FINITE-STATE CONTRACT THEORY WITH A PRINCIPAL AND A FIELD OF AGENTS

RENÉ CARMONA AND PEIQI WANG

ABSTRACT. We use the recently developed probabilistic analysis of mean field games with finitely many states in the weak formulation, to set-up a principal / agent contract theory model where the principal faces a large population of agents interacting in a mean field manner. We reduce the problem to the optimal control of dynamics of the McKean-Vlasov type, and we solve this problem explicitly in a special case reminiscent of the linear - quadratic mean field game models. The paper concludes with a numerical example demonstrating the power of the results when applied to a simple example of epidemic containment.

1. Introduction

In many real life situations, not only are we interested in understanding whether and how a system of interacting particles or individuals reach equilibria, but we also attempt to control or manage such equilibria so that the macroscopic behaviors of these systems reflect certain preferences. For example, how should the government control a flu outbreak by encouraging citizens to vaccinate? How should taxes be levied to influence people's consumption, saving and investment decisions? How should an employer compensate its employees in order to boost productivity? Present in all these scenarios are two parties: 1) the *principal* who devises a *contract*, according to which incentives are given to, or penalties are imposed on, 2) the *agents*. We shall follow the directives of mainstream models and assume that the agents are *rational* in the sense that they behave optimally to maximize their utilities, which translates the tradeoff between the rewards/penalties they received and the efforts they put in. The principal's problem is therefore to design a contract that maximizes the principal's own utility, which is often a function of the population's states and the transactions with the agents. To make the model even more realistic, we can consider a situation where the principal can only partially observe the agents' actions. We can also model the agents' choices of accepting or declining the contract by introducing a reservation utility.

Historically, principal / agent problems were first studied under the framework of economic contract theory, and most contributions deal with models involving a principal and a single agent. In the seminal work [17], the authors considered a discrete-time model in which the agent's effort influences the drift of the state process. By assuming that both the principal and the agent have a CARA utility function, the authors showed that the optimal contract is linear. This model was further extended in [26], [28], [21], and [16]. The breakthrough in understanding the continuous-time principal agent problem can be attributed to Yuli Sannikov (see [24] and [25]) who exploited the dynamic nature of the agent's value function stemming from the dynamic programming principle. This remarkable observation allows the principal's optimal contracting problem to be formulated as a tractable optimal control problem. This approach was further investigated and generalized in [8] and [9] with the help of second-order Backward Stochastic Differential Equations (BSDEs).

It is however a prevailing situation in real life applications that the principal faces multiple agents, and sometimes a large population of agents. Needless to say, the interactions and the competition between the agents are likely to play a key role in shaping their behaviors. This adds an extra layer of complexity since the principal not only needs to worry about the optimal responses of the individual agents, but also if an equilibrium resulting from the interactions among the agents is possible. Equilibria formed by a large population of agents with complex dynamics used to be notoriously intractable, not to mention how to solve an optimization problem on the equilibria which are parameterized by principal's contracts. To circumvent these difficulties, [10] borrowed the idea from the theory

of mean field game and considered the limit scenario of infinitely many agents. Relying on the weak formulation of stochastic differential mean field games developed in [3], the authors obtained a succinct representation of the equilibrium strategies of individual agents as well as the evolution of the equilibrium distribution across the population. Based on this convenient description of the equilibria, the author managed to obtain a tractable formulation of the optimal contracting problem, which can be solved by the technique of McKean-Vlasov optimal control.

Harnessing the probabilistic approach to finite state mean field games developed in [5], we believe that the approach proposed in [10] can be extended to principal agent problem in finite state spaces. In this paper, we consider a game involving one principal and a population of infinitely many agents whose states belong to a finite space. Leveraging the mean field game models proposed and analyzed in [5], we use continuous-time Markov chains to model the evolutions of the agents' states, and we assume that the transition rates between different states are functions of the agents' controls and the statistical distributions of all the agents' states. At the beginning of the game, the principal chooses a continuous reward rate and a terminal reward to be paid to each single agent. They are assumed to be functions of the past history of the agents' states. Each agent then maximizes their objective function by choosing their optimal effort, accounting for the principal's promised reward and the states of all the other agents. Finally, the principal records a utility which depends upon the total reward offered to the agents as well as the states of the population of the agents. Assuming that the population of agents reaches a Nash equilibrium, the problem of the principal is to optimally choose a contract which will induce an equilibrium among the agents which achieves the maximal payoff.

Relying on the weak formulation developed in [5], the Nash equilibrium can be readily characterized by a McKean-Vlasov BSDE. This BSDE is parameterized by the principal's continuous payment stream, as well as the principal's terminal payment (as terminal condition). Using ideas from [9] and [10], we show that controlling the terminal condition of the BSDE is indeed equivalent to controlling the initial condition and the martingale term of the corresponding Stochastic Differential Equation (SDE). This allows us to transform the rather intractable formulation of the principal's optimal contracting problem into an optimal control problem of a McKean-Vlasov SDE.

We then focus on a special case where the agents are risk-neutral in the terminal reward, the transition rates among their states are linear, and their instantaneous costs are quadratic in the control variable. In this case, we show that the optimal contracting problem can be further simplified to a deterministic control problem on the flow of distribution of the agents' states, and the agents use Markovian strategies when the principal announces the optimal contract. By applying the Pontryagin maximum principal, the optimal strategies of the agents and the resulting dynamics of the states under the optimal contract can be obtained by solving a system of forward-backward Ordinary Differential Equations (ODEs).

The rest of the paper is organized as follow. We introduce the model in Section 2, where we revisit some key elements of the weak formulation of finite state mean field games introduced and analyzed in [5]. We also formulate the principal's optimal contracting problem. Our first main result is stated in Section 3, where we show the equivalence between the optimal contracting problem and a McKean-Vlasov control problem. In Section 4, we solve explicitly the linear-quadratic model, where the search for the optimal contract can be reduced to a deterministic control problem on the flow of distribution of agents' states. Finally, we illustrate the above results in this linear-quadratic model with numerical computations performed on an example of epidemic containment in Section 5.

2. Model Setup

2.1. **Notations.** Throughout the paper, whenever M is a square real matrix, we denote by by M^* its transpose and by M^+ its Moore-Penrose pseudo inverse. For a column vector X, we denote by $diag(X)$ the square diagonal matrix of which the diagonal elements are given by X. The multiplication is understood as matrix multiplication.

2.2. **Agent's State Process.** We consider a game of duration T involving a principal and a population of infinitely many agents. We assume that at any time $0 \leq t \leq T$ each agent finds itself in one of m different states. We denote by $E := \{e_1, e_2, \dots, e_m\}$ the space of these possibles states, where the e_i's are the basis vectors in $\mathbb{R}^m$. We denote by $\mathcal{S}$ the m-dimensional simplex $\mathcal{S} := \{p \in \mathbb{R}^m; \sum p_i = 1, p_i \geq 0\}$, which we identify with the space of probability measures on E.

Let Ω be the space of càdlàg (right continuous with left limits) functions from $[0,T]$ to E which are left-continuous on T. Let $\mathbf{X} = (X_t)_{0\leq t\leq T}$ be the canonical process. We interpret X_t as a representative agent's state at time t. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t\leq T}$ with $\mathcal{F}_t := \sigma(\{X_s, s \leq t\})$ the natural filtration generated by $\mathbf{X}$ and $\mathcal{F} := \mathcal{F}_T$. Next, we fix an initial distribution $p^\circ \in \mathcal{S}$ for the agents' states. On $(\Omega, \mathbb{F}, \mathcal{F})$ we consider the probability measure $\mathbb{P}$ under which the law of X_0 is p° and the canonical process $\mathbf{X}$ is a continuous-time Markov chain where the transition rate between any two different states is 1. We refer to $\mathbb{P}$ as the *reference measure*. We recall that the process $\mathbf{X}$ has the following semi-martingale representation (See [11], [6], [7]):

$$X_t = X_0 + \int_0^t Q^0 X_{s-} ds + \mathcal{M}_t, \tag{1}$$

where $\boldsymbol{\mathcal{M}} = (\mathcal{M}_t)_{0\leq t\leq T}$ is a $\mathbb{R}^m$-valued $\mathbb{P}$-martingale, and Q^0 is the square matrix with diagonal elements all equal to $-(m-1)$ and off-diagonal elements all equal to 1. Moreover, the predictable quadratic variation of the martingale $\boldsymbol{\mathcal{M}}$ under $\mathbb{P}$ is $\langle \boldsymbol{\mathcal{M}}, \boldsymbol{\mathcal{M}} \rangle_t = \int_0^t \psi_t dt$, where the matrix ψ_t is defined as

$$\psi_t := diag(Q^0 X_{t-}) - Q^0 diag(X_{t-}) - diag(X_{t-}) Q^0.$$

It is easy to verify that ψ_t is a semi-definite positive matrix, and we define the corresponding (stochastic) seminorm $\|\cdot\|_{X_{t-}}$ on $\mathbb{R}^m$ by:

$$\|z\|^2_{X_{t-}} := z^* \psi_t z. \tag{2}$$

The seminorm $\|\cdot\|_{X_{t-}}$ can be written in a more explicit way. For $i \in \{1, \dots, m\}$, let us define the seminorm $\|\cdot\|_{e_i}$ on $\mathbb{R}^m$ by :

$$\|z\|^2_{e_i} := z^* [diag(Q^0 e_i) - Q^0 diag(e_i) - diag(e_i) Q^0] z = \sum_{j\neq i} |z_j - z_i|^2. \tag{3}$$

Then it is easy to see that $\|z\|_{X_{t-}} = \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i) \|z\|_{e_i}$.

Let A be a compact subset of $\mathbb{R}^l$ in which the agents can choose their controls. Denote by $\mathbb{A}$ the collection of $\mathbb{F}$-predictable processes $\boldsymbol{\alpha} = (\alpha_t)_{0\leq t\leq T}$ taking values in A. We introduce the transition rate function $q : [0,T] \times \{1, \dots, m\}^2 \times A \times \mathcal{S} \ni (t, i, j, \alpha, p) \to q(t, i, j, \alpha, p) \in \mathbb{R}$, and we denote by $Q(t, \alpha, p)$ the transition rate matrix $[q(t, i, j, \alpha, p)]_{1\leq i,j\leq m}$. In the rest of the paper, we make the following assumption on the transition rate function q:

Assumption 2.1. *(i) For all $(t, i, \alpha, p) \in [0,T] \times \{1, \dots, m\} \times A \times \mathcal{S}$, we have $\sum_{j=1}^m q(t, i, j, \alpha, p) = 0$.*
(ii) There exist constants $C > 0$ such that for all $(t, i, j, \alpha, p) \in [0,T] \times \{1, \dots, m\}^2 \times A \times \mathcal{S}$ with $i \neq j$, we have $0 < q(t, i, j, \alpha, p) \leq C$.
(iii) There exists a constant $C > 0$ such that for all $(t, i, j) \in [0,T] \times \{1, \dots, m\}^2$, $\alpha, \alpha' \in A$, $p, p' \in \mathcal{S}$, we have:

$$|q(t, i, j, \alpha, p) - q(t, i, j, \alpha', p')| \leq C(\|\alpha - \alpha'\| + \|p - p'\|). \tag{4}$$

We use the weak formulation to specify how the states of the agents evolve. This means that we specify how the agents' controls and the states of the other agents determine the probability distribution of any given agent's state via its stochastic transition rate, rather than the dynamics of the state process sample path. In addition, we assume that the interactions between the agents are *mean field*, in the sense that each agent feels the impact of the other agents via the statistical distribution of their states. Let us fix $\mathbf{p} = (p(t))_{0\leq t\leq T}$ a flow of probability

measures in E such that $p(0) = p^\circ$. $p(t)$ represents the statistical distribution of the agents' states at time t. We also fix a control $\boldsymbol{\alpha} \in \mathbb{A}$. By Girsanov's theorem (see Theorem III.41 in [23] or Lemma 4.3 in [27]), we construct a probability measure $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$ which is equivalent to $\mathbb{P}$ and such that $\mathbf{X}$ admits the following representation:

$$X_t = X_0 + \int_{(0,t]} Q^*(s, \alpha_s, p(s)) X_{s-} dt + \mathcal{M}_t^{(\boldsymbol{\alpha},\mathbf{p})}, \tag{5}$$

where $\mathcal{M}_t^{(\boldsymbol{\alpha},\mathbf{p})}$ is a martingale under the measure $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$. The Radon-Nikodym derivative of $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$ with regard to $\mathbb{P}$ is given by $\frac{d\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}}{d\mathbb{P}} = \mathcal{E}(\mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p})})_T$, where $\mathcal{E}(\mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p})})$ is the Doléans-Dade exponential of the martingale $\mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p})}$ defined as follows:

$$dL_t^{(\boldsymbol{\alpha},\mathbf{p})} = X_{t-}^*(Q(t, \alpha_t, p(t)) - Q^0)\psi_t^+ d\mathcal{M}_t, \quad L_0^{(\boldsymbol{\alpha},\mathbf{p})} = 0. \tag{6}$$

The representation of $\mathbf{X}$ in equation (5) means that under the measure $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$, the stochastic intensity rate of jumps for the state process $\mathbf{X}$ is given by the matrix $Q(t, \alpha_t, p(t))$. In addition, since $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$ and the reference measure $\mathbb{P}$ coincide on $\mathcal{F}_0$, the distributions of X_0 under $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$ and under $\mathbb{P}$ are both p°. In particular, when $\boldsymbol{\alpha}$ is a Markov control, i.e. of the form $\alpha_t = \phi(t, X_{t-})$ for some measurable feedback function ϕ, the canonical process $\mathbf{X}$ becomes a continuous-time Markov chain under the measure $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$ for which the transition rate from state i to state j at time t is $q(t, i, j, \phi(t, i), p(t))$. See [5] for more details about the construction of the probability measure $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}$.

2.3. **Principal's Contract and Agent's Objective Function.** The principal chooses a contract consisting of a payment stream $\mathbf{r} = (r_t)_{0 \le t \le T}$ which is a $\mathbb{F}-$predictable process and a terminal payment ξ which is a $\mathcal{F}_T-$measurable $\mathbb{P}-$square integrable random variable. Let $c : [0,T] \times E \times A \times \mathcal{S} \to \mathbb{R}$ be the running cost function of the typical agent. Also, let $u : \mathbb{R} \to \mathbb{R}$ (resp. $U : \mathbb{R} \to \mathbb{R}$) be the agent's utility function with regard to the continuous payments (resp. the terminal payment). If an agent accepts the contract $(\mathbf{r}, \xi)$ and the statistical distribution of his/her state at time t is denoted $p(t)$ for $0 \le t \le T$, the agent's expected total cost, $J^{\mathbf{r},\xi}(\boldsymbol{\alpha}, \mathbf{p})$, is defined as:

$$J^{\mathbf{r},\xi}(\boldsymbol{\alpha}, \mathbf{p}) := \mathbb{E}^{\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}} \left[\int_0^T [c(t, X_t, \alpha_t, p(t)) - u(r_t)] dt - U(\xi) \right]. \tag{7}$$

2.4. **Principal's Objective Function.** The principal's objective function depends on the distribution of the agents' states and the payments made to the agents. Let $c_0 : [0,T] \times \mathcal{S} \to \mathbb{R}$ be the running cost function and $C_0 : \mathcal{S} \to \mathbb{R}$ the terminal cost function resulting from the distribution of the agents' behaviors. Given that all the agents choose $\boldsymbol{\alpha}$ as their control strategy, the distribution of the agents' states is given by $\mathbf{p} = (p(t))_{t \in [0,T]}$, and the contract offered by the principal is $(\mathbf{r}, \xi)$, the principal records the expected total cost, $J_0^{\boldsymbol{\alpha},\mathbf{p}}(\mathbf{r}, \xi)$, defined as:

$$J_0^{\boldsymbol{\alpha},\mathbf{p}}(\mathbf{r}, \xi) := \mathbb{E}^{\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}} \left[\int_0^T [c_0(t, p(t)) + r_t] dt + C_0(p(T)) + \xi \right]. \tag{8}$$

2.5. **Agents' Mean Field Nash Equilibria.** We assume that the population of infinitely many agents reach a Nash equilibrium for a given contract $(\mathbf{r}, \xi)$ proposed by the principal. We recall the following definition of Nash equilibria in the *weak formulation* introduced in [5], and adapted to the present situation. See Definition 2.8 therein.

Definition 2.2. Let $(\mathbf{r}, \xi)$ be a contract, $\hat{\mathbf{p}} : [0,T] \to \mathcal{S}$ be a measurable mapping such that $\hat{p}(0) = p^\circ$, and $\hat{\boldsymbol{\alpha}} \in \mathbb{A}$ be an admissible control for the agents. We say that the couple $(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}})$ is a Nash equilibrium for the contract $(\mathbf{r}, \xi)$, in which case we use the notation $(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}}) \in \mathcal{N}(\mathbf{r}, \xi)$, if:

(i) $\hat{\boldsymbol{\alpha}}$ minimizes the cost when the agent is committed to the contract $(\mathbf{r}, \xi)$ and the distribution of all the agents is given by the flow $\hat{\mathbf{p}}$, i.e. if:

$$\hat{\boldsymbol{\alpha}} = \arg\inf_{\boldsymbol{\alpha}\in\mathbb{A}} \mathbb{E}^{\mathbb{Q}^{(\boldsymbol{\alpha},\hat{\mathbf{p}})}}\left[\int_0^T [c(t, X_t, \alpha_t, \hat{p}(t)) - u(r_t)]dt - U(\xi)\right]. \tag{9}$$

(ii) $(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}})$ satisfies the consistency conditions:

$$\forall t \in [0, T] \qquad \hat{p}(t) = \mathbb{E}^{\mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})}}[X_t]. \tag{10}$$

Note that equation (10) is equivalent to $\hat{p}_i(t) = \mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})}[X_t = e_i]$, for all $t \in [0, T]$ and $i \in \{1, \dots, m\}$.

2.6. **Principal's Optimal Contracting Problem.** We now turn to the principal's optimal choice of the contract. This amounts to minimizing the objective function computed based on the Nash equilibria formed by the agents. Of course we need to address the existence of Nash equilibria. However, in the following formulation of the optimal contracting problem, we shall avoid the problem of existence by only considering the contracts $(\mathbf{r}, \xi)$ that result in at least one Nash equilibrium. We call such contracts *admissible contracts*, and we denote by $\mathcal{C}$ the collection of all admissible contracts. In addition, among all the possible Nash equilibria, we would like to disregard the equilibria in which the agent's expected total cost is above a given threshold κ. The motivation for imposing this additional constraint is to model the *take-it-or-leave-it* behavior of the agents in contract theory: if the agent's expected total cost exceeds a certain threshold, it should be able to turn down the contract. Summarizing the constraints mentioned above, we propose the following optimization problem for the principal:

$$V(\kappa) := \inf_{(\mathbf{r},\xi)\in\mathcal{C}} \inf_{\substack{(\boldsymbol{\alpha},\mathbf{p})\in\mathcal{N}(\mathbf{r},\xi) \\ J^{\mathbf{r},\xi}(\boldsymbol{\alpha},\mathbf{p})\le\kappa}} \mathbb{E}^{\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}}\left[\int_0^T [c_0(t, p(t)) + r_t]dt + C_0(p(T)) + \xi\right], \tag{11}$$

where we adopt the convention that the infimum over an empty set equals $+\infty$.

3. A McKean-Vlasov Control Problem

The original formulation of the principal's optimal contract is intuitive but far from tractable. In this section we provide an equivalent formulation which turns out to be an optimal control problem of McKean-Vlasov type. To this end, we rely on the probabilistic characterization of the agents' Nash equilibria developed in [5].

3.1. **Agent's Optimization Problem.** We define the Hamiltonian $H : [0, T] \times E \times \mathbb{R}^m \times A \times \mathcal{S} \times \mathbb{R} \to \mathbb{R}$ for the agent's optimization problem by:

$$H(t, x, z, \alpha, p, r) := c(t, x, \alpha, p) - r + x^*(Q(t, \alpha, p) - Q^0)z. \tag{12}$$

Recall that the canonical process $\mathbf{X}$ representing the typical agent's state takes value in the set $E = \{e_1, \dots, e_m\}$. We define the reduced Hamiltonian by $H_i(t, z, \alpha, p, R) := H(t, e_i, z, \alpha, p, R)$ for $i = 1, dots, m$. It is straightforward to check:

$$H_i(t, z, \alpha, p, r) := c(t, e_i, \alpha, p) - r + \sum_{j\neq i}(z_j - z_i)(q(t, i, j, \alpha, p) - 1). \tag{13}$$

We make the following assumption on the Hamiltonian minimizarion. Clearly, minimizers of H_i should not depend on r.

Assumption 3.1. *For all $0 \le t \le T$, $i \in \{1, \dots, m\}$, $z \in \mathbb{R}^m$ and $p \in \mathcal{S}$, the mapping $\alpha \to H_i(t, z, \alpha, p, r)$ admits a unique minimizer. We denote the minimizer by $\hat{a}_i(t, z, p)$. In addition, for all $i \in \{1, \dots, m\}$, we assume that $\hat{a}_i$ is a measurable mapping on $[0, T] \times \mathbb{R}^m \times \mathcal{S}$, and there exists a constant $C > 0$ such that for all $i \in \{1, \dots, m\}$, $z, z' \in \mathbb{R}^m$ and $p \in \mathcal{S}$:*

$$\|\hat{a}_i(t, z, p) - \hat{a}_i(t, z', p)\| \le C\|z - z'\|_{e_i} \tag{14}$$

Remark 3.2. Assumption 3.1 holds, for example, when the cost function c is strongly convex in α and the transition rate function is linear in α. This can be proved by following the arguments in the proof of Lemma 3.3, Chapter 3 in [2], although the later is stated in the context of the stochastic optimal control of SDEs, which has a slightly different form of Hamiltonian.

We denote by $\hat{H}_i$ the minimum of the reduced Hamiltonian:

$$\hat{H}_i(t,z,p,r) := \hat{H}_i(t,z,\hat{a}_i(t,z,p),p,r). \tag{15}$$

Now we define the mappings $\hat{H}$ and $\hat{a}$ by:

$$\hat{H}(t,x,z,p,r) := \sum_{i=1}^{m} \hat{H}_i(t,z,p,r)\mathbb{1}(x=e_i), \tag{16}$$

$$\hat{a}(t,x,z,p) := \sum_{i=1}^{m} \hat{a}_i(t,z,p)\mathbb{1}(x=e_i). \tag{17}$$

Under Assumption 3.1, it is clear that $\hat{a}(t,x,z,p)$ is the unique minimizer of the mapping $\alpha \to H(t,x,z,\alpha,p,r)$, and the minimum is given by $\hat{H}(t,x,z,p,r)$. In addition, from Assumption 2.1, we can show the following result on the Lipschitz property of $\hat{H}$:

Lemma 3.3. *There exists a constant $C>0$ such that for all $(\omega,t)\in\Omega\times(0,T]$, $r\in\mathbb{R}$, $p\in\mathcal{S}$ and $z,z'\in\mathbb{R}^m$, we have:*

$$|\hat{H}(t,X_{t-},z,p,r)-\hat{H}(t,X_{t-},z',p,r)| \le C\|z-z'\|_{X_{t-}} \tag{18}$$

$$|\hat{a}(t,X_{t-},z,p)-\hat{a}(t,X_{t-},z',p)| \le C\|z-z'\|_{X_{t-}}. \tag{19}$$

The total expected cost of the agents and the value function of the agent's optimization problem can be characterized by BSDEs driven by continuous-time Markov chain. We refer the readers to [6] and [7] for the general theory for this type of BSDE. Let us fix a contract $(\mathbf{r},\xi)$ and a measurable function $\mathbf{p}:[0,T]\to\mathcal{S}$. Given $\boldsymbol{\alpha}\in\mathbb{A}$, we consider the following BSDEs:

$$Y_t = -U(\xi) + \int_t^T H(s,X_{s-},Z_s,\alpha_s,p(s),u(r_t))ds - \int_t^T Z_s^* d\mathcal{M}_s. \tag{20}$$

$$Y_t = -U(\xi) + \int_t^T \hat{H}(s,X_{s-},Z_s,p(s),u(r_t))ds - \int_t^T Z_s^* d\mathcal{M}_s. \tag{21}$$

In [5], we proved the following result:

Lemma 3.4. *For each fixed contract $(\mathbf{r},\xi)$, $\boldsymbol{\alpha}\in\mathbb{A}$ and measurable mapping $\mathbf{p}:[0,T]\to\mathcal{S}$,*
(i) the BSDE (20) admits a unique solution $(\mathbf{Y},\mathbf{Z})$ and we have $J^{\mathbf{r},\xi}(\boldsymbol{\alpha},\mathbf{p})=\mathbb{E}^{\mathbb{P}}[Y_0]$.
(ii) The BSDE (21) admits a unique solution $(\mathbf{Y},\mathbf{Z})$ and we have $\inf_{\boldsymbol{\alpha}\in\mathbb{A}} J^{\mathbf{r},\xi}(\boldsymbol{\alpha},\mathbf{p}) = \mathbb{E}^{\mathbb{P}}[Y_0]$. In addition, the optimal control of the agent is $\hat{a}(t,X_{t-},Z_t,p(t))$.

3.2. **Representation of Nash Equilibria Based on BSDEs.** Now we consider the following McKean-Vlasov BSDE:

$$Y_t = -U(\xi) + \int_t^T \hat{H}(s,X_{s-},Z_s,p(s),u(r_s))ds - \int_t^T Z_s^* d\mathcal{M}_s, \tag{22}$$

$$\mathcal{E}_t = 1 + \int_0^t \mathcal{E}_{s-} X_{s-}^* (Q(s,\alpha_s,p(s)) - Q^0)\psi_s^+ d\mathcal{M}_s, \tag{23}$$

$$\alpha_t = \hat{a}(t,X_{t-},Z_t,p(t)), \tag{24}$$

$$p(t) = \mathbb{E}^{\mathbb{Q}}[X_t], \quad \frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}_T. \tag{25}$$

Definition 3.5. We say that a tuple $(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\alpha}, \mathbf{p}, \mathbb{Q})$ is a solution to the McKean-Vlasov BSDE (22)-(25) if $\mathbf{Y}$ is an adapted càdlàg process such that $\mathbb{E}^{\mathbb{P}}[\int_0^T Y_t^2] < +\infty$ for all $t \in [0,T]$, $\mathbf{Z}$ is an adapted left-continuous process such that $\mathbb{E}^{\mathbb{P}}[\int_0^T \|Z_t\|^2_{X_{t-}}] < +\infty$, $\boldsymbol{\alpha} \in \mathbb{A}$, $\mathbf{p} : [0,T] \to \mathcal{S}$ is a measurable mapping, $\mathbb{Q}$ is a probability measure on Ω and equations (22)-(25) are satisfied $\mathbb{P}$-a.s. for all $0 \le t \le T$.

The following result links the solution of the McKean-Vlasov BSDE (22)-(25) to the Nash equilibria of the agents.

Theorem 3.6. *Let Assumption 2.1 and Assumption 3.1 hold, and let $(\mathbf{r}, \xi)$ be a contract. If the BSDE (22)-(25) admits a solution $(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\alpha}, \mathbf{p}, \mathbb{Q})$, then $(\boldsymbol{\alpha}, \mathbf{p})$ is a Nash equilibrium. Conversely if $(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}})$ is a Nash equilibrium, then the BSDE (22)-(25) admits a solution $(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\alpha}, \mathbf{p}, \mathbb{Q})$ such that $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$, $d\mathbb{P} \otimes dt$-a.e. and $p(t) = \hat{p}(t)$ dt-a.e.*

Proof. Let $(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\alpha}, \mathbf{p}, \mathbb{Q})$ be a solution to the system (22)-(25). We use Lemma 3.4 to check that $\boldsymbol{\alpha}$ is the optimal control of the agent when all the agents are committed to the contract $(\mathbf{r}, \xi)$ and the distribution of their states is given by $\mathbf{p}$. Indeed, from equations (23) and (25), we see that item (ii) in Definition 2.2 is verified. Therefore $(\boldsymbol{\alpha}, \mathbf{p})$ is a Nash equilibrium.

We now show the second part of the claim. Let $(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}})$ be a Nash equilibrium and set $\hat{\mathbb{Q}} := \mathbb{Q}^{(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}})}$. By item (ii) in Definition 2.2, we see that (23) and (25) are satisfied. Therefore it only remains to show (22) and (24). Let us consider the BSDE:

$$Y_t = -U(\xi) + \int_t^T H(s, X_{s-}, Z_s, \hat{\alpha}_s, \hat{p}(s), u(r_s))ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s. \tag{26}$$

By Lemma 3.4, the above BSDE admits a unique solution which we denote by $(\mathbf{Y}^0, \mathbf{Z}^0)$. In addition, we have $\mathbb{E}^{\mathbb{P}}[Y_0^0] = J^{\mathbf{r},\xi}(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{p}})$. On the other hand, we consider the following BSDE:

$$Y_t = -U(\xi) + \int_t^T \hat{H}(s, X_{s-}, Z_s, \hat{p}(s), u(r_s))ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s. \tag{27}$$

From Lemma 3.4, we see that BSDE (27) also admits a unique solution which we denote by $(\mathbf{Y}^1, \mathbf{Z}^1)$, and we have $\mathbb{E}^{\mathbb{P}}[Y_0^1] = \inf_{\boldsymbol{\alpha} \in \mathbb{A}} J^{\mathbf{r},\xi}(\boldsymbol{\alpha}, \hat{\mathbf{p}})$. By the definition of the Nash equilibrium, we have:

$$\hat{\boldsymbol{\alpha}} \in \arg\inf_{\boldsymbol{\alpha} \in \mathbb{A}} J^{\mathbf{r},\xi}(\boldsymbol{\alpha}, \hat{p}),$$

which implies that $\mathbb{E}^{\mathbb{P}}[Y_0^1] = \mathbb{E}^{\mathbb{P}}[Y_0^0]$. Note that Y_0^1 and Y_0^0 are $\mathcal{F}_0$-measurable and $\mathbb{P}$ coincides with $\hat{\mathbb{Q}}$ on $\mathcal{F}_0$. Therefore we have $\mathbb{E}^{\hat{\mathbb{Q}}}[Y_0^1 - Y_0^0] = 0$. Now we set $\hat{\alpha}'_t := \hat{a}(t, X_{t-}, Z_t^1, \hat{p}(t))$ which minimizes the mapping $\alpha \to H(s, X_{t-}, Z_t^1, \alpha, \hat{p}(t), u(r_t))$. Since $(\mathbf{Y}^1, \mathbf{Z}^1)$ solves the BSDE (27), we have:

$$Y_t^1 = -U(\xi) + \int_t^T H(s, X_{s-}, Z_s^1, \hat{\alpha}'_t, \hat{p}(s), u(r_s))ds - \int_t^T (Z_s^1)^* \cdot d\mathcal{M}_s. \tag{28}$$

Taking the difference of the BSDEs (28) and (26), we obtain:

$$\begin{aligned}
Y_0^0 - Y_0^1 &= \int_0^T \big[H(t, X_{t-}, Z_t^0, \hat{\alpha}_t, \hat{p}(t), u(r_t)) - H(t, X_{t-}, Z_t^1, \hat{\alpha}'_t, \hat{p}(t), u(r_t))\big]dt - \int_0^T (Z_t^0 - Z_t^1)^* \cdot d\mathcal{M}_t \\
&= \int_0^T \big[c(t, X_{t-}, \hat{\alpha}_t, \hat{p}(t)) - c(t, X_{t-}, \hat{\alpha}'_t, \hat{p}(t))\big]dt \\
&\quad + \int_0^T \big[X_{t-}^*(Q(t, \hat{\alpha}_t, \hat{p}(t)) - Q^0)Z_t^0 - X_{t-}^*(Q(t, \hat{\alpha}'_t, \hat{p}(t)) - Q^0)Z_t^1\big]dt - \int_0^T (Z_t^0 - Z_t^1)^* \cdot d\mathcal{M}_t \\
&= \int_0^T \big[c(t, X_{t-}, \hat{\alpha}_t, \hat{p}(t)) - c(t, X_{t-}, \hat{\alpha}'_t, \hat{p}(t)) + X_{t-}^*(Q(t, \hat{\alpha}_t, \hat{p}(t)) - Q(t, \hat{\alpha}'_t, \hat{p}(t)))Z_t^1\big]dt
\end{aligned}$$

$$-\int_0^T (Z_t^0 - Z_t^1)^* \cdot d\mathcal{M}_t^{(\hat{\alpha},\hat{p})}.$$

Using the optimality of $\hat{\alpha}'_t$ and taking the expectation under the measure $\mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})}$, we obtain:

$$0 = \mathbb{E}^{\mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})}}[H(t, X_{t-}, Z_t^1, \hat{\alpha}_t, \hat{p}(t), u(r_t)) - H(t, X_{t-}, Z_t^1, \hat{\alpha}'_t, \hat{p}(t), u(r_t))] \geq 0. \tag{29}$$

Assume that there exists a measurable subset N of $[0,T] \times \Omega$ with strictly positve $d\mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})} \otimes dt$ measure, such that $\hat{\alpha}'_t \neq \hat{\alpha}_t$ for $(\omega, t) \in N$. By Assumption 3.1, the mapping $\alpha \to H(t, X_{t-}, Z_t^1, \alpha, \hat{p}(t), u(r_t))$ admits a unique minimizer and therefore for all $(\omega, t) \in N$, we have:

$$H(t, X_{t-}, Z_t^1, \hat{\alpha}_t, \hat{p}(t), u(r_t)) > H(t, X_{t-}, Z_t^1, \hat{\alpha}'_t, \hat{p}(t), u(r_t)).$$

Piggybacking on the argument laid out above, we see that the second inequality is strict in (29) which leads to a contradiction.Therefore we have $\hat{\alpha}'_t = \hat{\alpha}_t$, $d\mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})} \otimes dt$-a.e. Since $\mathbb{Q}^{(\hat{\boldsymbol{\alpha}},\hat{\mathbf{p}})}$ and $\mathbb{P}$ are equivalent, we have $\hat{\alpha}'_t = \hat{\alpha}_t$, $d\mathbb{P} \otimes dt$-a.e.

Comparing BSDEs (26) and (28) and using the uniqueness of the solution, we get $\mathbb{E}[\int_0^T \|Z_t^1 - Z_t^0\|^2_{X_{t-}}] = 0$. Now by the regularity of $\hat{a}$ in Assumption 3.1, we have:

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \|\hat{\alpha}_t - \hat{a}(t, X_{t-}, Z_t^0, \hat{p}(t))\|^2 dt\right] &= \mathbb{E}\left[\int_0^T \|\hat{\alpha}'_t - \hat{a}(t, X_{t-}, Z_t^0, \hat{p}(t))\|^2 dt\right] \\
&= \mathbb{E}\left[\int_0^T \|\hat{a}(t, X_{t-}, Z_t^1, \hat{p}(t)) - \hat{a}(t, X_{t-}, Z_t^0, \hat{p}(t))\|^2 dt\right] \\
&\leq C\mathbb{E}[\int_0^T \|Z_t^1 - Z_t^0\|^2_{X_{t-}}] \\
&= 0.
\end{aligned}$$

Therefore we have $\hat{\alpha}_t = \hat{a}(t, X_{t-}, Z_t^0, \hat{p}(t))$, $d\mathbb{P} \otimes dt$-a.e., which immediately implies (22) and (24). This completes the proof. □

3.3. **Principal's Optimal Contracting Problem.** We now turn to the principal's optimal choice of the contract. Recall that we have defined $\mathcal{C}$ to be the collection of all contracts that result in at least one Nash equilibrium. In addition, in order to model the fact that agents are not accepting contracts that will incur a cost above a certain (reservation) threshold, we further restrict our optimization to the collection of Nash equilibria in which the agent's expected total cost is below a given threshold κ. Therefore we propose to solve the following optimization problem for the principal:

$$V(\kappa) := \inf_{(\mathbf{r},\xi)\in\mathcal{C}} \inf_{\substack{(\boldsymbol{\alpha},\mathbf{p})\in\mathcal{N}(\mathbf{r},\xi) \\ J^{\mathbf{r},\xi}(\boldsymbol{\alpha},\mathbf{p})\leq\kappa}} \mathbb{E}^{\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p})}}\left[\int_0^T [c_0(t,p(t)) + r_t]dt + C_0(p(T)) + \xi\right],$$

where we adopt the convention that the infimum over an empty set equals $+\infty$.

Formulated in this way, the problem seems rather intractable. However, thanks to the probabilistic characterization of agents' Nash equilibria stated in Theorem 3.6, it is possible to transform it into a McKean-Vlasov control problem. Let us denote by $\mathcal{H}_X^2$ the collection of $\mathbb{F}-$adapted and left-continuous processes $\mathbf{Z}$ such that $Z_t \in \mathbb{R}^m$ for $0 \leq t \leq T$ and $\mathbb{E}[\int_0^T \|Z_t\|_{X_{t-}} dt] < +\infty$. We also denote by $\mathcal{R}$ the collection of $\mathbb{F}-$predictable process taking values in $\mathbb{R}$. Given $\mathbf{Z} \in \mathcal{H}_X^2$, $\mathbf{r} \in \mathcal{R}$ and Y_0 a $\mathcal{F}_0$-measurable random variable, we consider the following SDE of McKean-Vlasov type:

$$Y_t = Y_0 - \int_0^t \hat{H}(s, X_{s-}, Z_s, p(s), u(r_s))ds + \int_0^t Z_s^* d\mathcal{M}_s, \tag{30}$$

$$\mathcal{E}_t = 1 + \int_0^t \mathcal{E}_{s-} X^*_{s-} (Q(s, \alpha_s, p(s)) - Q^0) \psi^+_s d\mathcal{M}_s, \tag{31}$$

$$\alpha_t = \hat{a}(t, X_{t-}, Z_t, p(t)), \tag{32}$$

$$p(t) = \mathbb{E}^{\mathbb{Q}}[X_t], \quad \frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}_T. \tag{33}$$

These are exactly the same equations as in the McKean-Vlasov BSDE (22)-(25), except that we write the dynamic of $\mathbf{Y}$ in the forward direction of time. Let us denote its solution by $(\mathbf{Y}^{\mathbf{Z},br,Y_0)}, \mathbf{Z}^{(\mathbf{Z},\mathbf{r},Y_0)}, \boldsymbol{\alpha}^{(\mathbf{Z},\mathbf{r},Y_0)}, \mathbf{p}^{(\mathbf{Z},\mathbf{r},Y_0)}, \mathbb{P}^{(\mathbf{Z},\mathbf{r},Y_0)})$ and the expectation under $\mathbb{P}^{(\mathbf{Z},\mathbf{r},Y_0)}$ by $\mathbb{E}^{(\mathbf{Z},\mathbf{r},Y_0)}$ and let us consider the following optimal control problem:

$$\tilde{V}(\kappa) := \inf_{\mathbb{E}^{\mathbb{P}}[Y_0] \le \kappa} \inf_{\substack{\mathbf{Z} \in \mathcal{H}^2_X \\ \mathbf{r} \in \mathcal{R}}} \mathbb{E}^{(\mathbf{Z},\mathbf{r},Y_0)} \left[\int_0^T [c_0(t, p^{(\mathbf{Z},\mathbf{r},Y_0)}(t)) + r_t] dt + C_0(p^{(\mathbf{Z},\mathbf{r},Y_0)}(T)) + U^{-1}(-Y_T^{(\mathbf{Z},\mathbf{r},Y_0)}) \right]. \tag{34}$$

As a direct consequence of Theorem 3.6, we have the following result:

Theorem 3.7. *Let Assumption 2.1 and Assumption 3.1 hold. Then* $\tilde{V}(\kappa) = V(\kappa)$.

4. Solving the Linear-Quadratic Model

Solving the contract theory problem in the full generality considered so far seems out of reach. So in this section, we focus on a special setup of the principal agent problem where the transition rate has a linear structure and the cost function is quadratic in the control. When the agents are risk-neutral in term of the utility of their terminal reward, we show that the principal's optimal contracting problem can be further reduced to a deterministic control problem on the space of probability distributions of the agents' states.

4.1. Model Setup.

We set the initial distribution of the agents to be $p^\circ \in \mathcal{S}$. Each agent is allowed to pick a control α which belongs to the bounded interval $A := [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}^+$. We assume that the transition rate is a linear function of the control and we define:

$$q(t, i, j, \alpha, p) := \bar{q}_{i,j}(t, p) + \lambda_{i,j}(\alpha - \underline{\alpha}), \quad \text{for } i \neq j, \tag{35}$$

$$q(t, i, i, \alpha, p) := -\sum_{j \neq i} q(t, i, j, \alpha, p), \tag{36}$$

where we assume that $\lambda_{i,j} \in \mathbb{R}^+$ for all $i \neq j$, $\sum_{j \neq i} \lambda_{i,j} > 0$ for all i, and $\bar{q}_{i,j} : [0, T] \times \mathcal{S} \to \mathbb{R}^+$ are continuous mappings for all $i \neq j$. We assume that the cost function of a typical agent (not including the utility derived from the payment stream) takes the following form:

$$c(t, e_i, \alpha, p) := c_1(t, e_i, p) + \frac{\gamma_i}{2} \alpha^2, \tag{37}$$

where $\gamma_i > 0$, and the mapping $(t, p) \to c_1(t, e_i, p)$ is continuous for all $i \in \{1, \dots, m\}$. Finally we define the agent's utility function of terminal reward to be $U(\xi) = \xi$ and the utility function of continuous reward u to be a continuous, concave and increasing function.

Under the setup outlined above, it is straightforward to compute the optimizer of the Hamiltonian H defined in (12). We get:

$$\hat{a}(t, e_i, z, p) = \hat{a}(e_i, z) = b\left(-\frac{1}{\gamma_i} \sum_{j \neq i} \lambda_{i,j}(z_j - z_i) \right), \tag{38}$$

for $i \in \{1, \dots, m\}$, where we have defined $b(z) := \min\{\max\{z, \underline{\alpha}\}, \bar{\alpha}\}$.

4.2. **Reduction to the Optimal Control on Flows of Probability Measures.** We now proceed to reduce the principal's optimal contracting problem to a deterministic control problem on the flow of the probability measures corresponding to a continuous-time Markov chain. From the SDE (30) and the definition of $\hat{H}$, we have:

$$Y_T^{(\mathbf{Z},\mathbf{r},Y_0)} = Y_0 - \int_0^T [c(t, X_t, \hat{a}(X_{t-}, Z_t), p^{(\mathbf{Z},\mathbf{r},Y_0)}(t)) - u(r_t)]dt + \int_0^T Z_t^* d\mathcal{M}_t^{(\mathbf{Z},\mathbf{r},Y_0)},$$

where
$bcM^{(Z,r,Y_0)}$ is a martingale under the measure $\mathbb{P}^{(\mathbf{Z},\mathbf{r},Y_0)}$. Now using $U^{-1}(y) = y$ and injecting the SDE into the objective function of the principal defined in (11), we may rewrite the principal's optimal contracting problem:

$$\begin{aligned} V(\kappa) = &\inf_{\mathbb{E}^{\mathbb{P}}[Y_0]\le\kappa} \inf_{\substack{\mathbf{Z}\in\mathcal{H}_X^2\\ \mathbf{r}\in\mathcal{R}}} \mathbb{E}^{(\mathbf{Z},\mathbf{r},Y_0)}\bigg[\int_0^T \big[c_0(t, p^{(\mathbf{Z},\mathbf{r},Y_0)}(t)) + c(t, X_t, \hat{a}(X_{t-}, Z_t), p^{(\mathbf{Z},\mathbf{r},Y_0)}(t))\big]dt \\ &\qquad + \int_0^T \big[r_t - u(r_t)\big]dt + C_0(p^{(\mathbf{Z},\mathbf{r},Y_0)}(T)) - Y_0\bigg] \\ = &-\kappa + \inf_{\substack{\mathbf{Z}\in\mathcal{H}_X^2\\ \mathbf{r}\in\mathcal{R}}} \mathbb{E}^{(\mathbf{Z},\mathbf{r},Y_0)}\bigg[\int_0^T \big[c_0(t, p^{(\mathbf{Z},\mathbf{r},Y_0)}(t)) + c(t, X_t, \hat{a}(X_{t-}, Z_t), p^{(\mathbf{Z},\mathbf{r},Y_0)}(t))\big]dt \\ &\qquad + \int_0^T \big[r_t - u(r_t)\big]dt + C_0(p^{(\mathbf{Z},\mathbf{r},Y_0)}(T))\bigg]. \end{aligned}$$

Here we have used the equality $\mathbb{E}^{\mathbb{P}}[Y_0] = \mathbb{E}^{(\mathbf{Z},\mathbf{r},Y_0)}[Y_0]$. Notice that both the transition rate matrix Q and the optimizer $\hat{a}$ do not depend on the reward $\mathbf{r}$ or the agent's expected total cost Y_0, so we can drop the dependency of $p^{(\mathbf{Z},\mathbf{r},Y_0)}$, $\mathbb{P}^{(\mathbf{Z},\mathbf{r},Y_0)}$ and $\alpha^{(\mathbf{Z},\mathbf{r},Y_0)}$ on $\mathbf{r}$ and Y_0. We can also isolate the optimal choice of $\mathbf{r}$ from the principal's optimization problem. Indeed, let $\hat{r}$ be a minimizer of the mapping $r \to r - u(r)$. It is immediately clear that for the principal it is optimal to choose $r_t = \hat{r}$ for $t \in [0, T]$. It follows that:

$$\begin{aligned} V(\kappa) = \inf_{\mathbf{Z}\in\mathcal{H}_X^2} \mathbb{E}^{\mathbf{Z}}\bigg[\int_0^T \big[c_0(t, p^{\mathbf{Z}}(t)) + \sum_{i=1}^m \mathbb{1}(X_t = e_i)[c_1(t, e_i, p^{\mathbf{Z}}(t)) + \frac{\gamma_i}{2}\hat{a}^2(e_i, Z_t)]\big]dt \\ + C_0(p^{\mathbf{Z}}(T))\bigg] - \kappa + T(\hat{r} - u(\hat{r})). \end{aligned}$$

We now focus on:

$$W := \inf_{\mathbf{Z}\in\mathcal{H}_X^2} \mathbb{E}^{\mathbf{Z}}\bigg[\int_0^T \big[c_0(t, p^{\mathbf{Z}}(t)) + \sum_{i=1}^m \mathbb{1}(X_t = e_i)[c_1(t, e_i, p^{\mathbf{Z}}(t)) + \frac{\gamma_i}{2}\hat{a}^2(e_i, Z_t)]\big]dt + C_0(p^{\mathbf{Z}}(T))\bigg],$$

where $p^{\mathbf{Z}}(t) = \mathbb{P}^{\mathbf{Z}}[X_t]$ and under $\mathbb{P}^{\mathbf{Z}}$, $\mathbf{X}$ has the following decomposition with $\boldsymbol{\mathcal{M}}^{(Z)}$ being a $\mathbb{P}^{\mathbf{Z}}$-martingale:

$$X_t = X_0 + \int_0^t Q^*(s, \hat{a}(X_{s-}, Z_s), p^{\mathbf{Z}}(s))X_{s-}ds + \mathcal{M}_t^{\mathbf{Z}}.$$

The key observation is that the control $\mathbf{Z}$ affects the value function only through the optimal control $\alpha_t^{\mathbf{Z}} := \hat{a}(t, X_{t-}, Z_t)$ and the mapping $\mathcal{H}_X^2 \ni \mathbf{Z} \to \alpha^{\mathbf{Z}} \in \mathbb{A}$ is surjective. For any $\mathbf{Z} \in \mathcal{H}_X^2$, it is clear from the definition of $\hat{a}$ that $\alpha^{\mathbf{Z}} \in \mathbb{A}$. On the other hand, given an arbitrary $\boldsymbol{\alpha} \in \mathbb{A}$, we set the j-th component of the process $\mathbf{Z}$ to be:

$$Z_t^j := \sum_{i,i\neq j} \mathbb{1}(X_{t-} = e_i)\frac{\gamma_i\alpha_t}{(m-1)\sum_{k,k\neq i}\lambda_{i,k}}. \tag{39}$$

By the boundedness of $\boldsymbol{\alpha}$ we have $\mathbf{Z} \in \mathcal{H}^2_X$ and it is plain to verify that $\hat{a}(t, X_{t-}, Z_t) = \alpha_t$. Therefore we can transform the optimization problem W into:

$$W = \inf_{\boldsymbol{\alpha} \in \mathbb{A}} I(\boldsymbol{\alpha}), \tag{40}$$

where we define:

$$I(\alpha) := \mathbb{E}^{\boldsymbol{\alpha}}\bigg[\int_0^T \big[c_0(t, p^{\boldsymbol{\alpha}}(t)) + \sum_{i=1}^m \mathbb{1}(X_t = e_i)[c_1(t, e_i, p^{\boldsymbol{\alpha}}(t)) + \frac{\gamma_i}{2}\alpha_t^2]\big]\, dt + C_0(p^{\boldsymbol{\alpha}}(T))\bigg]. \tag{41}$$

Here, with a mild abuse of notation, we denote by $\mathbb{E}^{\boldsymbol{\alpha}}$ the expectation under $\mathbb{P}^{\boldsymbol{\alpha})}$, and $\mathbb{P}^{\boldsymbol{\alpha}}$ is the probability measure defined by the following McKean-Vlasov SDE:

$$d\mathcal{E}^{\boldsymbol{\alpha}}_t = 1 + \int_0^t \mathcal{E}_{s-} X^*_{s-}(Q(s, \alpha_s, p^{\boldsymbol{\alpha}\boldsymbol{\alpha}}(s)) - Q^0)\psi^+_s d\mathcal{M}_s, \tag{42}$$

$$\frac{d\mathbb{P}^{\boldsymbol{\alpha}}}{d\mathbb{P}} = \mathcal{E}^{\boldsymbol{\alpha}}_T, \quad p^{\boldsymbol{\alpha}}(t) = \mathbb{E}^{\boldsymbol{\alpha}}[X_t]. \tag{43}$$

Recall that under $\mathbb{P}^{\boldsymbol{\alpha}}$, $\mathbf{X}$ has the decomposition:

$$X_t = X_0 + \int_0^t Q^*(s, \alpha_s, p^{\boldsymbol{\alpha}}(s)) X_{s-} ds + \mathcal{M}^{\boldsymbol{\alpha}}_t,$$

where $\boldsymbol{\mathcal{M}}^{\boldsymbol{\alpha}}$ is a $\mathbb{P}^{\boldsymbol{\alpha}}$-martingale. Taking expectation under $\mathbb{P}^{\boldsymbol{\alpha}}$, we have:

$$\begin{aligned} p^{\boldsymbol{\alpha}}(t) &= p^{\boldsymbol{\alpha}}(0) + \int_0^t \mathbb{E}^{\boldsymbol{\alpha}}[Q^*(s, \alpha_s, p^{\boldsymbol{\alpha}}(s)) X_{s-}]\, ds \\ &= p^\circ + \int_0^t \sum_{j=1}^m p^{\boldsymbol{\alpha}}_j(s) \mathbb{E}^{\boldsymbol{\alpha}}[Q^*(s, \alpha_s, p^{\boldsymbol{\alpha}}(s)) \cdot e_j | X_{s-} = e_j]\, ds. \end{aligned}$$

Rewriting this in terms of the coordinates $p^{\boldsymbol{\alpha}}_i(t)$ of $p^{\boldsymbol{\alpha}}(t)$ and using the linear structure of Q, we have:

$$\begin{aligned} p^{\boldsymbol{\alpha}}_i(s) &= p^\circ_i + \int_0^t \sum_{j=1}^m p^{\boldsymbol{\alpha}}_j(s) \mathbb{E}^{\boldsymbol{\alpha}}[e_j^* \cdot Q(s, \alpha_s, p^{\boldsymbol{\alpha}}(s)) e_i | X_{s-} = e_j]\, ds \\ &= p^\circ_i + \int_0^t \sum_{j=1}^m p^{\boldsymbol{\alpha}}_j(s) \mathbb{E}^{\boldsymbol{\alpha}}[q(s, j, i, \alpha_s, p^{\boldsymbol{\alpha}}(s)) | X_{s-} = e_j]\, ds \\ &= p^\circ_i + \int_0^t \sum_{j, j \neq i} p^{\boldsymbol{\alpha}}_j(s) \mathbb{E}^{\boldsymbol{\alpha}}[q(s, j, i, \alpha_s, p^{\boldsymbol{\alpha}}(s)) | X_{s-} = e_j]\, ds \\ &\qquad - \int_0^t \sum_{j, j \neq i} p^{\boldsymbol{\alpha}}_i(s) \mathbb{E}^{\boldsymbol{\alpha}}[q(s, i, j, \alpha_s, p^{\boldsymbol{\alpha}}(s)) | X_{s-} = e_i]\, ds \\ &= p^\circ_i + \int_0^t \sum_{j, j \neq i} p^{\boldsymbol{\alpha}}_j(s) \left[\lambda_{j,i}(\mathbb{E}^{\boldsymbol{\alpha}}[\alpha_s | X_{s-} = e_j] - \underline{\alpha}) + \bar{q}_{j,i}(s, p^{\boldsymbol{\alpha}}(s))\right] ds \\ &\qquad - \int_0^t \sum_{j, j \neq i} p^{\boldsymbol{\alpha}}_i(s) \left[\lambda_{i,j}(\mathbb{E}^{\boldsymbol{\alpha}}[\alpha_s | X_{s-} = e_i] - \underline{\alpha}) + \bar{q}_{i,j}(s, p^{\boldsymbol{\alpha}}(s))\right] ds. \end{aligned}$$

We see that the dynamics of $p^{\boldsymbol{\alpha}}$ are completely driven by the deterministic processes $t \to \mathbb{E}^{\boldsymbol{\alpha}}[\alpha_t | X_{t-} = e_i]$ for $i \in \{1, \dots, m\}$. From now on, we denote $\tilde{\alpha}^i_t := \mathbb{E}^{\boldsymbol{\alpha}}[\alpha_t | X_{t-} = e_i]$ and $\tilde{\alpha}_t := [\tilde{\alpha}^1_t, \dots, \tilde{\alpha}^m_t]$. By Jensen's inequality, for

all $\boldsymbol{\alpha} \in \mathbb{A}$ we have:

$$\begin{aligned} I(\alpha) = & \int_0^T \Big[c_0(t, p_t^{\boldsymbol{\alpha}}) + \sum_{i=1}^m \big(c_1(t, e_i, p_t^{\boldsymbol{\alpha}}) + \frac{\gamma_i}{2} \mathbb{E}^{\boldsymbol{\alpha}}[\alpha_t^2 | X_{t-} = e_i] \big) p_i^{\boldsymbol{\alpha}}(t) \Big] dt \\ & + C_0(p^{\boldsymbol{\alpha}}(T)) \\ \geq & \int_0^T \Big[c_0(t, p_t^{\boldsymbol{\alpha}}) + \sum_{i=1}^m \big(c_1(t, e_i, p_t^{\boldsymbol{\alpha}}) + \frac{\gamma_i}{2} (\mathbb{E}^{\boldsymbol{\alpha}}[\alpha_t | X_{t-} = e_i])^2 \big) p_i^{\boldsymbol{\alpha}}(t) \Big] dt + C_0(p^{(\alpha)}(T)). \end{aligned}$$

This leads to the deterministic control problem:

$$\tilde{W} := \inf_{\tilde{\boldsymbol{\alpha}} \in \tilde{\mathbb{A}}} \tilde{I}(\tilde{\alpha}), \tag{44}$$

where we denote by $\tilde{\mathbb{A}}$ the collection of all measurable mappings from $[0,T]$ to A^m, and $\pi^{\tilde{\boldsymbol{\alpha}}}$ the solution to the following system of coupled ODEs:

$$\begin{aligned} \frac{d\pi_i^{\tilde{\boldsymbol{\alpha}}}(t)}{dt} = & \sum_{j, j \neq i} \pi_j^{\tilde{\boldsymbol{\alpha}}}(t) \Big[\lambda_{j,i}(\tilde{\alpha}_t^j - \underline{\alpha}) + \bar{q}_{j,i}(t, \pi^{\tilde{\boldsymbol{\alpha}}}(t)) \Big] - \sum_{j, j \neq i} \pi_i^{\tilde{\boldsymbol{\alpha}}}(t) \left[\lambda_{i,j}(\tilde{\alpha}_t^i - \underline{\alpha}) + \bar{q}_{i,j}(t, \pi^{\tilde{\boldsymbol{\alpha}}}(t)) \right], \\ \pi^{\tilde{\boldsymbol{\alpha}}}(0) = & \ p^\circ, \end{aligned} \tag{45}$$

and finally the objective function $\tilde{I}$ is given by:

$$\tilde{I}(\tilde{\alpha}) := \int_0^T \Big[c_0(t, \pi^{\tilde{\boldsymbol{\alpha}}}(t)) + \sum_{i=1}^m [c_1(t, e_i, \pi^{\tilde{\boldsymbol{\alpha}}}(t)) + \frac{\gamma_i}{2} (\tilde{\alpha}_t^i)^2] \pi_i^{\tilde{\boldsymbol{\alpha}}}(t) \Big] dt + C_0(\pi^{\tilde{\boldsymbol{\alpha}}}(T)). \tag{46}$$

The following results show that W and $\tilde{W}$ are two equivalent optimization problems.

Proposition 4.1. *We have $W = \tilde{W}$. If $\tilde{\boldsymbol{\alpha}}$ is a solution to the optimization problem $\tilde{W}$, then the predictable process $\boldsymbol{\alpha}$ defined by $\alpha_t = \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i) \tilde{\alpha}_t^i$ is an optimal control for W. Moreover, under the probability measure at the optimum, the agent's state evolves as a continuous-time Markov chain.*

Proof. From the derivation above, we already see that $W \geq \tilde{W}$. We now show that $\tilde{W} \geq W$. Given any $\tilde{\boldsymbol{\alpha}} \in \tilde{\mathbb{A}}$, we set:

$$\alpha_t := \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i) \tilde{\alpha}_t^i. \tag{47}$$

Clearly, $\boldsymbol{\alpha} \in \mathbb{A}$. By standard results on ordinary differential equation, equation (45) admits a unique solution, say $(\pi^{\tilde{\boldsymbol{\alpha}}}(t))_{t \in [0,T]}$. Now let us consider the probability measure $\tilde{\mathbb{P}}$ defined by:

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \tilde{\mathcal{E}}_T, \tag{48}$$

$$d\tilde{\mathcal{E}}_t = \tilde{\mathcal{E}}_{t-} X_{t-}^* (Q(t, \alpha_t, \pi^{(\tilde{\alpha})}(t)) - Q^0) \psi_t^+ d\mathcal{M}_t, \ \ \tilde{\mathcal{E}}_0 = 1. \tag{49}$$

It is easy to see that under $\tilde{\mathbb{P}}$, the canonical process $\mathbf{X}$ has the decomposition:

$$X_t = X_0 + \int_0^t Q^*(s, \alpha_s, \pi^{\tilde{\boldsymbol{\alpha}}}(s)) X_{s-} ds + \tilde{\mathcal{M}}_t = X_0 + \int_0^t \tilde{Q}^*(s) X_{s-} ds + \tilde{\mathcal{M}}_t,$$

where $\tilde{\boldsymbol{\mathcal{M}}} = (\tilde{\mathcal{M}}_t)_{0\le t\le T}$ is a martingale and $\tilde{Q}(t)$ is the transition rate matrix with components $\tilde{Q}_{ij}(t) = q(t,i,j,\tilde{\alpha}^i_t,\pi^{\tilde{\boldsymbol{\alpha}}}(t))$. This implies that $\mathbf{X}$ is a continuous-time Markov chain under $\tilde{\mathbb{P}}$ and it is then straightforward to write the Kolmogorov equation satisfied by the marginal laws of $\mathbf{X}$ under $\tilde{\mathbb{P}}$. Comparing this Kolmogorov equation with the ODE (45), we conclude by the uniqueness of the solutions that $\tilde{\mathbb{E}}[X_t] = \pi^{\tilde{\boldsymbol{\alpha}}}(t)$, where $\tilde{\mathbb{E}}$ stands for the expectation under $\tilde{\mathbb{P}}$. Now in light of equations (48)-(49), we conclude that $\tilde{\mathbb{P}}$ is the solution to the McKean-Vlasov SDE defined in (42)-(43) corresponding to the control $\boldsymbol{\alpha}$. It follows that $\tilde{\mathbb{P}} = \mathbb{P}^{\boldsymbol{\alpha}}$ and $\pi^{\tilde{\boldsymbol{\alpha}}}(t) = p^{\boldsymbol{\alpha}}(t)$. We can now compute $I(\boldsymbol{\alpha})$:

$$\begin{aligned}
I(\boldsymbol{\alpha}) =& \ \mathbb{E}^{\boldsymbol{\alpha}}\left[\int_0^T \Big(c_0(t,p^{\boldsymbol{\alpha}}(t)) + \sum_{i=1}^m \mathbb{1}(X_t = e_i)[c_1(t,e_i,p^{\boldsymbol{\alpha}}(t)) + \frac{\gamma_i}{2}\alpha_t^2]\Big)dt + C_0(p^{\boldsymbol{\alpha}}(T))\right] \\
=& \int_0^T \left[c_0(t,p^{\boldsymbol{\alpha}}(t)) + \sum_{i=1}^m [c_1(t,e_i,p^{\boldsymbol{\alpha}}(t)) + \frac{\gamma_i}{2}(\tilde{\alpha}^i_t)^2]\mathbb{P}^{\boldsymbol{\alpha}}[X_t = e_i]\right] dt + C_0(p^{\boldsymbol{\alpha}}(T)) \\
=& \int_0^T \left[c_0(t,\pi^{\tilde{\boldsymbol{\alpha}}}(t)) + \sum_{i=1}^m [c_1(t,e_i,\pi^{\tilde{\boldsymbol{\alpha}}}(t)) + \frac{\gamma_i}{2}(\tilde{\alpha}^i_t)^2]\pi^{\tilde{\boldsymbol{\alpha}}}_i(t)\right] dt + C_0(\pi^{\tilde{\boldsymbol{\alpha}}}(T)) \\
=& \ \tilde{I}(\tilde{\boldsymbol{\alpha}}).
\end{aligned}$$

From this, we deduce that $\tilde{I}(\tilde{\boldsymbol{\alpha}}) = I(\boldsymbol{\alpha}) \ge W$ and finally $\tilde{W} \ge W$. Therefore we have $\tilde{W} = W$. Let $\tilde{\boldsymbol{\alpha}} \in \tilde{\mathbb{A}}$ be the optimizer of $\tilde{W}$ and define $\boldsymbol{\alpha} \in \mathbb{A}$ as in (47). Then from the computations above, we see that $\tilde{W} = \tilde{I}(\tilde{\boldsymbol{\alpha}}) = I(\boldsymbol{\alpha})$. This immediately implies $I(\boldsymbol{\alpha}) = W$ and $\boldsymbol{\alpha}$ is the optimal control of W. □

4.3. **Construction of the Optimal Contract.** We continue to investigate the deterministic control problem $\tilde{W}$ as defined in equations (44)-(46). Once we have identified the optimal strategy of the agents at the equilibrium from the control problem $\tilde{W}$, we can then provide a semi-explicit construction for the optimal contract.

Lemma 4.2. *An optimal control exists for the deterministic control problem $\tilde{W}$.*

Proof. It is straightforward to verify that: (1) The space of controls is convex and compact. (2) The right-hand side of the ODE (45) is $\mathcal{C}^1$ and is linear in $\tilde{\alpha}$. (3) The running cost is $\mathcal{C}^1$ and convex in α for all $(t,\pi) \in [0,T]\times\mathcal{S}$ and the terminal cost is $\mathcal{C}^1$. This allows us to apply Theorem I.11.1 in [13] and obtain the existence of the optimal control. □

Having verified that $\tilde{W}$ admits an optimal solution, we now apply the necessary part of the Pontryagin maximum principle (see Theorem I.6.3 [13]), and derive a system of ODEs that characterizes the optimal control and the corresponding flow of probability measures. The Hamiltonian $\tilde{H}$ of the control problem $\tilde{W}$ is a mapping from $[0,T]\times\mathcal{S}\times\mathbb{R}^m\times A^m$ to $\mathbb{R}$ defined by:

$$\begin{aligned}
\tilde{H}(t,\pi,y,\alpha) :=& \sum_{i=1}^m \sum_{j,j\neq i} y_i\left[\pi_j\left(\lambda_{j,i}(\alpha_j - \underline{\alpha}) + \bar{q}_{j,i}(t,\pi)\right) - \pi_i\left(\lambda_{i,j}(\alpha_i - \underline{\alpha}) + \bar{q}_{i,j}(t,\pi)\right)\right] \\
&+ c_0(t,\pi) + \sum_{i=1}^m \pi_i\left[c_1(t,e_i,\pi) + \frac{\gamma_i}{2}(\alpha_i)^2\right].
\end{aligned}$$

It is straightforward to obtain:

$$\partial_{\alpha_i}\tilde{H}(t,\pi,y,\alpha) = \pi_i\left(\sum_{k\neq i}(y_k - y_i)\lambda_{i,k} + \gamma_i\alpha_i\right),$$

and the minimizer of $\alpha \to \tilde{H}(t,\pi,y,\alpha)$ is

$$\hat{a}_i(y) = b(-\sum_{k\neq i}\lambda_{i,k}(y_k - y_i)/\gamma_i),$$

for the function b we already defined as $b(z) := \min\{\max\{z, \underline{\alpha}\}, \bar{\alpha}\}$. By the necessary condition of Pontryagin's maximum principle, if $(\pi(t))_{0\le t\le T}$ is the flow of measures associated with the optimal control, then $(\pi(t), y(t))_{t\in[0,T]}$ is the solution to the following system of forward-backward ODEs:

$$\frac{d\pi(t)}{dt} = \partial_y \tilde{H}(t,\pi(t),y(t),\hat{a}(y(t))), \quad \pi(0) = p^\circ, \tag{50}$$

$$\frac{dy(t)}{dt} = -\partial_\pi \tilde{H}(t,\pi(t),y(t),\hat{a}(y(t))), \quad y(T) = \nabla C_0(\pi(T)). \tag{51}$$

Summarizing the above discussion, as well as the arguments which allow us to reduce the optimal contracting problem to the deterministic control problem we just solved, we provide a semi-explicit construction of the optimal contract and the optimal strategy of the agents at the equilibrium.

Theorem 4.3. *Let $(\hat{\pi}, \hat{y})$ be the solution to the system (50)-(51), and let us define the processes $\hat{\alpha} \in \mathbb{A}$ and $\hat{\mathbf{Z}} \in \mathcal{H}^2_X$ by:*

$$\hat{\alpha}_t := \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i)\hat{a}_i(\hat{y}(t)), \tag{52}$$

$$\hat{Z}_t := [Z_t^1, \dots, Z_t^m]^*, \quad \hat{Z}_t^i := \sum_{j,j\neq i}\mathbb{1}(X_{t-} = e_j)\frac{\gamma_j \hat{a}_j(\hat{y}(t))}{(m-1)\sum_{k\neq j}\lambda_{j,k}}. \tag{53}$$

Now let $y_0 \in \mathbb{R}^m$ be such that $p_0^\circ \cdot y_0 \le \kappa$ and let $\hat{r}$ be the minimizer of the mapping $r \to r - u(r)$. We then define the random variable $\hat{\xi}$ almost surely by the following Stieltjes integral:

$$\hat{\xi} := -X_0^* y_0 + \int_0^T [c(t, X_{t-}, \hat{\alpha}_t, \hat{\pi}(t)) - \hat{r} + X_{t-}^* Q(t, \hat{\alpha}_t, \hat{\pi}(t))\hat{Z}_t]dt - \int_0^T \hat{Z}_t^* dX_{t-}. \tag{54}$$

Then $(\hat{\mathbf{r}}, \hat{\xi})$ is an optimal contract. Moreover, under the optimal contract, every agent adopts the Markovian strategy where they pick the control $\hat{a}_i(\hat{y}(t))$ when in the state e_i at time t, and the flow of distributions of agents' states is given by $(\hat{\pi}(t))_{0\le t\le T}$.

5. Application to a Model of Epidemic Containment

To illustrate the inner workings of the model completely solved above, we consider an example of epidemic containment. We imagine a disease control authority that aims at containing the spread of a virus over a time period $[0,T]$ within its own jurisdiction, which consists of two cities A and B. The state of each individual is encoded by whether it is infected (denoted by I) or healthy (denoted by H), and by its location (denoted by A or B). Therefore the state space is $E = \{AI, AH, BI, BH\}$, and we use $\pi_{AI}, \pi_{AH}, \pi_{BI}, \pi_{BH}$ to denote the proportion of individuals in each of these 4 states. We assume that each individual's state evolves as a continuous-time Markov chain, and our modeling of the transition rate accounts for the following set of mechanisms regarding the contraction of the virus and the possible migration of individuals between the two cities:

(1) Within each city, the rate of contracting the virus depends on the proportion of infected individuals in the city, and accordingly we assume that the transition rate from state AH to state AI is $\theta_A^-(\frac{\pi_{AI}}{\pi_{AI}+\pi_{AH}})$, while the transition rate from state BH to state BI is $\theta_B^-(\frac{\pi_{BI}}{\pi_{BI}+\pi_{BH}})$. Here θ_A^- and θ_B^- are two increasing, positive and differentiable mappings from $[0,1]$ to $\mathbb{R}^+$. They capture the quality of health care in city A and B, respectively.

(2) Likewise, the rate of recovery in each city is a function of the proportion of healthy individuals in the city. We thus assume that the transition rate from state AI to state AH is $\theta_A^+(\frac{\pi_{AH}}{\pi_{AI}+\pi_{AH}})$ and the transition rate from state BH to state BI is $\theta_B^+(\frac{\pi_{BH}}{\pi_{BI}+\pi_{BH}})$. Similarly, θ_A^+ and θ_B^+ are two increasing, positive and differentiable mappings from $[0,1]$ to $\mathbb{R}^+$, characterizing the quality of health care in city A and B respectively.

(3) Each individual can choose a level of effort α to move to the other city. We assume that the efficacy of that effort depends on whether the individual is healthy or infected. Accordingly, we set $\nu_I\alpha$ as the transition rates between the states AI and BI, and we set $\nu_H\alpha$ as the transition rates between the states AH and BH.

(4) To model the inflow of infection, we assume that each individual's status of infection does not change when it moves between cities. This means that we set the transition rates between the state AI and BH, and the transition rates between the state AH and BI to 0.

To summarize, we define the transition rate matrix Q to be:

$$Q(t,\alpha,\pi) := \begin{bmatrix} \cdots & \theta_A^+(\frac{\pi_{AH}}{\pi_{AI}+\pi_{AH}}) & \nu_I\alpha & 0 \\ \theta_A^-(\frac{\pi_{AI}}{\pi_{AI}+\pi_{AH}}) & \cdots & 0 & \nu_H\alpha \\ \nu_I\alpha & 0 & \cdots & \theta_B^+(\frac{\pi_{BH}}{\pi_{BI}+\pi_{BH}}) \\ 0 & \nu_H\alpha & \theta_B^-(\frac{\pi_{BI}}{\pi_{BI}+\pi_{BH}}) & \cdots \end{bmatrix}. \tag{55}$$

Here the transition rate matrix is written assuming the order AI, AH, BI, BH for the states. For simplicity of the notation, we also omit the diagonal elements. Indeed, since Q is a transition rate matrix, each diagonal element equals the negative of the sum of the off-diagonal elements in the same row.

We resume the description of our model in terms of the cost functions for the individuals and the disease control authority. We assume that individuals living in the same city incur the same cost, which depends on the proportion of infected individuals in that city. On the other hand, the cost for exerting the effort to move depends on the status of infection. Using the notations of the cost function for the linear quadratic model as in (37), we define:

$$c_1(t,AI,\pi) = c_1(t,AH,\pi) := \phi_A\left(\frac{\pi_{AI}}{\pi_{AI}+\pi_{AH}}\right), \tag{56}$$

$$c_1(t,BI,\pi) = c_1(t,BH,\pi) := \phi_B\left(\frac{\pi_{BI}}{\pi_{BI}+\pi_{BH}}\right), \tag{57}$$

$$\gamma_{AI} = \gamma_{BI} := \gamma_I, \quad \gamma_{AH} = \gamma_{BH} := \gamma_H, \tag{58}$$

where ϕ_A and ϕ_B are two increasing mappings on $\mathbb{R}$. For the authority, we propose to use the following running cost and terminal cost:

$$c_0(t,\pi) = \exp(\sigma_A\pi_{AI} + \sigma_B\pi_{BI}), \tag{59}$$

$$C_0(\pi) = \sigma_P\cdot(\pi_{AI}+\pi_{AH}-\pi_A^0)^2, \tag{60}$$

where π_A^0 is the population of city A at time 0. Intuitively speaking, the above cost function tries to encapsulate a form of trade-off between the control of the epidemic and population planning. On the one hand, the authority attempts to minimize the infection rate of both cities. On the other hand, as we shall see shortly in the numerical simulation, individuals tends to move away from the city with higher infection rate and poorer health care, which might result in the overpopulation of the other city. Therefore, the authority also wishes to maintain the population of both cities at a steady level. The coefficients σ_A, σ_B and σ_P reflects the relative importance the authority attributes to each of these objectives.

It can be easily verified that the setup outlined above satisfies the assumptions of the linear quadratic model studied in Section 4. Although the forward-backward system of ODEs (50) - (51) which characterizes the optimal contract can be readily derived, for the sake of completeness, we shall give the details of the equations to be solved in the appendix.

For the purpose of illustration, we give the results of numerical simulations for a scenario in which city A has a higher quality of health care than city B. Accordingly we set:

$$\theta_A^+(q) := 0.4q, \;\; \theta_A^-(q) := 0.1q, \;\; \theta_B^+(q) := 0.2q, \;\; \theta_B^-(q) := 0.2q. \tag{61}$$

This means that it is easier to recover and harder to get infected in city A than in city B. In addition, we assume that individuals suffer a higher cost associated with the epidemic in city B than in city A, and we set the cost function of individuals in each city to be:

$$\phi_A(q) := q, \;\; \phi_B(q) := 2q. \tag{62}$$

We set the maximal possible effort of individuals to $\bar{\alpha} := 10$ and the coefficients for quadratic cost of efforts to $\gamma_I := 2.0$ and $\gamma_H = 0.5$. Finally, the parameters for the cost of the authority in equations (59) and (60) are set to:

$$\sigma_A = \sigma_B := 1, \;\; \sigma_P := 0. \tag{63}$$

Notice that the authority gives the same importance to the infection rates of city A and city B while disregarding the problem of overpopulation.

In the following, we shall visualize the effect of the disease control authority's intervention by comparing the equilibrium computed from the principal agent problem with the equilibrium from the mean field game of *anarchy*. By the term *anarchy*, we refer to the situation where the states of individuals are still governed by the same transition rate, but the individuals do not receive any rewards or penalties from the authority. More specifically, the expected total cost of each individual is given by:

$$\mathbb{E}^{\mathbb{Q}^{(\boldsymbol{\alpha},\pi)}}\left[\int_0^T c(t, X_t, \alpha_t, \pi(t))dt\right],$$

with the instantaneous cost c is given by:

$$\begin{aligned} c(t,x,\alpha,\pi) := \; & \mathbb{1}(x \in \{AI, AH\}) \cdot \phi_A\left(\frac{\pi_{AI}}{\pi_{AI}+\pi_{AH}}\right) \\ & + \mathbb{1}(x \in \{BI, BH\}) \cdot \phi_B\left(\frac{\pi_{BI}}{\pi_{BI}+\pi_{BH}}\right) \\ & + (\mathbb{1}(x \in \{AI, BI\})\gamma_I + \mathbb{1}(x \in \{AH, BH\})\gamma_H) \cdot \frac{\alpha^2}{2}. \end{aligned}$$

Following the analytical approach of finite-state mean field games introduced in [14], it is straightforward to derive the system of forward-backward ODEs characterizing the Nash equilibrium (See the system of ODEs (12)-(13) in [14]). For the sake of completeness, we display the system of ODEs in the appendix.

The effect of the authority's intervention is conspicuous in diminishing the infection rate among the entire population, as is shown in the upper panel of Figure 1. However, when we visualize the respective infection rate of each city in the lower panels of Figure 1, we do observe a surge of infection in city A as a result of the authority's intervention, although the epidemic eventually dies down. This is due to the inflow of infected individuals from city B in the early stage of the epidemic. Indeed, since city A provides better health care and maintains a lower rate of infection compared to city B, an infected individual has a better chance of recovery if it moves to city A. The cost of moving prevents individuals from seeking better health care in the scenario of anarchy, whereas individuals seem to receive subsidy to move when the authority tries to intervene. This outcome of the authority's intervention is further corroborated when we visualize an individual's optimal strategy in Figure 2. We observe that individuals have a greater propensity to move when the authority provides incentives.

Recall that in the above numerical computation, by setting $\sigma_P = 0$ in its terminal cost function (60), the authority does not attempt to maintain the balance of population between the two cities. We now investigate how the behavior of the individuals changes when the authority seeks to prevent the occurrence of overpopulation. To this end, we rerun the computation with $\sigma_P = 1.5$ and all the other parameters unchanged. In Figure 3, we

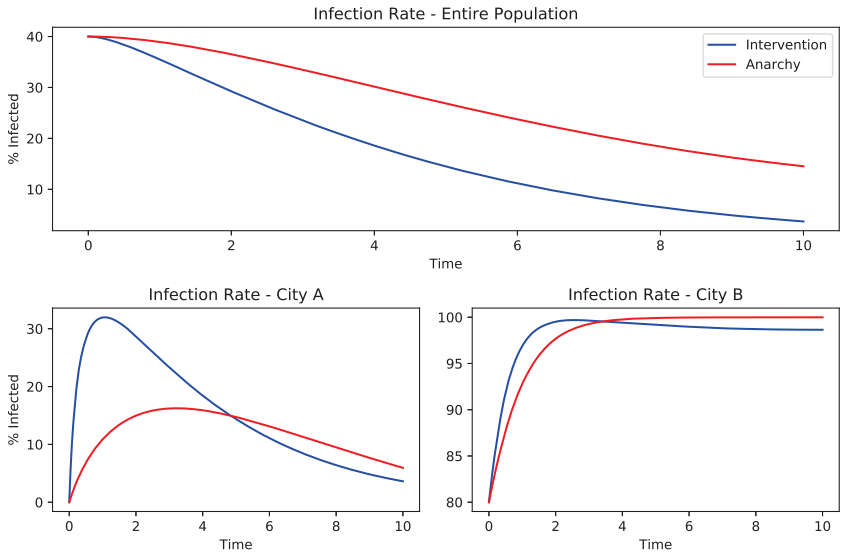

FIGURE 1. Evolution of infection rate with and without authority's intervention.

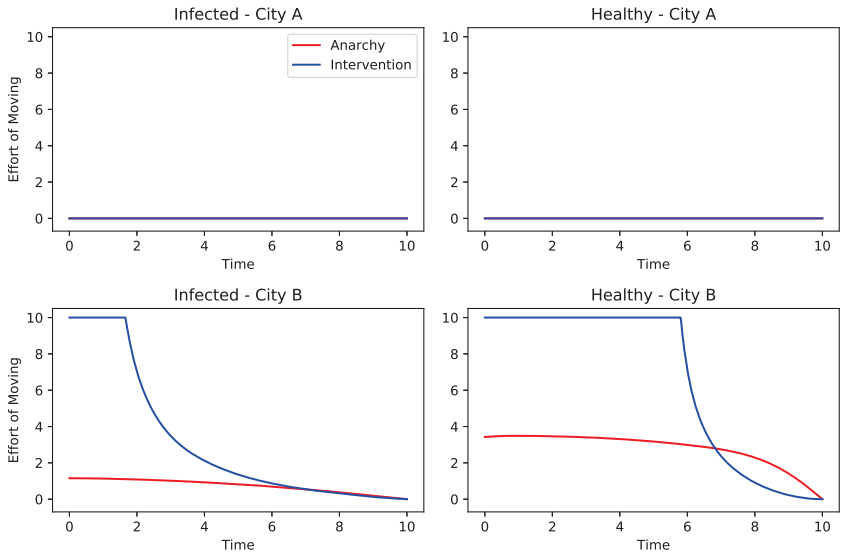

FIGURE 2. Optimal effort of moving for an individual in different states.

compare the evolution of the population in city A as well as the total infection rate with and without the population planning. When the authority does not try to control the flow of the population, the entire population ends up in city A. However, when a terminal cost related to population planning is introduced, we see a more balanced population distribution while the infection rate is still well managed. This can be explained by Figure 4, from which we have a more detailed perspective on the change of individual behavior when the authority implements the population planning. We see that healthy individuals in city A are now encouraged to move the city B, in

order to compensate the exodus caused by the epidemic in city B. On the other hand, healthy individuals in city B are now incentivized to stay in place.

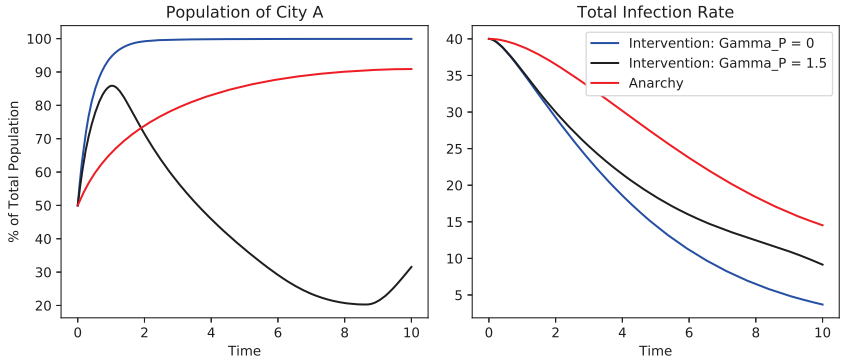

FIGURE 3. Evolution of population in city A and the total infection rate. The blue curve corresponds to intervention without population planning ($\sigma_P = 0$), the black curve corresponds to intervention with population planning ($\sigma_P = 1.5$) and the red curve corresponds to the absence of intervention.

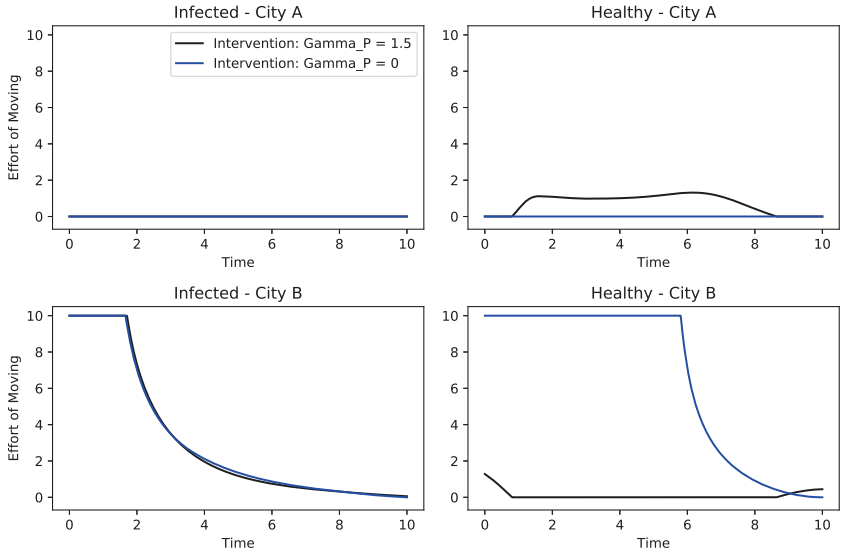

FIGURE 4. Individual's optimal effort of moving with and without population planning.

6. Appendix: System of ODEs for Epidemic Containment

6.1. **Authority's optimal planning.** Using the transition rate (55) and the cost functions (56)-(60), the system of ODEs (50)-(51) becomes:

$$\dot{y_0}(t) = -\frac{[y_0(t) - y_1(t)]\pi_1(t)}{(\pi_0(t) + \pi_1(t))^2} \cdot \left[\pi_1(t)(\theta_A^-)'\left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)}\right) + \pi_0(t)(\theta_A^+)'\left(\frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)}\right)\right]$$

$$
\begin{aligned}
&+ [y_0(t) - y_1(t)] \cdot \theta_A^+ \left(\frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) - y_2(t) \nu_I \hat{a}_0(y(t)) - \partial_{\pi_0} c_0(\pi(t)) \\
&- \frac{1}{2} \gamma_I ((\hat{a}_0(y(t)))^2 - \phi_A \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) - \phi_A' \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) \cdot \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)},
\end{aligned}
$$

$$
\begin{aligned}
\dot{y_1}(t) = &- \frac{[y_1(t) - y_0(t)]\pi_0(t)}{(\pi_0(t) + \pi_1(t))^2} \cdot \left[\pi_1(t) (\theta_A^-)' \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \pi_0(t) (\theta_A^+)' \left(\frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) \right] \\
&+ [y_1(t) - y_0(t)] \cdot \theta_A^+ \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) - y_3(t) \nu_H \hat{a}_1(y(t)) - \partial_{\pi_1} c_0(\pi(t)) \\
&- \frac{1}{2} \gamma_H ((\hat{a}_1(y(t)))^2 - \phi_A \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \phi_A' \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) \cdot \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)},
\end{aligned}
$$

$$
\begin{aligned}
\dot{y_2}(t) = &- \frac{[y_2(t) - y_3(t)]\pi_3(t)}{(\pi_2(t) + \pi_3(t))^2} \cdot \left[\pi_3(t) (\theta_B^-)' \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \pi_2(t) (\theta_B^+)' \left(\frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) \right] \\
&+ [y_2(t) - y_3(t)] \cdot \theta_B^+ \left(\frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) - y_0(t) \nu_I \hat{a}_2(y(t)) - \partial_{\pi_2} c_0(\pi(t)) \\
&- \frac{1}{2} \gamma_I ((\hat{a}_2(y(t)))^2 - \phi_B \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) - \phi_B' \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) \cdot \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)},
\end{aligned}
$$

$$
\begin{aligned}
\dot{y_3}(t) = &- \frac{[y_3(t) - y_2(t)]\pi_2(t)}{(\pi_2(t) + \pi_3(t))^2} \cdot \left[\pi_3(t) (\theta_B^-)' \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \pi_2(t) (\theta_B^+)' \left(\frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) \right] \\
&+ [y_3(t) - y_2(t)] \cdot \theta_B^+ \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) - y_1(t) \nu_H \hat{a}_3(y(t)) - \partial_{\pi_3} c_0(\pi(t)) \\
&- \frac{1}{2} \gamma_H ((\hat{a}_3(y(t)))^2 - \phi_B \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \phi_B' \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) \cdot \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)},
\end{aligned}
$$

$$
\begin{aligned}
\dot{\pi_0}(t) = &\ \pi_1(t) \theta_A^- \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) - \pi_0(t) \left[\theta_A^+ \left(\frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) + \nu_I \hat{a}_0(y(t)) \right] \\
&+ \pi_2(t) \nu_I \hat{a}_2(y(t)),
\end{aligned}
$$

$$
\begin{aligned}
\dot{\pi_1}(t) = &\ \pi_0(t) \theta_A^+ \left(\frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) - \pi_1(t) \left[\theta_A^- \left(\frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \nu_H \hat{a}_1(y(t)) \right] \\
&+ \pi_3(t) \nu_H \hat{a}_3(y(t)),
\end{aligned}
$$

$$
\begin{aligned}
\dot{\pi_2}(t) = &\ \pi_3(t) \theta_B^- \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) - \pi_2(t) \left[\theta_B^+ \left(\frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) + \nu_I \hat{a}_2(y(t)) \right] \\
&+ \pi_0(t) \nu_I \hat{a}_0(y(t)),
\end{aligned}
$$

$$
\begin{aligned}
\dot{\pi_3}(t) = &\ \pi_2(t) \theta_B^+ \left(\frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) - \pi_3(t) \left[\theta_B^- \left(\frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \nu_H \hat{a}_3(y(t)) \right] \\
&+ \pi_1(t) \nu_H \hat{a}_1(y(t)),
\end{aligned}
$$

where the optimal control is defined by:

$$\hat{a}_0(y) := b\left(\frac{\nu_I(y_0-y_2)}{\gamma_I}\right),\ \ \hat{a}_1(y) := b\left(\frac{\nu_H(y_1-y_3)}{\gamma_H}\right),$$
$$\hat{a}_2(y) := b\left(\frac{\nu_I(y_2-y_0)}{\gamma_I}\right),\ \ \hat{a}_3(y) := b\left(\frac{\nu_H(y_3-y_1)}{\gamma_H}\right),$$

and the terminal conditions are $\pi(0) = \mathbf{p}^\circ$ and $y(T) = \nabla C_0(\pi(T))$.

6.2. **Mean field equilibrium in the absence of the authority.** The system of ODEs characterizing the mean field game equilibrium consists of the Hamilton-Jacobi equation and the Kolmogorov equation.

$$\begin{aligned}\dot{v_0}(t) =\ & [v_1(t)-v_0(t)]\theta_A^+\left(\frac{\pi_1(t)}{\pi_0(t)+\pi_1(t)}\right) + [v_2(t)-v_0(t)]\nu_I\hat{a}_0(v(t))\\ & + \frac{1}{2}\gamma_I(\hat{a}_0(v(t)))^2 + \phi_A\left(\frac{\pi_0(t)}{\pi_0(t)+\pi_1(t)}\right),\end{aligned}$$

$$\begin{aligned}\dot{v_1}(t) =\ & [v_0(t)-v_1(t)]\theta_A^-\left(\frac{\pi_0(t)}{\pi_0(t)+\pi_1(t)}\right) + [v_3(t)-v_1(t)]\nu_H\hat{a}_1(v(t))\\ & + \frac{1}{2}\gamma_H(\hat{a}_1(v(t)))^2 + \phi_A\left(\frac{\pi_0(t)}{\pi_0(t)+\pi_1(t)}\right),\end{aligned}$$

$$\begin{aligned}\dot{v_2}(t) =\ & [v_3(t)-v_2(t)]\theta_B^+\left(\frac{\pi_3(t)}{\pi_2(t)+\pi_3(t)}\right) + [v_0(t)-v_2(t)]\nu_I\hat{a}_2(v(t))\\ & + \frac{1}{2}\gamma_I(\hat{a}_2(v(t)))^2 + \phi_B\left(\frac{\pi_2(t)}{\pi_2(t)+\pi_3(t)}\right),\end{aligned}$$

$$\begin{aligned}\dot{v_3}(t) =\ & [v_2(t)-v_3(t)]\theta_B^-\left(\frac{\pi_2(t)}{\pi_2(t)+\pi_3(t)}\right) + [v_1(t)-v_3(t)]\nu_H\hat{a}_3(v(t))\\ & + \frac{1}{2}\gamma_H(\hat{a}_3(v(t)))^2 + \phi_B\left(\frac{\pi_2(t)}{\pi_2(t)+\pi_3(t)}\right),\end{aligned}$$

$$\begin{aligned}\dot{\pi_0}(t) =\ & \pi_1(t)\theta_A^-\left(\frac{\pi_0(t)}{\pi_0(t)+\pi_1(t)}\right) - \pi_0(t)\left[\theta_A^+\left(\frac{\pi_1(t)}{\pi_0(t)+\pi_1(t)}\right) + \nu_I\hat{a}_0(v(t))\right]\\ & + \pi_2(t)\nu_I\hat{a}_2(v(t)),\end{aligned}$$

$$\begin{aligned}\dot{\pi_1}(t) =\ & \pi_0(t)\theta_A^+\left(\frac{\pi_1(t)}{\pi_0(t)+\pi_1(t)}\right) - \pi_1(t)\left[\theta_A^-\left(\frac{\pi_0(t)}{\pi_0(t)+\pi_1(t)}\right) + \nu_H\hat{a}_1(v(t))\right]\\ & + \pi_3(t)\nu_H\hat{a}_3(v(t)),\end{aligned}$$

$$\begin{aligned}\dot{\pi_2}(t) =\ & \pi_3(t)\theta_B^-\left(\frac{\pi_2(t)}{\pi_2(t)+\pi_3(t)}\right) - \pi_2(t)\left[\theta_B^+\left(\frac{\pi_3(t)}{\pi_2(t)+\pi_3(t)}\right) + \nu_I\hat{a}_2(v(t))\right]\\ & + \pi_0(t)\nu_I\hat{a}_0(v(t)),\end{aligned}$$

$$\dot{\pi_3}(t) = \pi_2(t)\theta_B^+\left(\frac{\pi_3(t)}{\pi_2(t)+\pi_3(t)}\right) - \pi_3(t)\left[\theta_B^-\left(\frac{\pi_2(t)}{\pi_2(t)+\pi_3(t)}\right) + \nu_H\hat{a}_3(v(t))\right] + \pi_1(t)\nu_H\hat{a}_1(v(t)),$$

where the optimal control is defined by:

$$\hat{a}_0(v) := b\left(\frac{\nu_I(v_0-v_2)}{\gamma_I}\right), \quad \hat{a}_1(v) := b\left(\frac{\nu_H(v_1-v_3)}{\gamma_H}\right),$$
$$\hat{a}_2(v) := b\left(\frac{\nu_I(v_2-v_0)}{\gamma_I}\right), \quad \hat{a}_3(v) := b\left(\frac{\nu_H(v_3-v_1)}{\gamma_H}\right),$$

and the terminal conditions are $\pi(0) = \mathbf{p}^\circ$ and $v(T) = 0$.

References

[1] A. Bensoussan, K. Sung, S. C. P. Yam, and S.-P. Yung, *Linear-quadratic mean field games*, Journal of Optimization Theory and Applications, 169 (2016), pp. 496–529.

[2] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer, 2018.

[3] R. Carmona and D. Lacker, *A probabilistic weak formulation of mean field games and applications*, The Annals of Applied Probability, 25 (2015), pp. 1189–1231.

[4] R. Carmona and P. Wang, *Finite state mean field games with major and minor players*, arXiv preprint arXiv:1610.05408, (2016).

[5] R. Carmona and P. Wang, *A probabilistic approach to extended finite state mean field games*, (2018).

[6] S. N. Cohen and R. J. Elliott, *Solutions of backward stochastic differen- tial equations on Markov chains*, Commun. Stoch. Anal., (2008), pp. 251–262.

[7] ———, *Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions*, Ann. Appl. Probab., 20 (2010), pp. 267–311.

[8] J. Cvitanić, D. Possamaï, and N. Touzi, *Dynamic programming approach to principal-agent problems*, arXiv preprint, (2015).

[9] J. Cvitanić, D. Possamaï, and N. Touzi, *Moral hazard in dynamic risk management*, Management Science, 63 (2016), pp. 3328–3346.

[10] R. Elie, T. Mastrolia, and D. Possamaï, *A tale of a principal and many many agents*, arXiv preprint arXiv:1608.05226, (2016).

[11] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, no. 29 in Applications of Mathematics, Springer, New York, 1995.

[12] S. N. Ethier and T. G. Kurtz, *Markov processes: characterization and convergence*, vol. 282, John Wiley & Sons, 2009.

[13] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, vol. 25, Springer Science & Business Media, 2006.

[14] D. A. Gomes, J. Mohr, and R. R. Souza, *Continuous time finite state mean field games*, Applied Mathematics & Optimization, 68 (2013), pp. 99–143.

[15] O. Guéant, *Existence and uniqueness result for mean field games with congestion effect on graphs*, Applied Mathematics and Optimization, 72 (2015), pp. 291–303.

[16] M. F. Hellwig and K. M. Schmidt, *Discrete–time approximations of the holmström–milgrom brownian–motion model of intertemporal incentive provision*, Econometrica, 70 (2002), pp. 2225–2264.

[17] B. Holmstrom and P. Milgrom, *Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design*, Journal of Law, Economics, & Organization, 7 (1991), pp. 24–52.

[18] V. Kolokoltsov and A. Bensoussan, *Mean-field-game model for botnet defense in cyber-security*, Applied Mathematics and Optimization, 74 (2016), pp. 669–692.

[19] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. i–le cas stationnaire*, Comptes Rendus Mathématique, 343 (2006), pp. 619–625.

[20] ———, *Jeux à champ moyen. ii–horizon fini et contrôle optimal*, Comptes Rendus Mathématique, 343 (2006), pp. 679–684.

[21] H. M. Müller, *The first-best sharing rule in the continuous-time principal-agent problem with exponential utility*, Journal of Economic Theory, 79 (1998), pp. 276–280.

[22] H. Pham and X. Wei, *Dynamic programming for optimal control of stochastic mckean–vlasov dynamics*, SIAM Journal on Control and Optimization, 55 (2017), pp. 1069–1101.

[23] P. E. Protter, *Stochastic differential equations*, in Stochastic integration and differential equations, Springer, 2005, pp. 249–361.

[24] Y. Sannikov, *A continuous-time version of the principal-agent problem*, The Review of Economic Studies, 75 (2008), pp. 957–984.

[25] ———, *Contracts: The theory of dynamic principal-agent relationships and the continuous-time approach*, in 10th World Congress of the Econometric Society, 2012.
[26] H. Schättler and J. Sung, *The first-order approach to the continuous-time principal–agent problem with exponential utility*, Journal of Economic Theory, 61 (1993), pp. 331–371.
[27] A. Sokol and N. R. Hansen, *Exponential martingales and changes of measure for counting processes*, Stochastic analysis and applications, 33 (2015), pp. 823–843.
[28] J. Sung, *Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems*, The RAND Journal of Economics, (1995), pp. 720–743.