rule prescribes a mechanism π^t_* that splits over the two posterior beliefs at the endpoints of this linear interval.

Lemma 4: For all $t \geq 1$, $p < W(r^t)$, suppose that $\pi^t_*(p)$ splits on $\mu^t_b \leq p < W(r^t)$ and $\mu^t_a \geq W(r^t)$, then $V^t_{\delta,\mathcal{Y}}(\underline{p}^t)$ is linear on $[\mu^t_b, \mu^t_a]$. Furthermore, for all $\mu^t_a \leq p \leq \mu^t_b$, $\pi^t_*(p)$ splits on μ^t_b and μ^t_a.

This lemma provides us with a powerful insight into the optimal disclosure rule. Under the optimal disclosure rule $\pi^t_*(\underline{p}^t)$, the posteriors μ^t_b and μ^t_a are fixed across different values of $\underline{p}^t < W(\underline{p}^t)$ since $V^t_{\delta,\mathcal{Y}}$ will have intervals where $V^t_{\delta,\mathcal{Y}}$ is concave and coincides with $\bar{V}^t_{\delta,\mathcal{Y}}$ – here, the non-informative mechanism is optimal. Otherwise, $V^t_{\delta,\mathcal{Y}}$ will have intervals where it is linear and the posteriors chosen by the optimal mechanism π^t_* (μ^t_b and μ^t_a) are identical across all $\underline{p}^t$ in that interval. There can only be one such linear segment since to the left of and to the right of $W(r^t)$, $\bar{V}^t_{\delta,\mathcal{Y}}$ is already concave so no linear segment is added wholly within $[0, W(r^t))$ or $[W(r^t), 1]$. Two such segments cannot exist since they would necessarily coincide which leads to a contradiction. Hence, the one set of posteriors μ^t_b and μ^t_a is sufficient to describe π^t_*.

To see this more concretely, observe that for the non-informative mechanism must be optimal for all $\underline{p}^t \geq W(r^t)$, $conc(\bar{V}^t_{\delta,\mathcal{Y}})$ must coincide with $\bar{V}^t_{\delta,\mathcal{Y}}$ on that interval. However, as depicted in Fig. 2 and in Equation (13), there is a sudden jump in the posterior value function $\bar{V}^t_{\delta,\mathcal{Y}}$ at $\underline{p}^t = W(r^t)$. Therefore, by definition of the concave envelope $V^t_{\delta,\mathcal{Y}} = conc(\bar{V}^t_{\delta,\mathcal{Y}})$, it must incorporate a non-trivial, *maximal* interval where it is linear and *strictly* above the posterior value function $\bar{V}^t_{\delta,\mathcal{Y}}$ – in fact, by Proposition 4 and Lemma 4, one end of the linear segment must lie in $\mathcal{W}(r^t)$ and the other end below it in order to cross this discontinuity jump and concavify $\bar{V}^t_{\delta,\mathcal{Y}}$. Thus, in order for the posterior value function to coincide with its concave envelope for $\underline{p}^t \geq W(r^t)$, the linear segment of $conc(\bar{V}^t_{\delta,\mathcal{Y}})$ must end precisely at the posterior belief $W(r^t)$. However, by Lemma 4, this right endpoint of the interval where $V^t_{\delta,\mathcal{Y}}$ is linear is the second posterior μ^t_a.

Proposition 6: If $W(r^t) \leq 1$ and $\underline{p}^t < W(r^t)$, there exists an optimal mechanism $\pi^t_*(\underline{p}^t)$ with $\mu^t_a = W(r^t)$ for all t.

To summarize the insights on the optimal disclosure rule thus far: (i) if $\underline{p}^t \geq W(r^t)$ or $W(r^t) = \infty$, we can choose the optimal signalling mechanism $\pi^t_*(\underline{p}^t) = \pi_{NI}(\underline{p}^t)$; (ii) if $\underline{p}^t < W(r^t)$, we need to only specify the posterior belief $\mu^t_b < \underline{p}^t$ as an optimal signalling mechanism exists that has at most two posteriors with one of them being $\mu^t_a = W(r^t)$. In fact, μ^t_b is fixed across the disclosure rule π^t_* since by Lemma 4, if $\underline{p}^t \leq \mu^t_b$ or if $\underline{p}^t \geq W(r^t)$, then $\pi^t_*(\underline{p}^t) = \pi_{NI}$, and otherwise if $\mu^t_b < \underline{p}^t < W(r^t)$, the disclosure policy chooses $\pi^t_*(\underline{p}^t) = \{(\frac{W(r^t) - \underline{p}^t}{W(r^t) - \mu^t_b}, \mu^t_b), (\frac{\underline{p}^t - \mu^t_b}{W(r^t) - \mu^t_b}, W(r^t)\}$. Hence, all that remains to determine the optimal disclosure rule across time is to solve for the optimal sequence $\{\mu^t_b\}_{t\geq 1}$.

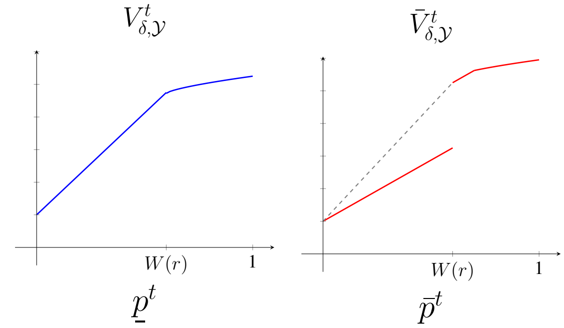

Fig. 2: $V^t_{\delta,\mathcal{Y}} = conc(\bar{V}^t_{\delta,\mathcal{Y}})$ (on left), $\bar{V}^t_{\delta,\mathcal{Y}}$ (on right).

B. Optimality of Greedy Disclosure Rule

Our results in Section IV-A establish that the planner seeks to disclose as little information as they can due to the concavity of their value functions. We now formalize the *greedy* disclosure rule at time t which simultaneously provides as little information as possible and maximizes the current period reward. Observe that this is achieved by splitting into posterior beliefs $\mu^t_b = 0$ and $\mu^t_a = W(r^t)$, as intuitively $W(r^t)$ is the *closest* belief we can induce to achieve the current period reward. Moreover, by pushing μ^t_b to 0, we maximize the probability q^t_a that we induce a posterior belief leading to a reward. [3]

Definition 1: The greedy disclosure rule $\pi^t_\dagger(\underline{p}^t)$ is such that if $\underline{p}^t \geq W(r^t)$ or $W(r^t) = \infty$ then $\pi^t_\dagger(\underline{p}^t) = \pi_{NI}(\underline{p}^t)$. If $\underline{p}^t < W(r^t)$, $\pi^t_\dagger(\underline{p}^t) = \{(\frac{\underline{p}^t}{W(r^t)}, W(r^t)), (\frac{W(r^t) - \underline{p}^t}{W(r^t)}, 0)\}$.

We seek to show that the greedy disclosure rule will become optimal after some finite amount of time. We can directly comment on the quality of the greedy disclosure rule against the optimal disclosure rule as specified by the optimal sequence $\{\mu^t_b\}_{t\geq 1}$. Observe the greedy disclosure rule chooses posterior beliefs 0 and $W(r^t)$ whenever $\underline{p}^t < W(r^t) \leq 1$.

Here, we only focus on the most interesting case where $p^{eq} < W(r^{eq})$. Observe that, if $p^{eq} > W(r^{eq})$, the Markovian drift on the beliefs under no information will translate $\underline{p}^1$ to $\underline{p}^t = \phi_\Gamma^{(n-1)}(\underline{p}^1)$ and, that for some finite n, $\underline{p}^{t'} \geq W(r^{eq})$ for all $t' \geq n$. Hence, without any intervention, the planner will eventually begin using no information and collecting the period rewards for perpetuity. The planner's problem when $p^{eq} < W(r^{eq})$ is consequently of more direct interest, and we show that the greedy disclosure

[3] This mechanism is unique and exists strictly because the state space of the uncertainty set is of size two.

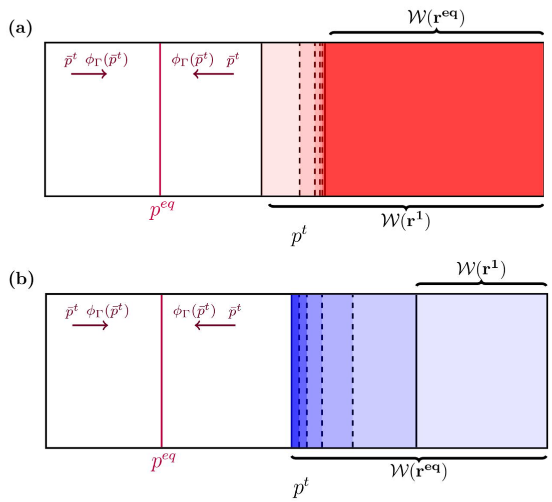

Fig. 3: Diagram of the two cases described in Theorem 1.

rule eventually becomes optimal in finite time. In fact, we can show that the greedy disclosure rule eventually becomes optimal once the planner cannot guarantee a reward in the next period by providing no information when he does not earn a reward in the current period by doing so (i.e. we require $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$).

Theorem 1: If $p^{eq} < W(r^{eq})$, subject to Assumption 1, the following holds:

(a) If r^t is increasing, then $\pi_*^t = \pi_\dagger^t$ for all $t > t_\dagger$ where $t_\dagger = \sup \{ t : W(r^t) < p^{eq} \} < \infty$.

(b) If r^t is decreasing, then $\pi_*^t = \pi_\dagger^t$ for all $t > t_\dagger$ where $t_\dagger = \sup \{ t : \phi_\Gamma(W(r^t)) > W(r^{t+1}) \} < \infty$.

We depict both cases graphically in Fig. 3. The theorem affirms that the optimal disclosure rule for a planner eventually becomes the greedy disclosure rule, which seeks to minimize information sharing while maximizing the current period reward. Specifically, this property holds when the beliefs $\underline{p}^t$ move *away* from the equilibrium belief set $W(r^{eq})$. In practice, when individuals grow more pessimistic about the booster's efficacy over time and converge to a belief that is more pessimistic than required for the desired outcome, the planner best maximizes their utility by greedily inducing the belief with maximum probability that achieves the *easiest* outcome in $\mathcal{Y}$ (i.e. achieving the threshold $y_t^* = x$).

The intuition of the optimality of the greedy mechanism is as follows. If a posterior μ_b^t slightly larger than 0 is chosen, the advantage is that if signal b is drawn by the mechanism, the planner starts the next period at $\underline{p}^{t+1} = \phi_\Gamma(\mu_b^t)$ where $\phi_\Gamma(0) < \phi_\Gamma(\mu_b^t) < W(r^t)$. However, because beliefs further away from p^{eq} converge to p^{eq} faster under Markovian dynamics, $|\phi_\Gamma(\mu_b^t) - \phi_\Gamma(0)| \leq |\mu_b^t - 0|$, so the positive effect is counteracted by the more rapid mixing of the Markov chain. The key disadvantage of choosing $\mu_b^t > 0$ is that we are limited to smaller weights q_a^t on beliefs μ_a^t that induce current period rewards. Therefore, intuition suggests that we should maximize the current period reward since not doing so comes at a minimal cost.

Observe that the greedy disclosure rule will only provide informative signals when $\underline{p}^t < W(r^t)$ and since $p^{eq} \leq W(r^{eq})$, eventually the beliefs $\phi_\Gamma(W(r^t))$ will always shift out of the beliefs necessary in the next period $\mathcal{W}(r^{t+1})$. If $\underline{p}^t < W(r^t)$, the greedy disclosure rule will optimistically generate posterior beliefs $\bar{p}^t = W(r^t)$ in any period t, and then the Markov chain updates the next period's initial belief to $\underline{p}^{t+1} = \phi_\Gamma(W(r^t)) \notin \mathcal{W}(r^{t+1})$. Thus, every period for perpetuity, the planner must provide informative signalling. This mimics something closer to practical settings where planners cannot abstain from providing information for perpetuity.

The previous theorem does not provide guidance on the optimal strategy π_*^t in the interim when $t \leq t_\dagger$. Generally, as we discuss in more detail in the proof of Theorem 1, since $t_\dagger$ is finite, this can be solved efficiently using dynamic programming since the value functions $V_{\delta,\mathcal{Y}}^{t_\dagger}$ can be easily computed and backtracking at each point in time $t \leq t_\dagger$, we need only consider a finite number of possible μ_b^t.

The main technical challenge preventing stronger guarantees on the optimality of greedy disclosure is that W is only locally Lipschitz without added structure imposed on G (which simplifies m and consequently W). Therefore, the relationship between the beliefs required in the next period compared to the current period is monotone, but lacking further insight. However, the following proposition shows that under some mild regularity on δ and the mixing rate imposed by ϕ_Γ, we can guarantee that the greedy disclosure rule is optimal even in the interim periods $t \leq t_\dagger$.

Proposition 7: If $\delta \leq \min_{j \leq t \leq t_\dagger} \frac{\phi_\Gamma^{-1}(W(r^{t+1}))}{W(r^t)}$, then $\pi_*^t = \pi_\dagger^t$ for all $t \geq j$.

Specifically, this proposition states that if the planner is sufficiently impatient and discounts significantly, then the greedy disclosure rule is optimal. This is intuitive as the planner should be more willing to take a gamble of only securing a current period reward with the maximal probability q_a^t rather than ensuring a reward next period by choosing non-informative signalling. Likewise, suppose that r^t is decreasing, then observe that if ϕ_Γ mixes faster (i.e. $1 - \nu_\Gamma\zeta_\Gamma - \zeta_\Gamma$ is small) then $\phi_\Gamma^{-1}(W(r^{t+1}))$ becomes larger. Therefore, when ϕ_Γ mixes fast, the possibility of the planner guaranteeing rewards in the next period dissipates since the population belief on the booster quickly moves close to p^{eq} which is outside $\mathcal{W}(r^t)$. Therefore, in practice, the viability of the greedy disclosure rule is most robust when the population beliefs quickly converge to stationary beliefs and planners are also operating on shorter time horizons.

V. Numerical Study

We present a brief numerical comparison of the total discounted reward sum $\tilde{V}(\underline{p}^1; \pi) := \sum_{t=1}^\infty \delta^{t-1}(1-\delta)v_\mathcal{Y}(\bar{p}^t)$ that the planner accrues under the greedy disclosure rule $\pi = \pi_\dagger$ as opposed to under the non-informative disclosure rule ($\pi^t = \pi_{NI}$ for all t) and under the fully informative disclosure rule ($\pi^t = \pi_{FI}$ for all t where $\pi_{FI} = \{(\underline{p}^t, 1), (1 - \underline{p}^t, 0)\}$). Fully-informative disclosure directly reveals the state of γ^t in each period t. In Figure 4, we plot the numerical estimates of $\tilde{V}(\underline{p}^1; \pi)$ under each of the three disclosure rules. Appealing to Proposition 7, the optimal disclosure rule coincides with the greedy disclosure rule in both these settings for $t \geq 1$, so $\tilde{V}(\underline{p}^1; \pi_\dagger)$ is equal to the value function $V^1(\underline{p}^1)$. In both settings presented $p^{eq} < W(r^{eq})$, but in case (i) r^t is increasing and in case (ii) r^t is decreasing. The greedy disclosure rule strictly dominates both benchmarks for all initial beliefs $\underline{p}^1$. Moreover, we can demonstrate that the greedy disclosure rule not only outperforms the benchmarks on expectation but can also do so on an instance level [19]. By design, the greedy disclosure rule is minimizing information sharing *subject to maximizing the current period reward*, so the greedy disclosure rule also outperforms these two mechanisms by achieving weakly higher period rewards in every period.

Observe that the greedy value function is piecewise linear in the initial belief $\underline{p}^t$, as the strategy involves splitting across $\mu_b^1 = 0$ and $\mu_a^1 = W(r^1)$. The dominance of the greedy disclosure rule across both settings in Fig. 4 affirm the value of greedy information disclosure.

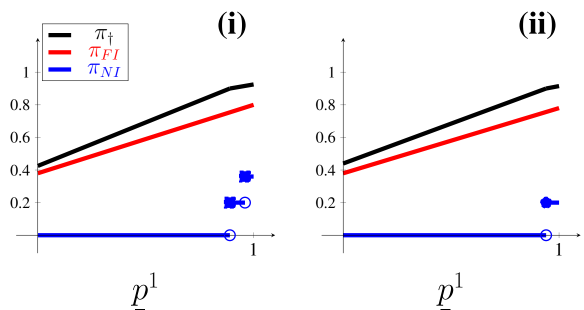

Fig. 4: Discounted sum of rewards, $\tilde{V}(\underline{p}^1; \pi)$, under the same parameters as Fig. 1 with $\delta = 0.8$, $p^{eq} = 0.2$, $x = 0.35$ and $(1 - \nu_\Theta\zeta_\Theta - \zeta_\Theta) = (1 - \nu_\Gamma\zeta_\Gamma - \zeta_\Gamma) = 0.8$. For case **(i)**, $r^1 = 0.5$; and for case **(ii)**, $r^1 = 0.8$.

VI. Concluding Remarks

In this paper, we introduced a model to study optimal dynamic information disclosure over booster efficacy amid non-stationary disease infectiousness. Our model captures two novel features: (a) heterogeneous agents making strategic decisions to trade-off activity against the infectious risks with or without a booster; (b) a non-stationary environment where the planner's value from inducing certain beliefs over booster efficacy varies as the infectiousness changes over time. We provided a complete description of the strategic equilibrium as a function of beliefs over booster efficacy. We also show that greedy disclosure eventually coincides with the optimal disclosure rule. Future work should examine how the optimal disclosure rule changes when the changes in disease infectiousness and booster efficacy are correlated.

References

[1] A. J. Zuckerman, "Effect of hepatitis B virus mutants on efficacy of vaccination," *The Lancet*, vol. 355, no. 9213, pp. 1382–1384, 2000, publisher: Elsevier.

[2] S. L. Wilson and C. Wiysonge, "Social media and vaccine hesitancy," *BMJ Global Health*, vol. 5, no. 10, p. e004206, Oct. 2020.

[3] A. Jelnov and P. Jelnov, "Vaccination policy and trust," *Economic Modelling*, vol. 108, p. 105773, Mar. 2022.

[4] E. Kamenica and M. Gentzkow, "Bayesian Persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, Oct. 2011.

[5] D. Bergemann and S. Morris, "Information Design: A Unified Perspective," *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, Mar. 2019.

[6] J. Ely, A. Galeotti, and J. Steiner, "Rotation as Contagion Mitigation," *Management Science*, vol. 67, no. 5, pp. 3117–3126, May 2021, publisher: INFORMS.

[7] A. Hernandez-Chanto, C. Oyarzun, and J. Hedlund, "Contagion Management through Information Disclosure," Social Science Research Network, Rochester, NY, SSRN, Dec. 2021.

[8] S. Shah, S. Amin, and P. Jaillet, "Optimal Information Provision for Strategic Hybrid Workers," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, Dec. 2022, pp. 3807–3814, iSSN: 2576-2370.

[9] I. Ball, "Dynamic Information Provision: Rewarding the Past and Guiding the Future," Rochester, NY, Jan. 2019.

[10] A. Smolin, "Dynamic Evaluation Design," *American Economic Journal: Microeconomics*, vol. 13, no. 4, pp. 300–331, 2021, publisher: American Economic Association.

[11] D. Orlov, A. Skrzypacz, and P. Zryumov, "Persuading the Principal to Wait," *Journal of Political Economy*, vol. 128, no. 7, pp. 2542–2578, Jul. 2020, publisher: The University of Chicago Press.

[12] J. C. Ely, "Beeps," *American Economic Review*, vol. 107, no. 1, pp. 31–53, Jan. 2017.

[13] J. Renault, E. Solan, and N. Vieille, "Optimal dynamic information provision," *Games and Economic Behavior*, vol. 104, pp. 329–349, Jul. 2017.

[14] L. Garrett, "The Next Pandemic?" *Foreign Affairs*, vol. 84, no. 4, pp. 3–23, 2005, publisher: Council on Foreign Relations.

[15] E. Chekol Abebe, M. Tiruneh G/Medhin, A. Behaile T/Mariam, and et al., "Mutational Pattern, Impacts and Potential Preventive Strategies of Omicron SARS-CoV-2 Variant Infection," *Infection and Drug Resistance*, vol. 15, pp. 1871–1887, Dec. 2022.

[16] G. McLean, J. Kamil, B. Lee, P. Moore, T. F. Schulz, A. Muik, U. Sahin, Türeci, and S. Pather, "The Impact of Evolving SARS-CoV-2 Mutations and Variants on COVID-19 Vaccines," *mBio*, vol. 13, no. 2, pp. e02 979–21, Apr. 2022.

[17] J. E. Marco-Franco, P. Pita-Barros, S. González-de Julián, and D. Vivas-Consuelo, "COVID-19, Fake News, and Vaccines: Should Regulation Be Implemented?" *International Journal of Environmental Research and Public Health*, Jan. 2021.

[18] W. Chen and C. Stoecker, "Mass media coverage and influenza vaccine uptake," *Vaccine*, vol. 38, no. 2, pp. 271–277, Jan. 2020.

[19] S. Shah, S. Amin, and P. Jaillet, "Informing Strategic Agents over Booster Efficacy," *LIDS Technical Report*, 2023. [Online]. Available: https://sshah95.github.io/assets/pdf/inforprovisioncdc2023.pdf

[20] D. Li and S. Liu, "Threshold dynamics and ergodicity of an SIRS epidemic model with Markovian switching," *Journal of Differential Equations*, vol. 263, no. 12, pp. 8873–8915, 2017.

[21] P. G. Smith, L. C. Rodrigues, and P. E. M. Fine, "Assessment of the Protective Efficacy of Vaccines against Common Diseases Using Case-Control and Cohort Studies," *International Journal of Epidemiology*, vol. 13, no. 1, pp. 87–93, 1984.

[22] R. Aumann and M. Maschler, "Repeated Games with Incomplete Information," *MIT Press*, 2005.

[23] A. Oberman, "The convex envelope is the solution of a nonlinear obstacle problem," *Proceedings of the American Mathematical Society*, vol. 135, no. 6, pp. 1689–1694, 2007.

[24] S. Sorin, *A First Course on Zero-Sum Repeated Games*. Springer, 2002.

[25] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1.

VII. Appendix

A. Proof of Prop. 1

First, we show that in equilibrium with isolated mass w, there cannot both be agents that choose ℓ_B and agents that choose ℓ_S. This is obvious from (1) and (2) since by independence of $\{\theta^t\}_{t\geq 1}$ and $\{\gamma^t\}_{t\geq 1}$ any agent with benefit

v choosing ℓ_B implies that:

$$\begin{aligned}\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_v(\ell_S,w;\theta^t,\gamma^t)] &= v-\tilde{\beta}(w;r^t)\\ &< v-\kappa-(1-\bar{p}^tE)\tilde{\beta}(w;r^t)\\ &= \mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_v(\ell_B,w;\theta^t,\gamma^t)]\end{aligned}$$

or equivalently $\tilde{\beta}(w;r^t) > \kappa + (1-\bar{p}^tE)\tilde{\beta}(w;r^t)$. Since the expected infectious cost is additive, this is true for all possible v. Hence, in an equilibrium, no rational agent with any benefit v' would choose ℓ_S. Therefore, for any equilibrium with isolated mass w, if $\tilde{\beta}(w;r^t) \geq \kappa + (1-\bar{p}^tE)\tilde{\beta}(w;r^t)$, only ℓ_B and ℓ_R are chosen in equilibrium, whereas if $\tilde{\beta}(w;r^t) < \kappa + (1-\bar{p}^tE)\tilde{\beta}(w;r^t)$ only ℓ_S and ℓ_R are chosen.

Next, observe that in any equilibrium with isolated mass w, agents choose ℓ_R if and only if their benefit $v < G^{-1}(w)$. Agents with $v = 0$ must choose ℓ_R since $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_0(\ell_S,y^t;\theta^t,\gamma^t)]$ and $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_0(\ell_B,y^t;\theta^t,\gamma^t)]$ are strictly less than 0 and $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_0(\ell_R,y^t;\theta^t,\gamma^t)] = 0$. If there exists some agent with benefit $v' > 0$ choosing ℓ_R in equilibrium, then $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_{v'}(\ell_S,y^t;\theta^t,\gamma^t)]$ and $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_{v'}(\ell_B,y^t;\theta^t,\gamma^t)]$ are strictly less than 0. However, observe that $u_v(\ell_B,y^t;\theta^t,\gamma^t)$ and $u_v(\ell_S,y^t;\theta^t,\gamma^t)$ are strictly increasing in v, so this implies that for any $0 < v'' < v'$, $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_{v''}(\ell_S,y^t;\theta^t,\gamma^t)]$ and $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_{v''}(\ell_B,y^t;\theta^t,\gamma^t)]$ are strictly less than 0 so any agents with benefit v'' must also choose ℓ_R. This implies that the set of agents choosing ℓ_R is precisely those whose benefits lie in some interval $[0,v')$, and since the total mass choosing ℓ_R is w, $v' = G^{-1}(w)$. Altogether, we have determined the actions of agents in equilibrium given the isolated mass as that which has the agents with the smallest benefit choose isolation.

Next, we show that an isolated mass w^* corresponds to the equilibrium of the previously specified form if and only if it satisfies (12). Suppose w^* does satisfy (12). Then, by verification of the expected utilities for agents with benefit $v < G^{-1}(w^*)$ and those with $v \geq G^{-1}(w^*)$, we see that no agent has an incentive to unilaterally deviate so it is an equilibrium. Suppose w^* does not satisfy (12). If $G^{-1}(w^*) > \min\{\tilde{\beta}(w^*;r^t), \kappa + (1-\bar{p}^tE)\tilde{\beta}(w^*;r^t)\}$ then by the continuity of m there exists $v \in [\min\{\tilde{\beta}(w^*;r^t), \kappa + (1-\bar{p}^tE)\tilde{\beta}(w^*;r^t)\}, G^{-1}(w^*)]$ such that $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_v(\ell_S,w^*;\theta^t,\gamma^t)]$ and $\mathbb{E}_{\theta^t\sim r^t,\gamma^t\sim\bar{p}^t}[u_v(\ell_B,w^*;\theta^t,\gamma^t)]$ are strictly less than 0 so agents with this benefit type will choose ℓ_R and deviate from the prescribed equilibrium. Therefore, the isolated masses that can correspond to equilibria are precisely those that solve (12).

Finally, observe that the solution to (12) has a unique solution w^* in $[0,1]$. This is obvious since $G^{-1}(w)$ is well-defined, strictly increasing and continuous over $w \in [0,1]$ since G has bounded nonzero density. Moreover, $\tilde{\beta}(w;r^t)$ is strictly decreasing over $w \in [0,1]$. Hence, the two arguments in the minimum are both continuous, strictly decreasing functions which makes the minimum continuous and strictly decreasing. Moreover, $G^{-1}(0) = 0 < \tilde{\beta}(0;r^t)$ and $G^{-1}(1) = M > 0 = \tilde{\beta}(1;r^t)$. Therefore, there must be a unique solution w^* to (12).

Together, this shows that there is a unique equilibrium with isolated mass specified by the solution to (12). Moreover, the actions of agents in this equilibrium are as described in the result. ■.

B. Proof of Lemma 1 and Lemma 2

The map $m(\bar{p}^t;r^t)$ is by definition bounded since the solution to the fixed point equation (12) is in $[0,1]$ by Proposition 1.

To show the map m is also weakly decreasing and continuous in $\bar{p}^t$ for any r^t, observe that it is equal to the solution to the following equation for fixed r^t and $\bar{p}^t$:

$$\begin{aligned}\tilde{m}(w,\bar{p}^t,r^t) &= G^{-1}(w) - \min\{\tilde{\beta}(w;r^t), \kappa + (1-\bar{p}^tE)\tilde{\beta}(w;r^t)\}\\ &= 0\end{aligned}$$

The function $\tilde{m}(w,\bar{p}^t,r^t)$ is jointly continuous, strictly increasing in w and weakly increasing in $\bar{p}^t$ (since $\kappa + (1-\bar{p}^tE)\tilde{\beta}(w;r^t)$ is decreasing in $\bar{p}^t$). Therefore, since $m(\bar{p}^t;r^t) = \sup\{w : \tilde{m}(w,\bar{p}^t,r^t) \leq 0\}$, by the Maximum Theorem, we know that $m(\bar{p}^t;r^t)$ is continuous in $\bar{p}^t$. Moreover, since the set $\{w : \tilde{m}(w,\bar{p}^t,r^t) \leq 0\}$ contracts as $\bar{p}^t$ increases, $m(\bar{p}^t;r^t)$ is weakly decreasing in $\bar{p}^t$. Likewise, since $\tilde{m}(w,\bar{p}^t,r^t)$ is strictly decreasing in r^t, we can analogously show that $m(\bar{p}^t;r^t)$ is continuous and weakly increasing in r^t.

We show that $m(\bar{p}^t;r^t)$ is Lipschitz in $\bar{p}^t$ by choosing any $p_1,p_2 \in [0,1]$ and observe that letting $w_1 = m(p_1;r^t)$ and $w_2 = m(p_2;r^t)$, $\tilde{m}(w_1,p_1,r^t) = 0 = \tilde{m}(w_2,p_2,r^t)$ so:

$$\tilde{m}(w_1,p_1,r^t) - \tilde{m}(w_2,p_1,r^t) = \tilde{m}(w_2,p_2,r^t) - \tilde{m}(w_2,p_1,r^t)$$

Using fact that $\min(x_2,y_2) - \min(x_1,y_1) \leq \max(x_2 - x_1, y_2 - y_1)$, observe that:

$$\begin{aligned}\tilde{m}(w_2,p_2,r^t) - \tilde{m}(w_2,p_1,r^t) &\leq E|p_2-p_1|\tilde{\beta}(w_2;r^t)\\ &\leq E|p_2-p_1|(\theta_H C_1 + C_2)\end{aligned}$$

Similarly, using fact that $0 \leq (1-\bar{p}^tE)\tilde{\beta}(w;r^t) \leq \tilde{\beta}(w;r^t)$ and that $y_2 > y_1, x_2 > x_1$ implies that $\min(x_2,y_2) - \min(x_1,y_1) \geq \min(x_2-x_1, y_2-y_1)$:

$$\frac{|w_2-w_1|}{\bar{g}} \leq \tilde{m}(w_1,p_1,r^t) - \tilde{m}(w_2,p_1,r^t)$$

Putting together, this implies that $\frac{|w_2-w_1|}{|p_2-p_1|} \leq E\bar{g}(\theta_H C_1 + C_2)$. Hence, $m(\bar{p}^t;r^t)$ is Lipschitz in $\bar{p}^t$ for all r^t.

Likewise, we can show that $m(\bar{p}^t;r^t)$ is $(\bar{g}(\theta_H-\theta_L)C_1)$-Lipschitz in r^t. This is immediate since by choosing $r_1,r_2 \in [0,1]$, with $w_1 = m(\bar{p}^t;r_1)$ and $w_2 = m(\bar{p}^t;r_2)$, $\tilde{m}(w_1,\bar{p}^t,r_1) = 0 = \tilde{m}(w_2,\bar{p}^t;r_2)$, we have:

$$\tilde{m}(w_2,\bar{p}^t,r_1) - \tilde{m}(w_1,\bar{p}^t,r_1) = \tilde{m}(w_2,\bar{p}^t,r_1) - \tilde{m}(w_2,\bar{p}^t,r_2)$$

As before:

$$\tilde{m}(w_2,\bar{p}^t,r_1) - \tilde{m}(w_1,\bar{p}^t,r_1) \geq \frac{|w_2-w_1|}{\bar{g}}$$

Then the result is implied since:

$$\tilde{m}(w_2, \bar{p}^t, r_2) - \tilde{m}(w_2, \bar{p}^t, r_1) \leq C_1(\theta_H - \theta_L)$$

$\blacksquare$.

C. Proof of Lemma 3

Recall that $W(r) = \inf\{\mathcal{W}(r)\}$ or equivalently $W(r) = \inf\{p : m(p;r) \in \mathcal{Y}\}$. By the continuity and monotonicity of m established in Lemma 2, $W(r)$ is well-defined and $\mathcal{W}(r) = [W(r), 1]$.

Consider any $0 \leq r_1 < r_2 \leq 1$. Then, if $W(r_2) = \infty$ (i.e. $\mathcal{W}(r_2) = \emptyset$), then trivially $W(r_2) \geq W(r_1)$. If $W(r_2) < \infty$, then $\mathcal{W}(r_2) \neq \emptyset$ so $W(r_2) \in [0,1]$ and $m(W(r_2); r_2) = x$. However, observe that $m(W(r_2); r_1) \leq m(W(r_2); r_2) = x$ by Lemma 2, so $W(r_2) \geq \inf\{\mathcal{W}(r_1)\} = W(r_1)$. Hence, it is proven that $W(r)$ is weakly increasing in r.

Analogously, choose any $r_1 \in dom(W)$, then there exists Δ such that for all r_2 with $|r_2 - r_1| < \Delta$, $r_2 \in dom(W)$. Without loss of generality, let $r_1 < r_2$, then $W(r_1) \leq W(r_2)$. Furthermore, observe that by continuity of m, $m(W(r_1); r_1) = m(W(r_2); r_2) = x$. We then observe that $m(W(r_2); r_1) < m(W(r_1); r_1) = x$, and observe that it must be that $\mathbb{E}_{\theta^t \sim r_1, \gamma^t \sim W(r_1)}[u_v(\ell_S, y^t; \theta^t, \gamma^t)] \leq \mathbb{E}_{\theta^t \sim r_1, \gamma^t \sim W(r_1)}[u_v(\ell_B, y^t; \theta^t, \gamma^t)]$ for all v. In other words, all active agents must be choosing ℓ_B since choosing booster efficacy belief $p < W(r_1)$ must yield $m(p; r_1) > x$ by definition of W so the infectious costs incurred by active population must be increasing with p which is only possible if active agents have chosen ℓ_B. Therefore, observe that by Lemma 2:

$$m(W(r_2); r_2) - m(W(r_2); r_1) \leq C_1 \bar{g}(\theta_H - \theta_L)(r_2 - r_1)$$

Moreover, since $0 < m(W(r_2); r_1) < x$:

$$m(W(r_1); r_1) - m(W(r_2); r_1) \geq \underline{g} E \tilde{\beta}(x; 0)(W(r_2) - W(r_1))$$

Putting together, this implies that for any $r_1, r_2 \in dom(W)$ with $W(r_1), W(r_2) \in (0,1)$:

$$\frac{W(r_2) - W(r_1)}{r_2 - r_1} \leq \frac{C_1 \bar{g}(\theta_H - \theta_L)}{\underline{g} E \tilde{\beta}(x; 0)}$$

Hence, since r_1 was arbitrary, W is locally Lipschitz continuous over $dom(W)$. $\blacksquare$

D. Proof of Corollary 1

Since $0 < W(r^{eq}) < 1$ and $r^{eq} \in dom(W)$, the proof is immediate by Lemma 3 since W is locally Lipschitz continuous over $dom(W)$, so limits are preserved under W. $\blacksquare$.

E. Proof of Proposition 2

This follows from the same logic presented in Proposition 2.2 in [24] and Lemma 1 in [13]. We sketch the proof briefly. It suffices to show that $V^t_{\delta,\mathcal{Y}}(p) \geq a_1 V^t_{\delta,\mathcal{Y}}(p_1) + a_2 V^t_{\delta,\mathcal{Y}}(p_2)$ whenever $p = a_1 p_1 + a_2 p_2$ for $a_1, a_2 \geq 0$ with $a_1 + a_2 = 1$. Consider that the planner chooses the following compound mechanism π^t_+. First, π_+ uses the mechanism corresponding to $\{(a_1, p_1), (a_2, p_2)\}$. Then, depending on the first signalling mechanism's realized signal, a second signalling mechanism is used to generate an *additional* signal according to $\pi^t_*(p_1)$ and $\pi^t_*(p_2)$. Therefore, this compound mechanism achieves a payoff of $a_1 V^t_{\delta,\mathcal{Y}}(p_1) + a_2 V^t_{\delta,\mathcal{Y}}(p_2)$. But, observe that the set of mechanisms is trivially closed under composition by taking the product set of the signals and the product of the signalling distributions, so $\pi_+ \in \Pi$. Hence, since $\pi_+ \in \Pi$, the value realized by π_+ lower bounds the true maximum over Π in (10), so $V^t_{\delta,\mathcal{Y}}(p) \geq a_1 V^t_{\delta,\mathcal{Y}}(p_1) + a_2 V^t_{\delta,\mathcal{Y}}(p_2)$. $\blacksquare$

F. Proof of Proposition 3

Fix any optimal mechanism $\pi^t_* = \{(q^t_i, \mu^t_i)\}_{i \in \mathcal{I}^t}$ and consider the value achieved by this mechanism:

$$\begin{aligned}
V^t_{\delta,\mathcal{Y}}(\underline{p}^t) &= \sum_{i \in \mathcal{I}^t} q^t_i (1-\delta)\mathbb{I}\{y^*_t(i^t) \in \mathcal{Y}\} + \delta \sum_{i \in \mathcal{I}^t} q^t_i V^{t+1}_{\delta,\mathcal{Y}}(\underline{p}^{t+1}) \\
&= \sum_{i \in \mathcal{I}^t} q^t_i (1-\delta)\mathbb{I}\{m(\mu^t_i; r^t) \in \mathcal{Y}\} \\
&\quad + \delta \sum_{i \in \mathcal{I}^t} q^t_i V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu^t_i)) \\
&= \sum_{i \in \mathcal{I}^t} q^t_i (1-\delta)\mathbb{I}\{\mu^t_i \geq W(r^t)\} + \delta \sum_{i \in \mathcal{I}^t} q^t_i V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu^t_i))
\end{aligned}$$

Observe that the first term is maximized at $(1-\delta)$ when choosing π_{NI} since $q^t_1 = 1$ and $\mu^t_1 = \underline{p}^t = \bar{p}^t \geq W(r^t)$. By concavity and linearity of ϕ_Γ, $V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\cdot))$ is concave. Hence, the second term is also maximized when choosing π_{NI} since π_{NI} achieves $\delta V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\underline{p}^t))$ and because by Jensen's inequality:

$$\begin{aligned}
\delta \sum_{i \in \mathcal{I}^t} q^t_i V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu^t_i)) &\leq \delta V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\sum_{i \in \mathcal{I}^t} q^t_i \mu^t_i)) \\
&= \delta V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\underline{p}^t))
\end{aligned}$$

Therefore, π_{NI} maximizes this equation and is optimal. If $W(r^t) = \infty$, the proof is analogous, since the first term is maximized at 0 when choosing π_{NI} as no mechanism can induce beliefs that generate any reward. Likewise, the second term is maximized by π_{NI}, so it is proven. $\blacksquare$

G. Proof of Proposition 4

Suppose $\underline{p}^t < W(r^t)$ and consider any signalling mechanism $\pi^t = \{(q^t_{i^t}, \mu^t_{i^t})\}_{i^t \in \mathcal{I}^t}$. We show that we can improve on π^t with one that is supported on at most two posteriors. Let $J = \{ i \in \mathcal{I}^t : \mu^t_i < W(r^t) \}$ and let $I = \mathcal{I}^t \setminus J$. Let $q_I = \sum_{i \in I} q^t_i$ and $q_J = \sum_{i \in J} q^t_i$. Similarly, let $\mu_I = \frac{\sum_{i \in I} q^t_i \mu^t_i}{q_I}$ and $\mu_J = \frac{\sum_{i \in J} q^t_i \mu^t_i}{q_J}$. Then observe by Jensen's inequality since $V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\cdot)))$ is concave, the value of the mechanism $\tilde{\pi} = \{(q_I, \mu_I), (q_J, \mu_J)\}$ is:

$$\begin{aligned}
q_I \bar{V}^t_{\delta,\mathcal{Y}}(\mu_I) + q_J \bar{V}^t_{\delta,\mathcal{Y}}(\mu_J) &= (1-\delta) q_I \\
&+ \delta(q_I V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu_I)) + q_J V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu_J))) \\
&\geq (1-\delta) \sum_{i \in I} q^t_i + \\
\delta \sum_{i \in I} q^t_i V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu^t_i)) &+ \delta \sum_{i \in J} q^t_i V^{t+1}_{\delta,\mathcal{Y}}(\phi_\Gamma(\mu^t_i))
\end{aligned}$$

The last quantity is exactly the value achieved by π^t. Hence, the mechanism $\tilde{\pi}$ proposed weakly improves on π^t, thus any mechanism can be weakly improved with one that is split into at most two posteriors with one in $\mathcal{W}(r^t)$ and one in the complement. Thus, an optimal mechanism can be chosen to be such a mechanism. Finally, observe that since the two posteriors must be mean-preserving and a weighted average of μ_1^t and μ_2^t must equal $\underline{p}^t$, one posterior μ_1^t must be less than or equal to $\underline{p}^t$ and the other posterior $\mu_{i^t}^t$ larger than or equal to $\underline{p}^t$. ■.

H. Proof of Proposition 5

The proof is immediate from Corollary 2 in [4]. ■

I. Proof of Lemma 4

By the optimality of the π_*^t, $V_{\delta,\mathcal{Y}}^t(\underline{p}^t) = q_b^t V_{\delta,\mathcal{Y}}^t(\mu_b^t) + q_a^t V_{\delta,\mathcal{Y}}^t(\mu_a^t)$. But, since $V_{\delta,\mathcal{Y}}^t(p)$ is concave on $[\mu_b^t, \mu_a^t]$, it is necessarily linear on this entire interval. Therefore, given any point $p \in [\mu_b^t, \mu_a^t]$, the value from splitting on μ_b^t and μ_a^t is precisely equal to $V_{\delta,\mathcal{Y}}^t(p)$, so doing so is optimal. ■.

J. Proof of Proposition 6

Observe that from Proposition 5 that $V_{\delta,\mathcal{Y}}^t(p)$ is continuous in p for all t. Therefore, from (10), observe that $\bar{V}_{\delta,\mathcal{Y}}^t(p)$ is piecewise continuous and concave on $[0, W(r^t))$ and $[W(r^t), 1]$. However, there is a discontinuity at $p = W(r^t)$, so $\bar{V}_{\delta,\mathcal{Y}}^t$ is not concave and therefore by the stated result from [23], there must be a maximal interval $[s_1, s_2]$ with:

$$s_1 = \sup\{p < W(r^t) : \bar{V}_{\delta,\mathcal{Y}}^t(p) < V_{\delta,\mathcal{Y}}^t(p)\}$$
$$s_2 = \inf\{p \geq W(r^t) : \bar{V}_{\delta,\mathcal{Y}}^t(p) < V_{\delta,\mathcal{Y}}^t(p)\}$$

Moreover, by the stated definition of the concave envelope $V_{\delta,\mathcal{Y}}^t$ is linear on $[s_1, s_2]$ with $0 \leq s_1 < W(r^t)$ and $W(r^t) \leq s_2 \leq 1$. Hence, $V_{\delta,\mathcal{Y}}^t(p) = \frac{s_2-p}{s_2-s_1} V_{\delta,\mathcal{Y}}^t(s_1) + \frac{p-s_1}{s_2-s_1} V_{\delta,\mathcal{Y}}^t(s_2)$ and $V_{\delta,\mathcal{Y}}^t(p) > \bar{V}_{\delta,\mathcal{Y}}^t(p)$ for all $p \in (s_1, s_2)$. If $s_2 > W(r^t)$, this would be a contradiction since by Lemma 4, it is strictly optimal at $p = W(r^t)$ to use the mechanism that mixes over s_1 and s_2 over the non-informative mechanism. However, this contradicts Proposition 3 since $W(r^t)$ would lie in (s_1, s_2). Therefore, $s_2 = W(r^t)$. Hence, for any $p \in (s_1, W(r^t))$, the optimal mechanism mixes over $\mu_b^t = s_1$ and $\mu_a^t = W(r^t)$. Since $V_{\delta,\mathcal{Y}}^t(p)$ was piecewise concave, there can be at most one linear segment where $V_{\delta,\mathcal{Y}}^t(p) > \bar{V}_{\delta,\mathcal{Y}}^t(p)$, so $V_{\delta,\mathcal{Y}}^t(p) = \bar{V}_{\delta,\mathcal{Y}}^t(p)$ is concave on $0 < p < s_1$. This implies that for $p \in (0, s_1)$, $\pi_*^t(p) = \pi_{NI}(p)$ as the non-informative mechanism is sufficient to achieve the optimum. Observe that for such $p < s_1$, choosing $\mu_b^t = p$ and $\mu_a^t = W(r^t)$, is equivalent to π_{NI} so the statement holds in full generality. ■.

K. Proof of Theorem 1

We first verify that $t_\dagger$ is finite. For case **(a)**, since Corollary 1 guarantees that $W(r^t)$ converges to $W(r^{eq}) > p^{eq}$, there must exist some finite $t' > 0$ such that for all $t \geq t'$, $W(r^t) > \frac{p^{eq}+W(r^{eq})}{2} > p^{eq}$. Therefore, $t_\dagger \leq t'$, so $t_\dagger$ is finite. For case **(b)**, since r^t is decreasing sequence, it must be that $r^t \geq r^{eq}$ for all t. Therefore, $W(r^t) \geq W(r^{eq})$ for all t which implies that $W(r^t) \geq p^{eq}$ for all t. Hence, by the Markov update being contractive (i.e. $0 < \zeta_\Gamma + \zeta_\Gamma \nu_\Gamma < 1$), $(W(r^t) - \phi_\Gamma(W(r^t))) \geq (W(r^{eq}) - \phi_\Gamma(W(r^{eq}))) > 0$. Let $\Delta = W(r^{eq}) - \phi_\Gamma(W(r^{eq})) > 0$. Then, using the convergence of $W(r^t)$ from Corollary 1, for $\frac{\Delta}{2} > 0$, there exists some finite t' such that for all $t > t'$, $|W(r^t) - W(r^{t+1})| \leq \frac{\Delta}{2}$. Therefore:

$$\begin{aligned}\phi_\Gamma(W(r^t)) &= \phi_\Gamma(W(r^t)) - W(r^t) + W(r^t)\\ &\leq \phi_\Gamma(W(r^{eq})) - W(r^{eq}) + W(r^t)\\ &\leq W(r^t) - \Delta\\ &\leq W(r^{t+1}) - \frac{\Delta}{2} \leq W(r^{t+1})\end{aligned}$$

Therefore, for all $t > t'$, $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$. Hence, $t_\dagger \leq t'$. Hence, in both cases we have shown that $t_\dagger$ is finite.

Next we show that for all $t > t_\dagger$, $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$. For case **(b)**, this is trivially true by definition. For case **(a)**, observe that for $t > t_\dagger$, $W(r^t) \geq p^{eq}$. Recall that for $p^{eq} \leq p$, $p^{eq} \leq \phi_\Gamma(p) \leq p$. Hence, since W is weakly increasing and r^t is increasing in t for case **(a)**:

$$\begin{aligned}p^{eq} &\leq \phi_\Gamma(W(r^t))\\ &\leq W(r^t)\\ &\leq W(r^{t+1})\end{aligned}$$

Therefore, it is proven that for all $t > t_\dagger$, $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$. Now, what remains to show is that if for all $t > t_\dagger$, $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$, then $\pi_*^t = \pi_\dagger^t$.

To show this claim, we first observe that we can reformulate the Markov decision process in 11 by specifying $\mu_b^t \in [0, W(r^t))$ as opposed to π^t as we have shown in Proposition 6 that an optimal mechanism π_*^t is fully specified by μ_b^t so this is sufficient to achieve optimality. Observe that since $\mu_a^t = W(r^t)$ by choosing μ_b^t, q_b^t and q_a^t are implicitly defined by μ_b^t. We drop the dependence for notational convenience.

$$\begin{aligned}V_{\delta,\mathcal{Y}}^t(\underline{p}^t) &= \max_{\mu_b^t \in [0, W(r^t))} \mathbb{E}[(1-\delta)v_\mathcal{Y}(\bar{p}^t) + \delta V_{\delta,\mathcal{Y}}^{t+1}(\underline{p}^{t+1})]\\ &= \max_{\mu_b^t \in [0, W(r^t))} (1-\delta)v_\mathcal{Y}(\mu_a^t)\\ &\quad + \delta q_a^t \bar{V}_{\delta,\mathcal{Y}}^{t+1}(\mu_a^t) + \delta q_b^t \bar{V}_{\delta,\mathcal{Y}}^{t+1}(\mu_b^t)]\\ &= \max_{\mu_b^t \in [0, W(r^t))} (1-\delta)v_\mathcal{Y}(\mu_a^t)\\ &\quad + \delta q_a^t V_{\delta,\mathcal{Y}}^{t+1}(\phi_\Gamma(\mu_a^t)) + \delta q_b^t V_{\delta,\mathcal{Y}}^{t+1}(\phi_\Gamma(\mu_b^t))]\end{aligned}$$

Therefore, all that remains to show by Proposition 6, is that in any such period t, when $\underline{p}^t < W(r^t)$, an agent achieves $V_{\delta,\mathcal{Y}}^t(\underline{p}^t)$ by choosing $\mu_b^t = 0$.

For any $t > t_\dagger$, suppose $V_{\delta,\mathcal{Y}}^{t+1}$ is linear on $[0, W(r^{t+1})]$ – we seek to show that this implies that $\mu_b^t = 0$. This implies that $\bar{V}_{\delta,\mathcal{Y}}^t$ being linear on $[0, W(r^t))$ since $\phi_\Gamma([0, W(r^t)]) \subseteq [0, W(r^{t+1})]$ since $\phi_\Gamma(0) > 0$ and $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$. If any $\mu_b^t \in (0, W(r^t))$ were chosen, then, this would be a contradiction since $\bar{V}_{\delta,\mathcal{Y}}^t = V_{\delta,\mathcal{Y}}^t$ and is linear on $[0, \mu_b^t]$ with slope which we denote by ρ. But since, $\bar{V}_{\delta,\mathcal{Y}}^t$, is linear on $[0, W(r^t)]$, it has slope ρ on $[\mu_b^t, W(r^t)]$ as well. Moreover, by the choice of μ_b^t, $V_{\delta,\mathcal{Y}}^t$ is linear on $[\mu_b^t, W(r^t)]$ with slope ρ' with $\bar{V}_{\delta,\mathcal{Y}}^t(W(r^t)) = V_{\delta,\mathcal{Y}}^t(W(r^t))$. If $\rho = \rho'$, then observe that $V_{\delta,\mathcal{Y}}^t$ is linear on $[0, W(r^t)]$ and this would imply that $\mu_b^t = 0$. Hence, by the concavity of $V_{\delta,\mathcal{Y}}^t$, $\rho' < \rho$. However, this would imply that:

$$\begin{aligned} V^t_{\delta,\mathcal{Y}}(\mu^t_b) - V^t_{\delta,\mathcal{Y}}(W(r^t)) &< \bar{V}^t_{\delta,\mathcal{Y}}(\mu^t_b) - \bar{V}^t_{\delta,\mathcal{Y}}(W(r^t)) \\ &= V^t_{\delta,\mathcal{Y}}(\mu^t_b) - V^t_{\delta,\mathcal{Y}}(W(r^t)) \end{aligned}$$

Therefore, if $V^{t+1}_{\delta,\mathcal{Y}}$ is linear on $[0, W(r^{t+1})]$, then $\mu^t_b = 0$. Hence, it suffices to show that for all $t > t_\dagger$, $V^{t+1}_{\delta,\mathcal{Y}}$ is linear on $[0, W(r^{t+1})]$.

Next, as is common in the dynamic programming literature, define the functional operator T_t which operates on functions $V : [0, 1] \to [0, 1]$. Formally, we define for any such V if $p < W(r^t)$:

$$\begin{aligned} T_t V(p) = \max_{\mu^t_b \in [0, W(r^t))} &(1-\delta) v_{\mathcal{Y}}(\mu^t_a) \\ &+ \delta q^t_a V(\phi_\Gamma(\mu^t_a)) + \delta q^t_b V(\phi_\Gamma(\mu^t_b)) \end{aligned}$$

Moreover, if $p \geq W(r^t)$:

$$T_t V(p) = (1-\delta) + \delta V(\phi_\Gamma(p))$$

Notationally, we will let $T_{[t_1,t_2]}$ for some $t_1 < t_2$ refer the joint operation $T_{t_1} \ldots T_{t_2}$. By definition, we note that $V^t_{\delta,\mathcal{Y}} = T_t V^{t+1}_{\delta,\mathcal{Y}}$.

We highlight the following fact from value iteration: misspecification of the value functions propagates error subexponentially when using dynamic programming. Namely, we state have the following lemma which states that by applying the value function at any time $j < \hat{t}$ has a bounded exponential difference.

Lemma 5: Given a function $V : [0, 1] \to [0, 1]$, for $j < t$, $\|T_{[j,t-1]} V^t_{\delta,\mathcal{Y}} - T_{[j,t-1]} V\|_\infty \leq \delta^{t-j}$.

Proof: For all p^t, $0 \leq V^t_{\delta,\mathcal{Y}}(p^t) \leq \sum_{j=0}^\infty (1-\delta)\delta^j = 1$. Hence, $\|V^t_{\delta,\mathcal{Y}} - V\|_\infty \leq 1$. Next, observe that T_t is a δ-contractive for all t, as is a standard dynamic programming result [25]. Then, this implies that for any $j < t$:

$$\begin{aligned} \|T_{[j,t-1]} V^t_{\delta,\mathcal{Y}} - T_{[j,t-1]} V\| &\leq \delta^{t-j} \|T_{[j,t-1]} V^t_{\delta,\mathcal{Y}} - T_{[j,t-1]} V\| \\ &\leq \delta^{t-j} \end{aligned}$$

■

Next, we make the following observation that when μ^t_b is chosen by applying T_t on any function V, it corresponds to the left endpoint on the concave envelope of the function $(1-\delta) v_{\mathcal{Y}}(\mu^t_a) + \delta q^t_a V(\phi_\Gamma(\mu^t_a)) + \delta q^t_b V(\phi_\Gamma(\mu^t_b))$ on the restricted domain $[0, W(r^t)]$. For $t > t_\dagger$, by the previous analysis, if V is linear over $[0, W(r^{t+1})]$ then $T_t V$ chooses $\mu^t_b = 0$ and $T_t V$ is linear over $[0, W(r^t)]$ as in Fig. 5.

Therefore, for any $t > t_\dagger$ we arbitrarily choose any linear V and compute $T_{[t+1:\ell]} V$ for some $\ell > t+1$. Therefore, since V was linear and $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$ for all $j \in [t+1, .., \ell]$, $T_{[t+1:\ell]} V$ is linear over $[0, W(r^{t+1})]$. But by Lemma 5 $\|V^{t+1}_{\delta,\mathcal{Y}} - T_{[t+1:\ell]} V\| \leq \delta^{\ell-t-1}$, and since ℓ was arbitrary we can take $\ell \to \infty$ and hence $T_{[t+1:\ell]} V = V^{t+1}_{\delta,\mathcal{Y}}$ for any V. But since our arbitary choice of V was linear, this shows that $V^{t+1}_{\delta,\mathcal{Y}}$ is also necessarily linear on $[0, W(r^{t+1})]$. Hence, $\mu^t_b = 0$ must be optimal for all $t > t_\dagger$ and π^t_* for all $t > t_\dagger$.

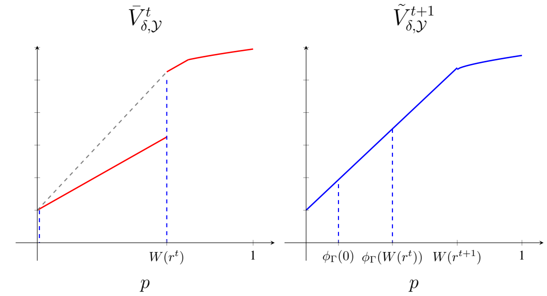

Fig. 5: $\bar{V}^t_{\delta,\mathcal{Y}}$ (red, on left) $V^t_{\delta,\mathcal{Y}} = conc(\bar{V}^t_{\delta,\mathcal{Y}})$ (dashed, on left), $\tilde{V}^{t+1}_{\delta,\mathcal{Y}}$ (on right).

L. Proof of Proposition 7

Suppose that $\delta \leq \min_{j \leq t \leq t_\dagger} \frac{\phi_\Gamma^{-1}(W(r^{t+1}))}{W(r^t)}$. Recall by Theorem 1, that $\pi^t_* = \pi^t_\dagger$ for all $t > t_\dagger$. To show the proposition statement, it suffices to show that $\mu^t_b = 0$ for all $j \leq t \leq t_\dagger$ or equivalently that $V^t_{\delta,\mathcal{Y}} = conc(\bar{V}^t_{\delta,\mathcal{Y}})$ is linear on $[0, W(r^t)]$. We proceed by induction from $t = t_\dagger + 1$ down to $t = j$. At $t = t_\dagger + 1$, the greedy strategy is optimal by Theorem 1, so $\mu^t_b = 0$ and $V^t_{\delta,\mathcal{Y}}$ is trivially linear on $[0, W(r^t)]$.

If at $j < t+1 \leq t_\dagger + 1$, suppose $\mu^{t+1}_b = 0$, we seek to show that $\mu^t_b = 0$. Notice that by the induction hypothesis, $V^{t+1}_{\delta,\mathcal{Y}}$ is linear on $[0, W(r^{t+1})]$. Therefore, observe that if $\phi_\Gamma(W(r^t)) \leq W(r^{t+1})$, by the same argument in Theorem 1, $\mu^t_b = 0$ and we are done. If $\phi_\Gamma(W(r^t)) > W(r^{t+1})$, then $\bar{V}^t_{\delta,\mathcal{Y}}$ is linear on $[0, \phi_\Gamma^{-1}(W(r^{t+1}))]$ since $\phi_\Gamma([0, \phi_\Gamma^{-1}(W(r^{t+1}))]) \subseteq [0, W(r^{t+1})]$. Therefore, since a linear segment on the concave envelope dominating $\bar{V}^t_{\delta,\mathcal{Y}}$ cannot start in the middle of a segment where $\bar{V}^t_{\delta,\mathcal{Y}}$ is linear, either $\mu^t_b = 0$ or $\mu^t_b \in [\phi_\Gamma^{-1}(W(r^{t+1}))), W(r^t))$. However, if the slope of $\bar{V}^t_{\delta,\mathcal{Y}}$ on $[0, \phi_\Gamma^{-1}(W(r^{t+1}))]$ is less than the average slope between $[\phi_\Gamma^{-1}(W(r^{t+1})), W(r^t)]$, it is necessary that $\mu^t_b = 0$.

We note that the average slope between $[\phi_\Gamma^{-1}(W(r^{t+1})), W(r^t)]$ is at least $\frac{\delta}{W(r^t) - \phi_\Gamma^{-1}(W(r^{t+1}))}$ and the slope between $\bar{V}^t_{\delta,\mathcal{Y}}$ on $[0, \phi_\Gamma^{-1}(W(r^{t+1}))]$ is upper bounded by $\frac{\delta}{\phi_\Gamma^{-1}(W(r^{t+1}))}$ since $|\bar{V}^t_{\delta,\mathcal{Y}}(\phi_\Gamma^{-1}(W(r^{t+1}))) - \bar{V}^t_{\delta,\mathcal{Y}}(0)| \leq 1$. Therefore, the following inequality implies that $\mu^t_b = 0$:

$$\frac{1-\delta}{W(r^t) - \phi_\Gamma^{-1}(W(r^{t+1}))} \geq \frac{\delta}{\phi_\Gamma^{-1}(W(r^{t+1}))}$$

Rearranging, this inequality becomes $\delta \leq \frac{\phi_\Gamma^{-1}(W(r^{t+1}))}{W(r^t)}$, which is satisfied by the assumption in the proposition, so $\mu^t_b = 0$.

Since $\mu^t_b = 0$ for all $t \geq j$, $\pi^t_* = \pi^t_\dagger$ for all $t \geq j$. ■.

We offer a brief discussion on how to solve for μ^t_b for $t \leq t_\dagger$ more generally, although this is not the focus of this paper. Backtracking from $t = t_\dagger$ where we know $V^{t+1}_{\delta,\mathcal{Y}}$ is piecewise linear for all such t. Therefore, $\bar{V}^t_{\delta,\mathcal{Y}}$ is piecewise-linear with a finite number of "kinks" (or non-differentiable points). It is known that the concave envelope of a piecewise-linear function can only add linear segments at these kinks. Furthermore, it can be shown that the concave envelope of $\bar{V}^t_{\delta,\mathcal{Y}}$ when $V^t_{\delta,\mathcal{Y}}$ is piecewise-linear with ℓ kinks, has at most $\ell+1$ kinks. Therefore, a backtracking dynamic programming solution exists that keeps at each point in time t tests which kink point is optimal and should be chosen as μ^t_b. Since $t_\dagger$ is finite, and the maximum number of kinks we can consider in any round is bounded by $t_\dagger + 2$, this algorithm has finite runtime.

M. Instance-level Dominance of Greedy Disclosure Rule

In this subsection, we verify the instance-level dominance of the greedy disclosure rule numerically. Specifically, we show that the reward obtained in each stage (1 if and only if $y^*_t \in \mathcal{Y}$) by the greedy disclosure rule weakly dominates that of the fully-informative and non-informative disclosure rules. Fig. 6 displays the trajectories of $y^*_t(i^t)$ under a single stochastic trajectory of states $\{\gamma^t\}_{t\geq 1}$ as an example. The figure illustrates that by minimizing the information shared while maximizing the current period reward, the greedy disclosure rule outperforms the outcomes (i.e. current period rewards) of the full information scheme and the non-informative scheme in each period. The very form of the construction that motivates the greedy disclosure rule consequently allows it to achieve the best possible outcome for each period, leading to superior performance on an instance level.

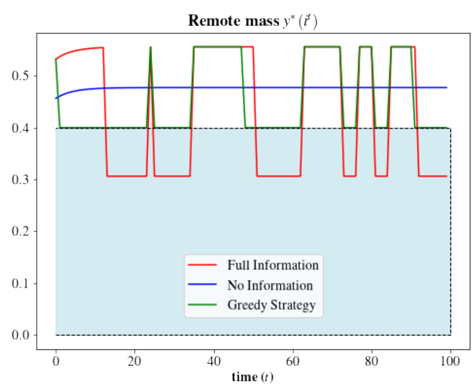

Fig. 6: Sample trajectory of $y^*_t(i^t)$ under a non-informative, fully-informative and greedy signalling. The parameters are identical Fig. 4 with $p^{eq} = 0.3$ and $\underline{p}^1 = 0.25$.