Beyond the Surface of Digital Contact Tracing: Delving into the Interconnected World of Technology, Individuals, and Society

Jiejun Hu-Bolz, Katayoun Farrahi, and Manuel Cebrian

***Abstract*—Digital Contact Tracing (DCT) is a great example of information systems assisting societal problems. However, privacy concerns lead to reduced DCT adoption rates. When dealing with critical societal issues, policymakers seek to use various strategies, such as interventions and financial subsidies, to steer the behavors of individuals. This causes individuals to face a sophisticated decision-making process when coping with the public health crisis and the adoption of DCT, i.e., giving up privacy and freedom to gain information to remain healthy. In this paper, we consider a scenario, where policymakers allocate rewards to individuals to motivate their compliance with the interventions; And the individuals decide the optimal compliance effort based on their health state, privacy loss, interaction with their neighbors, and rewards. To tackle the trade-off between a number of individuals and policymakers, in this paper, we propose a Leader-Followers Mean-Field Game model to analyze this time-dependent, dynamic, and large-scale decision-making problem. The numerical results demonstrate that by allocating appropriate rewards, policymakers can play a role in guiding the behavior of individuals in various scenarios.**

***Index Terms*—Digital Contact Tracing, Mean-Field Game, Dynamic systems**

I. Introduction

Digital Contact Tracing (DCT) is the best digital means available in containing an epidemic. However, DCT efficacy can only be maximized under certain settings, such as with unlimited testing resources, proper testing prioritizing strategies, and a considerable number of compliant app users [1]. DCT is designed to utilize information, such as location, test results, and contacts, to notify users of risks and recommendations. For DCT to be effective, users' cooperation and participation are paramount. However, there are some obstacles to the large-scale adoption of DCT. First, users have privacy concerns when using DCT. Then, it leads to extra costs if users follow the recommendations when high risk occurs, such as quarantine and testing. Lastly, some users doubt if the DCT works as promised in the wild, i.e., a complex heterogeneous network. Users may desire more incentives to participate in this large-scale information collection activity.

In this paper, we consider a scenario where society is challenged by active infectious diseases, such as coronavirus and influenza. For modeling purposes, we consider two main characters, namely policymakers and individuals, who interact with each other. Policymakers introduce multiple strategies, e.g., DCT, to mitigate the impact of infectious disease. The proposed scenario is depicted in Fig. 1. From the policymakers' point of view (the outer circle in Fig. 1), they continuously counterbalance the societal resources of a complex system that consists of social activities (such as education, sports events, and family visiting) and economic activities (such as producing, trading, and consuming products) under the pressure of active virus transmission in the population. Policymakers face a dilemma: On one hand, the seriousness of the disease impacts the population's well-being, straining the healthcare system and hindering social and economic activities; on the other hand, participation in these activities is vital, especially during challenging times, but it can also lead to higher infection rates. Policymakers aim to reduce the impact of the pandemic and restore order in society by strategically implementing interventions [2]. Intervention effectiveness is gauged by disease progression and collective efforts. Based on these indicators, policymakers dynamically adapt strategies for optimal outcomes. For individuals, the general trend we observe is that they desire normal activities for mental support and material needs, yet these interactions can affect health. Complying with interventions may incur economic costs and privacy concerns. Individuals face an intricate decision-making process of how to behave in this scenario. And their decisions are a direct signal of the effectiveness of the interventions.

It is important for policymakers to find effective strategies that can lead to positive effects on society while also respecting the needs of individuals. In this paper, we investigate the complex interactions in a simplified mathematical formulation between individuals and policymakers by modeling their respective decision-making processes. Various factors are used to characterize individuals, such as their preferences for using DCT, interactions with neighbors, and stubbornness. This enables us to evaluate the effectiveness of the policymakers' strategies in different scenarios. By analyzing the interactions among individuals and policymakers, we aim to unveil the essential criteria for achieving a balanced and harmonious dynamic within. Note, in this paper, we make some assumptions that may be oversimplified and only consider non-pharmaceutical interventions. However, this is valuable in the novel approach which can be easily extended in future work.

We propose a Leader-Follower Mean-Field Game (LFMFG) to investigate a time-dependent, dynamic, and large-scale decision-making problem for policymakers and individuals.

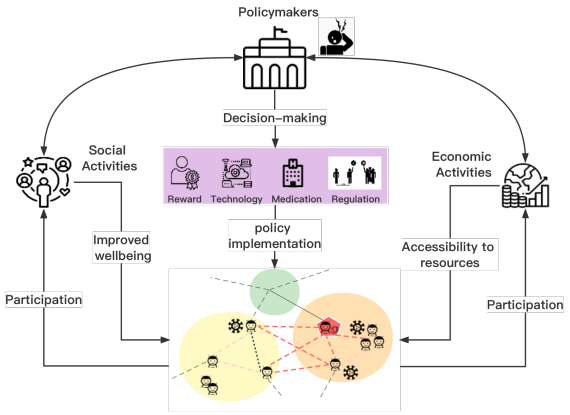

Figure 1. The complex relationship between policymakers and individuals

Particularly, we would like to shed light on the following issues: 1) Examine the interconnectedness of pandemics' progression, individual decision-making, and rewards; 2) Analyze the dynamic decision-making of individuals in scenarios involving influential neighbors and policymaker subsidies. Explore the equilibrium in multiple populations regarding dynamic exogenous factors, including policymakers' strategies; 3) Investigate policymakers' decision-making in implementing DCT and allocating subsidies and its impact on collective control efforts. Assess the effectiveness and challenges of policymakers' strategies.

It is essential to investigate the relationship between information-driven technology, individuals, and policymakers as a whole. The proposed scenario can be also adopted to analyze other complex systems, such as the influence of personalized advertisement on an individual's shopping preferences and market regulation; the influence of social media on individual politic views and democracy; the influence of communication technology on individual information transmission speed and cultural revolution.

II. Related work

Mean-Field Game (MFG) was first proposed by Lasry et. al [3]. MFG is unlike legacy game theory [4], which focuses on studying the equilibrium among a limited number of agents. It is often used to analyze large-scale interactions between non-cooperative agents. MFG is widely adopted to solve population mobility issues, such as emergency evacuation [5], opinion evolution in social networks [6], and market dynamics [7].

Game theory is adopted to address large-scale decision-making problems. [8]–[10]combine game theory and compartmental models in epidemiology, e.g., the SIR (Susceptible-Infected-Recovered) model and the SEIR (Susceptible-Exposed-Infected-Recovered) model, to study the dynamics of COVID-19 and investigate the optimal actions taken by the government and individuals. Hubert et. al [11] study moral hazard in an SIR-based principle-agent model. The optimal government subsidy is calculated considering the infection rate and optimal effort of the population. Elie et. al [12] adopt MFG in the compartmental model to inform the optimal contact behavior of the individuals. However, individuals and government are modeled separately. [13] reassembles most closely the proposed model in this paper. It proposes a Stackelberg-based extended MFG framework in a single population. The agents aim to control the rates of transition in the SIR model to minimize the individual cost. Monte Carlo simulations and machine learning tools are used to solve the problem. Although integrating compartmental models is beneficial in modeling the progress of the pandemic, it also oversimplifies the decision-making process of the individuals. Additionally, the transition of different states is in a Markov fashion, where an individual progresses from a susceptible state to an infected state. However, this is not the case in reality, since individuals can make more effort in protecting themselves.

We observe the emergence of information technology as a means to mitigate the impact of infectious diseases. Information technology, including data collection and dissemination techniques like DCT, gather data to optimize the utilities of the agents. Policymakers, in turn, proactively utilize information technology and interventions to address the adverse effects of infectious diseases. Consequently, it is crucial to investigate how these factors influence individuals' decision-making processes. Haw et. al [14] study the trade-off between economic, social, and health outcomes in the management of a pandemic in a macro view. López et. al [15] study the key factors that affect the effectiveness of DCT in France. They propose a network and SIR-based model to analyze different transmission settings, adoption rates, and ages. It informs policymakers in designing and implementing effective DCT strategies to control the spread of infectious diseases. Although several network structures are considered to demonstrate the various relationships of individuals, the decision-making processes of individuals are merely presented. Lorch et. al [16] propose a temporal point access modeling framework to study COVID-19 transmission in super spreading events. The proposed model combines an agent-based approach and a compartmental model together to introduce new human behavioral elements, namely appearance at the hotspots. However, people's subjective initiative is neglected, such as they may decide not to be present.

In comparison to existing work, this paper offers several distinct differences and contributions, as follows:

- Compared to the SIR and network-based models, our agent-based model introduces continuous health states and factors like compliance cost, interaction cost, information cost, and rewards. This comprehensive approach provides a detailed perspective on strategy dynamics and reward effectiveness, setting it apart from the previous study.
- This work considers the collective effort of multiple populations, integrating specific influential factors. The proposed model offers valuable insights into motivating interactions among heterogeneous populations using rewards.
- The proposed LFMFG captures time-dependent individual decisions and population health distribution, revealing the impact of policymakers' determination and rewards. It explains interventions' adjustments, providing a comprehensive understanding of the interplay between individual decisions, policymaker interventions, and population health.

III. MODELING

In this work, we consider the interactions between policymakers and individuals as a leader-followers game. The objective of policymakers is to optimize their *determination* in implementing rewards, which serve to motivate individuals' control efforts and encourage the adoption of DCT. A higher determination corresponds to a more robust implementation of the rewards, which can also be seen as encompassing interventions. Individuals make optimal control efforts based on factors, such as health state, interaction cost, privacy cost, and rewards provided by policymakers. A higher level of control effort can result in reduced infection risk, as indicated by a lower health state value. While maintaining a low infection risk is advantageous for both individuals and society, the increased control effort can impose significant costs on daily life due to the potential restrictions on social and economic activities.

We consider multiple populations in the proposed model over a period of time $t \in [0, T]$. Individuals are denoted as $i \in \mathcal{I} = \{1, ..., N\}$. We assume that individuals of different populations can interact freely. The contact set of individual i is defined as $\mathcal{N}_i = \{0, 1, ..., M\}$. The health state of individual i is represented as $r_i \in [0, 1]$ and it is influenced by its contacts $\mathcal{N}_i$. Each individual exerts a time-dependent control effort $u_i(r_i, t)$ based on its current health state r_i. In the following subsections, we first introduce the optimization problem faced by policymakers, followed by the individual decision-making dynamics. Lastly, we establish the LFMFG to solve this dynamic, time-dependent, large-scale game.

A. Leader's objective and constraints

Policymakers provide incentives to individuals based on the collective control effort of populations, with the objective of minimizing the overall cost of regulation implementation, DCT deployment, and total incentive expenditure. We quantify the policymakers' determination of rewards implementation and denote it as $D(t)$. A set of rewards for policymakers is denoted as $\mathcal{G}$, and the reward at time t is represented as $g(t)$. The value of $g(t)$ is influenced by the collective control effort $\mathbf{u}$ and the determination $D(t)$. Hence, we define the dynamics of the reward as a partial differential equation (PDE) [17]

$$dg(t) = \{\omega_1[D(t) - \mathbf{u}]g(t) + \omega_2 \sum_{p=1}^{P} \bar{\delta_p}(t)\}dt \tag{1}$$

where ω_1, ω_2 are positive weights. When $D(t) \geq \mathbf{u}$, the unit reward $g(t)$ increases, indicating that policymakers exert a strong determination to implement regulations and DCT. On the contrary, weak collective compliance requires generous rewards to be stimulated. The second term represents the average health state, where $\sum_{p=1}^{P} \bar{\delta_p}(t)$ represents the aggregated average state of population p. The aggregated average state will be explained in more detail in Section III-B. Subsequently, we define the cost of policymakers

$$J_0(D, g, \mathbf{u}) = \int_0^T [\mathbf{u}g(t) + \eta D(t) + \tau D(t)^2]dt \tag{2}$$

where η and τ represent the unit costs of data processing and system maintenance, respectively. $\mathbf{u}g(t)$ represents the total reward. $\eta D(t)$ denotes the data processing cost, which is proportional to the determination. $\tau D(t)^2$ represents the maintenance cost. The quadratic form of the maintenance cost ensures diminishing returns [18]. In summary, policymakers aim to minimize the cost by solving this PDE-constrained optimization problem

$$\min_{D(t)} \quad J_0 = \int_0^T [\mathbf{u}g(t) + \eta D(t) + \tau D(t)^2]dt \tag{3a}$$

$$\text{s.t.}$$

$$dg(t) = \{\omega_1[D(t) - \mathbf{u}]g(t) + \omega_2 \sum_{p=1}^{P} \bar{\delta_p}(t)\}dt \tag{3b}$$

B. Follower's objective and constraints

Individuals, as followers, face this complex decision-making process that requires balancing participation in social and economic activities, potential privacy loss, control effort, health preservation, and rewards. The total cost is bounded by the dynamics of their health states, which are influenced by the health condition of their contacts, control effort, and rewards. We first present the dynamics of an individual i's health state r_i. Health state is a time-dependent variable. The dynamics of health states can be defined as a PDE

$$dr_i(t) = [\frac{\sum_{n \in \mathcal{N}_i} \alpha_{i,n} r_{i,n}(t)}{\sum_{n \in \mathcal{N}_i} \alpha_{i,n}} - r_i(t) + \beta u_i(r_i, t) + \gamma g(t)]dt \tag{4}$$

where β, γ are positive weights. The health state difference over time depends on the interaction between one's contacts in the first term, one's current health state $r_i(t)$, control effort $u_i(t)$, and rewards $g(t)$ allocated by policymakers. The interaction between individual i and its neighbor n is denoted as a binary variable $\alpha_{i,n} \in \{0, 1\}$. $r_{i,n}$ represents the encounter risk between individual i and n, which can be further defined based on factors, e.g., contact duration, distance, and environment. Health state r_i can also be regarded as the risk factor, with a low value indicating low risk and vice versa. Throughout the rest of this paper, the term "health state" will be used as the default expression.

We then define the cost of individual i, which includes the control cost, interaction cost with its neighbors, privacy cost while using the DCT, penalty when the actual health state deviates from the ideal state, the termination cost, and the reward for the control effort. To simplify the interactions between individual i and its neighbors, we define a mean-field term that aims to replace the description of the many-to-many interactions with an aggregated term. We define the mean-field term of population i as $\delta_i(r, t)$, which enables us to indicate the population i's influence on its neighbor population n.

- **Control cost** f_c: The control cost of the current state r_i is defined as $f_c(r_i, u_i, t) = \frac{1}{2} a_1 u_i^2(r_i, t)$, where a_1 is a positive weight. The control cost is in quadratic form to ensure diminishing returns [4].
- **Interaction cost** f_a: The interaction cost is based on individual i and its neighbors. We then denote the interaction cost as $f_a(r_i, \delta_n, t) = r_i(t) + a_2 \sum_{n=1}^{M} \delta_n$, where a_2 is a positive weight and δ_n is the neighbor's health state distribution. The underlying assumption is that if individual

i interacts with its contacts, it will be affected by higher risk r_i. However, if individual i interacts with contacts who have low-risk, its state remains the same.

- **Privacy cost** f_p: DCT requires users' data to provide accurate warnings. We assume that individuals are selfish and have varying preferences for data sharing. For example, in the benchmark setting, individuals may feel hesitant or ashamed to share their information when the risk factor is high. We also introduce the concept of a "technology embracer" who is more inclined to share information when the risk is either very low or high. The privacy cost is defined as $f_p(r_i,t) = a_3\delta_i[(r_i - w_1)^2 + w_2]$, where a_3 is the unit value of privacy. w_1 is the differentiator between the benchmark setting and the technology embracer scenario.
- **Penalty** f_{pe}: We define the ideal health state as r_0, representing a state where there is no risk to one's health. Individual aims to achieve a health state as close to r_0 as possible to avoid the penalty. Therefore, the penalty is defined as $f_{pe}(r_i, r_0, t) = a_4[r_i(t) - r_0(t)]^2$, where a_4 represents the unit penalty for the deviation between the current health state and the ideal health state. The definition implies that the greater the deviation from the ideal state, the higher the penalty will be.
- **Termination cost** f_t: Termination cost is determined by the health state at the termination time T. We define the termination cost as $f_t(r(T)) = a_5 r(T)$, where a_5 is a positive weight.
- **Reward**: The individual's control effort is rewarded by policymakers, which is defined as $g(t)u_i(r_i,t)$.

According to the definitions of the individuals' costs, the cost function is denoted as L_i for individual i.

$$\begin{aligned} L_i(u_i,\delta,g,r_i) =& \frac{1}{2}a_1 u_i^2(r_i,t) + (r_i + a_2\sum_{n=1}^{M}\delta_n) \\ &+ a_3\delta_i[(r_i - w_1)^2 + w_2] + a_4(r_i - r_0)^2 \\ &+ a_5 r_i(T) - g(t)u_i(r_i,t) \end{aligned} \tag{5}$$

Then, we define the expected accumulated cost function of individual i as follows

$$J_i(u_i,\delta,g,r_i) = \int_0^T [L_i(u_i,\delta,g,r_i)]dt + a_5 r_i(T) \tag{6}$$

Since the cost function contains δ_i and δ_n, we only write δ in the cost function. However, it is not realistic to assume that an individual has only a limited number of contacts during social and economic activities, considering "next closed" encounters. Each individual aims to minimize the cost, and all the individuals in this population share the same objective. Hence, to analyze the decision-making process at a population level, we can adopt an individual i to represent the entire population to which they belong. It is important to note that in this paper, we replace notion i, which represents an individual, with population i to represent the corresponding population. We now define the health state dynamics of population i as

$$\begin{aligned} dr_i &= [\sum_{n=1}^{M}\epsilon_n\zeta_{i,p}\delta_n - r_i + \beta u_i(r_i,t) + \gamma g(t)]dt + \sigma_i dW_i(t) \\ &= f_i(u_i,\delta_n,g,r_i)dt + \sigma dW(t) \end{aligned} \tag{7}$$

where ϵ_n is the impact factor of population i's neighboring population n, which can be proportional to the size of the population or the mobility of the population. $\zeta_{i,n}$ is the correlation between population i and n. We can infer when two populations are geographically adjacent when the correlation is strong. δ_n is population n's distribution of health state. We capture the randomness by using Brownian motion, where σ_i is the diffusion constant of population i, and $W_i(t)$ is a standard Wiener process [3].

Hence, we generalize the individual's optimization problem to population i as the following PDE-constrained optimization problem

$$\min_{u_i,\,g} \quad J_i = \int_0^T [L_i(u_i,\delta,g,r_i)]dt + a_5 r_i(T) \tag{8a}$$

s.t.

$$dr_i = f_i(u_i,\delta_n,g,r_i)dt + \sigma_i dW_i(t) \tag{8b}$$

IV. Solution of LFMFG

To solve the optimal control in a dynamic environment, we use Hamilton-Jacobi-Bellman (HJB) equation for optimal control and Fokker-Planck-Kolmogorov (FPK) equation for system dynamics [3].

A. Problem reformulation

Any individual in population i aims to minimize the cost. The evolution of the optimal compliance u_i and the mean-field term δ_i can be captured by HJB and FPK equations, respectively. The fore-mentioned optimization problems in Eq. (3) and (8) can be represented in HJB and FPK equations. First, we define the value function $v_i(r_i,t)$ for population i

$$v_i(r_i,t) = \inf_{u_i^*\in U} J_i(u_i,\delta,g,r_i) \tag{9}$$

As indicated in Eq.(9), the value function represents the minimum value of the cost when choosing an optimal u_i^*. To obtain the optimal control, we adopt the HJB as follows

$$\partial_t v_i + H_i(r_i, D_r v_i, D_{rr}^2 v_i, \delta) = 0 \tag{10}$$

where value function v_i is twice differentiable with r_i and differentiable with t. $H_i(r_i, D_r v, D_{rr}^2 v, \delta)$ is Hamiltonian, which is defined as

$$H_i(r,p,q,\delta) = \inf_{u\in U}[L_i(u_i,\delta,g,r_i) + p f_i(u_i,\delta_n,g,r_i)] + \frac{\sigma^2}{2}q \tag{11}$$

where $p = D_r v_i$ is the adjoint variable and $q = D_{rr}^2 v_i$. Since u_i^* can be obtained from Hamiltonian, we also have the solution of the following equation

$$f_i(u_i^*,\delta,g,r_i) = \frac{\partial H_i}{\partial p}(r_i,\delta,p) \tag{12}$$

$\delta(r_i,t)$ is the distribution of state r_i at time t. The distribution δ changes over time, which is affected by the individuals changing into state r_i and individuals changing state r_i into other states. We could capture the dynamics by using the divergence of the distribution's gradient.

$$\begin{aligned} \partial_t\delta_i + div(D_p H_i(r_i,\delta,D_u)\delta_i) - \frac{\sigma^2}{2}\partial_{rr}^2\delta_i = 0 \\ \partial_t\delta_i + div(f_i(r_i,\delta,u_i)\delta_i) - \frac{\sigma^2}{2}\partial_{rr}^2\delta_i = 0 \end{aligned} \tag{13}$$

We re-formulate the optimization problem according to HJB and FPK. We aim to find the optimal control u_i^* and determination D^* as a Nash Equilibrium point for populations and policymakers, respectively.

Definition 1 (Equilibrium of populations and policymaker)**.** *For a given reward g, the population i reaches Nash Equilibrium, if and only if all the individuals in population i reach the relationship described by the following equations*

$$J_i(u_i^*, \delta, g, r_i) > J_i(u_i, \delta, g, r_i)$$
$$J_0(D^*, g, \mathbf{u}) > J_0(D, g, \mathbf{u})$$

According to the NE, no individuals in population i can benefit from deviating from the optimal control strategy u_i^*, while policymakers strive to exert the optimal determination D^* in implementing the rewards. To find the NE, we adopt a backward induction approach: First, the followers treat the reward g from policymakers as given and optimize the control effort u_i. Then, the leader takes the collective control effort into account and optimizes the policy deployment determination D.

B. Equilibrium of the populations

In Eq. (8b), the dynamics of the health state of individual i is based on the individual's control effort, neighbor's health state, rewards, and the current health state. To generalize the problem into a multiple population scale, we need to reformulate Eq. (8b) to present the mean-field term of population i, i.e., the dynamics of the health state distribution, according to Eq. (13).

$$\partial_t \delta_i + div(f_i(u_i, \delta_n, g, r)\delta_i) - \frac{\sigma_i^2}{2}\partial_{rr}^2 \delta_i = 0 \tag{14}$$

The divergence term can be derived as $div(f_i\delta_i) = \frac{\partial f_i(u_i,\delta_n,g,r)}{\partial r}\delta_i + f_i(u_i, \delta_n, g, r)\frac{\partial \delta_i}{\partial r}$. This term represents the diffusion rate of population i's health state distribution caused by neighboring populations. Therefore, we reformulate the optimization problem for the multiple-population scenario as follows

$$\min_{u_i, g} \quad J_i = \int_0^T [L_i(u_i, \delta, g, r_i)\delta_i]dt + a_5\delta_i(r_i, T)r_i(T) \tag{15a}$$

s.t.

$$\partial_t \delta_i = -div(f_i(u_i, \delta_n, g, r)\delta_i) + \frac{\sigma_i^2}{2}\partial_{rr}^2\delta_i \tag{15b}$$

The objective function in Eq. (15a) represents the expected cost of population i with the constraint of the mean-field term δ_i dynamics in Eq. (15b). We solve this optimization problem by adopting the method of Lagrangian multipliers. We then write the Lagrangian function as follows.

$$\begin{aligned}\mathcal{L}_i(u_i, \delta_i, v_i) = &\int_0^t \int_{r\in R^d} L_i(u_i, \delta, g, r_i)\delta_i dr dt \\ &+ \int_{r\in R^d} a_5\delta_i(r, T)r(T) + \sum_{i=1}^N \int_0^t \int_{r\in R^d} \\ &p_i[\partial_t\delta_i + div(f_i\delta_i) - \frac{\sigma_p^2}{2}\partial_{rr}^2\delta_i]drdt\end{aligned} \tag{16}$$

where $p_i(r,t) = \frac{\partial v_i(r,t)}{\partial r}$ is the Lagrangian multiplier. We assume that there exists a tuple (u_i, δ_i, v_i) to minimize the cost function, which satisfies $\frac{\partial \mathcal{J}_i}{\partial u_i} = 0$, $\frac{\partial \mathcal{J}_i}{\partial \delta_i} = 0$, and $\frac{\partial \mathcal{J}_i}{\partial v_i} = 0$. Therefore, we obtain

$$\frac{\partial \mathcal{L}_i}{\partial u_i} = \delta_i \frac{\partial L_i}{\partial u_i} + \frac{\partial v_i}{\partial r}\frac{\partial f_i}{\partial u_i}\frac{\partial \delta}{\partial r_i} \tag{17}$$

$$\frac{\partial \mathcal{L}_i}{\partial \delta_i} = \frac{\partial v_i}{\partial t} + H_i + a_5 r + \delta_i \frac{\partial L_i}{\partial \delta_i} + \delta_i \frac{\partial f_i}{\partial \delta_i}\frac{\partial v_i}{\partial r} - \frac{\sigma^2}{2}\frac{\partial^2 v_i}{\partial r^2} \tag{18}$$

$$\frac{\partial \mathcal{L}_i}{\partial v_i} = \frac{\partial \delta_i}{\partial t} + \frac{\partial f_i \delta_i}{\partial r} - \frac{\sigma^2}{2}\frac{\partial^2 \delta_i}{\partial r^2} \tag{19}$$

We adopt finite difference method (FDM) [19] to solve the Eq. (17), (18), and (19). We first discretize the time step $t \in [0, T]$ and health state step $j \in [0, 1]$. Then, we define the iteration step sizes of time and state space as $\Delta t = \frac{T}{X}$ and $\Delta r = \frac{1}{Y}$, where X and Y are the numbers of time and state step size, respectively. By applying the FTCS scheme, we first Hence, we denote the discrete mean-field term of population i's jth health state at time t as $\mathbf{d}_{i,j}^t$, the value function as $\mathbf{v}_{i,j}^t$, and control effort as $\mathbf{u}_{i,j}^t$.

$$\begin{aligned}\mathbf{d}_{i,j}^{t+1} =&\frac{1}{2}(\mathbf{d}_{i,j+1}^t + \mathbf{d}_{i,j-1}^t) - \frac{\Delta t}{2\Delta r}(f_{i,j+1}^t \mathbf{d}_{i,j+1}^t - f_{i,j-1}^t \mathbf{d}_{i,j-1}^t) \\ &+ \frac{\Delta t}{4\Delta r^2}\frac{\sigma_j^2}{2}(\mathbf{d}_{i,j+2}^t - \mathbf{d}_{i,j}^t + \mathbf{d}_{i,j-2}^t)\end{aligned} \tag{20}$$

$$\begin{aligned}\mathbf{v}_{i,j}^{t-1} =&\frac{1}{2}(\mathbf{v}_{i,j+1}^t + \mathbf{v}_{i,j-1}^t) + \Delta t(L_{i,j} + a_5 r + u_{i,j}^t \frac{\partial L_{i,j}}{\partial u_{i,j}^t}) \\ &- \frac{\Delta t}{2\Delta r}(f_{i,j} + \frac{\partial f_{i,j}}{\partial u})(\mathbf{v}_{i,j+1}^t - \mathbf{v}_{i,j-1}^t) \\ &+ \frac{\sigma_{i,j}^2}{8\Delta r}(\mathbf{v}_{i,j+2}^t - 2\mathbf{v}_{i,j}^t + \mathbf{v}_{i,j-2}^t)\end{aligned} \tag{21}$$

$$\mathbf{u}_{i,j}^t = \frac{1}{a_1}[g(t) - \beta \frac{\mathbf{v}_{i,j+1}^t - \mathbf{v}_{i,j-1}^t}{2\Delta r}\frac{\mathbf{d}_{i,j+1}^t - \mathbf{d}_{i,j-1}^t}{2\Delta r}] \tag{22}$$

C. Solution of the Leader's optimal determination

We solve the leader's optimization problem. The leader allocates reward $g(t) \in \mathcal{G}$ and expects the followers exert the best collective control effort ${u_{i,j}^t}^*$, which is defined as $\mathbf{u} = \sum_{i=1}^N \sum_{j=1}^{\mathcal{R}} {u_{i,j}^t}^* \delta_i^t$. The aggregating average health state is defined as $\delta_i^t = \sum_{i=1}^N \delta_i^t$. We can obtain the leader's optimal determination $D^*(t)$ according to HJB, which is defined as $H = L_0 + \lambda f_0$, where λ is the adjoint variable. Additionally, L_0 and f_0 are defined as

$$L_0 = \mathbf{u}g + \eta D(t) + \tau D(t)^2 \tag{23}$$

$$f_0 = \omega_1[D(t) - \mathbf{u}]g(t) + \omega_2\bar{\delta} \tag{24}$$

By solving the first derivation of H with respect to D, we obtain the optimal D^* at reward state k at time t

$$D^*(t) = -\frac{1}{2\tau}(\eta + \lambda\omega_1 g) \tag{25}$$

We define the value function of the leader's objective as $\mathbf{V}(g,t) = inf_{D(t)}J_0$, which also satisfies $\frac{\partial \mathbf{V}}{\partial t} + \frac{\partial H}{\partial g} = 0$. Hence, we obtain

$$-\frac{\partial \mathbf{V}}{\partial t} = \mathbf{u} + \lambda\omega_1[D(t) - \mathbf{u}] \tag{26}$$

We substitute the optimal determination in (25) into (26)

$$-\frac{\partial \mathbf{V}}{\partial t} = \mathbf{u} + \lambda\omega_1[-\frac{1}{2\tau}(\eta + \lambda\omega_1 g) - \mathbf{u}] \tag{27}$$

where $\lambda = \frac{\partial \mathbf{V}}{\partial g}$. We discretize the reward state space $[0, g_{max}]$ and the time interval $[0, T]$. Then, we define the step size of time and state space as $\Delta t = \frac{T}{Y}$ and $\Delta g = \frac{1}{Z}$, where Y and Z are the numbers of time and reward state steps, respectively.

$$\mathbf{V}_k^{t-1} = \frac{1}{2}(\mathbf{V}_{k+1}^t + \mathbf{V}_{k-1}^t) + \frac{\Delta t}{\Delta g}\omega_1(\frac{\eta}{2\tau} + \mathbf{u})(\mathbf{V}_{k+1}^t - \mathbf{V}_{k-1}^t) + \frac{\Delta t}{\Delta g^2}\frac{{\omega_1}^2 g}{2\tau}(\mathbf{V}_{k+1}^t - \mathbf{V}_{k-1}^t)^2 - \Delta t\mathbf{u} \quad (28)$$

By solving leaders value function $\mathbf{V}$, we obtain the leader's optimal strategy, namely the policy deployment determination, $\mathbf{D}_k(t)^*$ for the kth reward at time t according to (25)

$$\mathbf{D}_k^*(t) = -\frac{1}{2\tau}(\eta + \frac{\mathbf{V}_{k+1}^t - \mathbf{V}_{k-1}^t}{\Delta g}\omega_1 g) \quad (29)$$

D. Algorithm of LFMFG

We propose algorithm 1 to solve the proposed LFMFG. We first initiate the discrete mean-field term $\mathbf{d}_{i,j}^0$ and control effort $\mathbf{u}_{i,j}^0$, for population i at time 0; And value function $\mathbf{v}_{i,j}^T$ for population i at time T. We then initiate the discrete value function $\mathbf{V}_k^T$ of the kth reward at time T. In Algorithm 1, we first solve the population i's optimization problem by considering reward k. Then, the adjoint variable and health state distribution of population p are updated according to Eq. (21) and (20), respectively. The control effort $\mathbf{u}_{i,j}^t$ is updated according to Line 15, where α is the learning rate. The iteration runs until the error reaches the satisfaction or the maximum iteration limitation, where the maximum iteration is $I = 25$ and the error $\rho = 0.1$. The policymakers' optimization solution is presented from Line 21. After obtaining the collective control effort, the value function and the optimal determination of policymakers are calculated by HJB. The complexity of Algorithm 1 depends on the dimension of the problem, discrete step sizes, and the required accuracy. For the proposed FDM-based LFMFG, the current settings and the various simulated scenarios show stable performance and convergence. The details are presented in Section V-G.

Algorithm 1 LFMFG

1: **Initialize:** $\mathbf{d}_{i,j}^0, \mathbf{u}_{i,j}^0, \mathbf{v}_{i,j}^T, \mathbf{V}_k^T$, and err
2: **for** Reward $k = 0, .., K$ **do**
3: **for** Population $i = 0, .., N$ **do**
4: **while** $it \leq I$ or $err \geq \rho$ **do**
5: **for** $j = 0, ..., J$, $t = 0, ..., T$ **do**
6: **if** $\mathbf{u}_{i,j} = 0$ **then**
7: $\mathbf{d}_{i,j}^{t+1} = \mathbf{d}_{i,j}^t$
8: **else**
9: Solve $\mathbf{d}_{i,j}^t$ using (20)
10: **end if**
11: **end for**
12: **for** $j = 0, ..., J$, $t = T, ..., 0$ **do**
13: Solve $\mathbf{v}_{i,j}^t$ using (21)
14: **end for**
15: **for** $j = 0, ..., J$, $t = 0, ..., T$ **do**
16: Update $\mathbf{u}_{i,j}^t \leftarrow \alpha\mathbf{u}_{i,j}^t + (1-\alpha)\frac{\partial L_{i,j}^t}{\partial \mathbf{u}_{i,j}^t}$
17: **end for**
18: Compute $err = |L_{i,j}^t(it) - L_{i,j}^t(it-1)|$
19: **end while**
20: **end for**
21: **end for**
22: Leader's optimization problem
23: **for** $t = 0, ..., T$, $k = 0, ..., K$ **do**
24: Update $\mathbf{u}_k^t = \sum_{i=1}^N \sum_{j=1}^J {\mathbf{u}_{i,j}^t}^* \mathbf{d}_{i,j}^t$
25: **end for**
26: **for** $t = T, ..., 0$, $k = 0, .., K$ **do**
27: Update $\mathbf{V}_k^t$ using (28)
28: **end for**
29: **for** $t = 0, ..., T$, $k = 0, .., K$ **do**
30: Update $\mathbf{D}_k^t$ using (29)
31: **end for**

V. Simulation of the Proposed Model

We present simulation results for our proposed model in this section. We define scenarios to contextualize the parameters. We showcase results for these scenarios, including health state distribution, epidemic progression, individuals' optimal control evolution, policymakers' optimal determination, and model stability and equilibrium.

A. Evaluation setup

We first set the benchmark scenario with two populations. The initial distributions of population 1 and 2 are $(0.25, 0.2)$ and $(0.75, 0.5)$, respectively. The impact of populations 1 and 2 is denoted as $\epsilon_1 = 0.8$ and $\epsilon_2 = 0.5$. Additionally, the correlation factor of populations is set as $\zeta_{12} = 2$ to indicate population 1 has a greater influence on population 2. In the benchmark scenario, individuals have a higher overall cost of data contribution with a privacy weight $w_1 = 1.5$. These individuals can be referred to as data conservatives. We set $X = 30$, $Y = 50$, and $Z = 20$ as the steps of health states, time states, and rewards, respectively. The ideal health state is set as $r_0 = 0.1$. In the following simulations, we assess scenarios with influential populations, technology embracers, and intrinsic seekers, exploring their impact on policymakers' reward strategies. For the influential populations scenario, we raise ζ_{12}, enhancing population 1's influence on population 2. In the technology embracers scenario, we set $w_1 = 0.1$, lowering overall data contribution costs. In the intrinsic seekers scenario, individuals prioritize intrinsic costs with a smaller γ.

B. Health state dynamics and progression of epidemic

First, we study the dynamics of the health state distribution in Fig. 2 to understand the reward effectiveness. We demonstrate the impact of different rewards on the health state distribution of populations 1 and 2. The majority of the population naturally moves towards the low health states even if there are no rewards given, as shown in Fig. 2a (left). *This finding demonstrates that rational individuals in a natural setting are able to make sound decisions without additional influences.* Then, we introduce nudges from policymakers by allocating rewards in Fig. 2a (middle), where population 1 experiences fluctuation but eventually settles in a fairly ideal state. *This demonstrates that when regulation and rewards are enforced by policymakers, it initially leads to disruption, followed by a majority of the population settling at low health states.* On the other hand, the health state distribution of population 2 is influenced by population 1. With the aid of rewards, the population of the mediocre health state dominates

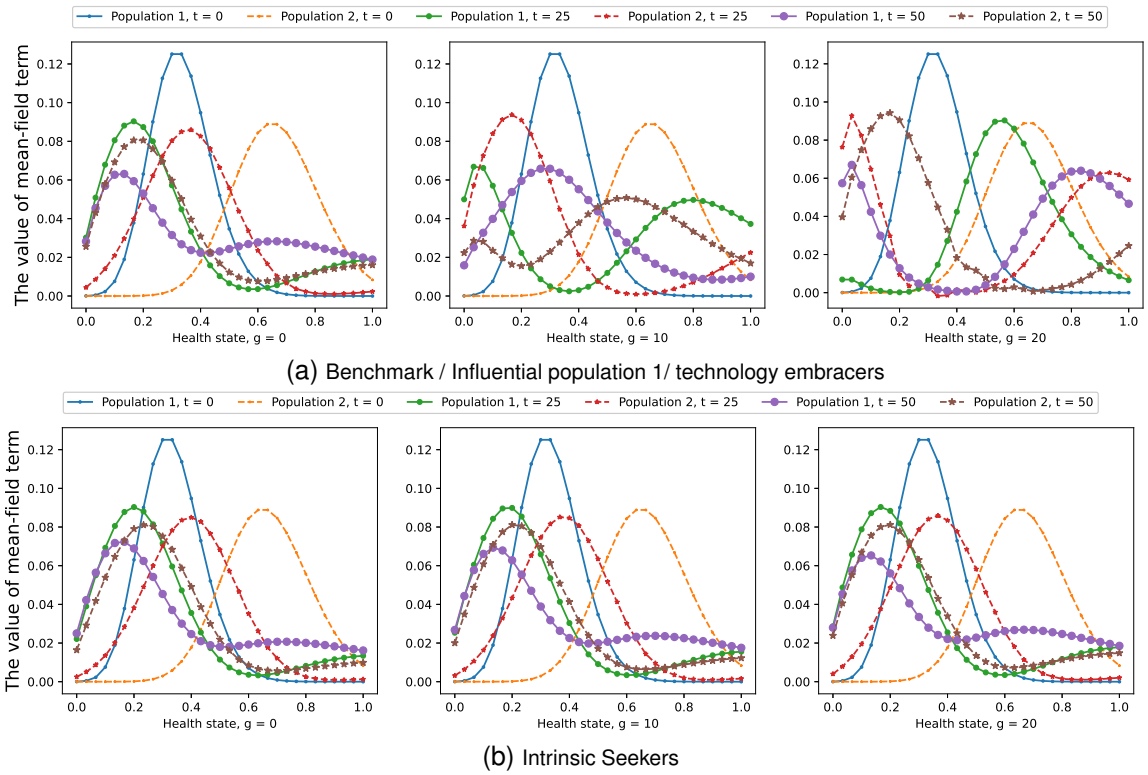

Figure 2. The health state distribution of populations 1 and 2 with respect to various rewards and health states: population 1 is represented by solid lines with dots, where the size of the dots increases with the time step. population 2 is depicted by dash lines with stars. The subfigures in 2a and 2b correspond to rewards $g = 0$ (left), 10 (middle), and 20 (right), respectively. Populations naturally tend towards low health states regardless of rewards. With increasing rewards, populations are initially disrupted but eventually settle into the ideal health states based on the given reward.

the majority compared to those without rewards. Surprisingly, when implementing $g = 20$, population 2 reaches the lowest health state and stabilizes in a fairly low health state. However, population 1 exhibits a highly polarised health distribution, namely a significant portion of the population is concentrated in very low and high health states. This indicates that the reward plays a substantial role in the decision marking process of individuals. Although being in high health states can lead to penalties, the reward can still compensate for the negative effects. It is astonishing that the maximum reward can encourage reckless behaviors and resistance to maintaining health. We also show that when considering influential population 1 and technology embracers, the distributions follow the same pattern. The scenario of intrinsic seekers is displayed in Fig. 2b. Although the population still has incentives to move towards lower health states, the results show resistance to the rewards.

SIR model [20] demonstrates epidemic dynamics with susceptible and infected populations changing over time. It lacks considerations for regulations, information, and individual subjectivity. In contrast, our proposed model captures epidemic progression dynamics along with other influential factors. In Fig. 3 (left), we observe a similar result to the SIR model: the low-risk group keeps increasing but eventually shows a decrease; the mediocre-risk group decreases initially, and then increases; and the high-risk group increases and then stabilizes. The deviations, such as the increasing mediocre-risk group, demonstrate a realistic decision-making process, namely individuals can afford mediocre risks in the current setting. This demonstrates a scenario where a mild infectious disease progresses naturally. The majority is able to keep well without any intervention by the government. Policymakers need to closely evaluate the progression of the situation and the impact on society, such as whether the available medical resource is enough to accommodate the increasing high-risk population. In Fig. 3 (middle), the high-risk and low-risk groups decrease significantly after reaching the peak, and the mediocre-risk group goes through periods of adjustment over time. These results show with the aid of rewards, the group moves to mediocre risks, allowing individuals to participate in social and economic activities while maintaining a balanced control and health cost. We can observe an improvement where the majority deviates from high risks when a greater reward is allocated. In Fig. 3 (left), we observe periodic dynamics with respect to different groups. Initially, low and high-risk groups dominate the population, followed by an increase in the mediocre-risk group. This situation echoes the COVID-19 scenario in Belgium, as policymakers deploy the most robust intervention and rewards, we observe the periodic dynamics of cases rising and declining [21].

C. Optimal control and precaution

We study the evolution of individuals' optimal control effort with a fixed reward $g = 10$. Fig. 4 illustrates a consistent

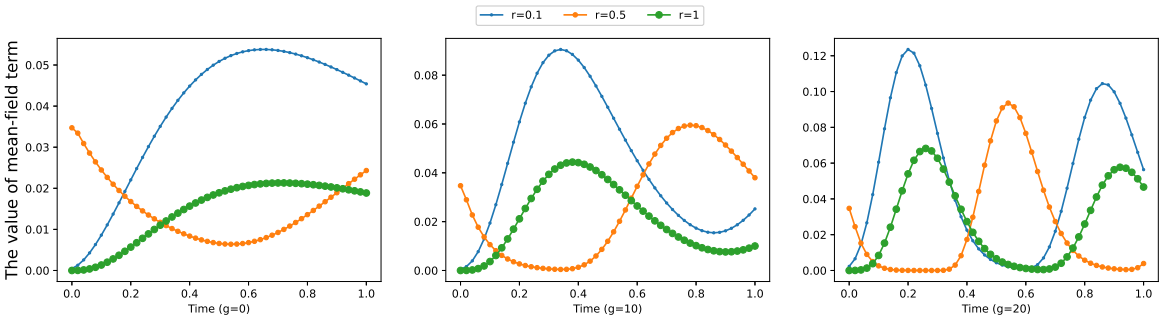

Figure 3. The health state distribution of population 1 with respect to various rewards and time in benchmark setting: Health states are used to demonstrate three distinct groups, namely the low risk (recovered), the mediocre risk (susceptible), and the high risk (infected). Health state $r = 0.1$ represents the low-risk group; health state $r = 0.5$ indicates the mediocre-risk group; and health state $r = 1$ stands for the high-risk group. With increasing rewards, the sizes of high, mediocre, and low-risk groups demonstrate a faster and periodic evolution.

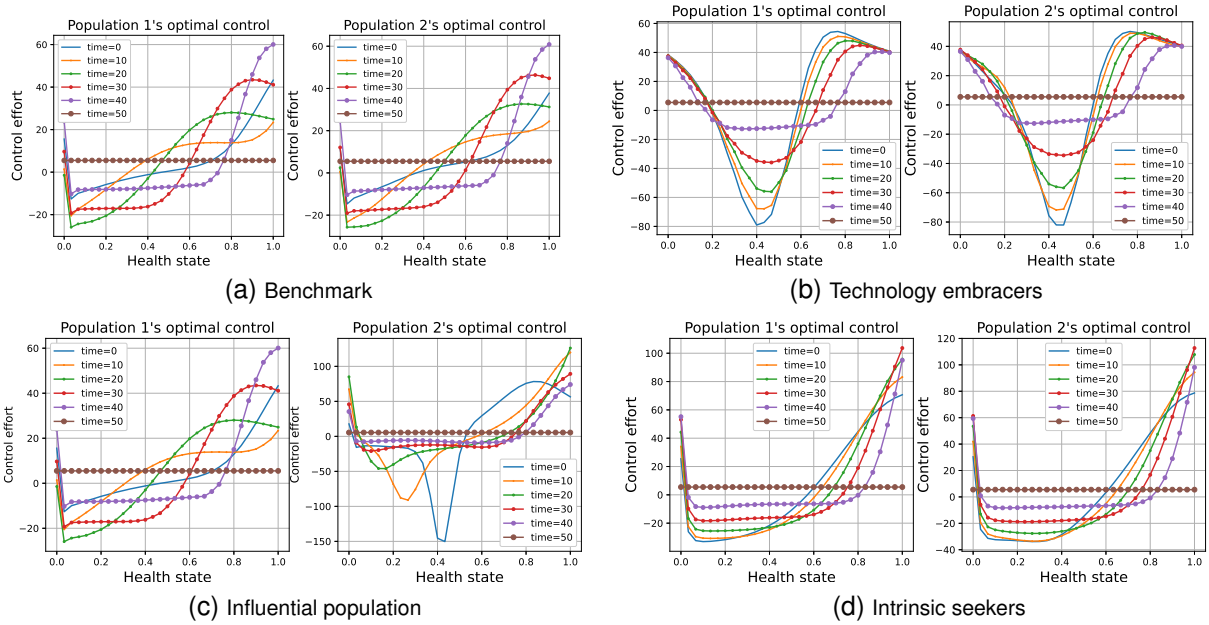

Figure 4. Optimal control effort of populations 1 and 2 with respect to health state given reward $g = 10$: The control effort evolution of populations 1 and 2 are depicted by solid lines with dots, where the size of the dots increases with time steps ($time = 0, 10, ..., 50$). In general, populations exert positive control efforts when health states are high. However, it is possible to alter this phenomenon by lowering the overall decision-making cost of individuals, as in the technology embracers scenario.

trend of the optimal control effort being amplified in situations where the health state is high (i.e., the risk is high). In our definition, a positive control effort signifies regulation compliance, while a negative control effort reflects a laid-back attitude. Note that due to the settings of the final state penalty parameter, when $t = 50$, there exists a uniformed optimal control effort. We also investigate the control effort with respect to rewards at $t = 30$ in Fig. 5. It shows that *the proposed rewards effectively encourage individuals to be preventive; larger rewards lead to early precautions (e.g., as low as $r = 0.4$ with reward $g = 20$)*. Subsidies/rewards, such as the 'State Aid' program by the European Commission, enable the setup of home-care equipment for the elderly.

The technology embracers scenario in Fig. 4b demonstrates a significantly distinct perspective. In this case, technology embracers show a greater inclination to share information in both low and high health states due to the lower data contribution cost compared to the benchmark, which leads to an enhancement of control efforts. Additionally, individuals in mediocre health states are motivated to avoid extreme rebellion over time. It's worth noticing that individuals exert positive control efforts over a larger spectrum of health states ($0 < r < 0.2$ and $0.7 < r < 1$), which showcases that *promoting the technology by making it cost-effective to use can promote precarious behaviors.* We then analyze the influential population scenario in Fig. 4c. In this case, population 1 has a greater impact on population 2. The laid-back attitude of population 2 at $r = 0.4$ reflects a scenario where interactions with individuals from a generally healthy region (population 1) make the mediocre-healthy individuals feel relaxed. Fortunately, this situation changes over time, namely the laid-back attitude eventually disappears. The laid-back attitude can cause temporary negative effects when gathering happens among different populations, such as superspreading

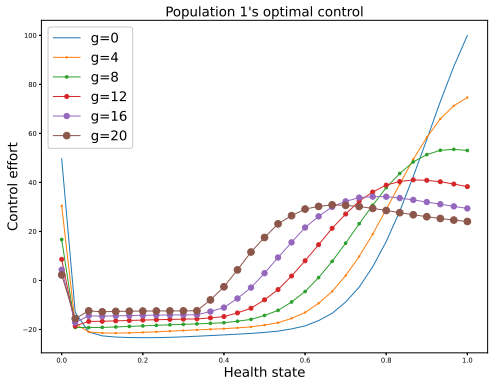

Figure 5. Optimal control effort of populations 1 with respect to health state at $t = 30$: The control effort evolution of population 1 is depicted by solid lines with dots, where the size of the dots increases with rewards. Rewards effectively encourage individuals to be preventive. With increasing rewards, it leads to early precautions when the health states are still relatively low.

events [22]. *Hence, policymakers need to be cautious when multiple populations interact together due to those temporary negative effects.*

D. Leader's determination

In Fig. 6, we explore policymakers' optimization. The optimal determination links to collective control efforts and rewards. Strong determination implies more intervention and rewards, fostering higher collective control effort by individuals. In Fig. 6a and 6c, we demonstrate the evolution of the optimal determination with respect to the reward state in the benchmark and influential population scenarios, respectively. In Stage 1, the optimal determination evolves drastically: when $t = 0$, the optimal determination mirrors the collective control effort in Fig. 7a and 7c where the collective control effort is at a low value regardless of the rewards. To stimulate the collective control effort, policymakers put more determination on the high reward in Stage 1. In Fig. 6c, policymakers increase the optimal determination in general to mitigate the impact of the influential population in Stage 1, which can be verified in Fig. 7c. In Stage 2, it illustrates the counter-effect of the determination and the collective control effort. There is an intensive degradation of high reward determination after acknowledging that collective control effort can be encouraged with lower rewards. The evolution finally settles down in Stage 3. With the collective control effort settling at an average value, the optimal determination reaches stability when allocating different rewards. However, the optimal determination severely drops to discourage individuals' prone to interest seeking.

In the benchmark and influential population scenarios, individuals are more sensitive to rewards due to additional overall costs. Policymakers show increased determination in Stages 1 and 2. They implement higher rewards to motivate the public because individuals need more compensation to take action. However, in Stage 3, policymakers lose motivation to implement high rewards. This is because there is a long-lasting effect on individuals who continue to exert appropriate control effort when the rewards are low. *This reflects the situation where, at the beginning of COVID-19, policymakers enforced various interventions and subsidies; however, once the public became accustomed to the situation, the intensity of publicity campaigns decreased.* We then analyze the scenarios of technology embracers and intrinsic seekers in Fig. 6b and 6d, respectively. Individuals from these groups are easily motivated due to the lower overall cost. As a result, the determination is disproportional to the collective control effort in Fig. 7b in Stage 1, which represents a cost-effective strategy for policymakers. The rewards effectively motivate the intrinsic seekers as the reward grows and do not cause any significant disturbance in Fig. 6d and 7d. Based on the results in Fig. 6b and 6d, it's worth noticing that the policymakers' determination demonstrates a similar evolution, where they show resistance to implementing high rewards, while individuals still exert fair control efforts. *This reflects a scenario where individuals behave cautiously and rationally based on the pandemic situation with only moderate government guidance.* Overall, the determination can counterbalance the collective control effort. When the collective control effort is high at a particular reward state, policymakers decrease the determination associated with that reward state.

E. Determination and epidemic progression

The proposed model can also demonstrate the relationship between epidemic progression and intervention implementation. *Policymakers exert strong determination when the low-risk group shrinks or a dominated high-risk group increases.* For instance, there is a decrease in the low-risk group as shown in Fig. 3 (right) at $t = 0.4 - 0.8$, policymakers increase the determination in Fig. 8a (right, $g = 20$). Similarly, when the dominated high-risk group increases in Fig. 3 (left) at $t = 0.2 - 0.6$, the determination increases drastically in Fig. 8a (left, $g = 0$). *Policymakers show higher determination to implement interventions that protect the low-risk groups from being infected. Simultaneously, they closely monitor the dynamics of the high-risk group to adjust the interventions accordingly, such as New Zealand [23] implemented lockdown when the confirmed cases significantly increased.* Policymakers demonstrate stronger determination with more drastic changes when allocating generous rewards. For instance, in Fig. 8a (right, $g = 20$), when $t = 0.2 - 0.4$, due to the decrease in the low-risk group in Fig. 3 (right), policymakers increase the determination drastically. Additionally, there is a further increase in determination as the low-risk group continues to shrink at $t = 0.4 - 0.6$. We can observe a periodic dynamic in these three groups in Fig. 3 (right). *With resolute determination and high rewards, policymakers implement a range of regulations, resulting in heightened control efforts by individuals. This, in turn, triggers a rapid evolution of the epidemic progression, manifesting as a periodic effect, which potentially speeds up the epidemic progression.* This dynamic mirrors the situation during COVID-19, where policymakers' strong determination gave rise to periodic fluctuations in the number of cases.

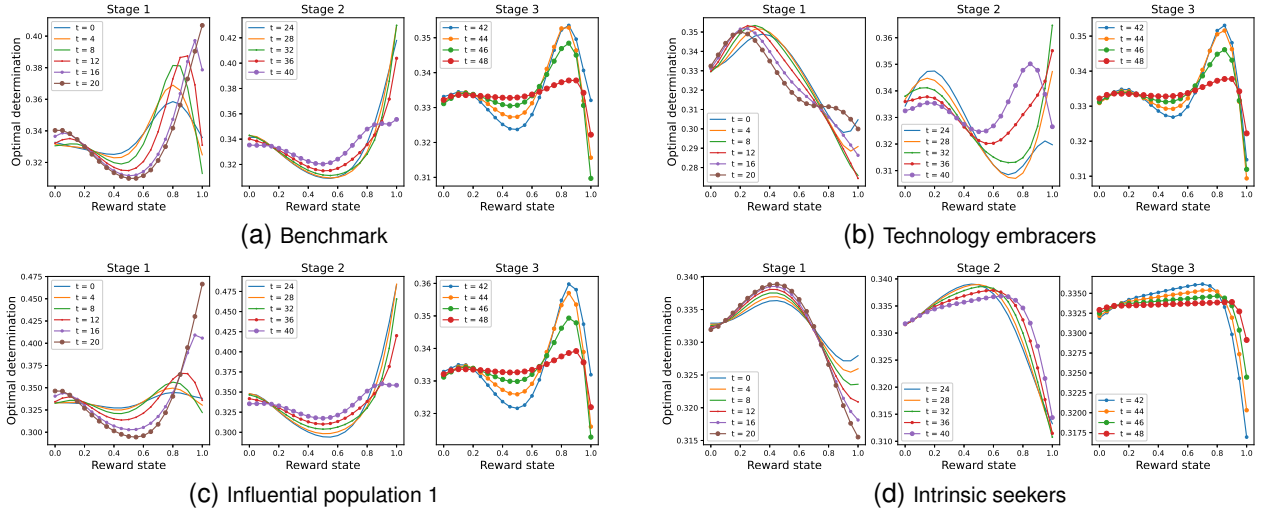

Figure 6. Optimal determination evolution with respect to reward state: Stage 1 is from $t = 0$ to $t = 20$; Stage 2 is from $t = 21$ to $t = 40$; Stage 3 is from $t = 41$ to $t = 48$. The optimal determination evolution is depicted by solid lines with dots, where the size of the dots increases with time steps. Policymakers' optimal determination demonstrates a similar evolution based on the individuals' overall costs. They show a willingness to implement high rewards when individuals can be effectively motivated.

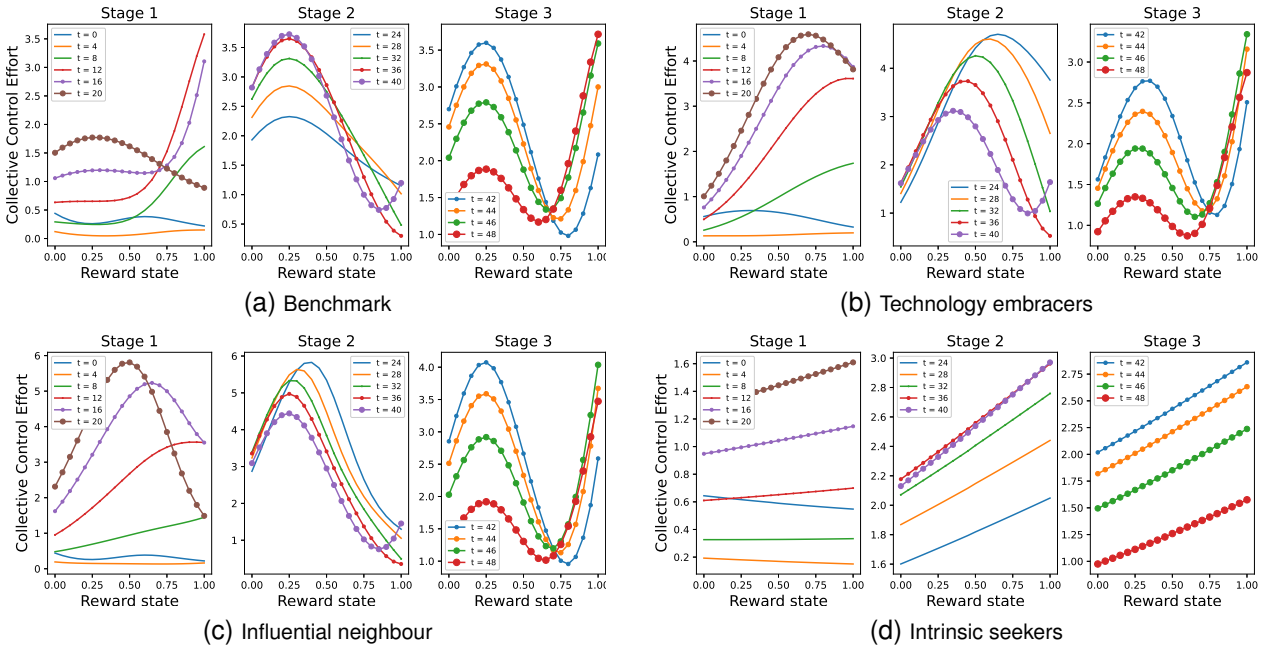

Figure 7. Collective control effort with respect to rewards: It demonstrates the key factor that influences the determination of policymakers.

F. Equilibrium and stability

To evaluate the equilibrium and stability of the model, we use entropy. According to the Second Law of Thermodynamics, an isolated system will naturally evolve to reach maximum entropy, which corresponds to thermodynamic equilibrium [24]. Furthermore, maximum entropy indicates system stability and can be represented by an entropy equation $E = -\kappa \sum_{i=1}^{N} p_i ln(p_i)$, where κ is the Boltzmann constant, and p_i represents the probability of the system being in the i-th micro-state. Here, we consider the individuals' health states as the micro-states of society. The strategies of policymakers can be viewed as external heat applied to the system. This external influence affects the behavior of individuals and drives them to adapt their strategies accordingly. Consequently, the changing dynamics of individual actions impact the overall entropy of society. In Fig. 9 (left, $g = 20$), we observe three stationary points, namely at $t = 0.4$, $t = 0.8$, and $t = 1$. Prior to reaching the first equilibrium at $t = 0.4$, the optimal determination increases at $t = 0.2 - 0.4$, as shown in Fig. 8a (right, $g = 20$). As the entropy starts to decrease from $t = 0.4$, indicating system instability, policymakers respond by exerting a higher level of determination to address the instability at $t = 0.4 - 0.6$. To maintain stability and promote entropy, policymakers seek more information on micro-states, which consequently leads to a further increase in determination at $t = 0.4 - 0.6$. It is noticeable that the dynamics of determination and entropy do not match perfectly, such as the decreasing determination at $t = 0.6 - 0.8$ with entropy reaching another peak at $t = 1$,

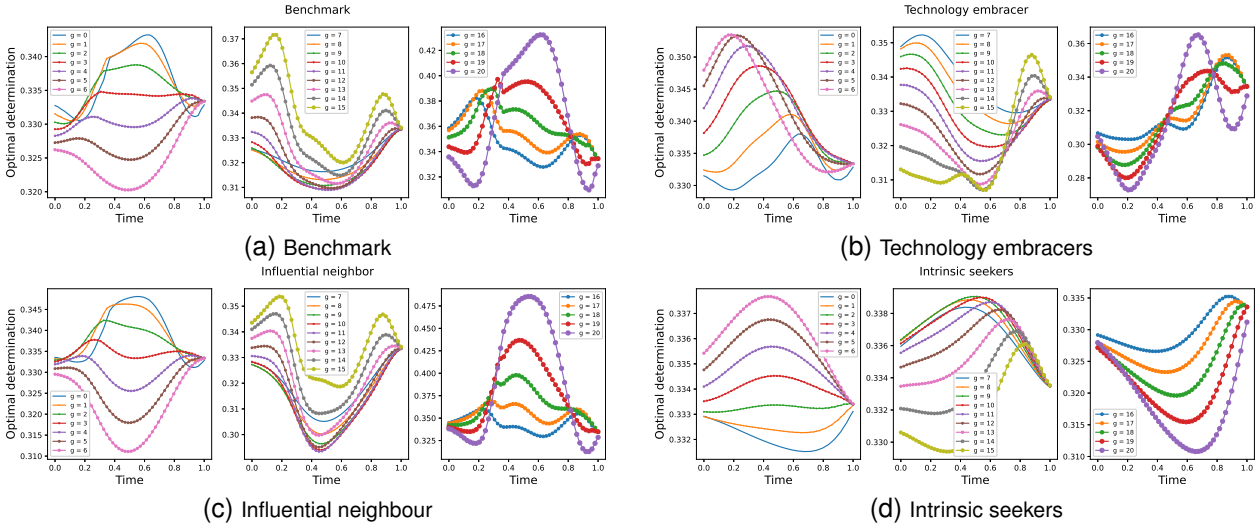

Figure 8. Optimal determination with respect to time: To clearly illustrate the evolution of the optimal determination, the reward states are divided into 3 sets, namely $g = [0, 6]$ (left), $g = [7, 15]$ (middle), and $g = [16, 20]$ (right). The size of the dots on the solid lines increases with respect to the reward states. By adjusting the optimal determination, policymakers can steer the epidemic progression and the stability of society. The optimal determination in Fig. 8b, 8c, and 8c shows similar evolution. However, the intrinsic seekers scenario demonstrates distinct evolution to achieve the same outcome.

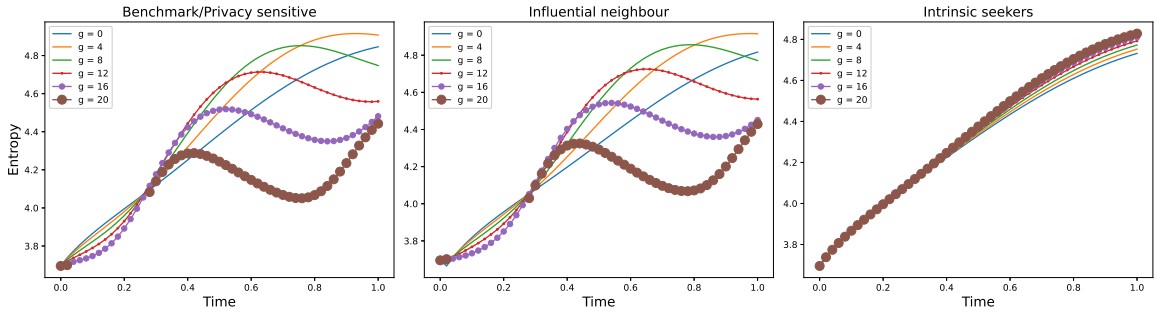

Figure 9. Entropy with respect to time: The stationary points in this figure indicate crucial points within society, with the local maxima suggesting stability and equilibrium. Policymakers adjust the optimal determination to maintain societal stability, in correspondence with the evolution of entropy. The scenario involving intrinsic seekers showcases significant differences compared to the other scenarios, driven by individual preferences and the lag effect in the optimal determination.

suggesting a potential time lag. If we consider the policymakers' determination as external heat applied to the system, this lag can be likened to a phenomenon called "thermal lag" in thermodynamics [25], This lag effect can persist for a considerable duration. In the case of low reward states (from $g = 0$ to $g = 14$), the determination initially drops for a period of time and then increases again. Higher rewards have a more significant impact on society, leading to more drastic and frequent policy changes. The technology embracers scenario mirrors the benchmark scenario in Fig. 8b (right, $g = 20$). However, the determination increases stably since individuals in this scenario contribute more information. In the influential population case, the results align with the benchmark but with slightly higher determination throughout the reward states in Fig. 8d. The increasing determination ensures that the system maintains its stability, since the interaction among populations is more difficult to infer compared to benchmark scenario. For intrinsic seekers scenario, the system entropy gradually approaches its maximum regardless of the reward. When studying the optimal determination with respect to time in Fig. 8d, we observe that with the increasing reward, the determination initially rises but then decreases as the system approaches stability. Conversely, When the reward is high, the determination drops at the beginning. To maintain the stability of the system, determination increases once again. Additionally, when the entropy reaches a peak, it indicates a well-balanced mixture of various health states. This can be a positive outcome, with the majority of the population being low/mediocre-risk individuals. In Fig. 9 (left), we observed that the peak occurs when $g = 20$ at $t = 0.4$. Referring to Fig. 3 (right), we notice a well-balanced distribution of individuals with different health states. *It is not ideal for any health state, but a low-risk group, to dominate the population. Hence, if the statistics suggest that the entropy is decreasing regardless of the reason, it is time for policymakers to enforce stricter interventions.* The ultimate goal is to reduce the number of infections and minimize the impact of the disease on the population. Public health efforts aim to lower the number of susceptible individuals (through vaccination or other means) and reduce the transmission of the disease to control the

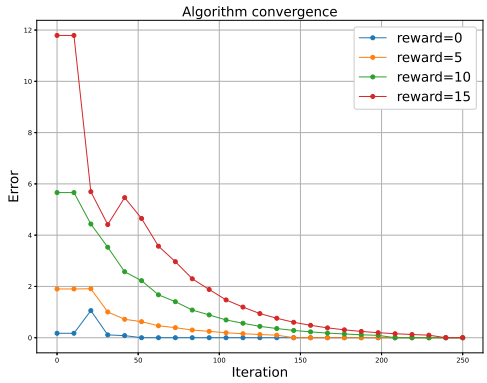

Figure 10. Convergence of the proposed algorithm: The performance and convergence of the proposed algorithm are highly related to the dimension of the problem. However, in the current setting, the proposed algorithm can converge in a reasonable number of iterations.

outbreak, which leads to an increase in determination at $t = 0.4 - 0.6$, $g = 20$ in Fig. 8a, to eliminate the susceptible. This corresponds to the decrease in entropy in Fig. 9.

G. Convergence

In Fig. 10, we show the convergence of our LFMFG model with $\rho = 0.1$. The complexity of our model, encompassing health states, population, reward, determination, control effort, and time, can pose challenges for convergence. Although machine learning methods can offer insights [26], this paper focuses on rewards' impact on decision dynamics, excluding machine learning approaches. The algorithm's convergence in Fig. 10 varies with rewards, with higher rewards causing more disturbance and requiring more iterations for accuracy. Notably, our current setting allows the algorithm to converge reasonably quickly.

VI. Conclusion

This paper introduces an LFMFG exploring the interplay of technology, individuals, and society during emergencies like pandemics. The model captures the time-dependent dynamic decision-making of individuals and policymakers. Through varied scenario analysis, we demonstrate rewards' and interventions' effectiveness, individuals' dynamic decision-making, and the link between collective efforts and epidemic progression. Future work will consider innovative reward mechanisms to enhance societal stability.

Acknowledgments

Manuel Cebrian was supported by the Ministry of Universities of the Government of Spain under the program 'Convocatoria de Ayudas para la recualificacion del sistema universitario espanol para ~ 2021-2023' from the Universidad Carlos III de Madrid, dated July 1, 2021.

References

[1] Q. Kong, M. Garcia-Herranz, I. Dotu, and M. Cebrian, "Contact tracing: Computational bounds, limitations and implications," 2021.
[2] M. Cebrian, "The past, present and future of digital contact tracing," *Nature Electronics*, vol. 4, no. 1, pp. 2–4, 2021.
[3] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese journal of mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
[4] J. Hu, M. Reed, N. Thomos, M. F. Al-Naday, and K. Yang, "A dynamic service trading in a dlt-assisted industrial iot marketplace," *IEEE Transactions on Network and Service Management*, vol. 19, no. 4, pp. 4691–4705, 2022.
[5] B. Djehiche, A. Tcheukam, and H. Tembine, "A mean-field game of evacuation in multilevel building," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5154–5169, 2017.
[6] R. A. Banez, H. Gao, L. Li, C. Yang, Z. Han, and H. V. Poor, "Belief and opinion evolution in social networks based on a multi-population mean field game approach," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.
[7] A. Lachapelle, J.-M. Lasry, C.-A. Lehalle, and P.-L. Lions, "Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis," *Mathematics and Financial Economics*, vol. 10, pp. 223–262, 2016.
[8] W. Lee, S. Liu, H. Tembine, W. Li, and S. Osher, "Controlling propagation of epidemics via mean-field control," *SIAM Journal on Applied Mathematics*, vol. 81, no. 1, pp. 190–207, 2021.
[9] B. Gaujal, J. Doncel, and N. Gast, "Vaccination in a large population: mean field equilibrium versus social optimum," in *NETGCOOP 2020-10th International Conference on NETwork Games, COntrol and OPtimization*, 2021, pp. 1–9.
[10] H. Liu and X. Tian, "Data-driven optimal control of a seir model for covid-19," *arXiv preprint arXiv:2012.00698*, 2020.
[11] E. Hubert, T. Mastrolia, D. Possamaï, and X. Warin, "Incentives, lockdown, and testing: from thucydides' analysis to the covid-19 pandemic," *Journal of mathematical biology*, vol. 84, no. 5, p. 37, 2022.
[12] R. Elie, E. Hubert, and G. Turinici, "Contact rate epidemic control of covid-19: an equilibrium view," *Mathematical Modelling of Natural Phenomena*, vol. 15, p. 35, 2020.
[13] A. Aurell, R. Carmona, G. Dayanikli, and M. Lauriere, "Optimal incentives to mitigate epidemics: a stackelberg mean field game approach," *SIAM Journal on Control and Optimization*, vol. 60, no. 2, pp. S294–S322, 2022.
[14] D. J. Haw, G. Forchini, P. Doohan, P. Christen, M. Pianella, R. Johnson, S. Bajaj, A. B. Hogan, P. Winskill, M. Miraldo *et al.*, "Optimizing social and economic activity while containing sars-cov-2 transmission using daedalus," *Nature Computational Science*, vol. 2, no. 4, pp. 223–233, 2022.
[15] J. A. Moreno López, B. Arregui García, P. Bentkowski, L. Bioglio, F. Pinotti, P.-Y. Boëlle, A. Barrat, V. Colizza, and C. Poletto, "Anatomy of digital contact tracing: Role of age, transmission setting, adoption, and case detection," *Science advances*, vol. 7, no. 15, p. eabd8750, 2021.
[16] L. Lorch, H. Kremer, W. Trouleau, S. Tsirtsis, A. Szanto, B. Schölkopf, and M. Gomez-Rodriguez, "Quantifying the effects of contact tracing, testing, and containment measures in the presence of infection hotspots," *ACM TSAS*, vol. 8, no. 4, pp. 1–28, 2022.
[17] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE constraints*. Springer Science & Business Media, 2008, vol. 23.
[18] J. Hu, K. Yang, K. Wang, and K. Zhang, "A blockchain-based reward mechanism for mobile crowdsensing," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 178–191, 2020.
[19] J. Kiusalaas, *Numerical methods in engineering with Python 3*. Cambridge university press, 2013.
[20] Y.-C. Chen, P.-E. Lu, C.-S. Chang, and T.-H. Liu, "A time-dependent sir model for covid-19 with undetectable infected persons," *Ieee transactions on network science and engineering*, vol. 7, no. 4, pp. 3279–3294, 2020.
[21] Z. Desson, E. Weller, P. McMeekin, and M. Ammi, "An analysis of the policy responses to the covid-19 pandemic in france, belgium, and canada," *Health Policy and Technology*, vol. 9, no. 4, pp. 430–446, 2020.
[22] D. Majra, J. Benson, J. Pitts, and J. Stebbing, "Sars-cov-2 (covid-19) superspreader events," *Journal of Infection*, vol. 82, no. 1, pp. 36–40, 2021.
[23] S. Jefferies, N. French, C. Gilkison, G. Graham, V. Hope, J. Marshall, C. McElnay, A. McNeill, P. Muellner, S. Paine *et al.*, "Covid-19 in new zealand and the impact of the national response: a descriptive epidemiological study," *The Lancet Public Health*, vol. 5, no. 11, pp. e612–e623, 2020.
[24] J. Rifkin, "Entropy: a new world view.[social and political implications of the second law of thermodynamics]," 1980.
[25] D. Y. Tzou, *Macro-to microscale heat transfer: the lagging behavior*. John Wiley & Sons, 2014.
[26] L. Ruthotto, S. J. Osher, W. Li, L. Nurbekyan, and S. W. Fung, "A machine learning framework for solving high-dimensional mean field game and mean field control problems," *Proceedings of the National Academy of Sciences*, vol. 117, no. 17, pp. 9183–9193, 2020.