A Mean Field Competition

Marcel Nutz[*]   Yuchong Zhang[†]

August 3, 2017

Abstract

We introduce a mean field game with rank-based reward: competing agents optimize their effort to achieve a goal, are ranked according to their completion time, and paid a reward based on their relative rank. First, we propose a tractable Poissonian model in which we can describe the optimal effort for a given reward scheme. Second, we study the principal–agent problem of designing an optimal reward scheme. A surprising, explicit design is found to minimize the time until a given fraction of the population has reached the goal.

Keywords Mean field game; rank-based reward; optimal contract; R&D competition
AMS 2010 Subject Classification 91A13; 91B40; 93E20

1 Introduction

In this paper, we introduce two game-theoretic problems. The first one is a mean field game with infinitely many players that compete to obtain a reward. The second one is a principal–agent problem where the principal interacts with these agents; namely, the principal aims to distribute a given reward budget to the different ranks such as to minimize the time until the agents complete their task.

Let us think of the agents as independent research teams trying to develop a result or product in the same field. Following the literature on dynamic research and development (R&D) detailed below, attaining a result will be modeled as a binary event. At any time t, each agent chooses a research effort λ for which a quadratic instantaneous cost $c\lambda^2$ is to be paid, where $c > 0$ is assumed to be constant for the purpose of this Introduction. In a Poissonian

1

fashion, the agents' probability of reaching the goal in a small time interval Δt is then given by $\lambda \Delta t + o(\Delta t)$; in the R&D literature, λ is sometimes interpreted as the accumulation of private knowledge and the goal is to file a patent. The agents are ranked according to their completion times and paid a reward $R(r)$ for rank r, where the reward scheme R is a given decreasing[1] function. At any time, the agents observe the fraction $\rho(t)$ of players that have already completed the task and thus which portion of R is still available; more precisely, the agents use feedback controls $\lambda(\rho(t))$. This rank-based coupling of the agents' optimization problems is a non-standard example of a mean field interaction. We shall show that given R, this game has a unique Nash equilibrium when agents optimize the expectation of reward minus cost. In fact, this setting turns out to be very tractable: Theorem 2.2 provides explicit formulas for the equilibrium optimal control λ^* and the agents' value function. These quantities are independent of the cost c which, in the body of the paper, is also allowed to depend on the state $\rho(t)$ to model that the cost may diminish as more results become available.

The second problem is built on top of the first: as we have seen that any reward scheme R leads to a unique equilibrium between the agents, we can study the problem of a manager or policy maker who would like to advance research. More precisely, we aim to minimize the time T_α^* until a given fraction α of the population has completed their task. The principal has a fixed reward budget $B = \int_0^1 R(r)\,dr$ but may choose the shape of the decreasing function R; that is, how much reward to allocate to each rank. Quite surprisingly, the principal's optimization problem has an explicit but nontrivial solution (Theorem 3.1):

$$R^*(r) = \frac{B}{C'} \left\{ \frac{1}{\sqrt{2-r}} + \frac{1}{2} \log \frac{(1+\sqrt{2-\alpha})(1-\sqrt{2-r})}{(1-\sqrt{2-\alpha})(1+\sqrt{2-r})} \right\} \mathbf{1}_{[0,\alpha]}(r),$$

where C' is a constant such that the budget constraint is saturated. As can be seen in Figure 1, this function has two main features. The first one is a discontinuity at $r = \alpha$: a substantial amount is awarded to the last few relevant agents, but it is optimal to pay zero reward to the ranks after α. While it is clearly important to incentivize the last agents that will complete the α fraction, these agents are not too discouraged by the fact that they may miss the rewarding ranks. The second feature is the shape of R^* on $[0, \alpha]$. A priori, it may not even be obvious if it is better to provide a strictly decreasing reward, compared to paying the same amount to the first α ranks. It turns out that R^* is decreasing, even if not very much so, and moreover

---

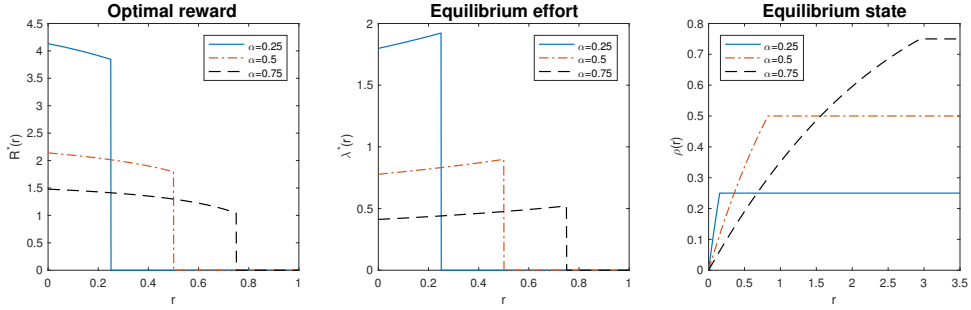[1] Decreasing and increasing are understood in the non-strict sense throughout the paper.

Figure 1: The optimal reward scheme R^* and the corresponding equilibrium effort λ^* and state process ρ for three different cut-off values α, with $B = 1$ and $c \equiv 1$.

it is concave. Thus, the difference in reward between two equidistant ranks increases later in the game, apparently to incentivize the remaining agents to choose a higher effort as shown on the second panel of Figure 1.

While formulating our game with a continuum of agents is convenient to get directly to our main results, a more fundamental justification for the mean field game is to study an N-player game for $N \to \infty$. Indeed, we establish that the N-player versions of our two problems have unique solutions, albeit somewhat less explicit than in the mean field case. We show that the N-player equilibrium converges to the mean field limit; that is, the value functions and the optimal feedback controls converge (Theorem 5.1) if the given reward schemes converge. Moreover, the optimal reward schemes for the principal and the corresponding expected completion times converge (Theorem 5.3). The analysis for finite N also allows us to study size effects which cannot be observed in the mean field limit and thus are rarely addressed for large population games. In particular, we shall observe in Section 4.3 that an increase in population size adversely affects the principal: the minimal expected completion time for a fixed target proportion α is increasing in N.

Many of our results make crucial use of the explicit or semi-explicit formulas that can be obtained in our model; in fact, much of our work has been devoted to finding a setting that is tractable. On the other hand, one may suspect that qualitative features of the results, in particular the shape of the optimal reward scheme R^*, would be similar even if the precise mechanism of the competition were different. At this time, it seems that we do not have the tools to address this question which therefore remains an interesting direction for future research.

3

1.1 Literature

Dynamic competitions (also called races) are classical in the Economics literature. An early reference related to our paper is Reinganum [30] which discusses an N-player dynamic game of R&D and patent protection. Rewards are paid at a fixed time horizon and two cases are studied: either only the first-ranked player is rewarded (perfect patent protection), or the subsequent ranks receive a positive, smaller reward; however, only the case of an identical reward for all "imitators" is considered. Malueg and Tsutsui [26] extend Reinganum's setting with a hazard rate that represents changes in the difficulty of the research project, an aspect we have incorporated differently by using a state-dependent cost that can model how the increase in public knowledge affects the project. Harris and Vickers [17] and Grossman and Shapiro [15] focus on the strategic interactions between agents in multiperiod 2-player games. A recent work in this area is Cao [6] which studies a continuous-time, continuous-state version of a model in [17]. We refer to [6, 26] for further references in this literature.

Mean field games were introduced by Lasry and Lions [22, 23, 24] and Huang, Malhamé, and Caines [19, 20] to study Nash equilibria in the limiting regime where the number of players tends to infinity and interactions take place through the empirical distribution of the private states; we refer to Guéant, Lasry and Lions [16], Bensoussan, Frehse and Yam [4] and Carmona and Delarue [8, 9] for background on mean field games. Since the impact of an individual player on the aggregate distribution is negligible, finding a Nash equilibrium reduces to solving a stochastic optimization problem for a representative player against a fixed environment, together with a consistency condition. This can be justified rigorously by formulating a game with a continuum of players and by showing that it is the limit of an N-player game. Convergence is often shown backwards; i.e., the mean field equilibrium is shown to provide an ε-Nash equilibrium for the N-player game. The forward convergence of the N-player equilibrium to the mean field equilibrium is typically more difficult to prove. For standard, diffusion-driven mean field games, this has been accomplished recently in the seminal work of Cardaliaguet, Delarue, Lasry and Lions [7]. While our game is of a different form, its tractability allows us to give an elementary yet nontrivial proof of the forward convergence; one feature in common with [7] is that we use feedback controls. In a finite-state setting, forward convergence was shown by Gomes, Mohr and Souza [14] for a small time horizon and recently by Bayraktar and Cohen [1] for an arbitrary finite time horizon.

Competitions, i.e., rank-based rewards, are a classical topic in contract

theory, dating back to the work of Lazear and Rosen [25]. With applications from school grades to sports and business competitions, rank-order prize allocation is one of the most widely used relative performance evaluation criteria; we refer to Vojnović [34] for a detailed introduction and an extensive list of references. To the best of our knowledge, the only existing work on mean field games with rank-based reward is Bayraktar and Zhang [2] where players are ranked according to their terminal positions. The main aim of that work is to obtain abstract existence results for games with common noise via translation invariance, whereas in the present work players are ranked according to exit times and the focus is on specific properties of solutions (and, of course, the principal's problem). In a different way, exit times are used in the toy example "When does the meeting start?" of [16]. Rank-based features are also studied in the literature on (uncontrolled) particle systems; one example is Shkolnikov [31]. Nadtochiy and Shkolnikov [27] consider particles interacting through hitting times. A different but related recent literature studies mean field games of timing where players directly choose stopping times; see Carmona and Lacker [10], Bertucci [5] and Nutz [28].

Continuous-time principal–agent problems with multiple agents have been studied by Koo, Shim and Sung [21] and Elie and Possamaï [13], and extended to the mean field setting by Elie, Mastrolia and Possamaï [12] and Bensoussan, Chau and Yam [3]. While these works have not considered rank-based rewards, a common feature is the Stackelberg equilibrium: the principal designs a reward scheme which the agents take as an external input to form a Nash equilibrium among themselves. By contrast, in mean field games with a major player as in Huang [18] or Carmona and Wang [11], a Nash equilibrium is formed collectively by the major and minor players. To the best of our knowledge, the convergence of the N-player principal–agent problem to the mean field limit has only been established in a simple example where the equilibrium controls are independent of N; see [12].

The remainder of this paper is structured as follows. In Section 2, we determine the unique Nash equilibrium of the mean field competition for a continuum of players with a given reward scheme. On the strength of this result, Section 3 solves the associated principal–agent problem where the principal designs the reward scheme. In Section 4 we study the corresponding N-player problems and Section 5 establishes their convergence as $N \to \infty$. Proofs are gathered in Appendix A, whereas Appendix B provides background on the Exact Law of Large Numbers used in Section 2.

5

2   Mean Field Game

Let $(I, \mathcal{I}, \mu)$ be an atomless probability space; each $i \in I$ is thought of as an agent. Moreover, let $(\Omega, \mathcal{F}, P)$ be another probability space, to be used as the sample space. Let $(Z^i)_{i \in I}$ be a family of exponential(1)-distributed random variables on Ω which is essentially pairwise independent; that is, for μ-almost all $i \in I$, Z^i is independent of Z^j for μ-almost all $j \in I$. We assume that this family is defined on an extension of the product $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \mu \otimes P)$ for which the Exact Law of Large Numbers holds, as detailed in Appendix B. Given a locally Lebesgue-integrable function $\theta : \mathbb{R} \to [0, \infty)$, we define

$$\tau_\theta^i = \inf \left\{ t : \int_0^t \theta(s)\, ds = Z^i \right\}; \tag{2.1}$$

then $(\tau_\theta^i)_{i \in I}$ are essentially pairwise independent and their distribution corresponds to the first jump time of an inhomogeneous Poisson process with intensity θ. Below, the function θ is of the form $\theta = \lambda \circ \rho$ where λ is a function chosen by the agent and ρ is a given function, and we shall find it convenient to write τ_λ^i for τ_θ^i despite the abuse of notation. If $\tau_\lambda^i \leq t$, we shall say that agent i has "arrived" by time t.

We define an *admissible (feedback) control* as a piecewise Lipschitz continuous[2] function $\lambda : [0, 1) \to \mathbb{R}_+$. The next lemma introduces the state process that emerges if all agents use the control λ.

Lemma 2.1. *Let $\lambda \in \Lambda$ be an admissible feedback control. There exists a unique continuous function $\rho : \mathbb{R}_+ \to [0, 1)$ satisfying*

$$\rho(t) = \int_0^t \lambda(\rho(s))(1 - \rho(s))\, ds, \quad t \geq 0. \tag{2.2}$$

If all agents use the feedback control λ, then $\rho(t) = \mu\{i : \tau_\lambda^i(\omega) \in [0, t]\}$ P-a.s. as well as $\rho(t) = P\{\tau_\lambda^i \in [0, t]\}$ μ-a.s.; that is, $\rho(t)$ is both the proportion of agents that have arrived by time t and the probability that any given agent i has arrived by time t.

Next, we fix a *cost coefficient* $c : [0, 1] \to (0, \infty)$ which is assumed to be Lipschitz continuous (thus, c and $1/c$ are bounded). Moreover, we fix a *reward scheme* $R : [0, 1] \to \mathbb{R}_+$ which is assumed to be decreasing, piecewise Lipschitz continuous, and left-continuous at $r = 1$. We interpret $R(r)$ as the reward paid to an agent arriving at rank r.

---

[2]That is, $[0, 1)$ is the union of finitely many intervals on which λ is Lipschitz.

Let us now consider the control problem of a given agent i before arriving, assuming that a proportion r of the population has already arrived and that all other agents transition according to a deterministic function ρ. Then, our agents' value function is

$$v(r) = \sup_{\lambda \in \Lambda} E\left[R(\rho(\tau_\lambda)) - \int_0^{\tau_\lambda} c(\rho(t))\lambda(\rho(t))^2\, dt\,\Big|\,\rho(0) = r\right], \qquad (2.3)$$

where $\tau_\lambda = \tau_\lambda^i$ is the arrival time of the given agent for the control λ. Here, we use the convention that $\rho(\infty) := 1$, meaning that agents who never arrive are paid the reward $R(1)$. If $\lambda \in \Lambda$ attains the supremum in (2.3), we say that λ is an *optimal control given* ρ.

If λ is an optimal control given the induced function ρ defined by (2.2), we say that λ is an *equilibrium optimal control* and ρ is the corresponding *equilibrium state process*. This is a Nash equilibrium: if all other players use the feedback control λ, the state evolves according to ρ by Lemma 2.1 (recall that μ is atomless), and then λ is an optimal control for our fixed player.

Theorem 2.2. *Let R be a reward scheme. Then there exists a unique (a.e.) equilibrium optimal control $\lambda^* \in \Lambda$, given by*

$$\lambda^*(r) = \frac{R(r) - \frac{1}{2\sqrt{1-r}}\int_r^1 \frac{R(y)}{\sqrt{1-y}}\, dy}{2c(r)}, \quad r \in [0,1) \qquad (2.4)$$

and the corresponding equilibrium state process ρ is determined by (2.2) with $\lambda = \lambda^*$. In equilibrium, the value function of any agent before arriving is

$$v(r) = \frac{1}{2\sqrt{1-r}}\int_r^1 \frac{R(y)}{\sqrt{1-y}}\, dy, \quad r \in [0,1). \qquad (2.5)$$

Let us remark that while the piecewise Lipschitz requirement is mainly for convenience, the continuity of R at $r = 1$ is more essential in providing existence. The following example exhibits a phenomenon that is familiar in infinite-horizon optimal stopping problems.

Example 2.3. Suppose that $c \equiv 1$ and $R = \mathbf{1}_{[0,1)}$; that is, the reward is one for agents arriving in finite time and zero for those who never arrive. Then, using the constant control $\lambda \equiv \varepsilon > 0$ yields an exponential arrival time with $E[\tau] = 1/\varepsilon$ and thus the expected reward is

$$E\left[R(\rho(\tau)) - \int_0^\tau \varepsilon^2\, dt\right] = 1 - \varepsilon.$$

7

As a result, the value function satisfies $v(r) = 1 = R(r)$ for all $r < 1$. But since $\lambda \equiv 0$ yields zero reward and any other control has positive cost, this value is not attained: there is no optimal control, and thus no equilibrium in the above sense.

Remark 2.4. We see from (2.5) that the equilibrium value function is independent of the cost coefficient c. This can also be understood directly by expressing (2.3) as an integral over the ranks, using the Kolmogorov equation (2.2) and the change-of-variable formula:

$$v(r_0) = \sup_{\lambda \in \Lambda} E\left[R(\rho(\tau_\lambda)) - \int_{\rho(0)}^{\rho(\tau_\lambda)} c(r)\lambda(r)\,\frac{dr}{1-r}\,\Big|\,\rho(0) = r_0\right].$$

Indeed, $\rho(\tau_\lambda)$ is independent of λ—when all agents use the same control, their ranking is given by the ranking of the Z^i. On the other hand, $c\lambda \in \Lambda$ if and only if $\lambda \in \Lambda$, and hence v is independent of c in equilibrium. Intuitively, a higher cost leads to a smaller optimal effort, but since this holds for all agents, the equilibrium state ρ is slowed down to the extent that the reduced effort results the same reward.

The equilibrium value function of any agent has a surprising interpretation: it can be compared to a deal where the agent pays no cost for his (constant) effort but is given the handicap of running at half the intensity of the competitors.

Proposition 2.5. *The equilibrium value function v of (2.5) coincides with the value function of an agent whose effort is fixed at $\lambda \equiv \lambda_0 \in (0, \infty)$ and is charged zero cost, while all other agents use $\lambda \equiv 2\lambda_0$:*

$v(r) = E[R(\rho(\tau))|\rho(0) = r], \quad where \ \tau \sim Exp(\lambda_0) \ and \ \rho'(t) = 2\lambda_0(1 - \rho(t)).$

In particular, $v(0) = E[R(1 - e^{-2\tau})]$ for $\tau \sim Exp(1)$.

The following result shows that the unique equilibrium of Theorem 2.2 is stable with respect to the reward scheme.

Proposition 2.6. *Let R_n, R be reward schemes such that $R_n \to R$ pointwise. Then the corresponding equilibrium optimal controls also converge pointwise, whereas the equilibrium value functions and state processes converge uniformly.*

2.1 Examples with Closed-Form Solutions

In this section, we present a family of explicitly solvable examples. Given a total reward budget $B = \int_0^1 R(r)\,dr \geq 0$, the family has a cut-off parameter $\alpha \in (0,1]$ indicating that no reward will be paid to agents ranked lower than α, as well as a shape parameter $q \geq 0$. The general form is then given by

$$R(r) = \kappa(1-r)^q \mathbf{1}_{[0,\alpha]}(r), \qquad \kappa = \frac{B(1+q)}{1-(1-\alpha)^{1+q}};$$

the constant κ is chosen such that $B = \int_0^1 R(r)\,dr$. We notice that a larger value of q indicates that a larger portion of the reward budget is paid to highly ranked player, whereas $q = 0$ corresponds to a uniform distribution of the reward among the top α ranks. For such a reward, the value function and the optimal effort of Theorem 2.2 admit closed-form solutions:

$$v(r) = \frac{\kappa}{(1+2q)} \left((1-r)^q - (1-\alpha)^q \sqrt{\frac{1-\alpha}{1-r}} \right)^+,$$

$$\lambda^*(r) = \mathbf{1}_{\{r \leq \alpha\}} \frac{\kappa}{2c(r)(1+2q)} \left(2q(1-r)^q + (1-\alpha)^q \sqrt{\frac{1-\alpha}{1-r}} \right).$$

In the boundary case $\alpha = 1$, the Lipschitz assumption of Theorem 2.2 is not satisfied when $0 < q < 1$. However, one can check the indicated formulas by direct computation in this case.

In general, the c.d.f. $F_{\tau_{\lambda^*}}(t) = \rho(t)$ of any agent's equilibrium completion time can be computed numerically by solving the Kolmogorov equation (2.2) for ρ. Inverting the equilibrium state process also gives rise to the quantile $T_\beta = \inf\{t : \rho(t) \geq \beta\}$; that is, the time until a β-proportion of the players has reached the goal. In the following special cases, these quantities can be obtained in closed form.

2.1.1 Power Reward Without Cut-off

This case corresponds to $\alpha = 1$, where we also assume that the cost c is constant. Then the above formulas specialize to

$$v(r) = \frac{B(1+q)}{1+2q}(1-r)^q, \qquad \lambda^*(r) = \frac{Bq(1+q)}{c(1+2q)}(1-r)^q,$$

and we can also solve for $F_{\tau_{\lambda^*}}(t) = \rho(t) = 1 - \left(1 + \frac{Bq^2(1+q)}{c(1+2q)}t\right)^{-\frac{1}{q}}$ and its β-quantile $T_\beta = \frac{c(1+2q)}{Bq^2(1+q)}\left[(1-\beta)^{-q} - 1\right]$. We see that the equilibrium value

9

v is decreasing in q. That is, each individual is worse off if the reward scheme heavily favors the highly-ranked players; this can be attributed to the cost caused by the large effort level λ^* in the beginning of the competition. We also observe that λ^* is decreasing in r so that agents decrease their effort once the higher ranks are filled.

2.1.2 Uniform Reward with Cut-off

This case corresponds to $q = 0$, and we again assume that the cost c is constant. The general formulas now specialize to

$$v(r) = \frac{B}{\alpha}\left(1 - \sqrt{\frac{1-\alpha}{1-r}}\right)^{+}, \qquad \lambda^*(r) = \mathbf{1}_{\{r \leq \alpha\}}\frac{B}{2c\alpha}\sqrt{\frac{1-\alpha}{1-r}}.$$

We also have $F_{\tau_{\lambda^*}}(t) = \rho(t) = 1 - \left(1 - \frac{B\sqrt{1-\alpha}}{4c\alpha}t\right)^2$ for $t \leq T_\alpha$ and $F_{\tau_{\lambda^*}}(t) = \rho(t) = \alpha$ for $t > T_\alpha$, where

$$T_\alpha = \frac{4c\alpha(1 - \sqrt{1-\alpha})}{B\sqrt{1-\alpha}}, \tag{2.6}$$

and then the general quantile is $T_\beta = \frac{4c\alpha(1-\sqrt{1-\beta})}{B\sqrt{1-\alpha}}$ for $\beta \leq \alpha$ and $T_\beta = \infty$ for $\beta > \alpha$. In contrast to the case $\alpha = 1$, we see that λ^* is increasing in r for $r \leq \alpha$: as the race progresses, the agents compete for the remaining reward and increase their effort up to the time when an α-proportion of agents has reached the goal, and then the remaining players give up; cf. Figure 2.
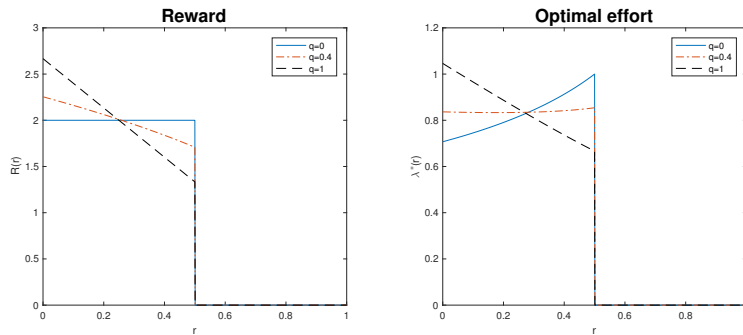

Figure 2: Optimal effort under power reward with cut-off $\alpha = 0.5$, assuming $B = 1$ and $c \equiv 1$.

2.2 Staircase Reward

Consider a reward scheme R and a cost coefficient c of the staircase form

$$R = R_1 \mathbf{1}_{[r_0, r_1]} + \sum_{j=2}^{n} R_j \mathbf{1}_{(r_{j-1}, r_j]}, \qquad c = c_1 \mathbf{1}_{[r_0, r_1]} + \sum_{i=2}^{n} c_j \mathbf{1}_{(r_{j-1}, r_j]},$$

where $R_1 \geq R_2 \geq \cdots \geq R_n \geq 0$ and $0 = r_0 < r_1 < \cdots < r_n = 1$ are constants. The formulas (2.4) and (2.5) then yield that for $r_{j-1} < r \leq r_j$,

$$v(r) = R_j + \frac{1}{\sqrt{1-r}} \left[-R_j \sqrt{1-r_j} + \sum_{k=j+1}^{n} R_k \left(\sqrt{1-r_{k-1}} - \sqrt{1-r_k} \right) \right],$$

$$\lambda^*(r) = \frac{1}{2c_j \sqrt{1-r}} \left[R_j \sqrt{1-r_j} - \sum_{k=j+1}^{n} R_k \left(\sqrt{1-r_{k-1}} - \sqrt{1-r_i} \right) \right].$$

We claim that the equilibrium state ρ is given by

$$\rho(t) = 1 - \left(\sqrt{1-r_{j-1}} - \frac{A_j}{4c_j}(t-t_{j-1}) \right)^2, \quad t_{j-1} \leq t \leq t_j, \qquad (2.7)$$

where $A_j = R_j \sqrt{1-r_j} - \sum_{k=j+1}^{n} R_k (\sqrt{1-r_{k-1}} - \sqrt{1-r_k})$ and t_j is recursively defined by $t_j = t_{j-1} + \frac{4c_j}{A_j}(\sqrt{1-r_{j-1}} - \sqrt{1-r_j})$ and $t_0 = 0$. This is to be read with the convention that $1/0 = \infty$; indeed, we have $A_j = 0$ and $t_j = \infty$ if (and only if) $R_j = R_{j+1} = \cdots = R_n$. To see (2.7), we may solve the ODE (2.2) successively for each interval $[r_{j-1}, r_j]$. Let $t_0 = 0$. Suppose we have already found $t_0, \ldots, t_{j-1}$ and $\rho(t)$ for $t \in [0, t_{j-1}]$. Then the ODE on the jth interval reads $\rho'(t) = \frac{A_j}{2c_j} \sqrt{1 - \rho(t)}$ with initial condition $\rho(t_{j-1}) = r_{j-1}$ and the solution is given by (2.7), whereas t_j is determined through the condition $\rho(t_j) = r_j$.

Finally, let $\beta \in (0, 1]$ be given. By adding β to the grid if necessary, we may assume without loss of generality that $\beta = r_{j_0}$ for some $j_0 \in \{1, \ldots, n\}$, and then the β-quantile is $T_\beta = t_{j_0} = \sum_{j=1}^{j_0} \frac{4c_j}{A_j}(\sqrt{1-r_{j-1}} - \sqrt{1-r_j})$.

3 Mean Field Principal–Agent Problem

We have seen that for a given reward scheme R, there exists a unique (deterministic) equilibrium state ρ and thus for $\alpha \in (0, 1]$, the time

$$T_\alpha(R) = \inf\{t \geq 0 : \rho(t) \geq \alpha\} \in (0, \infty]$$

is deterministic and well-defined. This is the time until an α-proportion of the population has reached the goal, or equivalently, T_α is the α-quantile of the distribution of the equilibrium arrival time τ^*.

In this section, we fix $\alpha \in (0, 1)$ and the total reward budget $B > 0$, and ask for a reward scheme R which minimizes $T_\alpha(R)$ subject to the constraint that $\int_0^1 R(r)\, dr \leq B$. This corresponds to a principal–agent problem in the second-best sense: the planner can set the reward for the agents, but cannot dictate their choice of controls. The principal considers her project completed when an α-proportion of the agents have reached their goal, and aims to find the minimal completion time

$$T_\alpha^* = \inf_{R \in \mathcal{R}: \int R(r)\, dr \leq B} T_\alpha(R), \tag{3.1}$$

where $\mathcal{R}$ is the set of all reward schemes. We remark that for $\alpha = 1$, we have $T_\alpha(R) = \infty$ for all R, whence we do not consider this case. On the other hand, $T_\alpha^* < \infty$ for all $\alpha \in (0, 1)$ as this is already accomplished by the uniform reward R with cut-off at α; cf. (2.6).

An additional assumption on the cost coefficient c is needed for our result:

$$r \mapsto \frac{c(r)(1 - r)}{2 - r} \quad \text{is decreasing.} \tag{3.2}$$

This assumption is discussed in more detail in Remark 3.2 below. The solution to the principal's problem is then given as follows.

Theorem 3.1. *Let c satisfy (3.2). Given a reward budget $B > 0$ and $\alpha \in (0, 1)$, there is an a.e. unique optimal reward scheme R^* attaining the minimal completion time T_α^* of (3.1), given by*

$$R^*(r) = \frac{B}{C} \left\{ \sqrt{\frac{c(r)}{2 - r}} + \frac{1}{2} \int_r^\alpha \frac{1}{1 - s} \sqrt{\frac{c(s)}{2 - s}}\, ds \right\} \mathbf{1}_{[0,\alpha]}(r), \tag{3.3}$$

and the minimal completion time is

$$T_\alpha^* = \frac{4C^2}{B}, \quad \text{where} \quad C = \frac{1}{2} \int_0^\alpha \frac{\sqrt{c(r)(2 - r)}}{1 - r}\, dr. \tag{3.4}$$

The corresponding equilibrium effort is

$$\lambda^*(r) = \frac{B}{2C} \frac{1}{\sqrt{(2 - r)c(r)}} \mathbf{1}_{[0,\alpha]}(r).$$

In the particular case where the cost c is constant, we have

$$R^*(r) = \frac{B}{C'} \left\{ \frac{1}{\sqrt{2-r}} + \frac{1}{2} \log \frac{(1+\sqrt{2-\alpha})(1-\sqrt{2-r})}{(1-\sqrt{2-\alpha})(1+\sqrt{2-r})} \right\} \mathbf{1}_{[0,\alpha]}(r),$$

$$T_\alpha^* = \frac{4cC'^2}{B},$$

$$C' = \frac{C}{\sqrt{c}} = \sqrt{2} - \sqrt{2-\alpha} + \frac{1}{2} \log \frac{(1+\sqrt{2-\alpha})(1-\sqrt{2})}{(1-\sqrt{2-\alpha})(1+\sqrt{2})}.$$

Figure 1 (rendered in the Introduction) shows R^*, λ^* and ρ for constant cost coefficient c. As discussed, the key features of R^* are the strict decrease and concavity on $[0, \alpha]$ and the discontinuity at α. The equilibrium effort λ^* is strictly increasing on $[0, \alpha]$. For general c, the product $\sqrt{c}\lambda^*$ is increasing, but λ^* need not be.

Remark 3.2. Assumption (3.2) is satisfied in particular if c is decreasing, which certainly holds in the applications we have in mind. If we suppose for simplicity that c is differentiable, the assumption is equivalent to the derivative c' satisfying $c'(r) \leq \frac{c(r)}{(2-r)(1-r)}$, for which a sufficient condition is that $c'(r) \leq c(r)/2$. Thus, an increase of c is permissible if it is limited relative to the level of c. When the assumption is not satisfied, (3.3) no longer describes the solution to the principal–agent problem; in fact, (3.3) is not a decreasing function and hence not a reward scheme. The proof of Theorem 3.1 shows that finding an optimal reward scheme can still be phrased as a convex optimization problem; however, the monotonicity constraint on R is now binding which is an obstruction to finding an explicit solution.

We may also ask the reverse question: given $\alpha \in (0, 1)$ and a desired completion time $T > 0$, what is the minimal budget enabling the principal to achieve T? The answer follows from Theorem 3.1 by inverting (3.4).

Corollary 3.3. *Let c satisfy* (3.2). *Given $\alpha \in (0, 1)$ and $T > 0$, the minimal budget enabling the principal to achieve a completion time $T_\alpha^* \leq T$ is $B^* = 4C^2/T$, where C is given by* (3.4).

4 N-Player Problems

In this section, we study a version of the competition with finitely many players as well as a corresponding version of the principal–agent problem. The connections between these and the mean field formulations will be established in Section 5.

4.1 The N-Player Game

We consider a game with N players, where $N \geq 1$ is a fixed integer. At any time t, each player i observes the number n of players that have already arrived, and chooses an effort level $\lambda_i(n) \in \mathbb{R}_+$. Suppose that player i uses the feedback control λ_i and all other players use a feedback control λ_{-i}. We denote by $\xi_{\lambda_i,\lambda_{-i}}(t)$ the number of players that have arrived by time t; i.e.,

$$\xi_{\lambda_i,\lambda_{-i}}(t) = \sum_{i=1}^{N} \mathbf{1}_{\{\tau^i \leq t\}},$$

where $\tau^i = \inf\{t \geq 0 : \int_0^t \lambda_i(\xi_{\lambda_i,\lambda_{-i}}(s))\,ds = Z^i\}$ is the arrival time of player i and $\tau^j = \inf\{t \geq 0 : \int_0^t \lambda_{-i}(\xi_{\lambda_i,\lambda_{-i}}(s))\,ds = Z^j\}$ is the arrival time of player $j \neq i$, for some independent exponential random variables $\{Z^1, \ldots, Z^N\}$ with unit rate. The existence of the state process is clear in this case; we may see $(\mathbf{1}_{\{\tau^i \leq t\}}, \xi_{\lambda_i,\lambda_{-i}})$ as a Markov pure jump process with values in $\{0,1\} \times \{0,1,\ldots,N\}$. We emphasize that we now use the number rather than the fraction of arrived players as the state variable.

Let $R_n \in \mathbb{R}_+$ be the reward for finishing at the n-th place; as before, we assume that $(R_n)_{1 \leq n \leq N}$ is decreasing, and we convene that R_N is paid to players that never arrive. Moreover, let $c_n > 0$ be the cost coefficient when n players have arrived. Then the objective of player i is to maximize

$$J_i(\lambda_i; \lambda_{-i}) = E\left[R_{\xi_{\lambda_i,\lambda_{-i}}(\tau^i)} - \int_0^{\tau^i} c_{\xi_{\lambda_i,\lambda_{-i}}(s)} \lambda_i^2(\xi_{\lambda_i,\lambda_{-i}}(s))ds \right]$$

and λ is a (symmetric) equilibrium optimal control if $\arg\max_{\lambda_i} J_i(\lambda_i; \lambda) = \lambda$ for all i. For $0 \leq n \leq N-1$, the value function of player i before arriving is

$$v_n := \sup_{\lambda_i} E\left[R_{\xi(\tau^i)} - \int_0^{\tau^i} c_{\xi(s)} \lambda_i^2(\xi(s))ds \,\middle|\, \xi(0) = n \right]$$

where $\xi := \xi_{\lambda_i,\lambda_{-i}}$. We also convene that $v_N := R_N$.

Remark 4.1. The last player never arrives. Indeed, once $N-1$ players have arrived, the remaining player achieves the optimal value v_{N-1} by using the control $\lambda_i \equiv 0$, and in fact $v_{N-1} = R_N = v_N$. This is due to the convention that R_N is paid to players that never arrive. On the other hand, this convention is necessary in order to have existence of an equilibrium, since the same value is asymptotically achieved by using the control $\lambda_i \equiv \varepsilon$ with small $\varepsilon > 0$.

Proposition 4.2. *The N-player game has a unique Nash equilibrium. The equilibrium value function $(v_n)_{0 \le n \le N}$ is the unique solution of the backward recursion*

$$v_n = \frac{R_{n+1} + 2(N - n - 1)v_{n+1}}{1 + 2(N - n - 1)}, \quad 0 \le n \le N - 1; \quad v_N = R_N. \quad (4.1)$$

The unique equilibrium optimal control is

$$\lambda^*(n) = \frac{R_{n+1} - v_n}{2c_n}, \quad 0 \le n \le N - 1. \quad (4.2)$$

4.2   The N-Player Principal–Agent Problem

Next, we consider the N-player version of the mean field principal–agent problem introduced in Section 3. Given $n_0 \in \{1, \dots, N-1\}$ and a (nonnegative, decreasing) reward scheme (R_n), let

$$T_{n_0} = \inf\{t \ge 0 : \xi_{\lambda^*}(t) = n_0\}$$

be the (random) time until n_0 players have arrived, where the players use the unique equilibrium optimal control λ^* for (R_n) and the (fixed, positive) cost coefficients (c_n); cf. (4.2). Below, we shall find it useful to write λ_n^* rather than $\lambda^*(n)$ whenever we are in the N-player setting.

Given the per capita[3] reward budget $B > 0$, the principal chooses a reward scheme (R_n) such as to minimize the expected completion time ET_{n_0}, subject to the budget constraint $\sum_{n=1}^{N} R_n \le NB$. In analogy to (3.2), we shall assume that c^N satisfies

$$c_n^N \le c_{n-1}^N \frac{(2N - 2n + 1)^2(2N - n - 1)}{4(N - n - 1)(N - n + 1)(2N - n)}, \quad n < n_0; \quad (4.3)$$

again, this is satisfied e.g. when $n \mapsto c_n^N$ is constant or decreasing.

Theorem 4.3. *Let c^N satisfy (4.3) and define $y_{n_0} = 0$,*

$$y_n = \sqrt{\frac{c_n N(N - n - 1)}{(N - n)(2N - n - 1)}}, \quad n < n_0.$$

---

[3] A normalization of the budget is necessary for a convergence result as in the subsequent section. We have done that implicitly by seeing B as the total budget in the mean field limit and as per capita budget in the N-player setting. Equivalently, one could normalize the mass of the population by assigning mass $1/N$ to each agent in the N-player game and see B as the total budget.

Given a per capita reward budget $B > 0$, there is a unique optimal reward scheme (R_n^*) attaining the minimal expected completion time $ET_{n_0}^*$, given by

$$R_n^* = \frac{B}{C}\left\{ y_{n-1} + \frac{1}{2} \sum_{k=n-1}^{n_0-1} \frac{y_k}{N-k-1} \right\} \mathbf{1}_{\{n \le n_0\}},$$

and the minimal expected completion time is

$$ET_{n_0}^* = \frac{4C^2}{B}, \quad where \quad C = \frac{1}{2\sqrt{N}} \sum_{n=0}^{n_0-1} \sqrt{\frac{c_n(2N-n-1)}{(N-n)(N-n-1)}}.$$

The corresponding equilibrium optimal control is

$$\lambda_n^* = \frac{B}{2C} \sqrt{\frac{N(N-n-1)}{c_n(N-n)(2N-n-1)}} \mathbf{1}_{\{n < n_0\}}.$$

4.3 Size Effects

We conclude this section with a brief discussion of the influence of the population size N on the principal's problem, in two different ways. To make the problems comparable, we assume that $c_n^N \equiv c$ is a constant independent of N.

(i) First, we consider as above a principal with a given per capita budget B aiming to minimize the expected time until an α proportion of the population has arrived. The left panel in Figure 3 shows a negative size effect: the minimum expected completion time $ET_{n_0}^N$ for $n_0 = \lceil \alpha N \rceil$ is increasing in N; that is, an increase in population size *adversely* affects the principal.

(ii) Second, we fix the total budget $K = NB$ and a completion *head count* n_0. Then, $ET_{n_0}^N = \frac{1}{NK} \left(\sum_{n=0}^{n_0-1} \left[\frac{c}{1-n/N}\left(1 + \frac{1}{1-(n+1)/N}\right) \right]^{1/2} \right)^2$ is strictly decreasing in N with limit equal to zero (right panel in Figure 3). That is, the principal aiming for a fixed number of completions benefits from an increase in the population size.

5 Convergence to the Mean Field

In this section, we show that the N-player competition and principal–agent problems converge to their mean field counterparts as $N \to \infty$.
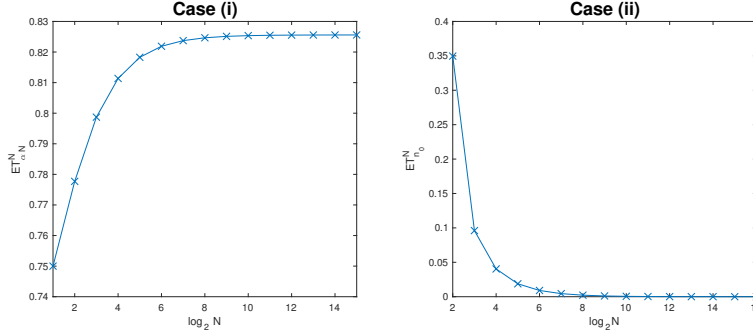
Figure 3: The left panel shows a negative size effect when fixing $\alpha = 0.5$ and $B = 1$. The right panel shows a positive size effect when fixing $n_0 = 3$ and $K = 2^5$. In both panels, $c \equiv 1$ and the minimal expected completion time is plotted against $\log_2 N$.

5.1 Convergence of the N-Player Equilibrium

We consider the mean field setting of Section 2 with a fixed reward scheme R and cost coefficient c, as well as an N-player game with reward (R_n^N) and cost (c_n^N). Our aim is to show that if $R^N \to R$ and $c^N \to c$ in a suitable sense, then the corresponding equilibria converge.

If we start with a reward scheme $R : [0,1] \to \mathbb{R}_+$ in the mean field setting, an obvious choice for R^N is the sampling $R_n^N = R\left(\frac{n}{N}\right)$. Since R is decreasing, we have $\frac{1}{N}\sum_{n=1}^N R_n^N \leq \int_0^1 R(r)\,dr$; i.e., this discretization may (and typically will) reduce the cumulative reward. Another choice is the moving average $R_n^N = N\int_{(n-1)/N}^{n/N} R(y)\,dy$ which preserves the reward. Next, we introduce a condition designed to cover either of these choices, and more.

Recall that $R : [0,1] \to \mathbb{R}_+$ is decreasing, piecewise Lipschitz continuous and left-continuous at $r = 1$. Let $0 = r_0 < r_1 < \ldots < r_m < r_{m+1} = 1$ be a finite partition of $[0,1]$ such that R is Lipschitz on each interval (r_{i-1}, r_i). For our results, we shall assume that

$$\sup_{r\in\bigcup_{i=1}^{m+1} I_i^N} \left| R_{\lceil rN \rceil}^N - R(r) \right| \leq \frac{K}{N} \tag{5.1}$$

for some constant K independent of N, where $I_i^N = [r_{i-1} + 1/N, r_i - 1/N]$

for $i = 1, \ldots, m$ and $I^N_{m+1} = [r_m + 1/N, 1]$. Similarly, we assume that[4]

$$\sup_{r \in [0,1]} \left| c^N_{\lceil rN \rceil} - c(r) \right| \leq \frac{K}{N}. \tag{5.2}$$

The next result establishes that the N-player equilibrium converges to the mean field equilibrium as $N \to \infty$. Thus, it gives a second justification to the mean field formulation, apart from the direct derivation with a continuum of players as in Section 2.

Theorem 5.1. *Let R^N, R and c^N, c satisfy (5.1) and (5.2), and let v^N, v and λ^N, λ^* be the corresponding value functions and equilibrium optimal controls, respectively. Then*

$$\sup_{r \in [0,1]} |v^N_{\lfloor rN \rfloor} - v(r)| = O(1/\sqrt{N}), \qquad \sup_{r \in \bigcup_{i=1}^{m+1} I^N_i} |\lambda^N_{\lfloor rN \rfloor} - \lambda^*(r)| = O(1/\sqrt{N}).$$

Remark 5.2. If $R(r) = 0$ on $(\alpha, 1]$ for some $\alpha < 1$, then the ODE (A.2) is nondegenerate up to the boundary and the convergence rates in Theorem 5.1 can be improved to $O(1/N)$; cf. the proof of the theorem. In particular, this holds for the optimal reward scheme R^* of the principal's problem in Theorem 3.1.

5.2 Convergence and ε-Optimality for the Principal

Following Section 3, we fix a cost coefficient c satisfying (3.2), a target proportion $\alpha \in (0, 1)$ and a budget $B > 0$. We have seen in Theorem 3.1 that in the mean field setting, there exists a unique reward scheme R^* which attains the minimal (deterministic) time T^*_α until an α-proportion of the population has arrived. For the N-player situation with a given cost c^N satisfying (4.3) and per capita budget B, we have seen in Theorem 4.3 that there exists a unique reward scheme R^N minimizing the expected time $ET^N_{n_0}$ until n_0 players have arrived. In the following result, we consider $n_0 = \lceil \alpha N \rceil$, so that the proportion n_0/N tends to α, and show that if $c^N \to c$, the expected completion time as well as the corresponding reward schemes converge as $N \to \infty$.

Theorem 5.3. *Let $c^N \to c$ in the sense of (5.2). Then*

$$\left| ET^N_{\lceil \alpha N \rceil} - T^*_\alpha \right| = O(1/N), \qquad \sup_{r \in [0,\alpha]} \left| R^N_{\lceil rN \rceil} - R^*(r) \right| = O(1/N).$$

---

[4]This covers, e.g., $c^N_n = c(n/N)$. We do not consider a more general convergence as in (5.1) since that would lead to more complicated statements below without substantially enhancing the scope.

The next result addresses the convergence of the principal's problem in a different sense: it shows that if the principal applies (a discretization of) the optimal reward scheme R^* from the mean field setting in an N-player game with large N, rather than the precise optimal scheme R^N of Theorem 4.3, then R^* is still ε-optimal for the minimization of the expected completion time.

Corollary 5.4. *Let $c^N \to c$ in the sense of (5.2) and let $R^{(N)}$ be a discretization of R^* satisfying (5.1). Then the completion time $T^{(N)}_{\lceil \alpha N \rceil}$ of the N-player game under $R^{(N)}$ satisfies*

$$\left| ET^{(N)}_{\lceil \alpha N \rceil} - ET^N_{\lceil \alpha N \rceil} \right| = O(1/N);$$

that is, the reward scheme $R^{(N)}$ is $O(1/N)$-optimal for the N-player principal–agent problem.

A Proofs

A.1 Proofs for Section 2

Proof of Lemma 2.1. The piecewise Lipschitz property of λ implies that (2.2) has a unique continuous solution ρ. This function is nonnegative, increasing, globally Lipschitz continuous, and continuously differentiable except at finitely many points (corresponding to the jumps to λ) where the left and right derivatives of ρ may disagree. Let $\bar\rho(t) = 1 - \exp(-\int_0^t \lambda(\rho(s))\,ds)$. By the Exact Law of Large Numbers (Proposition B.2) and (2.1),

$$\mu\{i :\ \tau_\lambda^i(\omega) \in [0,t]\} = P\{\omega :\ \tau_\lambda^i(\omega) \in [0,t]\} = \bar\rho(t)$$

holds almost-surely. On the other hand, the derivative of $\bar\rho$ is seen to satisfy $\bar\rho'(t) = \lambda(\rho(t))(1-\bar\rho(t))$ a.e., so that the uniqueness of the exponential ODE yields $\bar\rho = \rho$. □

Proof of Theorem 2.2. (i) Suppose $\bar\lambda \in \Lambda$ is an equilibrium optimal control, let ρ be the corresponding equilibrium state process and let $\bar v$ be the corresponding value function of (2.3) for any given player before arrival. Using the constant control $\lambda \equiv 0$ shows that $R(1) \le \bar v(r) \le R(r)$ and hence $\bar v(1-) = R(1)$; recall that R is left-continuous at $r = 1$. Let

$$r_0 = \inf\{r :\ R(r) = R(1)\}.$$

Using $\lambda \equiv 0$ also shows that $\bar{v} \equiv R(1)$ on $[r_0, 1]$, whereas $\bar{v} < R$ on $[0, r_0)$ because the only control with zero cost merely attains $R(1) < R(r)$. The inequality $R(1) < R(r)$ on $[0, r_0)$ also implies that $\bar{\lambda} > 0$ a.e. on $[0, r_0)$ and in fact, recalling that $\bar{\lambda} \in \Lambda$, even that $\bar{\lambda}$ is a.e. uniformly bounded away from zero on any interval $[0, r_1]$ with $r_1 < r_0$.

A control-theoretic argument shows that $\bar{v}$ is Lipschitz on any such interval. Indeed, let $0 \le r < r_1$ and choose $h > 0$ such that $r + 2h < r_1$. We may compare the optimal control $\bar{\lambda}$ started at r with a control λ such that

$$\lambda = \begin{cases} 0 & \text{on } [r, r+h), \\ a\bar{\lambda} & \text{on } [r+h, r+2h), \\ \bar{\lambda} & \text{on } [r+2h, 1), \end{cases}$$

where the constant $a \ge 1$ is chosen such that the integral of λ along ρ over $[r, r+2h]$ coincides with the integral of $\bar{\lambda}$ over the same interval. Then, both the extra cost of λ and the loss of expected reward are bounded by a constant (depending only on r_1) times h, because if $\bar{\lambda}$ achieves a rank in $[r + 2h, 1)$, then λ achieves the same rank, whereas the probability of $\bar{\lambda}$ ranking in $[r, r + 2h)$ is bounded by a constant times h. Since λ is an admissible control for the control problem started at $r + h$, it follows that $0 \le \bar{v}(r) - \bar{v}(r+h) \le Ch$ as claimed. In particular, $\bar{v}$ is absolutely continuous and a.e. differentiable.

By dynamic programming, the value function $\bar{v}$ must then a.e. satisfy the Hamilton–Jacobi equation

$$\sup_{l \ge 0}\{l[R(r) - v(r)] - c(r)l^2\} + \bar{\lambda}(r)(1-r)v'(r) = 0 \text{ on } [0,1), \quad v(1) = R(1).$$

Moreover, the optimal control $\bar{\lambda}$ must attain the supremum a.e. Recalling that $\bar{v} = R$ on $[r_0, 1]$ and $\bar{v} < R$ on $[0, r_0)$, it follows that

$$\bar{\lambda}(r) = \frac{R(r) - \bar{v}(r)}{2c(r)} \quad \text{a.e.} \tag{A.1}$$

and that $\bar{v}$ satisfies

$$R(r) - v(r) + 2(1-r)v'(r) = 0 \text{ a.e. on } [0, r_0), \quad v \equiv R(1) \text{ on } [r_0, 1]. \tag{A.2}$$

(ii) Using the regularity of R, we see that the function v of (2.5) is Lipschitz on $[0, 1)$ and satisfies $v(1-) = R(1)$. We extend v to a Lipschitz function on $[0, 1]$ by setting $v(1) = R(1)$. A direct calculation shows that v satisfies (A.2) and that λ^* of (2.4) is the corresponding maximizer (A.1).

Moreover, $\lambda^* \in \Lambda$. A verification argument then yields that v is the value function and λ^* is an optimal control.

(iii) It remains to show that (A.2) has at most one absolutely continuous solution. Indeed, if v_1 and v_2 are solutions, then $w = v_1 - v_2$ is absolutely continuous and satisfies

$$2(1 - r)w'(r) = w(r) \text{ a.e. on } [0, r_0), \quad w \equiv 0 \text{ on } [r_0, 1].$$

If $r_0 < 1$, this is a Lipschitz ODE and it follows directly that $w \equiv 0$ is the unique solution. If $r_0 = 1$, we set $u(t) = w(1 - e^{-2t})$ for $t \geq 0$; then u satisfies $u'(t) = u(t)$ on $[0, \infty)$ and hence $u(t) = u(0)e^t$. But $u(\infty) = w(1) = 0$ then yields that $u \equiv 0$ and thus $w \equiv 0$ as desired. $\square$

Proof of Proposition 2.5. Let $V(r) = E[R(\rho(\tau))|\rho(0) = r]$. The ODE for ρ has the unique solution $\rho(t) = 1 - (1 - r)e^{-2\lambda_0 t}$ for $t \geq 0$. Hence,

$$V(r) = E[R(1 - (1 - r)e^{-2\lambda_0 \tau})] = \int_0^\infty \lambda_0 e^{-\lambda_0 x} R(1 - (1 - r)e^{-2\lambda_0 x}) \, dx.$$

A change of variables then shows that $V(r)$ coincides with (2.5). $\square$

Proof of Proposition 2.6. Note that R_n, R have a uniform upper bound given by $\sup_n R_n(0)$. Moreover, by monotonicity, the convergence $R_n \to R$ is uniform on each interval of continuity of R. It follows directly from (2.5) that the value functions v_n converge uniformly to their counterpart v. Similarly, (2.4) yields that the optimal controls λ_n^* converge pointwise to their counterpart λ^*, and uniformly on each interval of Lipschitz continuity of R. Moreover, there is a uniform upper bound for the sequence (λ_n^*). By the ODE (2.2), this entails an upper bound for the Lipschitz constants of (ρ_n). Thus, after passing to a subsequence, (ρ_n) converges uniformly to a limit $\bar{\rho}$. To verify that $\bar{\rho} = \rho$, it suffices to show that $\bar{\rho}$ solves the ODE (2.2) defining ρ on each of the mentioned intervals, and that follows from the uniform convergence of (λ_n^*). Finally, by a subsequence argument, the entire sequence (ρ_n) must converge to ρ. $\square$

A.2 Proofs for Section 3

As a preparation for the proof of Theorem 3.1, we first show that no reward should be distributed to ranks below α—this is quite intuitive since the planner does not care about agents arriving after rank α. The converse is also true.

Lemma A.1. *The value T_α^* of (3.1) does not change if the infimum is restricted to $R \in \mathcal{R}$ satisfying $R > 0$ on $[0, \alpha)$ and $R = 0$ on $(\alpha, 1]$.*

Proof. For the first property, suppose that $R \in \mathcal{R}$ vanishes at $r \in [0, \alpha)$, and hence on $[r, 1]$. Then the equilibrium effort λ of (2.4) also vanishes at r, which by the Kolmogorov equation (2.2) implies that the state $\rho(t)$ never exceeds r, and hence that $T_\alpha(R) = \infty$.

To see the second property, let $R \in \mathcal{R}$ and set $\hat{R} = R\mathbf{1}_{[0,\alpha]}$; then $\hat{R} \in \mathcal{R}$. For $r \in [0, \alpha]$, the corresponding equilibrium efforts λ and $\hat{\lambda}$ of (2.4) satisfy

$$\hat{\lambda}(r) = \frac{\hat{R}(r) - \frac{1}{2\sqrt{1-r}} \int_r^1 \frac{\hat{R}(y)}{\sqrt{1-y}}\,dy}{2c(r)} \geq \frac{R(r) - \frac{1}{2\sqrt{1-r}} \int_r^1 \frac{R(y)}{\sqrt{1-y}}\,dy}{2c(r)} = \lambda(r),$$

and the inequality is strict if $\int_\alpha^1 \frac{R(y)}{\sqrt{1-y}}\,dy > 0$. As a result, if R does not vanish on $(\alpha, 1]$, then $\hat{R}$ produces a strictly larger equilibrium effort on $[0, \alpha]$, and hence $T_\alpha(\hat{R}) < T_\alpha(R)$ whenever $T_\alpha(R) < \infty$, by (2.2). $\qquad\square$

We can now show the theorem through a calculus of variations argument.

Proof of Theorem 3.1. Let

$$\mathcal{R}' = \big\{R \in \mathcal{R} : \textstyle\int R(r)\,dr \leq B,\ T_\alpha(R) < \infty,\ R\mathbf{1}_{(\alpha,1]} = 0\big\}.$$

As noted, $\mathcal{R}' \neq \emptyset$ by (2.6), and by Lemma A.1 it is sufficient to show that R^* is the unique optimizer in $\mathcal{R}'$. Moreover, we have seen in the proof of Lemma A.1 that λ^* is a.e. strictly positive on $[0, \alpha)$ for $R \in \mathcal{R}'$. Hence, ρ is strictly increasing and we have $T_r(R) = \rho^{-1}(r)$ for $r \in [0, \alpha)$. Differentiating $\rho^{-1}(r)$ and using the ODE (2.2) for ρ and (2.4), we then obtain that

$$T_\alpha(R) = \int_0^\alpha \frac{1}{(1-r)\lambda^*(r)}\,dr$$

$$= \int_0^\alpha \frac{2c(r)}{\sqrt{1-r}\left(R(r)\sqrt{1-r} - \int_r^\alpha \frac{R(s)}{2\sqrt{1-s}}\,ds\right)}\,dr, \quad R \in \mathcal{R}'. \quad (A.3)$$

We see from (A.3) that $R \mapsto T_\alpha(R)$ is strictly convex on $\mathcal{R}'$, up to a.e. equivalence. This implies that there is at most one optimal $R \in \mathcal{R}'$.

Next, we derive a sufficient condition for optimality. We first reparametrize the optimization problem: for $R \in \mathcal{R}'$, we consider

$$f(r) = f_R(r) = R(r)\sqrt{1-r} - \int_r^\alpha \frac{R(s)}{2\sqrt{1-s}}\,ds, \quad r \in [0, \alpha].$$

The mapping $R \mapsto f_R$ is one-to-one on $\mathcal{R}'$ since R can be recovered from f via

$$R(r) = \frac{f(r)}{\sqrt{1-r}} + \int_r^\alpha \frac{f(s)}{2(1-s)^{3/2}}\, ds. \tag{A.4}$$

Indeed, if R is differentiable, then $f'(r) = \sqrt{1-r}\, R'(r)$ and integration by parts yields

$$R(r) = R(\alpha) - \int_r^\alpha \frac{f'(s)}{\sqrt{1-s}}\, ds$$

$$= \frac{f(\alpha)}{\sqrt{1-\alpha}} - \int_r^\alpha \frac{f'(s)}{\sqrt{1-s}}\, ds = \frac{f(r)}{\sqrt{1-r}} + \int_r^\alpha \frac{f(s)}{2(1-s)^{3/2}}\, ds,$$

and now the claim for general $R \in \mathcal{R}'$ follows by approximation. Fubini's Theorem shows that $\int_0^\alpha R(r)\, dr = \frac{1}{2}\int_0^\alpha \frac{(2-r)f(r)}{(1-r)^{3/2}}\, dr$. Thus, recalling that $\alpha < 1$, the image $\mathcal{F}$ of $\mathcal{R}'$ under $R \mapsto f_R$ is the convex set of all piecewise Lipschitz, nonnegative, decreasing functions $f : [0, \alpha] \to \mathbb{R}$ such that (A.6) is finite and the budget constraint

$$\frac{1}{2}\int_0^\alpha \frac{(2-r)f(r)}{(1-r)^{3/2}}\, dr \leq B \tag{A.5}$$

is satisfied. We write $T_\alpha(f_R)$ for $T_\alpha(R)$ by a slight abuse of notation; then by (A.3) we have

$$T_\alpha(f) = \int_0^\alpha \frac{2c(r)}{\sqrt{1-r}\, f(r)}\, dr, \quad f \in \mathcal{F}. \tag{A.6}$$

The mapping $f \mapsto T_\alpha(f)$ is convex and finite-valued, and clearly, $f^* \in \mathcal{F}$ is optimal if an only if

$$\varepsilon \mapsto \phi(\varepsilon) = T((1-\varepsilon)f^* + \varepsilon f), \quad [0,1] \to \mathbb{R}$$

attains a minimum at $\varepsilon = 0$ for all $f \in \mathcal{F}$. By convexity, this function has a right derivative $\phi'(0)$ at $\varepsilon = 0$, and ϕ attains a minimum at $\varepsilon = 0$ if and only if $\phi'(0) \geq 0$. Note that for any convex function φ, the right difference quotient $(\varphi(x+\varepsilon) - \varphi(x))/\varepsilon$ satisfies $\varphi'(x) \leq (\varphi(x+\varepsilon) - \varphi(x))/\varepsilon \leq \varphi(x+1) - \varphi(x)$ for $\varepsilon \leq 1$. Using these bounds and dominated convergence, we see that $\phi'(0)$ can be computed by differentiating under the integral:

$$\phi'(0) = \int_0^\alpha \frac{-2c(r)(f(r) - f^*(r))}{\sqrt{1-r}\, f^*(r)^2}\, dr. \tag{A.7}$$

23

Let R^* be as in (3.3); then the corresponding function $f^* = f_{R^*}$ is given by

$$f^*(r) = \frac{B}{C}\sqrt{\frac{c(r)(1-r)}{2-r}}, \quad r \in [0, \alpha]. \tag{A.8}$$

Recalling that $\alpha < 1$ and that $\frac{c(r)(1-r)}{2-r}$ is decreasing, we verify directly that $f^* \in \mathcal{F}$. Moreover, the budget constraint (A.5) is satisfied with equality.

Fix an arbitrary $f \in \mathcal{F}$. Using the expression (A.8) in the denominator of (A.7), we have

$$\phi'(0) = \frac{-2C^2}{B^2} \int_0^\alpha \frac{(2-r)(f(r) - f^*(r))}{(1-r)^{3/2}} \, dr.$$

Since f^* satisfies (A.5) with equality and f satisfies the same with inequality, the above integral is nonpositive. Thus, $\phi'(0) \geq 0$, showing that $f^* \in \mathcal{F}$ is optimal and hence that $R^* \in \mathcal{R}$ is optimal. The formula (3.4) for T_α^* is then obtained from (A.6) and (A.8), and the formula for λ^* follows via (2.4). $\square$

A.3 Proofs for Section 4

Proof of Proposition 4.2. Fix player i and suppose that all other players use a control λ_{-i}. By dynamic programming, the value function v_n of player i before arrival satisfies

$$\sup_{\lambda_i \geq 0} \left\{ \lambda_i [R_{n+1} - v_n] - c_n \lambda_i^2 \right\} + \lambda_{-i}(n)(N - n - 1)(v_{n+1} - v_n) = 0$$

for $0 \leq n \leq N - 2$. For $n = N - 1$, the same holds by Remark 4.1 and our convention that $v_N = R_N$. Thus, (4.2) is the optimal control for player i and it follows that

$$\frac{(R_{n+1} - v_n)^2}{4c_n} + \lambda_{-i}(n)(N - n - 1)(v_{n+1} - v_n) = 0.$$

Assuming inductively that $v_{n+1} \leq R_{n+1}$, this quadratic equation has a unique nonnegative root v_n, and v_n satisfies $0 \leq v_n \leq R_{n+1} \leq R_n$. In a given equilibrium, the consistency condition $\lambda_i = \lambda_{-i}$ implies that

$$R_{n+1} - v_n + 2(N - n - 1)(v_{n+1} - v_n) = 0, \quad n = 0, \ldots, N - 1, \tag{A.9}$$

or equivalently (4.1), which clearly has a unique solution. Conversely, we can verify directly that (4.1) and (4.2) define an equilibrium. $\square$

Proof of Theorem 4.3. We first observe that T_{n_0} is a sum of independent exponential random variables (whenever finite) and thus

$$ET_{n_0} = \sum_{n=0}^{n_0-1} \frac{1}{(N-n)\lambda_n}. \tag{A.10}$$

Moreover, similarly as in Lemma A.1, it suffices to consider reward schemes with $R_n > 0$ for $n \leq n_0$ and $R_n = 0$ for $n > n_0$. Indeed, the first claim is immediate from (4.2). To obtain the second, we argue by contradiction and compare (R_n) with the scheme defined by $\hat{R}_n = R_n \mathbf{1}_{\{n \leq n_0\}}$. Proposition 4.2 implies that $\hat{R}_n$ leads to a strictly larger equilibrium control before n_0 and hence a strictly smaller completion time. Thus, we only consider reward schemes up to $n = n_0$ in what follows.

From (4.1), (4.2) and (A.10), we see that $ET_{n_0} : \mathbb{R}_+^{n_0} \to [0, \infty]$ is a strictly convex, continuous function of $(R_1, \ldots, R_{n_0})$. Moreover, the feasible set defined by $R_1 \geq R_2 \geq \cdots \geq R_{n_0} \geq 0$ and $\sum_{n=1}^{N} R_n \leq NB$ is nonempty, convex and compact. As a result, there exists a unique optimal reward scheme. In the remainder of the proof, we determine this reward scheme explicitly. To that end, define $x_{n_0} = 0$ and $x_n = R_{n+1} - v_n$ for $n < n_0$. By (4.2) and (A.10), the objective function can then be expressed as

$$ET_{n_0} = \sum_{n=0}^{n_0-1} \frac{2c_n}{(N-n)x_n}.$$

From (4.1) we obtain that $R_{n+1} - R_{n+2} = \frac{1+2(N-n-1)}{2(N-n-1)}x_n - x_{n+1}$ and thus the total reward $\sum_{n=1}^{n_0} R_n = \sum_{n=1}^{n_0} n(R_n - R_{n+1})$ can be expressed as

$$\sum_{n=1}^{n_0} R_n = \sum_{n=1}^{n_0} n \left\{ \frac{1+2(N-n)}{2(N-n)} x_{n-1} - x_n \right\} = \sum_{n=0}^{n_0-1} \frac{2N-n-1}{2(N-n-1)} x_n.$$

The constrained optimization problem of finding $x_n \in \mathbb{R}$ which

$$\text{minimize} \quad \sum_{n=0}^{n_0-1} \frac{2c_n}{(N-n)x_n} \quad \text{subject to} \quad \sum_{n=0}^{n_0-1} \frac{2N-n-1}{2(N-n-1)} x_n \leq NB$$

can be solved using the method of Lagrange multipliers. One finds that the optimal x_n are given by

$$x_n = 2\sqrt{\frac{c_n(N-n-1)}{\theta(N-n)(2N-n-1)}}, \quad n = 0, \ldots, n_0 - 1$$

25

where the Lagrange multiplier θ satisfies

$$\sqrt{\theta} = \frac{1}{NB} \sum_{n=0}^{n_0-1} \sqrt{\frac{c_n(2N-n-1)}{(N-n)(N-n-1)}}.$$

Note that θ and x_n are related to C and y_n of Theorem 4.3 by $C = B\sqrt{N\theta}/2$ and $x_n = 2y_n/\sqrt{N\theta} = By_n/C$. We have that $R_{n_0} \geq 0$ as $x_{n_0-1} \geq 0$, and $R_n - R_{n+1} = \frac{B}{C}\left(\frac{1+2(N-n)}{2(N-n)} y_{n-1} - y_n\right) \geq 0$ for $n < n_0$; here the last inequality is equivalent to (4.3). Thus, the reward scheme (R_n^*) associated with the optimizer (x_n) is indeed nonnegative and decreasing, and it follows that (R_n^*) is the optimal reward scheme. The formulas for R^*, $T_{n_0}^*$ and λ^* follow by direct calculation. $\qquad\square$

A.4 Proofs for Section 5

Proof of Theorem 5.1. Let $N \geq 4$ be large enough such that $\delta := 1/N$ satisfies $r_m < 1 - \sqrt{\delta} - \delta$. This implies, in particular, that $1 - r_m > \delta + \sqrt{\delta} > 2\delta$. We may rewrite the recursion (4.1) for v_n^N as

$$v_n^N = g\left(\frac{n}{N}\right) v_{n+1}^N + \left(1 - g\left(\frac{n}{N}\right)\right) R_{n+1}^N \tag{A.11}$$

where

$$g(r) = \frac{2(1-r-\delta)}{\delta + 2(1-r-\delta)}.$$

Using (2.5), we write the mean field value in a similar form:

$$v\left(\frac{n}{N}\right) = f\left(\frac{n}{N}\right) v\left(\frac{n+1}{N}\right) + \left(1 - f\left(\frac{n}{N}\right)\right) R_{n+1}^N + E_{1,n} \tag{A.12}$$

where

$$f(r) := \sqrt{\frac{1-r-\delta}{1-r}} \quad \text{and} \quad E_{1,n} := \frac{1}{2\sqrt{1-n\delta}} \int_{n\delta}^{(n+1)\delta} \frac{R(y) - R_{\lceil yN\rceil}^N}{\sqrt{1-y}}\, dy.$$

Next, we estimate $\Delta_n := \left|v_n^N - v\left(\frac{n}{N}\right)\right|$. For the last rank, we have $\Delta_N = \left|R_N^N - R(1)\right| \leq K\delta$ by (5.1). For the second-to-last rank, since $1 - \delta \geq 1 - \sqrt{\delta} > r_m + \delta$, (5.1) again implies that

$$\Delta_{N-1} = \left|R_N^N - v(1-\delta)\right| \leq \frac{1}{2\sqrt{\delta}} \int_{1-\delta}^1 \frac{\left|R_{\lceil yN\rceil}^N - R(y)\right|}{\sqrt{1-y}}\, dy \leq K\delta.$$

26

For $n \le N-2$, subtracting (A.12) from (A.11) and using that $g(r) \le 1$ when $0 \le r \le 1-\delta$, we obtain that

$$\Delta_n \le \Delta_{n+1} + E_{1,n} + E_{2,n}, \tag{A.13}$$

where

$$E_{2,n} = \left| g\left(\frac{n}{N}\right) - f\left(\frac{n}{N}\right) \right| \cdot \left| v\left(\frac{n+1}{N}\right) - R_{n+1}^N \right|.$$

Next, we estimate $E_{1,n}$ and $E_{2,n}$. To that end, it will be useful to define $n_0 := N - \lceil \sqrt{N} \rceil$ and note that $n \le n_0$ if and only if $n\delta \le 1 - \sqrt{\delta}$.

Estimation of $E_{1,n}$. If $n_0 + 1 \le n \le N - 2$, then $r_m + \delta < n\delta \le 1 - 2\delta$, and (5.1) implies

$$E_{1,n} \le K\delta \left(1 - \sqrt{\frac{1 - n\delta - \delta}{1 - n\delta}} \right) \le K\delta.$$

On the other hand, if $0 \le n \le n_0$ and $[n\delta, (n+1)\delta] \subseteq I_i^N$ for some i, then (5.1) and $1 - n\delta \ge \sqrt{\delta}$ imply

$$E_{1,n} \le K\delta \frac{\frac{\delta}{1-n\delta}}{1 + \sqrt{\frac{1-n\delta-\delta}{1-n\delta}}} \le K\delta^{3/2}.$$

We observe that there are at most $3m$ indices n for which $n \le n_0$ and $[n\delta, (n+1)\delta]$ is not contained in any I_i^N. For these n, we use boundedness of R and the inequality $1 - n\delta \ge 1 - r_m - \delta > (1 - r_m)/2$ to obtain

$$E_{1,n} \le (R(0) + K\delta) \left(1 - \sqrt{\frac{1 - n\delta - \delta}{1 - n\delta}} \right) < \frac{2(R(0) + K/4)\delta}{(1 - r_m)} =: C_0\delta.$$

In summary,

$$E_{1,n} \le \begin{cases} K\delta, & \text{if } n_0 + 1 \le n \le N - 2, \\ K\delta^{3/2}, & \text{if } n \le n_0 \text{ and } [n\delta, (n+1)\delta] \subseteq I_i^N \text{ for some } i, \\ C_0\delta, & \text{otherwise (at most } 3m \text{ instances).} \end{cases} \tag{A.14}$$

Estimation of $E_{2,n}$. Taylor's theorem implies that for all $x \in [0, 1]$,

$$0 \le \sqrt{x} - \frac{2x}{1+x} \le \frac{1}{2}(1-x)^2 \left(\frac{1}{4x^{3/2}} + \frac{4}{(1+x)^3} \right).$$

Using this with $x = 1 - \frac{\delta}{1-r}$ yields

$$|g(r) - f(r)| \leq \frac{\delta^2}{(1-r)^2} h \left(1 - \frac{\delta}{1-r}\right),$$

where $h(x) := \frac{1}{8x^{3/2}} + \frac{2}{(1+x)^3}$. For $0 \leq n \leq Nr_m$, we have $\frac{\delta}{1-n\delta} \leq \frac{\delta}{1-r_m} \leq \frac{1}{2}$ and thus

$$E_{2,n} \leq (R(0) + K/4) \frac{\delta^2}{(1-r_m)^2} h \left(\frac{1}{2}\right) =: C_1 \delta^2.$$

For $Nr_m < n \leq N - 2$, we have $(n+1)\delta > r_m + \delta$ and it follows that

$$\left| v\left(\frac{n+1}{N}\right) - R_{n+1}^N \right|$$

$$= \left| \frac{1}{2\sqrt{1-(n+1)\delta}} \int_{(n+1)\delta}^1 \frac{R(y) - R((n+1)\delta)}{\sqrt{1-y}} \, dy + R((n+1)\delta) - R_{n+1}^N \right|$$

$$\leq K(1 - (n+1)\delta) + K\delta = K(1 - n\delta).$$

Consequently,

$$E_{2,n} \leq K(1 - n\delta) \frac{\delta^2}{(1-n\delta)^2} h \left(1 - \frac{\delta}{1-n\delta}\right) \leq \frac{K}{2} h \left(\frac{1}{2}\right) \delta < \frac{K}{2} \delta,$$

where we have used that $1 - n\delta \geq 2\delta$.

As in the estimation of $E_{1,n}$, the bound can be improved if $Nr_m < n \leq n_0$ and thus $1 - n\delta \geq \sqrt{\delta}$. In this case, we have

$$E_{2,n} \leq Kh \left(\frac{1}{2}\right) \delta^{3/2} < K\delta^{3/2}.$$

In summary,

$$E_{2,n} \leq \begin{cases} \frac{1}{2} K\delta, & \text{if } n_0 + 1 \leq n \leq N - 2, \\ K\delta^{3/2}, & \text{if } Nr_m < n \leq n_0, \\ C_1 \delta^2, & \text{if } n \leq Nr_m. \end{cases} \tag{A.15}$$

Combining the Estimates. Combining (A.13), (A.14) and (A.15) and recalling $\Delta_{N-1} \leq K\delta$, we obtain for $n_0 + 1 \leq n \leq N - 2$ that

$$\Delta_n \leq \Delta_{N-1} + \frac{3}{2} K\delta(N - 1 - n) \leq K\delta + \frac{3}{2} K\sqrt{\delta} < \frac{5}{2} K\sqrt{\delta}$$

and for $n \le n_0$ that

$$\Delta_n \le \Delta_{n_0+1} + (2K + C_1)\delta^{3/2}(n_0 + 1 - n) + 3mC_0\delta$$
$$\le \frac{5}{2}K\sqrt{\delta} + (2K + C_1)\sqrt{\delta}(1 - \sqrt{\delta} + \delta) + 3mC_0\delta < C_2\sqrt{\delta}.$$

Putting everything together, we have $\sup_{0 \le n \le N} \Delta_n \le \frac{\max(5K/2, C_2)}{\sqrt{N}}$. It remains to note that

$$|v'(r)| = \frac{|R(r) - v(r)|}{2(1 - r)} \le \begin{cases} \frac{K}{2}, & r_m < r < 1, \\ \frac{R(0)}{2(1-r_m)}, & 0 \le r \le r_m, \end{cases}$$

and consequently,

$$\left| v_{\lfloor rN \rfloor}^N - v(r) \right| \le \Delta_{\lfloor rN \rfloor} + \left| v\left(\frac{\lfloor rN \rfloor}{N}\right) - v(r) \right| \le \Delta_{\lfloor rN \rfloor} + \frac{\|v'\|_\infty}{N} < \frac{C_3}{\sqrt{N}}.$$

Since $\lambda^*(r) = \frac{R(r) - v(r)}{2c(r)}$ and $\lambda_n^N = \frac{R_{n+1}^N - v_n^N}{2c_n^N}$, the convergence of λ^N follows from the uniform convergence of the value functions and the cost coefficients, the almost uniform convergence of the reward scheme, and the uniform boundedness of $1/c$. $\qquad\square$

Proof of Theorem 5.3. We first observe the convergence of the Riemann sum

$$C^N := \frac{1}{2N} \sum_{n=0}^{\lceil \alpha N \rceil - 1} \sqrt{\frac{c_n^N(2 - \frac{n+1}{N})}{(1 - \frac{n}{N})(1 - \frac{n+1}{N})}} \to \frac{1}{2} \int_0^\alpha \frac{\sqrt{c(r)(2 - r)}}{(1 - r)} \, dr =: C.$$

The rate of convergence is $O(1/N)$ since c^N converges to c (which is bounded away from zero) uniformly at rate $O(1/N)$ and $\frac{\sqrt{2-r}}{1-r}$ is Lipschitz continuous on $[0, \alpha]$. Using the formulas of Theorems 3.1 and 4.3, we conclude that

$$\lim_{N \to \infty} ET_{\lceil \alpha N \rceil}^N = \lim_{N \to \infty} \frac{4(C^N)^2}{B} = \frac{4C^2}{B} = T_\alpha^*$$

with rate $O(1/N)$. Turning to the convergence of the reward schemes, we similarly observe that for $r \le \alpha$,

$$y_{\lceil rN \rceil - 1} = \sqrt{\frac{c_{\lceil rN \rceil - 1}^N N(N - \lceil rN \rceil)}{(N - \lceil rN \rceil + 1)(2N - \lceil rN \rceil)}} \to \sqrt{\frac{c(r)}{2 - r}} = y(r)$$

and

$$\frac{1}{N} \sum_{k=\lceil rN \rceil - 1}^{\lceil \alpha N \rceil - 1} \frac{y_k}{1 - \frac{k+1}{N}} \to \int_r^\alpha \frac{y(s)}{1 - s} \, ds,$$

uniformly in $r \in [0, \alpha]$ with rate $O(1/N)$. Hence, the formulas in Theorems 3.1 and 4.3 yield that

$$
\begin{aligned}
R^N_{\lceil rN \rceil} &= \frac{B}{C^N} \left\{ y_{\lceil rN \rceil - 1} + \frac{1}{2} \sum_{k=\lceil rN \rceil - 1}^{\lceil \alpha N \rceil - 1} \frac{y_k}{N - k - 1} \right\} \\
&\to \frac{B}{C} \left\{ y(r) + \frac{1}{2} \int_r^\alpha \frac{y(s)}{1 - s} ds \right\} = R^*(r)
\end{aligned}
$$

uniformly in $r \in [0, \alpha]$ with rate $O(1/N)$. $\qquad \square$

Proof of Corollary 5.4. Theorem 5.1 and Remark 5.2 imply that the N-player equilibrium control $\lambda^{(N)}_{\lfloor rN \rfloor}$ for $R^{(N)}$ converges uniformly to the mean field equilibrium control $\lambda^*(r)$ for R^*, at rate $O(1/N)$. It follows that

$$
f_N(r) := (1 - \lfloor rN \rfloor) \lambda^{(N)}_{\lfloor rN \rfloor} \to (1 - r)\lambda^*(r) =: f(r)
$$

uniformly at rate $O(1/N)$. Since λ^* is uniformly bounded away from zero on $[0, \alpha]$, the same holds for f and f_N with large N. Thus, using (A.10) and (A.3),

$$
\begin{aligned}
\left| ET^{(N)}_{\lceil \alpha N \rceil} - T^*_\alpha \right| &= \left| \sum_{n=0}^{\lceil \alpha N \rceil - 1} \frac{1}{(N - n)\lambda^{(N)}_n} - \int_0^\alpha \frac{dr}{(1 - r)\lambda^*(r)} \right| \\
&\le \int_0^{\frac{\lceil \alpha N \rceil}{N}} \frac{|f(r) - f_N(r)|}{f_N(r)f(r)} dr = O(1/N).
\end{aligned}
$$

As Theorem 5.3 shows that $|T^*_\alpha - ET^N_{\lceil \alpha N \rceil}| = O(1/N)$, the claim follows. $\quad \square$

B   Exact Law of Large Numbers

In this section, we detail a setting such that the exact law of large numbers holds for a continuum of essentially pairwise independent random variables. Let $(I, \mathcal{I}, \mu)$ be an atomless (hence uncountable) probability space and let $(\Omega, \mathcal{F}, P)$ be another probability space.

Definition B.1. A family $(f_i)_{i \in I}$ of random variables on $(\Omega, \mathcal{F}, P)$ is *essentially pairwise independent* if for μ-almost all $i \in I$, f_i is independent of f_j for μ-almost all $j \in I$. The family is *essentially pairwise i.i.d.* if, in addition, all f_i have the same distribution.

In what follows, we need to work on a probability space that is larger than the usual product[5] $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \mu \otimes P)$, because the latter does not support relevant families of i.i.d. random variables (see, e.g., [32, Proposition 2.1]). Following Sun [32], a probability space $(I \times \Omega, \Sigma, \nu)$ is an *extension* of the product $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \mu \otimes P)$ if Σ contains $\mathcal{I} \otimes \mathcal{F}$ and the restriction of ν to $\mathcal{I} \otimes \mathcal{F}$ coincides with $\mu \otimes P$. It is a *Fubini extension* if, in addition, any ν-integrable[6] function $f : I \times \Omega \to \mathbb{R}$ satisfies the assertion of Fubini's theorem; that is,

(i) for μ-almost all $i \in I$, the function $f(i, \cdot)$ is P-integrable,

(ii) for P-almost all $\omega \in \Omega$, the function $f(\cdot, \omega)$ μ-integrable,

(iii) $i \mapsto \int f(i, \cdot) \, dP$ is μ-integrable, $\omega \mapsto \int f(\cdot, \omega) \, d\mu$ is P-integrable, and

$$\int f \, d\nu = \iint f(i, \omega) \, P(d\omega) \, \mu(di) = \iint f(i, \omega) \, \mu(di) \, P(d\omega).$$

Let $(I \times \Omega, \Sigma, \nu)$ be a Fubini extension of $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \mu \otimes P)$. Then, essentially pairwise independent families satisfy an exact version of the Law of Large Numbers. The following is a special case of [32, Corollary 2.9].

Proposition B.2 (Exact Law of Large Numbers)**.** *Let* $f : I \times \Omega \to \mathbb{R}$ *be* ν*-integrable. If* $f(i, \cdot)$*,* $i \in I$ *are essentially pairwise i.i.d. with a distribution having mean* m*, then* $\int f(\cdot, \omega) \, d\mu = m$ *for* P*-almost all* $\omega \in \Omega$*.*

Next, we turn to the existence of such a setting. The space $(I \times \Omega, \Sigma, \nu)$ is called *rich* if there exists a Σ-measurable function $f : I \times \Omega \to \mathbb{R}$ such that $f(i, \cdot)$, $i \in I$ are essentially pairwise i.i.d. with a uniform distribution on $[0, 1]$. By composing f with a suitable function, it then follows that $(I \times \Omega, \Sigma, \nu)$ supports essentially pairwise i.i.d. families with any given distribution.

Lemma B.3. *There exist atomless probability spaces* $(I, \mathcal{I}, \mu)$ *and* $(\Omega, \mathcal{F}, P)$ *such that* $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \mu \otimes P)$ *admits a rich Fubini extension.*

This is part of the assertion of [32, Proposition 5.6] which also shows that one can take $I = [0, 1]$ and $\Omega = \mathbb{R}^{[0,1]}$. The main result of Sun and Zhang [33] shows that, in addition, one can take μ to be an extension of the Lebesgue measure (but not the Lebesgue measure itself). A different construction is presented by Podczeck [29].

---

[5] Here we use the convention that the product σ-field $\mathcal{I} \otimes \mathcal{F}$ is completed.

[6] That is, f is measurable for the ν-completion of Σ and $\int |f| \, d\nu < \infty$.

References

[1] E. Bayraktar and A. Cohen. Analysis of a finite state many player game using its master equation. *Preprint arXiv:1707.02648v1*, 2017.

[2] E. Bayraktar and Y. Zhang. A rank-based mean field game in the strong formulation. *Electron. Commun. Probab.*, 21(72):1–12, 2016.

[3] A. Bensoussan, M. H. M. Chau, and S. C. P. Yam. Mean field games with a dominating player. *Appl. Math. Optim.*, 74(1):91–128, 2016.

[4] A. Bensoussan, J. Frehse, and S. C. P. Yam. *Mean field games and mean field type control theory.* Springer Briefs in Mathematics. Springer, New York, 2013.

[5] C. Bertucci. Optimal stopping in mean field games, an obstacle problem approach. *Preprint arXiv:1704.06553v1*, 2017.

[6] D. Cao. Racing under uncertainty: boundary value problem approach. *J. Econom. Theory*, 151:508–527, 2014.

[7] P. Cardaliaguet, F. Delarue, J. M. Lasry, and P. L. Lions. The master equation and the convergence problem in mean field games. *Preprint arXiv:1509.0250v1*, 2015.

[8] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I.* Springer, 2017.

[9] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications II.* Springer, 2017.

[10] R. Carmona, F. Delarue, and D. Lacker. Mean field games of timing and models for bank runs. *Appl. Math. Optim.*, 76(1):217–260, 2017.

[11] R. Carmona and P. Wang. Finite state mean field games with major and minor players. *Preprint arXiv:1610.05408v1*, 2016.

[12] R. Elie, T. Mastrolia, and D. Possamai. A tale of a principal and many many agents. *Preprint arXiv:1608.05226v1*, 2016.

[13] R. Elie and D. Possamai. Contracting theory with competitive interacting agents. *Preprint arXiv:1605.08099v1*, 2016.

[14] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. *Appl. Math. Optim.*, 68(1):99–143, 2013.

[15] G. M. Grossman and C. Shapiro. Dynamic R & D competition. *Econ. J.*, 97(386):372–387, 1987.

[16] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 205–266. Springer, Berlin, 2011.

[17] C. Harris and J. Vickers. Racing with uncertainty. *Rev. Econom. Stud.*, 54(1):1–21, 1987.

[18] M. Huang. Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.*, 48(5):3318–3353, 2009/10.

[19] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007.

[20] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006.

[21] H. K. Koo, G. Shim, and J. Sung. Optimal multi-agent performance measures for team contracts. *Math. Finance*, 18(4):649–667, 2008.

[22] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[23] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.

[24] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

[25] E. Lazear and S. Rosen. Rank-order tournaments as optimum labor contracts. *J Polit Econ.*, 89(5):841–64, 1981.

[26] D. A. Malueg and S. O. Tsutsui. Dynamic R & D competition with learning. *RAND J. Econ.*, 28(4):751–772, 1997.

[27] S. Nadtochiy and M. Shkolnikov. Particle systems with singular interaction through hitting times: application in systemic risk modeling. *Preprint arXiv:1705.00691v1*, 2017.

[28] M. Nutz. A mean field game of optimal stopping. *Preprint arXiv:1605.09112v1*, 2016.

[29] K. Podczeck. On existence of rich Fubini extensions. *Econom. Theory*, 45(1-2):1–22, 2010.

[30] J. F. Reinganum. A dynamic game of R and D: patent protection and competitive behavior. *Econometrica*, 50(3):671–688, 1982.

[31] M. Shkolnikov. Large systems of diffusions interacting through their ranks. *Stochastic Process. Appl.*, 122(4):1730–1747, 2012.

[32] Y. Sun. The exact law of large numbers via Fubini extension and characterization of insurable risks. *J. Econom. Theory*, 126(1):31–69, 2006.

[33] Y. Sun and Y. Zhang. Individual risk and Lebesgue extension without aggregate uncertainty. *J. Econom. Theory*, 144(1):432–443, 2009.

[34] M. Vojnović. *Contest Theory.* Cambridge University Press, 2015.