# Strategic Exploration: Preemption and Prioritization*

Qingmin Liu† Yu Fu Wong‡

October 29, 2020

**Abstract**

This paper provides a model of strategic exploration in which two competing players simultaneously explore a set of alternatives over time to study search dynamics, payoff divisions, and distributions of discovery time. The strategic tension is between *preemption*, i.e., the incentive to explore alternatives before the opponent explores them in future, and *prioritization*, i.e., the incentive to explore alternatives with the highest success probabilities. When players are symmetric in their speed of exploration, each player randomizes to level his opponent's posterior belief down, making greedy strategies best responses. In the asymmetric case, the weak player's strategy is greedy, but the strong player randomizes over alternatives with different posteriors and captures a share of payoff disproportionately larger than his share of exploration capacity. The weak player conducts *extensive* instead of *intensive* exploration, i.e., he covers many alternatives as the strong player does but never explores any alternative with cumulative probability one. The overall discovery time decreases in asymmetry in the first-order stochastic dominance sense.

*We thank Dilip Abreu, Nageeb Ali, Arjada Bardhi, Kalyan Chatterjee, Boyan Jovanovic, Navin Kartik, Rohit Lamba, George Mailath, Jacopo Perego, Weijie Zhong, and seminar participants at Columbia, Penn, NYU, and Stony Brook for helpful comments. This research is supported by NSF grant SES-1824328.

†Department of Economics, Columbia University. E-mail: ql2177@columbia.edu.

‡Department of Economics, Columbia University. E-mail: yw2942@columbia.edu.

# Contents

# 1 Introduction

This paper presents and analyzes a model of strategic exploration in which players compete over time to explore a set of alternatives in order to find good candidates. The benchmark model is in its simplest form. The set of alternatives is the unit interval, and at most one alternative is good. Two players sharing a common prior explore the set of alternatives independently and simultaneously in continuous time, without directly observing each other's exploration activities. They each face a capacity constraint on the measure of the alternatives explored per unit of time. Whoever finds the good alternative first will claim its return exclusively; they split the return equally in the case of simultaneous discovery.[1]

The model captures a trade-off between preemption and prioritization. Players are incentivized by competition to preempt their opponents' future explorations, but the presence of multiple available alternatives with different success probabilities incentivizes them to prioritize over the more promising ones. The trade-off remains the driving force of equilibrium behavior in more complex settings where the tools and concepts developed for the benchmark model are still applicable, such as asymmetric capacity constraints, general spaces of alternatives, multiple (or a continuum of) good alternatives, more than two players, and general time preferences. We will also consider gradual learning where the outcome of each alternative arrives at a Poisson rate controlled by resources allocated to the alternative. We show that the distributional strategy in the benchmark model of instantaneous learning materializes as a pure strategy in effort allocations over alternatives. This extension is an instance of tractable strategic bandit problems with a continuum of arms; in contrast, conventional strategic bandit models with more than two arms are often intractable.

The trade-off of preemption and prioritization appears in many dynamic economic search problems. These alternatives can be scientific experiments, research ideas, product designs, job opportunities, dating partners, etc. The primary objective of this paper is to study the strategic and payoff implications of this trade-off in a simple and flexible model and to develop useful analytical tools.

Preemption motives under capacity constraint necessitate randomization. In the model of continuous time and a continuous space of alternatives, the conventional approach of defining both pure and mixed strategies based on intuitions from discrete models is no longer useful (and a fully discrete model is not tractable either), as it does not take advantage of the continua to simplify the analysis. We redefine the pure strategy on the outcome space, which

[1] The analysis and the results remain the same if, instead, the first finder enjoys a larger return, and they split the return in an arbitrary way in the case of simultaneous discovery.

stipulates whether or not an alternative will be tried by each moment in time. We then define a notion of *distributional* strategies, and show that it is an appropriate representation of a mixture over pure strategies. The simple language of distributional strategies in continuous time and a continuous action space is convenient for describing the evolution of posterior beliefs and facilitating equilibrium analysis that is intractable in discrete problems.

The model has a unique Nash equilibrium in distributional strategies. Both players randomize over unexplored alternatives with the same highest posteriors from an expanding set, in a way so that the posteriors are "leveled off" gradually over time. The equilibrium strategy is "greedy" in that it searches only alternatives with the highest posterior. Myopic optimization happens to be a best response, but the equilibrium is not a result of myopic optimization. Players must consider not only the myopic value of each alternative (measured by its posterior density) but also the option value (which is determined by how intensively his opponent will explore certain alternatives in the future). Indeed, the expanding set of alternatives and its expansion over time are determined by the dynamic equilibration of preemption and prioritization. The necessity of randomization drives a wedge between equilibrium exploration and optimal exploration that minimizes the time of discovery. Without concerns of preemption, the latter would be achieved by a coordinated exploration that prioritizes alternatives according to their prior probabilities. The unique equilibrium in distributional strategies, determined fully by the level of the highest posterior beliefs over time, remains the same when the space of alternatives is multidimensional, or when there are multiple good alternatives that are independent and identically distributed, even though the implementation of mixed strategies necessarily depends on the space of alternatives.

The unique Nash equilibrium is symmetric if players have symmetric exploration capacity, i.e., the measure of alternatives a player explores per unit of time. With asymmetric capacities, the Nash equilibrium in distributional strategies is again unique. But the equilibrium posteriors facing the two players can no longer be the same due to their different speeds of exploration. The strategy of the strong player (she) is not greedy because she randomizes over alternatives with different posteriors. Her strategy levels off the posterior for the weak player (he) who plays a greedy strategy, i.e., he randomizes over alternatives with the highest posterior. The weak player covers all alternatives, but never explores any single one with cumulative probability one, even though he has the capacity to do so for a subset of them. In other words, the weak player conducts *extensive* exploration instead of *intensive* exploration. In contrast, the strong player always explores all alternatives with probability one. The unique equilibrium has an interesting payoff implication. We give a simple formula for the equilibrium payoffs, showing that the strong player enjoys a larger share in payoff than in exploration capacity. If the exploration capacity of the weak player is fraction $\alpha < 1$ of that of his opponent, the strong player enjoys fraction $(1-\alpha)+\frac{1}{2}\alpha$ of the surplus in equilibrium. It is as if the strong player exclusively enjoys fraction $1-\alpha$, before the two split the remainder equally. If we fix the total capacity and vary its division between the players, the overall discovery time decreases in asymmetry in the first-order stochastic dominance sense, but the preemption incentive plays a non-vanishing role in slowing down discovery even when the weak player's capacity is vanishingly small.

## Related literature

The paper relates to several branches of active research. Optimal exploration of an unknown area is a classic problem in operations research and computer science because of its applications in navigation algorithms and robotics, where inefficiency typically arises from the path dependence of exploration of physical locations.[2] This literature has so far neglected game-theoretic aspects of explorations, although many applications involve interactions of multiple agents. We do not consider the path-dependence in exploring physical locations. Instead, the alternatives can be research ideas, scientific experiments, job opportunities, etc., all of which are of interest in economics.

Fershtman and Rubinstein (1997) study a discrete-time finite-alternative search problem with preemption. As their analysis demonstrates, the discrete problem is intractable once we go beyond the case of a uniform prior. Under a uniform prior, however, there does not exist a trade-off between prioritization and preemption, which is the central strategic issue we study in this paper.[3] Matros, Ponomareva, Smirnov, and Wait (2019) consider a discrete-time continuum-alternative search model, but it is qualitatively different because the preemption incentive is assumed away so that an equilibrium can be found in pure strategies; furthermore, simultaneous discoveries are assumed to destroy the prize, so the rent is dissipated completely in the pure-strategy equilibrium by a Bertrand-style competition.[4]

Chatterjee and Evans (2004) embed a two-alternative model of treasure hunting in a dy-

[2] See, for example, the surveys by Kleinberg (1994) and Megow, Mehlhorn, and Schweitzer (2012).

[3] Fershtman and Rubinstein (1997) discuss the case where exactly one alternative has a different probability of success from the rest. Establishing an equilibrium is intricable even in this case. They make the observation that, under some parameter values, this alternative cannot be searched first with probability 1.

[4] Their uniqueness fails, however, without the assumption of symmetric or Markovian strategies. In fact, there exists a continuum of equilibria, one for each welfare level between the social optimal and zero with arbitrary division of rent. Matros and Smirnov (2016) and de Roos, Matros, Smirnov, and Wait (2018) look into variants of this model with observable actions and with/without coordination.

namic R&D game with Poisson bandits. Klein and Rady (2011) analyze a continuous-time model of a negative correlated bandit, in which one of the two arms contains a prize and two players share a common value instead of competing with each other. Again there is no preemption–prioritization trade-off. The canonical models of strategic experimentation by Keller, Rady, and Cripps (2005) and Bolton and Harris (1999), which capture a trade-off of exploration and exploitation, are very useful for studying dynamic search, learning, and innovation. These models are also used to investigate incentive designs and information revelation for experimentation in various economic applications.[5] Most models in this literature are variants of a one-armed bandit, with one risky alternative and a safe default option, which preclude the rich set of experiments available in many search and innovation processes. In addition, the issue of dynamic prioritization is assumed away in a model with only one risky arm. Multiple-armed strategic bandit problems, even when the arms are independent, are largely intractable. This paper focuses on exploration with the trade-off of preemption and prioritization. With an appropriate formulation of strategies, a continuous-time continuum-armed strategic bandit introduced in this paper overcomes these analytical difficulties and hence has a potential for applied research.[6]

# 2 Model

## 2.1 Setup

A good alternative $x$, if it exists, is in $X := [0, 1]$ endowed with the Lebesgue measure. The prior density $f$ over $X$ for it is bounded and strictly positive almost everywhere.[7] Let $\pi := \int_X f(x)dx \in (0, 1]$ denote the prior probability that the good alternative exists. Two players compete to find it in continuous time $t \in T := [0, 1]$ without observing each other's activities. Each player can explore a subset of alternatives at each moment in time, and the good alternative is discovered if it is contained in the set. The players face an identical capacity constraint: the Lebesgue measure of the set of alternatives explored by

[5] See, Halac, Kartik, and Liu (2016), McClellan (2019), Awaya and Krishna (2019), etc.

[6] There are several other novel alternative models of experimentation with multiple arms. Kremer, Mansour, and Perry (2014) present a model of two-alternative experimentation with a continuum of payoff levels. Jovanovic and Rob (1990) and Callander (2011) offer models with a rich set of alternatives where individuals draw observations on a sample path of a Brownian motion, and Wong (2020) extends the model to forward-looking agents, but they do not consider strategic interactions of multiple players. Chen (2020) introduces a game of experimentation in which players have random opportunities to revise their actions subject to probabilistic breakdowns.

[7] The boundedness of the prior density can be relaxed at the cost of additional notations.

each player per unit time is constant and normalized to 1. The first player to find the good alternative exclusively claims its return—a payoff of 1—and the two split it equally in the case of simultaneous discovery. In all other cases, their payoffs are normalized to 0. There is no temporal discounting. Once the good alternative is found, the discovery is publicly announced and the game is over (alternatively, we can assume that the discovery is not made public but the prize is taken away from the good alternative once discovered).

## 2.2 Strategy

The definition of strategies exploits two ideas: an *outcome function approach* to overcoming the indeterminacy of continuous-time strategies and a *distributional approach* to handling randomization.

Before introducing the formal definition, it is useful to explain why the intuition of "exploring one alternative per period" inherited from a discrete problem does not work. First, this measure-preserving bijection between continuous time and continuous alternative space is not a tractable object. Secondly, a good definition should naturally cover the case of an abstract space of alternatives, but the existence of a measure-preserving bijection is not always ensured, let alone equilibrium analysis in such bijections. Thirdly, the exploration activity should be a correspondence that specifies for each moment in time a *set* of alternatives to be explored. Instead of defining such a correspondence and dealing with an uncountable union (over time) of measurable sets, it is more useful to specify outcome functions that ambiguously determines the play of the game.

We shall formally define pure and mixed strategies, explaining the outcome function approach. An eager reader may skip directly to Section 2.2.3 for the definition of distributional strategies.

### 2.2.1 Pure Strategy

A pure strategy is a function $\sigma : T \times X \to \{0, 1\}$ which specifies that an alternative $x \in X$ is explored at or before time $t \in T$ if and only if $\sigma(t, x) = 1$.

**Definition 1.** *A function $\sigma : T \times X \to \{0, 1\}$ is a* ***pure strategy*** *if it satisfies the following four conditions:*

1. *Initial condition: $\sigma(0, \cdot) = 0$;*

2. *Monotonicity and right-continuity: $\sigma(\cdot, x)$ is non-decreasing and right continuous for all $x \in X$;*

*3. Measurability:* $\sigma(t, \cdot)$ *is measurable for all* $t \in T$*;*

*4. Capacity constraint:* $\int_X \sigma(t, x) dx = t$ *for all* $t \in T$*.*

The four conditions are obvious requirements. The *initial condition* states that none of the alternatives has been explored at the beginning of the game. The *monotonicity* condition requires that, once an alternative has been explored, it will have been explored in the future as well. The right continuity property, similar to that of a cumulative distribution function, guarantees that the time at which an alternative $x \in X$ is explored,

$$\tau(x) := \min\{t : \sigma(t, x) = 1\}, \tag{2.1}$$

is well defined. The *measurability* condition furthers that $\tau : X \to T$ is a measurable function and $\tau^{-1}(t)$, the set of alternatives to be explored at time $t$, is a measurable set. It is this induced map $\tau^{-1}$ that instructs how the player should actually search, and hence the strategy $\sigma$ is operational. The set of alternatives explored up to any time $t$ is given by

$$\{x : \sigma(t, x) = 1\} = \bigcup_{s \in [0,t]} \tau^{-1}(s), \tag{2.2}$$

which is measurable.[8] Lastly, the *capacity constraint* describes how quickly a player can explore the space of alternatives. The measure of alternatives explored per unit of time is normalized to 1. It implies the terminal condition: $\sigma(1, \cdot) = 1$ almost everywhere. Similar to the study of Lebesgue-measurable functions, we identify a strategy $\sigma$ up to a stationary null set of $X$.[9]

### 2.2.2 Mixed Strategy

We introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to describe randomization following Aumann (1964). In addition, weak measurability is used to accommodate continuous time and continuous space. This allows the usage of the (Gelfand–Pettis) weak integral, which extends

[8] The measurability of the uncountable union of measurable sets in Equation (2.2) is an example of why it is more convenient to define a pure strategy in this way than to define $\tau^{-1}$ directly as in discrete problems. Note also that the measurability of the mapping $\tau^{-1}$ is irrelevant because it does not concern outcomes and payoffs.

[9] Simon and Stinchcombe (1989) point out the *indeterminacy* of the continuous-time strategy when it is written as a function of histories (including a player's own past actions). The same issue is present here as available actions at each moment in time depends on previous exploration. This issue is overcome by the definition of pure strategies as paths of outcomes, and the assumption of unobservability of the opponent's actions.

Lebesgue integral to functional spaces.[10]

**Definition 2.** *A* ***mixed strategy*** *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a function* $\sigma : \Omega \times T \times X \to \{0, 1\}$ *that satisfies the following conditions:*

1. *Initial condition:* $\sigma(\omega, 0, \cdot) = 0$ *for all* $\omega \in \Omega$*;*

2. *Monotonicity and right-continuity:* $\sigma(\omega, \cdot, x)$ *is non-decreasing and right continuous for all* $\omega \in \Omega$ *and* $x \in X$*;*

3. *Measurability: mapping* $x \mapsto \sigma(\cdot, t, x) \in L^2(\Omega)$ *is weakly measurable for all* $t \in T$*;*

4. *Capacity constraint:* $\int_X \sigma(\omega, t, x)dx = t$ *for all* $t \in T$*, where the integral is the weak integral.*

The initial condition and the monotonicity and right-continuity condition are the realization-by-realization generalizations of their counterparts in Definition 1 of pure strategies. The measurability condition and the capacity constraint in Definition 2 for mixed strategies, however, are weaker than their counterparts. They must hold when averaged over any measurable event in $\Omega$ that has a positive probability under $\mathbb{P}$, but not necessarily at each $\omega \in \Omega$. If $\Omega$ is a singleton, a mixed strategy reduces to a pure strategy as defined in Definition 1.

With realization $\omega \in \Omega$, an alternative $x \in X$ is explored at or before time $t \in T$ if and only if $\sigma(\omega, t, x) = 1$, analogously to the pure strategy case. The stochastic time at which alternative $x$ is searched is a random variable on $\Omega$ given by $\tau(\omega, x) = \min\{t : \sigma(\omega, t, x) = 1\}$.

### 2.2.3 Distributional Strategy

A mixed strategy in Definition 2 is explicit about randomization in continuous time and continuous action space and hence is an operational instruction that the players can follow, but it is not amenable to analysis. The essence of randomization is the induced distributions that span the space of all feasible outcomes and payoffs.[11] We therefore define a notion of distributional strategy that specifies a distribution for each time $t \in T$.

A distributional strategy is a function $\rho : T \times X \to [0, 1]$ which specifies that an alternative $x \in X$ is explored by time $t \in T$ with probability $\rho(t, x)$.

**Definition 3.** *A function* $\rho : T \times X \to [0, 1]$ *is a* ***distributional strategy*** *if it satisfies the following conditions:*

[10] See Talagrand (1984) for an exposition of Gelfand–Pettis integral.

[11] This insight is already evident in Milgrom and Weber (1985) for continuous type spaces in games with incomplete information.

1. *Initial condition:* $\rho(0, \cdot) = 0$;

2. *Monotonicity and right-continuity:* $\rho(\cdot, x)$ *is non-decreasing and right continuous for all* $x \in X$;

3. *Measurability:* $\rho(t, \cdot)$ *is measurable for all* $t \in T$;

4. *Capacity constraint:* $\int_X \rho(t, x)dx = t$ *for all* $t \in T$.

The four natural conditions need no further explanation. A distributional strategy $\rho$ reduces to a pure strategy if $\rho(t, x) \in \{0, 1\}$, following a comparison of Definition 1 and Definition 3. The language of distributional strategies allows us to describe and analyze the play of the game, but it does not offer a specific instruction to the players on how to play it. The representation theorem below fills the gap. It shows that mixed strategies and distributional strategies are outcome equivalent for all $(t, x) \in T \times X$, and hence it is without loss of generality to study equilibria and payoffs in terms of distributional strategies.

**Theorem 1.** *1. For every mixed strategy* $\sigma : \Omega \times T \times X \to \{0, 1\}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, the function* $\rho$*, defined by* $\rho(t, x) := \mathbb{E}[\sigma(\cdot, t, x)]$ *for* $t \in T$ *and* $x \in X$*, is a distributional strategy that represents* $\sigma$*, i.e., the probability of an alternative* $x \in X$ *being explored by* $t \in T$ *under* $\sigma$ *is* $\rho(t, x)$*.*

*2. For every distributional strategy* $\rho$*, there exists a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a mixed strategy* $\sigma : \Omega \times T \times X \to \{0, 1\}$ *that implements* $\rho$*, i.e.,* $\mathbb{E}[\sigma(\cdot, t, x)] = \rho(t, x)$ *for all* $t \in T$ *and* $x \in X$*.*

Theorem 1 is proved in Appendix A.1 by construction. We remark that the mixed-strategy implementation is not unique.

**Example 1.** *Under the Lebesgue probability space on* $\Omega = [0, 1]$*, both mixed strategies* $\sigma_1(\omega, x, t) := \mathbf{1}_{\{\mathrm{frac}(x-\omega) \le t\}}$ *and* $\sigma_2(\omega, x, t) := \mathbf{1}_{\{\mathrm{frac}(x+\omega) \le t\}}$*, where* $\mathbf{1}$ *denotes the indicator function and* $\mathrm{frac}(y) = y - \lfloor y \rfloor$ *denotes the fractional part of* $y$*, implement the same distributional strategy* $\rho(t, x) = t$*. Intuitively, according to the mixed strategy* $\sigma_1$*, a player searches to the right starting from* $x = \omega$*, where* $\omega$ *is drawn uniformly from the interval* $[0, 1]$*, and continues at* $x = 0$ *after reaching* $x = 1$*, while according to* $\sigma_2$*, a player searches in the other direction starting from the same starting point.*

## 2.3 Payoff

We abuse the notion by using $-i$ to denote player $i$'s opponent. Given a profile of distributional strategies $(\rho_i, \rho_{-i})$, player $i$'s expected payoff $u_i(\rho_i, \rho_{-i})$ is

$$\int_X \int_T f(x)(1-\rho_{-i}(t,x))d_t\rho_i(t,x)dx + \frac{1}{2}\int_X f(x)\sum_{t\in D_x}\Delta_t\rho_i(t,x)\Delta_t\rho_{-i}(t,x)dx. \tag{2.3}$$

The first term in (2.3) is player $i$'s expected payoff from discovering the good alternative before her opponent. The time integral is the Lebesgue–Stieltjes integral with respect to the non-decreasing and right-continuous function $t \mapsto \rho_i(t,x)$.[12] $f(x)$ is the probability that $x$ is the good alternative, $1-\rho_{-i}(t,x)$ is the probability density that player $-i$ has not explored $x$ by time $t$, and $d_t\rho_i(t,x)$ is the instantaneous probability that player $i$ explores $x$ at time $t$. The second term in (2.3) is player $i$'s expected payoff from simultaneously discovering the good alternative with her opponent. For each $x$, the set $D_x \subset T$ is the at most countable set of discontinuity points of both $\rho_i(\cdot,x)$ and $\rho_{-i}(\cdot,x)$. The function $\Delta_t\rho_i(t,x) := \rho_i(t,x) - \rho_i(t^-,x)$ is the jump measure of the distributional strategy $\rho_i(\cdot,x)$ on $T$, where $\rho_i(t^-,x) := \lim_{s\uparrow t}\rho_i(s,x)$. Thus, $\sum_{t\in D_x}\Delta_t\rho_i(t,x)\Delta_t\rho_{-i}(t,x)$ is the probability of simultaneous exploration of the alternative $x$. The second integral is well defined as the integrand can be written as the limit of measurable functions.

## 2.4 Belief Updating

Starting with a prior density $f$ and knowing player $-i$ adopts a distributional strategy $\rho_{-i}$, player $i$'s posterior density that $x$ is a good alternative right after $t$ is

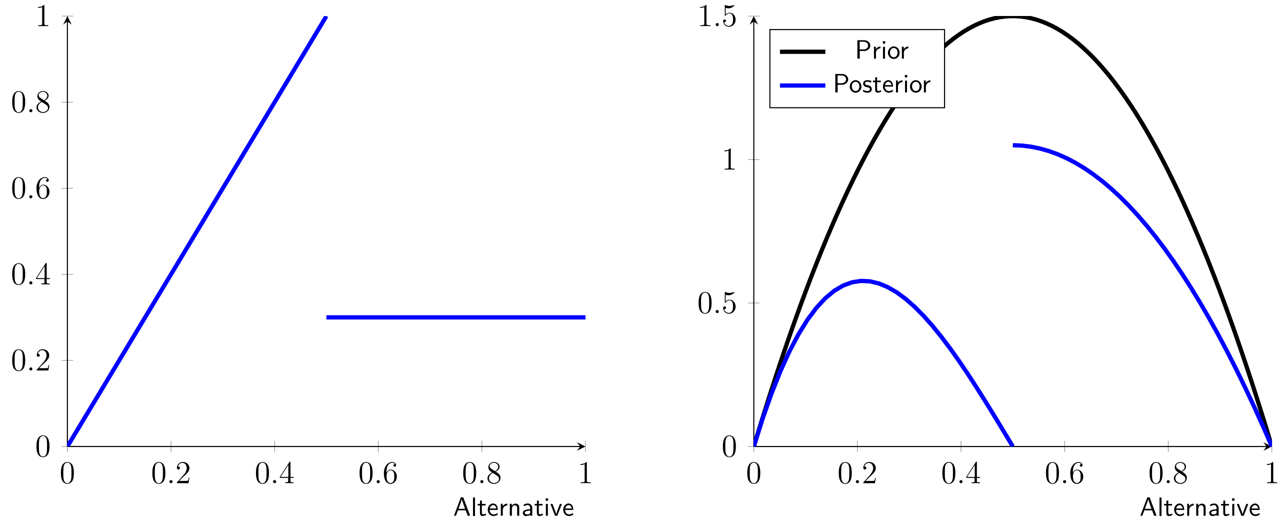
$$g_{-i}(t,x) := (1-\rho_{-i}(t,x))f(x). \tag{2.4}$$

We call $g_{-i}(t,x)$ player $i$'s (unnormalized) **posterior distribution** over $X$ at time $t$. We use the subscript "$-i$" because the posterior conditions *only* on the strategy of player $-i$, but, of course, player $i$ also knows her own search outcome which is not taken into account in computing $g_{-i}$.

By definition, $g_{-i}(0,x) = f(x)$ and $g_{-i}(t,x)$ is non-increasing in $t$ for each $x \in X$. Intuitively, as the alternatives are explored over time by her opponent $-i$, the posterior distribution is pushed lower and lower, until it vanishes at $t = 1$. If an alternative $x$ is explored with higher probability by time $t$, it will have a lower density at $t$. Figure 2.1

[12] The Lebesgue–Stieltjes measure is obtained from $\mu((s,t]) := \rho_i(t,x) - \rho_i(s,x)$ for all $0 \le s < t \le 1$.

illustrates the relationship between the distributional strategy, the prior distribution, and the posterior distribution.




(a) distributional strategy $\rho_{-i}$ at a fixed time

(b) prior $f$ and posterior $g_{-i}$ at a fixed time

Figure 2.1: A distributional strategy $\rho_{-i}$ and the posterior $g_{-i}$, at a fixed time.

# 3 Equilibration of Preemption and Prioritization

A profile of distributional strategies $(\rho_i, \rho_{-i})$ is a **Nash equilibrium** if $u_i(\rho_i, \rho_{-i}) \geq u_i(\rho_i', \rho_{-i})$ for each $i \in \{1, 2\}$ and distributional strategy $\rho_i'$.

## 3.1 Necessity of Randomization

A pure strategy is a distributional strategy with $\rho_i(t, x) \in \{0, 1\}$. We shall now argue that, due to preemption motives, no player can play a pure strategy in a Nash equilibrium. Facing any pure strategy $\rho_{-i}$ (e.g., it can be a strategy that prioritizes alternatives according to their prior densities), player $i$ can stay "one-step ahead" of her opponent. More precisely, for $\epsilon > 0$, let $A^\epsilon := \{x : \rho_{-i}(\epsilon, x) = 1\}$ be the alternatives that will be explored by player $-i$ by time $\epsilon$. Consider the following pure-strategy response for player $i$:

$$\rho_i^\epsilon(t, x) := \begin{cases} \rho_{-i}(t - \epsilon, x), & \text{if } x \notin A^\epsilon, \\ \rho_{-i}((t - (1 - \epsilon))^+, x), & \text{if } x \in A^\epsilon, \end{cases} \tag{3.1}$$

where $(\cdot)^+$ is the positive part. With $\rho_i^\epsilon$, player $i$ will beat her opponent by $\epsilon$ time to the prize if it is not in $A^\epsilon$. When $\epsilon$ is close to 0, player $i$'s payoff from this response is close to $\pi$ and her opponent's payoff is close to 0. Thus in the putative equilibrium, player $-i$'s payoff is 0. However, player $-i$ can always imitate player $i$'s equilibrium strategy to guarantee a payoff $\pi/2$. Therefore, equilibrium must involve randomization. We shall show that the trade-off between prioritization and preemption is resolved in such a way that equilibrium randomization levels off posterior densities over time.

## 3.2 Leveling Strategy

We first construct a Nash equilibrium in distributional strategies where the probability of simultaneous discovery is zero, and then show that this is the unique Nash equilibrium. When the probability of simultaneous discovery is zero, it follows from (2.3) that

$$u_i(\rho_i, \rho_{-i}) = \int_X \int_T (1 - \rho_{-i}(t,x)) f(x) d_t\rho_i(t,x)dx = \int_X \int_T g_{-i}(t,x) d_t\rho_i(t,x)dx. \quad (3.2)$$

Notice the posterior $g_{-i}(t,x)$ is also player $i$'s expected **flow payoff** from exploring an alternative $x$ at a given time $t$ if she has not explored $x$ yet.

In the candidate equilibrium, the posterior distribution $g_i(t,x)$ levels the prior $f(x)$ over time as illustrated in Figure 3.1. As such, we shall call it the *leveling strategy*. To construct

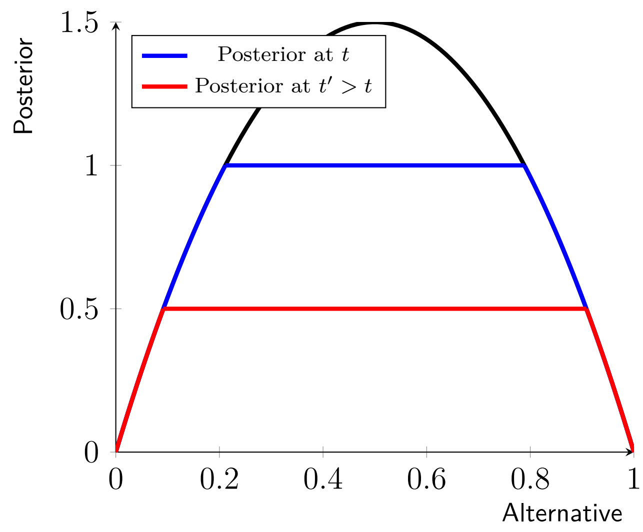



Figure 3.1: The equilibrium posterior $g_i(t,x)$ levels the prior $f(x)$ over time.

this strategy, it is instructive to first pin down the highest posterior as a function of time. By the definition of posterior distribution in Equation (2.4), player $-i$'s strategy $\rho_{-i}$ and its induced posterior $g_{-i}$ have the following relationship:

$$\rho_{-i}(t,x) = 1 - \frac{g_{-i}(t,x)}{f(x)} \tag{3.3}$$

for all $t \in T$ and $x \in X$. A function $\bar{g} : T \to [0, \sup f]$ is called the **leveling function** if it satisfies

$$\int_X \left(1 - \frac{\bar{g}(t)}{f(x)}\right) \mathbf{1}_{\{f(x) \geq \bar{g}(t)\}}(x)dx = t \tag{3.4}$$

for all $t \in T$. The motivation for Equation (3.4) is as follows. At time $t$, $\bar{g}(t)$ is the highest level of the posterior $g_{-i}(t, \cdot)$ across $x \in X$. It is achieved on $\{x \in X : f(x) \geq \bar{g}(t)\}$ as the posterior is bounded above by the prior. Noting that the distributional strategy is related to the posterior by $1 - \frac{\bar{g}(t)}{f(x)}$ in Equation (3.3), Equation (3.4) corresponds to the capacity constraint in Definition 3.

**Lemma 1.** *The leveling function exists and is unique, absolutely continuous, strictly decreasing, and convex.*

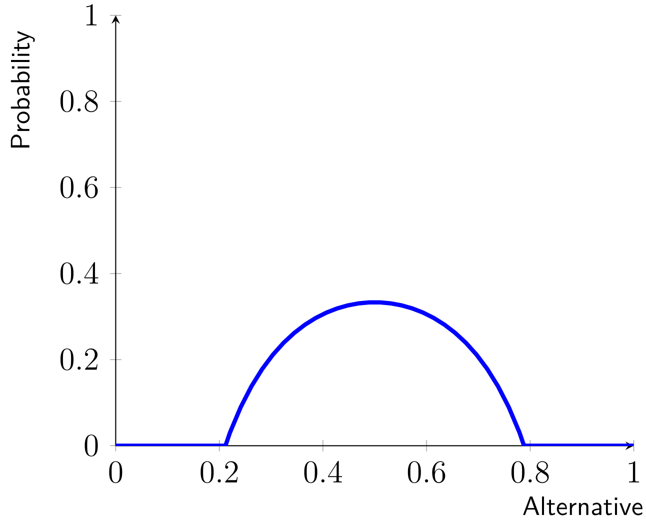
The **leveling strategy** $\bar{\rho} : T \times X \to [0, 1]$ is defined by the leveling function $\bar{g}$ through

$$\bar{\rho}(t,x) := \left(1 - \frac{\bar{g}(t)}{f(x)}\right) \mathbf{1}_{\{f(x) \geq \bar{g}(t)\}}(x) \tag{3.5}$$

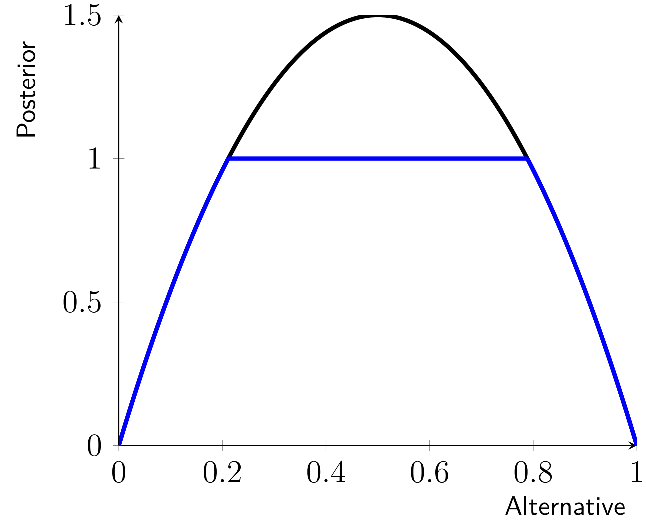
for all $t \in T$ and $x \in X$. It remains to verify that $\bar{\rho}$ is a well-defined distributional strategy. It is straightforward to check that $\bar{\rho}$ satisfies the initial condition. The function $(x, y) \mapsto \left(1 - \frac{y}{f(x)}\right) \mathbf{1}_{\{f(x) \geq y\}}$ is continuous and decreasing in $y$. Together with the continuity and monotonicity of $\bar{g}$, this property implies that $\bar{\rho}$ is continuous in $t$ and satisfies the monotonicity and right-continuity condition. The function is also measurable in $x$ and hence $\bar{\rho}$ satisfies the measurability condition. Finally, $\bar{\rho}$ respects the capacity constraint by Equation (3.4).

With an abuse of notation, we denote the posterior density induced by the leveling strategy at time $t$ as $\bar{g}(t, x) := (1 - \bar{\rho}(t, x))f(x)$ and call it the **leveling posterior** at $t$. We reiterate that it is player $i$'s leveling strategy $\bar{\rho}$ that levels player $-i$'s posterior $\bar{g}$.

The relationship between the leveling strategy, the prior, and the leveling posterior is demonstrated in Figure 3.2. The implementation of exploration over time is illustrated in Figure 3.3.

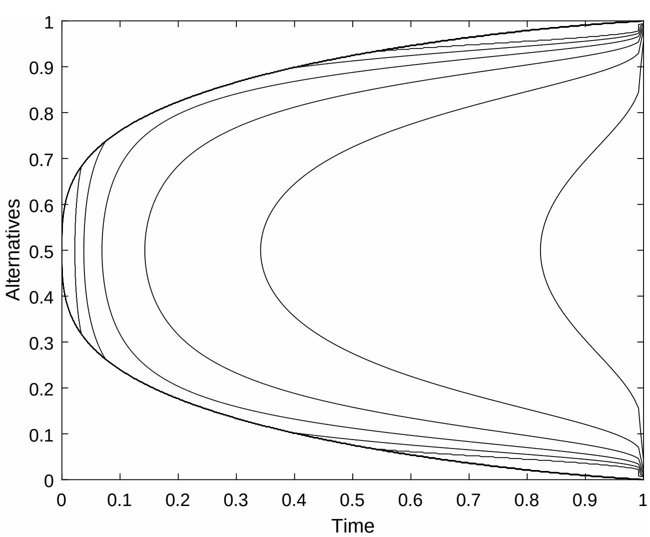



(a) leveling strategy

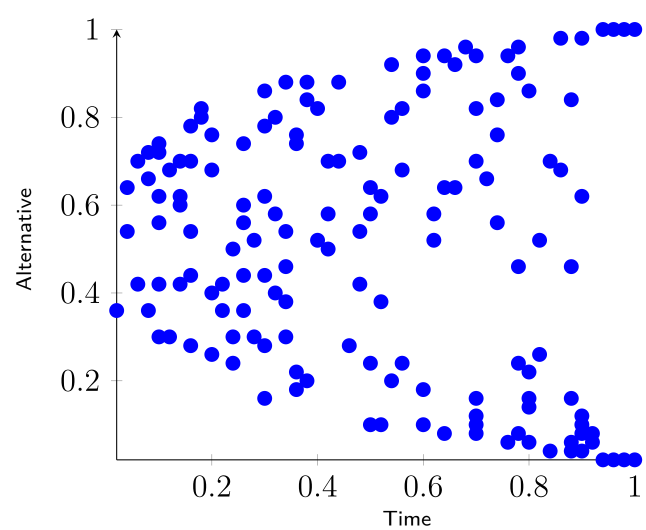



(b) leveling posterior

Figure 3.2: The leveling strategy $\bar{\rho}$ and the posterior density $\bar{g}$, at a fixed time




(a) contour plot of the intensity of exploration $\partial_t \bar{\rho}$




(b) a discretized realization of exploration

Figure 3.3: Exploration over time according to the leveling strategy $\bar{\rho}$

### 3.3 Unique Equilibrium

The special randomization by the leveling strategy $\bar{\rho}$ constitutes the unique Nash equilibrium of the game.

**Theorem 2.** *The profile $(\bar{\rho}, \bar{\rho})$ is the unique Nash equilibrium in distributional strategies.*

We remark on several additional features of this equilibrium.

**Remark 1.** (**Greedy Strategy**) The equilibrium strategy is greedy in that a player will explore only alternatives $x$ with the highest posterior (or flow payoff) $\bar{g}(t)$ at each time $t$. Indeed, given the leveling posterior, there are multiple greedy strategies, but the equilibrium strategy arises from the preemption–prioritization trade-off instead of myopic optimization. As we shall see in Section 4, the unique equilibrium of the asymmetric problem is that the less capable player plays a greedy but not leveling strategy, and the more capable player plays a leveling but not greedy strategy. The latter levels the posteriors for her opponent but she randomizes over alternatives with different posteriors.

**Remark 2.** (**Discovery Time**) Because of the greedy strategy, the highest level of posterior at $t$, $\bar{g}(t)$, is also the instantaneous probability with which each player makes a discovery at $t$. Therefore, the probability that a discovery is made by time $t$ is

$$P(t) = \int_0^t 2\bar{g}(s)ds. \tag{3.6}$$

Due to randomization, in terms of distributions of discovery time, this equilibrium first-order stochastically dominates (i.e., is slower than) a coordinated exploration that prioritizes alternatives according to their prior densities without the preemption motive.

**Remark 3.** (**Payoff Sharing Rule**) Since the unique equilibrium strategy is $t$-continuous, Theorem 2 remains true for arbitrary payoff-sharing in the case of simultaneous discovery, and it continues to hold *verbatim* even if the discoverer enjoys a larger, but not exclusive, share of the prize.

**Remark 4.** (**Space of Alternatives**) By Equation (2.4), there is a one-to-one correspondence between functions of posterior densities and distributional strategies. Since only the *level* of the highest posterior belief matters for the equilibrium characterization in distributional strategies, Theorem 2 continues to hold *verbatim* when the space of alternatives is multidimensional, with the same leveling strategy given by Equation (3.5) and the same leveling function given by Equation (3.4) (the actual exploration activity of course depends on the space of alternatives).

**Remark 5.** (**Multiple Good Alternatives**) Theorem 2 continues to hold *verbatim* if there are multiple good alternatives that are independent and identically distributed accordingly to $f$, assuming a player's payoff is a weighted sum of the payoffs he receives from each alternative. In fact, the result continues to hold with a continuum of good alternatives if in addition we assume that all good alternatives corresponding to the same position $x$ are discovered at once when $x$ is searched (this has a zero probability if the number of good alternatives is finite). A formal treatment of a continuum of independent random variables again requires the notion of weak integral.

### 3.3.1 Verification

We shall show that $(\bar{\rho}, \bar{\rho})$ is a Nash equilibrium. Suppose that player $-i$ plays the leveling strategy $\bar{\rho}$. The probability of simultaneous discovery is zero since the strategy is $t$-continuous. By Equation (3.2), the payoff of player $i$ with strategy $\rho_i$ is

$$u_i(\rho_i, \bar{\rho}) = \int_X \int_T \bar{g}(t, x) d_t \rho_i(t, x) dx.$$

We shall show that $u_i(\rho_i, \bar{\rho}) \leq u_i(\bar{\rho}, \bar{\rho})$ for any strategy $\rho_i$. By construction in Equation (3.5), the integrand $\bar{g}(t, x)$ is bounded from above by the leveling function $\bar{g}(t)$. Therefore,

$$u_i(\rho_i, \bar{\rho}) \leq \int_X \int_T \bar{g}(t) d_t \rho_i(t, x) dx. \tag{3.7}$$

For $x \in X$, let $\kappa_x \in \Delta(T)$ be the Lebesgue–Stieltjes measure induced by $\rho_i(\cdot, x)$. Then $\kappa_x([0, t]) = \rho_i(t, x)$. For any $t \in T$,

$$\int_X \kappa_x([0, t]) dx = \int_X \rho_i(t, x) dx = t, \tag{3.8}$$

where the last equality follows from the capacity constraint of the distributional strategy $\rho_i$. Thus $\int_X \kappa_x dx \in \Delta(T)$ is the Lebesgue measure by the Caratheodory extension theorem. Therefore,

$$\int_X \int_T \bar{g}(t) d_t \rho_i(t, x) dx = \int_X \int_T \bar{g}(t) d\kappa_x dx = \int_T \bar{g}(t) d \left( \int_X \kappa_x dx \right) = \int_T \bar{g}(t) dt, \tag{3.9}$$

where the first equality is by the definition of Lebesgue–Stieltjes integration and the second equality follows from Fubini's theorem. Combining (3.7) and (3.9), player $i$'s payoff of playing $\rho_i$ is bounded by

$$u_i(\rho_i, \bar{\rho}) \leq \int_T \bar{g}(t)dt. \tag{3.10}$$

The payoff of playing $\bar{\rho}$ is

$$u_i(\bar{\rho}, \bar{\rho}) = \int_X \int_T \bar{g}(t,x) d_t\bar{\rho}(t,x)dx = \int_X \int_T \bar{g}(t) d_t\bar{\rho}(t,x)dx = \int_T \bar{g}(t)dt, \tag{3.11}$$

where the first equality is due to Equation (3.2), the second equality holds because $\bar{\rho}(t,x) > 0$ only if $\bar{g}(t,x) = \bar{g}(t)$, and the third equality follows from Equation (3.9).

Combining (3.10) and (3.11), we have shown that $u_i(\rho_i, \bar{\rho}) \leq u_i(\bar{\rho}, \bar{\rho})$. By symmetry, the profile of leveling strategies $(\bar{\rho}, \bar{\rho})$ is a Nash equilibrium and each player obtains an expected equilibrium payoff

$$u_i(\bar{\rho}, \bar{\rho}) = \frac{1}{2}\int_X f(x)dx = \pi/2. \tag{3.12}$$

### 3.3.2 Uniqueness

The formal proof of uniqueness is contained in Appendix A.3. The intuition can be understood as follows. First, note that $u_i(\rho_i, \bar{\rho}) \leq \pi/2$ for any $\rho_i$ and hence, $u_i(\bar{\rho}, \rho_i) \geq \pi/2$ in the constant-sum game. The latter inequality means that the leveling strategy $\bar{\rho}$ guarantees a payoff of $\pi/2$. It therefore suffices to show that any non-leveling strategy cannot ensure $\pi/2$. For such a strategy, there must exist an interval of time over which the posterior declines faster than the leveling posterior in a positive measure set and slower in another positive measure set, due to the capacity constraint. One can then modify the leveling strategy to preempt this strategy by prioritizing the former set at the expense of the latter, in the spirit of the "one-step-ahead" strategy in Equation (3.1), to achieve a higher payoff.

# 4 Impact of Asymmetric Capacities

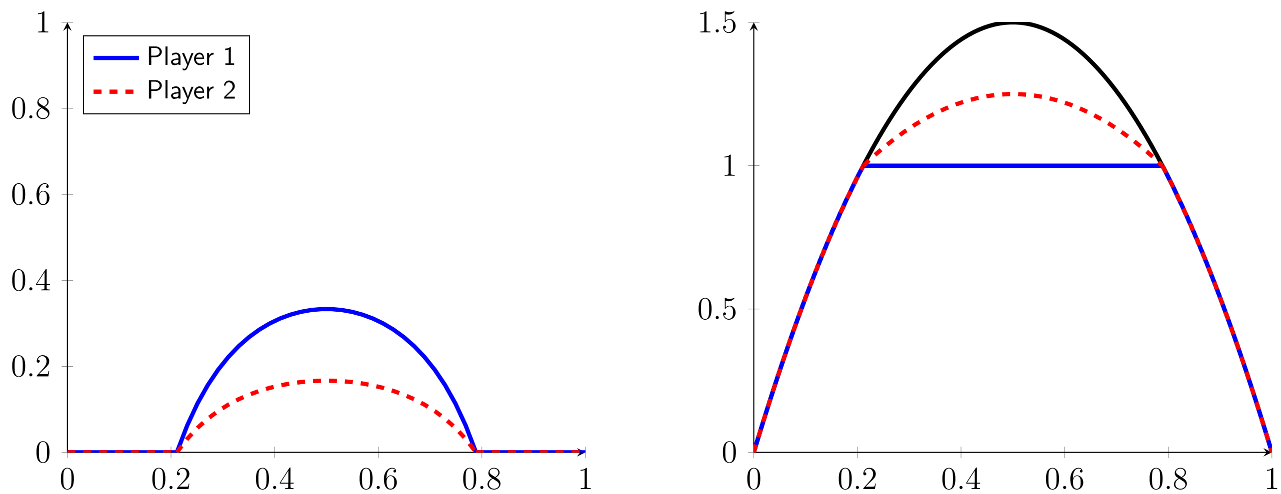
Suppose that the two players have different capacities: player 1 can explore measure 1 of alternatives per unit of time, while player 2 explores measure $\alpha \in (0,1]$ of alternatives per unit of time. That is, player 1 (the "strong" player, she) is more capable or more resourceful than player 2 (the "weak" player, he) at exploration. We will refer to $\alpha$ as player 2's *capacity*. Thus player 2's distributional strategy $\rho_2^\alpha : T \times X \to [0,1]$ should satisfy the new capacity constraint

$$\int_X \rho_2^\alpha(t,x)dx = \alpha t \tag{4.1}$$

for all $t \in T$, in addition to the first three conditions in Definition 3. Note that the strong player, player 1, will have explored all alternatives $x \in X$ by time $t = 1$ and so exploration will end by then regardless of $\alpha$. Hence, the time domain $T = [0,1]$ remains the relevant one.

## 4.1 Unique Equilibrium

Consider the leveling strategy $\bar{\rho}$ in the symmetric case. Then $\alpha\bar{\rho}$ is a distributional strategy for player 2 that satisfies the new capacity constraint of Equation (4.1). It is special in that any given alternative $x$ is explored with probability $\alpha$. The strategy $\alpha\bar{\rho}$ no longer levels the posterior as $\bar{\rho}$ does. Figure 4.1 demonstrates the difference.




(a) equilibrium strategies

(b) posterior densities

Figure 4.1: The strategy profile $(\bar{\rho}, \alpha\bar{\rho})$ and the corresponding posterior densities, at a fixed time. Player 2 is half as capable as player 1, i.e., $\alpha = 1/2$.

**Theorem 3.** *The profile of distributional strategies $(\bar{\rho}, \alpha\bar{\rho})$ is the unique Nash equilibrium of the game with asymmetric players. In equilibrium, player 1's payoff is $\left(1 - \frac{1}{2}\alpha\right)\pi$ and player 2's payoff is $\frac{1}{2}\alpha\pi$.*

Both the formal proof and the proof idea for Theorem 3 are relegated to Appendix A.4. The dynamics and payoff distributions in this unique equilibrium have several interesting features.

**Remark 6.** (**Extensive vs. Intensive Exploration**) Although the two players differ in their speeds of exploration, they randomize over the same expanding set of alternatives, modulo the alternatives they have respectively explored. This is clear as $\bar{\rho}$ and $\alpha\bar{\rho}$ have the same support. So the weak player explores as extensively as the strong player does. In addition, although *a priori* player 2 can choose to explore a subset of alternatives with probability greater than $\alpha$, he will not do so in equilibrium. Therefore, the weak player conducts *extensive* exploration instead of *intensive* exploration when facing a disadvantageous capacity constraint.[13]

**Remark 7.** (**Greedy vs. Non-Greedy**) As in the symmetric case, the strong player 1's leveling strategy $\bar{\rho}$ leads to a leveling posterior $\bar{g}(t,x) = (1-\bar{\rho}(t,x))f(x)$ that flattens the prior density over time. This posterior makes the greedy strategy a best response for the weak player 2. In contrast to the symmetric case, the posterior $(1-\alpha\bar{\rho}(t,x))f(x)$ induced by the equilibrium strategy $\alpha\bar{\rho}$ of player 2 is not flat over the set of alternatives $\{x \in X : f(x) \geq \bar{g}(t)\}$ as observed in Figure 4.1; the strong player is always more optimistic than the weak player. Thus player 1 cannot play a greedy strategy. However, player 1's posterior decreases at a constant rate across these alternatives. The equal option values allow her to randomize. Therefore, the strong player's equilibrium strategy is leveling but not greedy, and the weak player's strategy is greedy but not leveling.

**Remark 8.** (**Disproportionate Payoff Division**) The ratio of players' equilibrium payoffs $(2-\alpha):\alpha$ is greater than the ratio of capacities $1:\alpha$. It is as if player 1 monopolizes a fraction $1-\alpha$ of the total surplus and splits the remaining fraction evenly with player 2. For example, if $\alpha = \frac{1}{2}$, i.e., the strong player is twice as fast as the weak player, the payoff share is $(\frac{3}{4}, \frac{1}{4})$. The strong player's payoff is three times as much as the weak player's. In comparison, in a three-player game in which the more resourceful player is split into two equal selves, the payoff share will be $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ in a symmetric three-player equilibrium, as we shall show in the next section. The excess payoff of player 1 beyond the sum of her two selves is due to the pooled information of the two: knowing which alternatives have been explored by herself, player 1 does better as one big player than an ensemble of smaller selves who may duplicate explorations by their peers.

[13] If we model each alternative as a Poisson process with the arrival rate controlled by resources allocated to the alternative (see Section 5.1), the weak player would cover as many alternatives as the strong player does, but would spend less resource on each alternative.

## 4.2 Discovery Time

We now investigate the impact of asymmetry from a different angle. We fix the total exploration capacity to be 2 (as in the symmetric case), and varies the asymmetry between the two players. Its payoff impact has been clarified in Remark 8. The question is how this asymmetry affects the distribution of discovery time of the good alternative, which is a more relevant measurement of social value of exploration. We show that asymmetry speeds up the discovery.

Formally, let $\beta \in [1, 2)$ and consider the game in which the strong player has capacity $\beta$ and the weak player has capacity $2 - \beta \in (0, 1]$. Theorem 3 implies that in the unique equilibrium, player 1 plays the strategy $\rho_1(t, x) = \bar{\rho}(\beta t, x)$ with a leveling posterior $\bar{g}(\beta t)$ and player 2 plays $\rho_2(t, x) = \frac{1-\beta}{\beta}\bar{\rho}(\beta t, x)$.

For $t \geq 1/\beta$, the strong player has exhausted all alternatives so the probability of discovery by $t$ is $P_\beta(t) = \pi$. For $t \in [0, 1/\beta]$, the probability of discovery by $t$ is given by

$$P_\beta(t) = \int_X (f(x) - \bar{g}(\beta t)) \mathbf{1}_{\{f(x) \geq \bar{g}(\beta t)\}} dx + (2 - \beta)t\bar{g}(\beta t).$$

The probability is computed using the following idea: first let the strong player search until she levels the posterior down to $\bar{g}(\beta t)$, and then let the weak player search for a period with length $t$ conditional on the strong player's failure to make a discovery.

**Theorem 4.** *The distribution of discovery time is decreasing in $\beta$ in the first-order stochastic dominance sense.*

Figure 4.2 demonstrates how the the distribution of discovery time varies with $\beta$. The symmetric division of exploration capacity (when competition, and hence the preemptive incentive, is the most intense) leads to the slowest discovery time $P_1$, and the coordinated exploration (when only prioritization incentive prevails) has the fastest discovery time $P_2$. Distributions of discovery time under asymmetric divisions of search capacity lie in between the two extremes.

**Remark 9.** (**Discontinuity**) It is interesting to note that $P_\beta$ converges to $P_1$ pointwise as $\beta \to 1$, but it does not converge to $P_2$ as $\beta \to 2$. The discontinuity at $\beta = 2$ arises because the strong player must level the posteriors (with leveling posterior $\bar{g}(2t)$ in the limit) for the weak player to randomize. So the incentive of preemption does not vanish even in the limit. But there is no discontinuity in payoffs as the strong player explores all alternatives when the weak player's capacity vanishes. Indeed, in the limit of one monopolistic player with capacity 2, all exploration strategies are optimal.

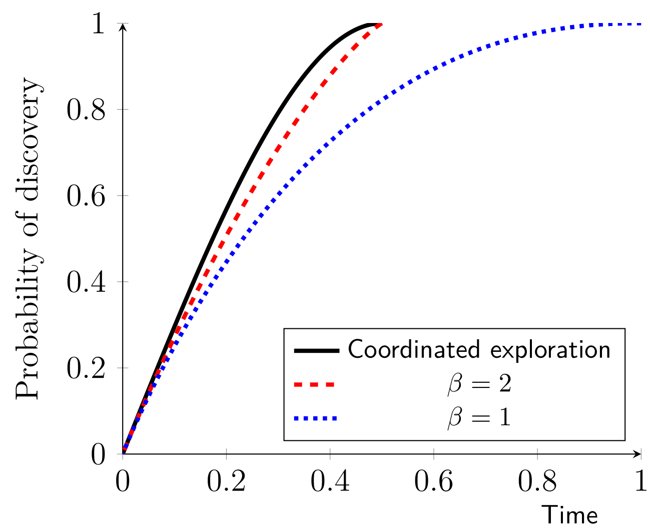



Figure 4.2: Distribution of discovery time for the coordinated exploration and equilibria with different divisions of capacity.

# 5 Further Results

The main model studies strategic exploration in the simplest setting, and its analysis and results are invariant to seemingly non-vacuous modifications as seen in Remarks 3–5. The model is amenable to enrichment to study search and learning in various settings, and we shall work out several cases that are of interest for applications, assuming symmetric capacities. We shall also establish comparative statics results with respect to the number of players and variations of prior distributions.

## 5.1 Poisson Learning

In the main model, players learn whether a given alternative is good instantaneously upon exploration. In this section, we consider the case where discovery is not immediate; instead, each alternative $x \in X$ represents a Poisson process with conclusive signals, and the arrival rate is independent of whether $x$ is a good alternative or not.[14] The arrival of signals is controlled by each player: conditional on the player not receiving any signals from a given alternative, the more resources she has spent on the alternative, the higher the probability of signal arrival from that alternative is for her. Specifically, we assume that the arrival rate of signals from an alternative is proportional to the flow rate of resources spent on the alternative. A player cannot explore all alternatives instantly; the total amount of resources

[14]See, e.g., Akcigit and Liu (2015) for a model of a binary confirmatory process with independent arrival rate.

available per unit of time is normalized to 1.

Formally, let $r(t,x)$ be the cumulative amount of resources a player spends on the alternative $x$ by time $t$ *conditional on receiving no signal before* $t$. The probability of signal arrival, which reveals the state of $x$, by time $t$ is $1-e^{-r(t,x)}$. If $r(\cdot,x)$ is differentiable in $t$, the partial derivative $\partial_t r(t,x)$ is the arrival rate of the potentially non-stationary Poisson process associated with alternative $x$.

The function $r : T \times X \to \mathbb{R}_+ \cup \{\infty\}$ is a *resource allocation strategy* if the following four conditions are satisfied:

1. Initial condition: $r(0,x) = 0$ for all $x \in X$;
2. Monotonicity and right-continuity: $r(\cdot,x)$ is increasing and right-continuous for all $x \in X$;
3. Measurability: $r(t,\cdot)$ is measurable for all $t \in T$;
4. Capacity constraint: $\int_X \left(1-e^{-r(t,x)}\right) dx = t$ for all $t \in T$.

The first three conditions need no further explanations. The capacity constraint warrants an elaboration. The strategy $r(t,x)$ may not be the actual amount of resources spent on alternative $x$ by time $t$, because it conditions on no signal arrival. No more resources will be spent once a signal arrives. The expected amount of resources spent on alternative $x$ by time $t$ is given by

$$\int_0^{r(t,x)} e^{-q} dq = 1-e^{-r(t,x)},$$

where $e^{-q}$ is the probability of no signal being received given the cumulative amount of resources $q$. As the arrival of Poisson signals is independent across alternatives, the law of large numbers applies and the capacity constraint applies to the expected amount of resources.[15]

Given player $i$'s resource allocation strategy $r_i$, we write the probability of signal arrival from alternative $x$ by time $t$ as $\rho_i(t,x) := 1-e^{-r_i(t,x)} \in [0,1]$. With this one-to-one relationship between $r_i$ and $\rho_i$, it is immediate that $r_i$ is a resource allocation strategy if and only if $\rho_i$ is a distributional strategy that satisfies the four conditions in Definition 3.

[15]As more signals arrive, the player concentrates his resource on the remaining alternatives, expediting the arrival of Poisson signals over those alternatives. At $t=1$, the capacity constraint can be satisfied if and only if $r(1,x)=\infty$ almost everywhere. The relevant temporal domain for this game is equal to the expected resources spent to obtain a signal, i.e., $T=[0,1]$.

Player $i$'s expected payoff from a profile $(r_i, r_{-i})$ is the same as $u_i(\rho_i, \rho_{-i})$ as defined in Equation (2.3). Therefore, the game with Poisson learning is isomorphic to the main model with instantaneous arrival. Define a resource allocation strategy

$$\bar{r} := -\log(1-\bar{\rho}),$$

where $\bar{\rho}$ is the leveling strategy. The following result is immediate from Theorem 2.

**Theorem 5.** *With Poisson learning, the profile of resource allocation strategy $(\bar{r}, \bar{r})$ is the unique Nash equilibrium.*

We should note that Remarks 3–5 are valid here. The equilibrium characterization in resource allocation strategies is invariant to the space of alternatives, payoff-sharing rules, and the multiplicity of good alternatives.

## 5.2 Multiple Players

Suppose that there are $n$ symmetric players, where $n \geq 2$. The set of strategies available to each of them is still given by Definition 3. Denote a distributional strategy of player $i$ by $\rho_i$, and define $\rho_{-i} := 1 - \prod_{j \neq i}(1-\rho_j)$. Then $\rho_{-i}(t,x)$ is the probability that $x$ is searched up to time $t$ by at least one of player $i$'s opponents. The posterior induced by $\rho_{-i}$ is given by $g_{-i}(t,x) := (1-\rho_{-i}(t,x))f(x)$. With this notation, the payoff of player $i$ is again given by Equation (2.3). The probability of simultaneous discovery involving player $i$ is zero if $\rho_i$ is continuous in $t$. In that case, the payoff can be simplified to Equation (3.2).

We consider the symmetric strategy profile such that the posterior $g_{-i}$ is leveling for every player $i$. More precisely, let $\bar{g} : T \to [0, \sup f]$ be the leveling function defined implicitly by

$$\int_X \left(1 - \left(\frac{\bar{g}(t)}{f(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{f(x) \geq \bar{g}(t)\}}(x) dx = t \tag{5.1}$$

for all $t \in T$. The proof of existence and uniqueness of $\bar{g}$ is analogous to the proof of Lemma 1. By Equation (5.1), the leveling strategy

$$\bar{\rho}(t,x) := \left(1 - \left(\frac{\bar{g}(t)}{f(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{f(x) \geq \bar{g}(t)\}}(x) \tag{5.2}$$

satisfies the capacity constraint.

**Theorem 6.** *The profile of distributional strategies $(\bar{\rho}, ..., \bar{\rho})$ is the unique symmetric Nash equilibrium of the game with $n$ players.*

Theorem 6 characterizes the unique symmetric equilibrium, but does not establish the uniqueness of *Nash equilibrium* which does not hold for $n > 2$. We present an example of an asymmetric equilibrium in which equilibrium payoffs are also unequal for symmetric players.

**Example 2.** *Take a uniform prior $f \equiv 1$ and consider $n = 5$. Partition the space $X = [0, 1]$ into two halves: $X_1 := [0, \frac{1}{2})$ and $X_2 := [\frac{1}{2}, 1]$. Define the strategies $\rho_i$ for player $i \in \{1, 2\}$ as follows:*

$$\rho_i(t, x) = \begin{cases} \min\{2t, 1\}, & \text{if } x \in X_i, \\ \max\{2t - 1, 0\}, & \text{if } x \in X_{i+1}, \end{cases}$$

*where $X_3 := X_1$. Player 1 searches uniformly over the left half $X_1$ until the alternatives are exhausted at $t = \frac{1}{2}$, and then the other half $X_2$. Player 2 searches in the reverse order. It can be verified that $(\rho_1, \rho_2, \rho_1, \rho_2, \rho_1)$ is a Nash equilibrium. Since the discovery must occur before $t = \frac{1}{2}$, the distribution of discovery time in this equilibrium is different from the one described in Theorem 6. Moreover, despite symmetric capacities, the equilibrium payoffs are asymmetric: player 1, 3, and 5 enjoy an expected payoff of $\frac{1}{6}$ while player 2 and 4 have an expected payoff of $\frac{1}{4}$.*

As the number of players increase, the effect of preemption will increase and there will be more duplicated search. But overall discovery is hastened.

**Theorem 7.** *In this class of symmetric equilibria, the distribution of discovery time is decreasing in $n$ in the first-order stochastic dominance sense.*

## 5.3 Impact of Prior Beliefs

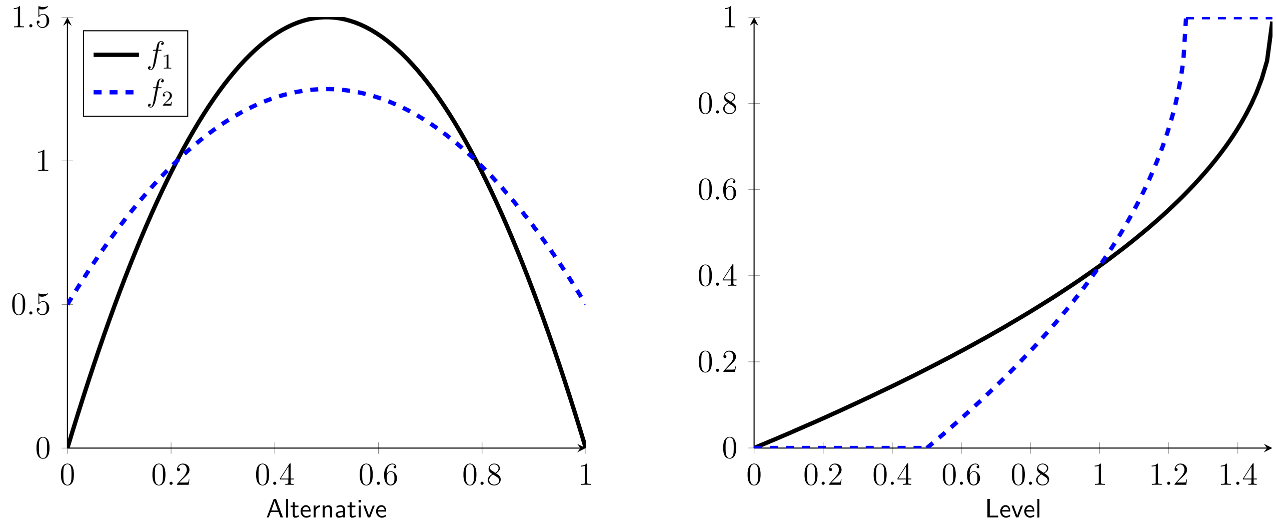
With the probability of existence of a good alternative $\pi$ fixed, how does the prior distribution $f$ affect the equilibrium discovery? In equilibrium, players concentrate their effort on the area with the highest equilibrium posteriors. If the prior distribution is more spread out or evenly distributed over $X$, the good alternative should be discovered later. We shall now formalize this notion of comparative statics.

Denote $\lambda$ as the Lebesgue measure. For any prior distribution $f$, let $\lambda \circ f^{-1}$ be the pushforward measure over $\mathbb{R}_+$. Note that $\lambda \circ f^{-1}(\mathbb{R}_+) = \lambda([0, 1]) = 1$, $\lambda \circ f^{-1}$ is a probability measure.

The distribution induced by $\lambda \circ f^{-1}$ has an expectation of $\int_{\mathbb{R}_+} y d\lambda \circ f^{-1} = \int_{[0,1]} f(x) d\lambda = \pi$. By definition, the pushforward measure is the distribution of prior density. For example, when the good alternative, if exists, is uniformly distributed over $X$, i.e., $f(x)$ is a constant, then $\lambda \circ f^{-1}$ assigns probability 1 to a single point. This is the case where the good alternative is most evenly distributed over $X = [0, 1]$.

**Definition 4.** *Let $f_1$ and $f_2$ be two prior distributions. We say that $f_2$ is* ***more even than*** *$f_1$ if $\lambda \circ f_1^{-1}$ is a mean-preserving spread of $\lambda \circ f_2^{-1}$.*

Figure 5.1 below illustrates the partial order of evenness.




(a) prior density functions

(b) pushforward measures

Figure 5.1: Two distributions $f_1$ and $f_2$, where $f_2$ is more even than $f_1$.

**Theorem 8.** *Consider the symmetric Nash equilibrium $(\bar{\rho}, ..., \bar{\rho})$ for the game with $n$ players. If $f_2$ is more even than $f_1$, then the distribution of equilibrium discovery time associated with $f_2$ first-order stochastically dominates that associated with $f_1$, i.e., the good alternative is discovered later with $f_2$ than with $f_1$.*

The comparative statics result on discovery times remains true if $f_2$ second-order/first-order stochastically dominates $f_2$, but the comparison is more meaningful holding the mean $\pi$ fixed.

## 5.4 Time Preferences

There are two ways to model time preferences. Suppose that both players have a common time preference function $\delta : T \to (0, +\infty)$. That is, a discovery at $t$ is worth $\delta(t)$ at time

0. The main model and all of its variants correspond to the case where $\delta$ is constant. The equilibria we characterized previously remain equilibria for any time preference function $\delta$, but we do not establish a general uniqueness result. We can actually show that the equilibrium is unique if $\delta$ is strictly increasing, i.e., the prize is getting bigger over time. This class of time preference describes, for example, two competing companies work on a drug with a rising price. In the unique equilibrium, the probability of discovery by each player $\bar{g}$ decreases over time. The players fail to coordinate and wait till the prize becomes larger—a manifestation of destructive preemption incentives—despite their preference for delayed discoveries.

The second way to model time preferences is to introduce a flow cost of time $c > 0$. We need to make the explicit assumption that a discovery is made public immediately and ends the game. It is readily verified that if $\pi = 1$ and $c < 1/2$, the equilibrium we have constructed in Theorem 2 remains an equilibrium.

# References


Akcigit, Ufuk and Qingmin Liu (2015). "The role of information in innovation and competition". In: *Journal of the European Economic Association* 14.4, pp. 828–870.

Aumann, Robert J (1964). "Mixed and behavior strategies in infinite extensive games". In: *Advances in Game Theory* 52, pp. 627–650.

Awaya, Yu and Vijay Krishna (2019). "Startups and Upstarts: Disadvantageous Information in R&D". In: *Working Paper, University of Rochester and Penn State University.*

Bolton, Patrick and Christopher Harris (1999). "Strategic experimentation". In: *Econometrica* 67.2, pp. 349–374.

Callander, Steven (2011). "Searching and learning by trial and error". In: *American Economic Review* 101.6, pp. 2277–2308.

Chatterjee, Kalyan and Robert Evans (2004). "Rivals' search for buried treasure: competition and duplication in R&D". In: *RAND Journal of Economics*, pp. 160–183.

Chen, Yi (2020). "A revision game of experimentation on a common threshold". In: *Journal of Economic Theory* 186, p. 104997.

de Roos, Nicolas, Alexander Matros, Vladimir Smirnov, and Andrew Wait (2018). "Shipwrecks and treasure hunters". In: *Journal of Economic Dynamics and Control* 90, pp. 259–283.

Fershtman, Chaim and Ariel Rubinstein (1997). "A simple model of equilibrium in search procedures". In: *Journal of Economic Theory* 72.2, pp. 432–441.

Halac, Marina, Navin Kartik, and Qingmin Liu (2016). "Optimal contracts for experimentation". In: *The Review of Economic Studies* 83.3, pp. 1040–1091.

Jovanovic, Boyan and Rafael Rob (1990). "Long waves and short waves: Growth through intensive and extensive search". In: *Econometrica*, pp. 1391–1409.

Keller, Godfrey, Sven Rady, and Martin Cripps (2005). "Strategic experimentation with exponential bandits". In: *Econometrica* 73.1, pp. 39–68.

Klein, Nicolas and Sven Rady (2011). "Negatively correlated bandits". In: *The Review of Economic Studies* 78.2, pp. 693–732.

Kleinberg, Jon M (1994). "On-line Search in a Simple Polygon". In: *SODA*. Vol. 94, pp. 8–15.

Kremer, Ilan, Yishay Mansour, and Motty Perry (2014). "Implementing the wisdom of the crowd". In: *Journal of Political Economy* 122, pp. 988–1012.

Matros, Alexander, Natalia Ponomareva, Vladimir Smirnov, and Andrew Wait (2019). "Search without observability". In: *Working Paper, University of Sydney.*

Matros, Alexander and Vladimir Smirnov (2016). “Duplicative search”. In: *Games and Economic Behavior* 99, pp. 1–22.

McClellan, Andrew (2019). “Experimentation and approval mechanisms”. In: *Working Paper, University of Chicago.*

Megow, Nicole, Kurt Mehlhorn, and Pascal Schweitzer (2012). “Online graph exploration: New results on old and new algorithms”. In: *Theoretical Computer Science* 463, pp. 62–72.

Milgrom, Paul R and Robert J Weber (1985). “Distributional strategies for games with incomplete information”. In: *Mathematics of operations research* 10.4, pp. 619–632.

Simon, Leo K. and Maxwell B. Stinchcombe (1989). “Extensive form games in continuous time: Pure strategies”. In: *Econometrica* 57, pp. 1171–1214.

Talagrand, Michel (1984). “Pettis integral and measure theory”. In: *Memoirs of the American Mathematical Society.*

Wong, Yu Fu (2020). “Spatial Experimentation”. In: *Working Paper, Columbia University.*

# A Proofs

## A.1 Proof of Theorem 1

*Proof.* The first part of Theorem 1 follows directly from the definitions of the weak measurability and the weak integral so its proof is omitted.

We show the second part by construction. By the Kolmogorov extension theorem, there exists a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ in which random variables $r_x \sim U(0,1)$ are i.i.d. across $x \in X$. Define candidate mixed strategy $\sigma(\omega, t, x) := \mathbf{1}_{\{r_x(\omega) \leq \rho(t,x)\}}(\omega, t, x)$. By construction, it satisfies the initial condition and the monotonicity and right-continuity condition, and implements the search density.

Fix $t \in [0,1]$. We shall show that $x \mapsto \sigma(\cdot, t, x) \in L^2(\Omega)$ is weak-integrable over the Lebesgue measure with integral $t$. The dual space of $L^2(\Omega)$ is isomorphic to itself by the Riesz representation theorem. Every element $Z \in L^2(\Omega)$ operates on $Y \in L^2(\Omega)$ via $ZY = \mathbb{E}[ZY]$.

Since $\sigma(\cdot, x, t) \in \{0, 1\}$, its variance is bounded by $1/4$. The pairwise independence of $\{\sigma(\cdot, t, x) : x \in X\}$ implies that $\{\sigma(\cdot, t, x) - \rho(t, x) : x \in X\}$ is an orthogonal set in $L_2(\Omega)$. By the Bessel theorem, we have that for any countable collection $\{x_n\}$

$$\frac{1}{4}\mathbb{E}[Z^2] \geq \sum_{n=1}^{\infty} \frac{\left(\mathbb{E}[Z(\sigma(\cdot, t, x_n) - \rho(t, x_n))]\right)^2}{4Var[\sigma(\cdot, t, x)]} \geq \sum_{n=1}^{\infty} \left(\mathbb{E}[Z(\sigma(\cdot, t, x_n) - \rho(t, x_n))]\right)^2$$

which implies that $\mathbb{E}[Z(\sigma(\cdot, t, x) - \rho(t, x))] = 0$, or $\mathbb{E}[Z\sigma(\cdot, t, x)] = \mathbb{E}[Z]\rho(t, x)$, everywhere except a countable set. Therefore, the function $\sigma(\omega, t, \cdot)$ is weakly measurable and has weak-integral $\int_X \rho(t, x) dx = t$, satisfying the capacity constraint. □

## A.2 Proof of Lemma 1

*Proof.* For $y \in [0, \sup f]$, let

$$h(y) := \int_X \left(1 - \frac{y}{f(x)}\right) \mathbf{1}_{\{f(x) \geq y\}}(x) dx. \tag{A.1}$$

For $x \in X$, the integrand $\left(1 - \frac{y}{f(x)}\right) \mathbf{1}_{\{f(x) \geq y\}}(x)$ is decreasing in $y$ and strictly so for $f(x) > y$, which has positive measure for $y < \sup f$. Thus, $h$ is strictly decreasing. In addition, the integrand is continuous, and therefore $h$ is continuous by the dominated convergence theorem. The convexity of $h$ also follows from that of the integrand.

The function $h$ is continuous and strictly decreasing with $h(0) = 1$ and $h(\sup f) = 0$.

Therefore, there exists a unique, continuous, and strictly decreasing function $\bar{g} = h^{-1}$ that solves Equation (3.4). Since $h$ is strictly decreasing, its inverse $\bar{g}$ is also convex. The absolute continuity of $\bar{g}$ follows from its continuity and convexity. □

## A.3 Proof of Theorem 2

**Overview.** The verification of the Nash equilibrium is provided in the main text. The uniqueness follows from three lemmas in this section. Lemma 2 states that, in equilibrium, each player can only search over the set of alternatives, which we call the upper contour set of $f$, that the leveling strategy randomizes over. Otherwise, he will enjoy payoff lower than $\pi/2$ against the leveling strategy.

Lemma 3 is key to Theorem 2. It states that, in equilibrium, the posterior declines fastest on the upper contour set. If instead the posterior declined slower in some subset of the upper contour set over the other during a period of time, the opponent could devise a *modified leveling strategy* that searches the former in place of the latter just before the period, and vice versa just after the period. The modification generalizes the "one-step ahead" strategy in Equation (3.1). The opponent's strategy would then preempt the player's strategy and yield higher payoff than the leveling strategy, which cannot be true in equilibrium in a constant-sum game.

Lemma 3 has two useful implications. Corollary 1 establishes the $t$-continuity of the equilibrium strategy. According to Lemma 3, the posterior decreases fastest over the upper contour set. Therefore, the posterior for those alternatives must all drop discontinuously. This violates the capacity constraint. Corollary 2 states that the decrease in posterior must be equal over the upper contour set. This is immediate by applying Lemma 3 twice.

Lemma 4 computes the equilibrium posterior within the upper contour set. As the decrease in posterior is the constant across the set of alternatives according to Corollary 2, the posterior is pinned down by the capacity constraint and the initial condition, which is exactly the leveling posterior defined by the leveling strategy.

**Proof of the Theorem.** Let $R_i(\rho_i, \rho_{-i}) := \frac{1}{2}\int_X f \sum_{t\in D_x} \Delta_t\rho_i \Delta_t\rho_{-i} dx$ be the payoff from simultaneous discoveries. We write $\rho_{-i}^-(t,x) := \lim_{s\uparrow t} \rho_{-i}(s,x)$ in place of $\rho_{-i}(t^-,x)$ when the arguments are abbreviated. The payoff function can be written as

$$u_i(\rho_i, \rho_{-i}) = \int_X f \int_T (1-\rho_{-i}) d_t\rho_i dx + R_i(\rho_i, \rho_{-i}) \le \int_X \int_T f \left(1-\rho_{-i}^-\right) d_t\rho_i dx,$$

where the inequality follows from an identity of Stieltjes integral

$$\int_T \rho_{-i}^{-} d_t\rho_i + \sum_{t\in D_x} \Delta_t\rho_i\Delta_t\rho_{-i} = \int_T \rho_{-i} d_t\rho_i.$$

In particular, it holds as an equality whenever one of the strategies is $t$-continuous.

As shown in Equation (3.11) and Equation (3.12) in the main text, the Nash equilibrium in leveling strategies gives a payoff

$$\int_X \int_T \bar{g} d_t\bar{\rho} dx = \frac{1}{2}\int_X f dx = \frac{1}{2}\pi.$$

As the game is constant-sum, the equilibrium strategies in any Nash equilibrium must be achieve the above payoff.

For $t \in T$, denote $H(t) := \{x \in X : f(x) \geq \bar{g}(t)\}$ as the upper contour set of $f$ and $H^C(t) := X\backslash H(t)$ as its complement.

**Lemma 2.** *Let $(\rho_1, \rho_2)$ be a Nash equilibrium. Then, for all $t_0 \in T$ and $i \in \{1, 2\}$, $\rho_i(t_0, x) = 0$ for $x \in H^C(t_0)$ almost everywhere.*

*Proof.* The statement for $t_0 = 0$ follows from the initial condition. Suppose there exists time $t_0 \in (0, 1]$, positive-measure set $A \subset H^C(t_0)$ such that $\rho_i(t_0, x) > 0$ for all $x \in A$. Then the payoff of player $i$ against a leveling strategy of player $-i$ is strictly below the equilibrium payoff:

$$\begin{aligned}
&u_i(\rho_i, \overline{\rho}) - u_i(\overline{\rho}, \overline{\rho}) \\
=& \int_X \int_T g_{-i}(t, x) d_t\rho_i(t, x) dx - \int_X \int_T \bar{g}(t, x) d_t\bar{\rho}(t, x) dx \\
=& \int_X \int_T g_{-i}(t, x) d_t\rho_i(t, x) dx - \int_X \int_T \bar{g}(t) d_t\rho_i(t, x) dx \\
\leq& \int_A \int_{[0,t_0]} \left(g_{-i}(t, x) - \overline{g}(t)\right) d_t\rho_i(t, x) dx \\
<& \ 0.
\end{aligned}$$

The second equality is due to Equation (3.9). The weak inequality follows from $g_{-i}(t, x) \leq \overline{g}(t)$ for all $t \in T$ and $x \in X$, and the strict one from the fact that $g_{-i}(t, x) < \overline{g}(t)$ for all $t \leq t_0$ and $\rho_i(t_0, x) > 0$ for all $x \in A$. ☐

For $i \in \{1, 2\}$, $t \in T$, and $x \in X$, denote $g_i(t^-, x) := \lim_{s\uparrow t} g_i(s, x)$.

**Lemma 3.** *Let $(\rho_1, \rho_2)$ be a Nash equilibrium. For $0 < t_0 < t_1 \leq t_2 < 1$ and $i \in \{1, 2\}$,*

$$g_i(t_2, x_A) - g_i(t_1^-, x_A) \geq g_i(t_2, x_B) - g_i(t_1^-, x_B)$$

*for $x_A \in X$ and $x_B \in H(t_0)$ almost everywhere.*

*Proof.* For $x_A \in H^C(t_2)$ almost everywhere, Lemma 2 implies $g_i(t_2, x_A) - g_i(t_1^-, x_A) = 0$ so the inequality follows from the monotonicity condition.

Suppose there exists positive-measure sets $A \subset H(t_2)$ and $B \subset H(t_0)$ such that $g_i(t_2, x_A) - g_i(t_1^-, x_A) < g_i(t_2, x_B) - g_i(t_1^-, x_B)$ for all $x_A \in A, x_B \in B$. Without loss of generality, assume $g_i(t_2, x_A) - g_i(t_1^-, x_A) < a < g_i(t_2, x_B) - g_i(t_1^-, x_B)$ for some $a < 0$, $\operatorname{ess\,inf}_{A \cup B} f > 0$, and $\int_A \frac{1}{f} = \int_B \frac{1}{f} > 0$; if this is not the case, replace the sets respectively by some positive-measure subsets. Fix $\epsilon \in (0, 1)$. We proceed in two steps.

Step 1: A Modified Leveling Strategy.

Let $\epsilon_1 := (t_1 - t_0)\epsilon > 0$. Let $\epsilon_2 > 0$ be a solution to $\bar{g}(t_1 - \epsilon_1) - \bar{g}(t_1) = \bar{g}(t_2) - \bar{g}(t_2 + \epsilon_2)$. For sufficiently small $\epsilon$, it exists and is unique by the continuity and monotonicity of $\bar{g}$. In addition, $0 < t_1 - \epsilon_1$ and $t_2 + \epsilon_2 < 1$. The left- and right-differentiability of $\bar{g}$ from Lemma 1 imply that $\epsilon_1 \partial_t^- \bar{g}(t_1) = \epsilon_2 \partial_t^+ \bar{g}(t_2) + o(\epsilon)$. Let $\Delta_1 t := [t_1 - \epsilon_1, t_1]$ and $\Delta_2 t := [t_2, t_2 + \epsilon_2]$.

Consider the following *modified leveling strategy* $\tilde{\rho}_{-i}$, which, comparing to the leveling strategy, searches $A$ in expense of $B$ over $\Delta_1 t$, and vice versa over $\Delta_2 t$.

If $t \in \Delta_1 t$, let

$$\tilde{\rho}_{-i}(t, x) := \begin{cases} \bar{\rho}(t, x) + \frac{\bar{g}(t_1 - \epsilon_1) - \bar{g}(t)}{f(x)}, & \text{if } x \in A; \\ \bar{\rho}(t_1 - \epsilon_1, x), & \text{if } x \in B. \end{cases}$$

If $t \in (t_1, t_2)$, let

$$\tilde{\rho}_{-i}(t, x) := \begin{cases} \bar{\rho}(t, x) + \frac{\bar{g}(t_1 - \epsilon_1) - \bar{g}(t_1)}{f(x)}, & \text{if } x \in A; \\ \bar{\rho}(t, x) - \frac{\bar{g}(t_1 - \epsilon_1) - \bar{g}(t_1)}{f(x)}, & \text{if } x \in B. \end{cases}$$

If $t \in \Delta_2 t$, let

$$\tilde{\rho}_{-i}(t, x) := \begin{cases} \bar{\rho}(t_2 + \epsilon_2, x), & \text{if } x \in A; \\ \bar{\rho}(t, x) - \frac{\bar{g}(t_1 - \epsilon_1) - \bar{g}(t_1)}{f(x)} + \frac{\bar{g}(t_2) - \bar{g}(t)}{f(x)}, & \text{if } x \in B. \end{cases}$$

If $x \notin A \cup B$ or $t \notin [t_1 - \epsilon_1, t_2 + \epsilon_2]$, let $\tilde{\rho}_{-i}(t, x) := \bar{\rho}(t, x)$.

Note that the modified strategy $\tilde{\rho}_{-i}$ is a strategy for player $-i$, and in particular that it satisfies the capacity constraint because $\int_A \frac{1}{f} = \int_B \frac{1}{f}$. It can be verified to be $t$-continuous.

Step 2: Payoffs from the Modified Leveling Strategy.

Observe that the difference in strategies is $d_t\tilde{\rho}_{-i} - d_t\bar{\rho} = -\frac{1}{f}d_t\bar{g}$ on $A$ and $d_t\tilde{\rho}_{-i} - d_t\bar{\rho} = \frac{1}{f}d_t\bar{g}$ on $B$ over $\Delta_1 t$, and vice versa over $\Delta_2 t$. It is zero otherwise. The change in utility of the modified leveling strategy comparing to the leveling strategy, $u_{-i}(\tilde{\rho}_{-i}, \rho_i) - u_{-i}(\bar{\rho}, \rho_i)$, is

$$-\int_A \frac{1}{f}\int_{\Delta_1 t} g_i d_t\bar{g}dx + \int_B \frac{1}{f}\int_{\Delta_1 t} g_i d_t\bar{g}dx + \int_A \frac{1}{f}\int_{\Delta_2 t} g_i d_t\bar{g}dx - \int_B \frac{1}{f}\int_{\Delta_2 t} g_i d_t\bar{g}dx. \tag{A.2}$$

For the first term in (A.2), we perform a change of variable to get

$$\begin{aligned}
&-\int_A \frac{1}{f(x)}\int_{\Delta_1 t} g_i(t,x)d\bar{g}(t,x)dx \\
=&-\epsilon_1\int_A \frac{1}{f(x)}\int_{[0,1]} g_i(t_1 - s\epsilon_1)\partial_t^-\bar{g}(t_1 - s\epsilon_1)dsdx \\
=&-\epsilon_1\partial_t^-\bar{g}(t_1)\int_A \frac{1}{f(x)}g_i(t_1^-, x)dx + o(\epsilon),
\end{aligned}$$

where the second equality is due to the dominated convergence theorem. The equation states that, over short time interval $\Delta_1 t$, both $g_i$ and $\partial_t^-\bar{g}_{-i}$ can be taken as constants with respect to time. The same can be applied to the other three terms.

The payoff difference $u_{-i}(\tilde{\rho}_{-i}, \rho_i) - u_{-i}(\bar{\rho}, \rho_i)$ can thus be written as

$$\begin{aligned}
&-\left(\int_A \frac{1}{f(x)}g_i(t_1^-, x)dx - \int_B \frac{1}{f(x)}g_i(t_1^-, x)dx\right)\epsilon_1\partial_t^-\bar{g}(t_1) \\
&+\left(\int_A \frac{1}{f(x)}g_i(t_2, x)dx - \int_B \frac{1}{f(x)}g_i(t_2, x)dx\right)\epsilon_2\partial_t^+\bar{g}(t_2) + o(\epsilon) \\
=&\left(-\int_A \frac{1}{f(x)}\Big(g_i(t_2,x) - g_i(t_1^-, x)\Big)dx + \int_B \frac{1}{f(x)}\Big(g_i(t_2,x) - g_i(t_1^-, x)\Big)dx\right)\partial_t^+\bar{g}(t_2)\epsilon_2 + o(\epsilon).
\end{aligned}$$

By supposition,

$$\begin{aligned}
&-\int_A \frac{1}{f(x)}\Big(g_i(t_2,x) - g_i(t_1^-, x)\Big)dx + \int_B \frac{1}{f(x)}\Big(g_i(t_2,x) - g_i(t_1^-, x)\Big) \\
>&-a\int_A \frac{1}{f(x)}dx + a\int_B \frac{1}{f(x)}dx \\
=&\ 0.
\end{aligned}$$

Therefore, there exists $\epsilon > 0$ sufficiently small such that, against $\rho_i$, the modified leveling strategy $\tilde{\rho}_{-i}$ yields strictly higher payoff than leveling strategy, which guarantees the maxmin payoff. $\square$

The first corollary below establishes the $t$-continuity of the equilibrium strategy. According to Lemma 3, the posterior decreases fastest over the upper contour set $H$. Therefore, the posterior for those alternatives must all drop discontinuously. This violates the capacity constraint.

**Corollary 1.** *In any Nash equilibrium, player i's strategy $\rho_i$ is $t$-continuous.*

*Proof.* The statement for $t = 0$ follows from the monotonicity and right-continuity condition, and that for $t = 1$ is without loss because the set $\{x \in X : \rho_i(1, x) - \rho_i(1^-, x) > 0\}$ is null.

Suppose there exists positive-measure set $B \subset X$ such that $\rho_i$, or equivalently $g_i$, is not $t$-continuous on $(0, 1) \times B$. Without loss of generality, there exists $b < 0$ and $\epsilon \in (0, 1)$ such that, for all $x \in B$, there is $t_x \in (\epsilon, 1)$ satisfying

$$g_i(t_x, x) - g_i(t_x^-, x) \leq b.$$

The compactness of $[\epsilon, 1]$ implies that, for any $\delta > 0$, there exists $\underline{t}_\delta, \bar{t}_\delta \in (\epsilon, 1)$ with $\underline{t}_\delta < \bar{t}_\delta$ and $\bar{t}_\delta - \underline{t}_\delta < \delta$, and positive-measure subset $B_\delta \subset B$ such that

$$g_i(\bar{t}_\delta, x) - g_i(\underline{t}_\delta^-, x) \leq b.$$

Lemma 3 implies that $\rho_i(\bar{t}_\delta, x) - \rho_i(\underline{t}_\delta^-, x) \geq -b/f(x)$ for all $x \in H(\epsilon)$, a positive-measure set. The capacity constraint reads

$$\delta = \int_X \rho_i(\bar{t}_\delta, x) - \rho_i(\underline{t}_\delta^-, x)dx \geq \int_{H(\epsilon)} \rho_i(\bar{t}_\delta, x) - \rho_i(\underline{t}_\delta^-, x)dx \geq \int_{H(\epsilon)} \frac{-b}{f(x)}dx,$$

which yields a contradiction as $\delta \downarrow 0$. $\square$

**Corollary 2.** *In any Nash equilibrium, for $0 < t_1 < t_2 \leq 1$,*

$$g_i(t_2, x_A) - g_i(t_1, x_A) = g_i(t_2, x_B) - g_i(t_1, x_B)$$

*for $x_A, x_B \in H(t_1)$ almost everywhere.*

*Proof.* Assume $t_2 < 1$. For any $t \in (t_1, t_2)$, $H(t_1) \subset H(t)$. Lemma 3 thus gives the equality

$$g_i(t_2, x_A) - g_i(t, x_A) = g_i(t_2, x_B) - g_i(t, x_B)$$

for $x_A, x_B \in H(t_1)$ almost everywhere. The statement is obtained by taking a countable sequence $t \uparrow t_1$, noting that $g_i$ is $t$-continuous by Corollary 1.

The boundary case $t_2 = 1$ follows similarly by taking a countable sequence $t_2 \uparrow 1$. $\square$

**Lemma 4.** *In any Nash equilibrium, for $i \in \{1, 2\}$ and $t \in T$, $g_i(t, x) = \bar{g}(t)$ for $x \in H(t)$ almost everywhere.*

*Proof.* Once the statement is proven for $t \in (0, 1]$, it extends to the endpoint $t = 0$ because of the monotonicity and right-continuity condition.

For $t \in (0, 1]$, define $\tilde{g}_i(t) := \sup_{x \in H(t)} g_i(t, x)$. The terminal condition implies $\tilde{g}_i(1) = 0$. Corollary 2 with $t_2 = 1$ reads

$$\tilde{g}_i(t) = g_i(t, x) - g_i(1, x) = g_i(t, x) \tag{A.3}$$

for all $x \in H(t)$ almost everywhere.

We now derive the right-derivative $\partial_t^+ \tilde{g}_i$. For $0 < t_1 < t_2 < 1$, the capacity constraint gives

$$\begin{aligned} t_2 - t_1 &= \int_{H(t_1)} \rho_i(t_2, x) - \rho_i(t_1, x) dx + \int_{H(t_2) \backslash H(t_1)} \rho_i(t_2, x) - \rho_i(t_1, x) dx \\ &= -\left(\tilde{g}_i(t_2) - \tilde{g}_i(t_1)\right) \int_{H(t_1)} \frac{dx}{f(x)} - \int_{H(t_2) \backslash H(t_1)} \frac{g_i(t_2, x) - g_i(t_1, x)}{f(x)} dx \\ &\leq -\left(\tilde{g}_i(t_2) - \tilde{g}_i(t_1)\right) \left( \int_{H(t_1)} \frac{dx}{f(x)} + \int_{H(t_2) \backslash H(t_1)} \frac{dx}{f(x)} \right). \end{aligned} \tag{A.4}$$

The inequality is due to Lemma 3. Rearranging the terms,

$$0 \geq \frac{\tilde{g}_i(t_2) - \tilde{g}_i(t_1)}{t_2 - t_1} \geq -\left( \int_{H(t_1)} \frac{dx}{f(x)} + \int_{H(t_2) \backslash H(t_1)} \frac{dx}{f(x)} \right)^{-1} \geq -\frac{\bar{g}(t_2)}{|H(t_2)|} > -\infty,$$

where the third inequality is due to the definition of $H$. The function $\tilde{g}_i$ is Lipschitz and thus absolutely continuous on $(0, 1)$.

Take $t_2 \downarrow t_1$. Since $H(t_2) \downarrow H(t_1)$ in the set-inclusion sense, the dominated convergence theorem states that $|H(t_2) \backslash H(t_1)| \downarrow 0$. The second term in Equation (A.4) is dominated by

$$\int_{H(t_2) \backslash H(t_1)} \frac{|g_i(t_2, x) - g_i(t_1, x)|}{f(x)} dx \leq \frac{|\tilde{g}_i(t_2) - \tilde{g}_i(t_1)| \, |H(t_2) \backslash H(t_1)|}{\bar{g}\left(\frac{1}{2}(t_1 + 1)\right)} = o(t_2 - t_1).$$

The right-derivative of $\tilde{g}_i$ is thus given by

$$\partial_t^+ \tilde{g}_i(t_1) = \lim_{t_2 \downarrow t_1} \frac{\tilde{g}_i(t_2) - \tilde{g}_i(t_1)}{t_2 - t_1} = -\int_{H(t_1)} \frac{dx}{f(x)}.$$

Since $\bar{g}$ also satisfies the first two lemmas and the two corollaries, an analogous calculation shows that

$$\partial_t^+ \bar{g}(t) = -\int_{H(t)} \frac{dx}{f(x)} = \partial_t^+ \tilde{g}_i(t).$$

Therefore, $\bar{g} = \tilde{g}_i + C$ for some constant $C \in \mathbb{R}$. The boundary condition at $t = 1$ is $\lim_{t \uparrow 1} \tilde{g}_i(t) = \bar{g}(1) = 0$ which implies $C = 0$. □

Lemma 2 and Lemma 4 imply that, for all $t \in T$, $g_i(t, \cdot) = \bar{g}(t, \cdot)$ almost everywhere. There exists a full measure set over which the inequality holds for all $t \in [0, 1] \cap \mathbb{Q}$. The theorem then follows from the monotonicity and right-continuity condition.

## A.4 Proof of Theorem 3

### A.4.1 Proof Strategy

The idea of the proof of Theorem 3 is as follows. For every strategy $\rho_2^\alpha$ of player 2, define $\rho_2 := \rho_2^\alpha / \alpha$. It is easy to verify that $\rho_2 : T \times X \to [0, 1/\alpha]$ satisfies the four conditions of Definition 3. It differs from a distributional strategy in its codomain $[0, 1/\alpha]$ instead of $[0, 1]$. We shall call $\rho_2 : T \times X \to [0, 1/\alpha]$ a *normalized strategy*. Players' payoffs from the strategy profile $(\rho_1, \rho_2^\alpha)$ can be rewritten as payoffs from $(\rho_1, \rho_2)$ as follows:

$$u_1(\rho_1, \rho_2^\alpha) = (1 - \alpha)\pi + \alpha u_1(\rho_1, \rho_2); \tag{A.5}$$

$$u_2(\rho_2^\alpha, \rho_1) = \alpha u_2(\rho_2, \rho_1). \tag{A.6}$$

Therefore, the payoff functions under asymmetric capacity are increasing affine transformations of those with a normalized strategy of player 2. Thus, the game with asymmetric capacity is strategically equivalent to the one with a normalized strategy, and the existence and uniqueness of the Nash equilibrium in the game with asymmetric players will follow from their counterparts in the game with normalized strategies. But the latter is not quite the same as the symmetric game because of the codomain of the normalized strategy $\rho_2$, i.e., it is not a priori clear that $\rho_2(1, \cdot) = 1$ in equilibrium. This gap is closed using the following proof strategy. We decompose the maximization over normalized strategies into two components: the (normalized) probability of exploration by the end of the game $\rho_2(1,\cdot)$, and the implementation of the exploration given this probability. We shall show that, for any probability of exploration, a generalized leveling strategy is optimal for player 2, and his payoff is uniquely maximized at $\rho_2(1,\cdot)=1$ given the leveling strategy.

### A.4.2 Formal Proof

We first derive Equation (A.5) and Equation (A.6) as follows.

$$\begin{aligned}
u_1(\rho_1,\rho_2^\alpha) &= \int_X\int_T (1-\alpha\rho_2(t,x))\, f(x) d_t\rho_1(t,x)dx + \frac{1}{2}\int_X f(x)\sum_{t\in D_x}\Delta_t\rho_1(t,x)\Delta_t\alpha\rho_2(t,x)dx\\
&= (1-\alpha)\int_X\int_T f(x)d_t\rho_1(t,x)dx\\
&\quad +\alpha\int_X\int_T (1-\rho_2(t,x))\, f(x)d_t\rho_1(t,x)dx + \frac{\alpha}{2}\int_X f(x)\sum_{t\in D_x}\Delta_t\rho_1(t,x)\Delta_t\rho_2(t,x)dx\\
&= (1-\alpha)\pi + \alpha u_1(\rho_1,\rho_2).\\
u_2(\rho_2^\alpha,\rho_1) &= \int_X\int_T \alpha\,(1-\rho_1(t,x))\, f(x)d_t\rho_2(t,x)dx + \frac{\alpha}{2}\int_X f(x)\sum_{t\in D_x}\Delta_t\rho_1(t,x)\Delta_t\rho_2(t,x)dx\\
&= \alpha u_2(\rho_2,\rho_1).
\end{aligned}$$

**Definition 5** (Timing game)**.** *Let the (normalized) probability of exploration be a function $\Delta\rho_2 : X\to[0,1/\alpha]$ such that $\int_X\Delta\rho_2=1$. The corresponding timing game is a game in which player 1 plays a distributional strategy, player 2 plays a normalized strategy subject to the terminal condition $\rho_2(1,\cdot)=\Delta\rho_2(\cdot)$, and the payoff function is given by Equation (2.3).*

Denote $\Delta\rho_1 := 1$. For any timing game, define $\Delta\rho_{\min} := \min\{\Delta\rho_1,\Delta\rho_2\}$ and $t^* := \int_X\Delta\rho_{\min}dx>0$.

**Definition 6** (Leveling strategy in timing game)**.** *The leveling function in a timing game $\bar g:[0,t^*]\to[0,\sup_x f\Delta\rho_{\min}]$ is the unique solution to*

$$\int_X\left(\Delta\rho_{\min}(x)-\frac{\bar g(t)}{f(x)}\right)\mathbf{1}_{\{f(x)\Delta\rho_{\min}(x)\ge\bar g(t)\}}(x)dx = t \textit{ for all } t\in[0,t^*].$$

*The leveling strategy in a timing game is defined as*

$$\bar\rho(t,x) := \left(\Delta\rho_{\min}(x)-\frac{\bar g(t)}{f(x)}\right)\mathbf{1}_{\{f(x)\Delta\rho_{\min}(x)\ge\bar g(t)\}}(x) \textit{ for all } t\in[0,t^*] \textit{ and } x\in X.$$

The proof of the existence, uniqueness, and absolute continuity of the leveling function in a timing game is analogous to that of Lemma 1. So we omit it here. With abuse of notation, a strategy $\rho$, normalized or otherwise, in the timing game is called a leveling strategy if $\rho|_{[0,t^*]\times X} = \bar{\rho}$. It is obvious that a leveling strategy exists. Since the characterization only applies to $t \in [0, t^*]$, it is unique only when $t^* = 1$ or equivalently $\Delta\rho_2 = 1$. In that case, the leveling strategy coincides with the benchmark one. We denote posterior distribution as $g_i(t,x) := f(x)\left(\Delta\rho_{\min}(x) - \rho_i(t,x)\right)$, and the upper contour set $H(t) := \{x \in X : f(x)\Delta\rho_{\min}(x) \geq \bar{g}(t)\}$. Note that the definitions agree with the benchmark case in which the probability of exploration is one, i.e., $\Delta\rho_2 = 1$.

Theorem 3 is shown by two further results. The key idea is that any strategy of player 2 can be decomposed to the probability of exploration and its corresponding timing of exploration. For fixed probability of exploration, Lemma 5 characterizes the set of Nash equilibrium in the timing game as the set of leveling strategy profiles. Its proof is analogous to that of Theorem 2. Lemma 6 concludes that the symmetric leveling profile in the benchmark case is the unique Nash equilibrium in the normalized game. Over all probabilities of exploration, $\Delta\rho_{\min} = 1$ uniquely achieves the highest minimum payoff for player 2.

**Lemma 5.** *For any $\Delta\rho_2$, the profile $(\rho_1, \rho_2)$ is a Nash equilibrium in the timing game if and only if it is a leveling strategy profile.*

*Proof.* The proof is similar to that for Theorem 2. We hereby comment on the three instances in which it requires modifications.

The payoff function can be written as

$$\begin{aligned} u_i(\rho_i, \rho_{-i}) &= \int_X f(1-\Delta\rho_{\min})\Delta\rho_i dx + \int_X \int_T f\left(\Delta\rho_{\min} - \rho_{-i}\right) d_t\rho_i dx + R_i(\rho_i, \rho_{-i}) \\ &\leq \int_X f(1-\Delta\rho_{\min})\Delta\rho_i dx + \int_X \int_T f\left(\Delta\rho_{\min} - \rho_{-i}^-\right) d_t\rho_i dx. \end{aligned}$$

The integrand in the second term motivates the more general definition of the posterior distribution. The myopic argument is applied to $t \in [0, t^*]$ during which the maximum posterior is $\bar{g}(t)$ attained on $H(t)$, and then to $t \in (t^*, 1]$ during which the maximum posterior 0 is attained on $\{x \in X : \Delta\rho_{\min} = 0\}$.

The equilibrium payoff of a leveling strategy profile $(\bar{\rho}_i, \bar{\rho}_{-i})$ is given by

$$\begin{aligned} u_i(\bar{\rho}_i, \bar{\rho}_{-i}) &= \int_X f(1-\Delta\rho_{\min})\Delta\rho_i dx + \int_X \int_{[0,t^*]} \bar{g} d_t\bar{\rho} dx \\ &= \int_X f(1-\Delta\rho_{\min})\Delta\rho_i dx + \frac{1}{2}\int_X f(\Delta\rho_{\min})^2 dx \end{aligned}$$

$$= \int_X f\left(1 - \Delta\rho_{\min}\Delta\rho_i + \frac{1}{2}(\Delta\rho_{\min})^2\right) dx. \tag{A.7}$$

In the benchmark case, the three lemmas and the two corollaries apply to $T \times X$; in the timing game, they carry through with restriction to $[0, t^*] \times X$.

For the result analogous to Lemma 4, the function $\tilde{g}_i$ is defined more generally as $\sup_{x \in H(t)} g_i(t, x) - g_i(t^*, x)$, because $g_2(t^*, \cdot)$ may not zero in the timing game. As in the benchmark case, the boundary condition at $t = t^*$ implies that the constant of integration is zero $C = 0$. On $H(t^*) = \{x \in X : \Delta\rho_{\min}(x) > 0\}$ almost surely, the other boundary condition at $t = \bar{g}^{-1}(f(x)\Delta\rho_{\min}(x)) < t^*$ shows that

$$g_i(t^*, x) = \lim_{s \downarrow t} g_i(s, x) - \bar{g}(t, x) = f(x)\Delta\rho_{\min}(x) - f(x)\Delta\rho_{\min}(x) = 0.$$

This establishes the desired result. $\square$

**Lemma 6.** *Let $\bar{\rho}$ be the leveling strategy in the benchmark case. In the normalized game, the symmetric leveling profile $(\bar{\rho}, \bar{\rho})$ is the unique Nash equilibrium.*

*Proof.* We first argue that the candidate is a Nash equilibrium in the normalized game. Recall that the leveling strategy in the benchmark case is the unique leveling strategy in the timing game with $\Delta\rho_2 = 1$. Player 1 has the same set of strategies in both the timing game with $\Delta\rho_2 = 1$ and the normalized game. Since the profile is a Nash equilibrium in the former game, he has no profitable deviation in the latter. Player 2 has no profitable deviations by the myopic argument because of the leveling posterior $g_1$.

We now show the uniqueness. Any equilibrium of the normalized game is an equilibrium of the timing game with the corresponding normalized probability of exploration, since the strategy set in the former game is a superset of that in the latter. For each $\Delta\rho_2$, Lemma 5 characterizes the set of Nash equilibria of the timing game as the set of leveling profiles, with equilibrium payoff given by Equation (A.7). For $\Delta\rho_1 = 1$, the function of $\Delta\rho_2(x)$

$$1 - \min\{1, \Delta\rho_2(x)\}\Delta\rho_2(x) + \frac{1}{2}\left(\min\{1, \Delta\rho_2(x)\}\right)$$

is uniquely maximized at $\Delta\rho_2(x) = 1$. Therefore, the equilibrium payoff $\pi/2$ can only be achieved with $\Delta\rho_2 = 1$ almost everywhere with the corresponding strategy profile $(\bar{\rho}, \bar{\rho})$.

Let $(\rho_1, \rho_2)$ be an equilibrium strategy profile of a timing game with $\Delta\rho_2(\cdot) \not\equiv 1$. The player 2's payoff of playing $\bar{\rho}_2$ is

$$u_2(\bar{\rho}, \rho_1) \geq u_2(\bar{\rho}, \bar{\rho}) > u_2(\rho_2, \rho_1).$$

The weak inequality is because $\bar{\rho}$ is a best response to itself for player 1. The strict inequality is due to the fact that $\Delta\rho_2(\cdot) \not\equiv 1$. Therefore, $\bar{\rho}$ is a profitable deviation for player 2 in the normalized game. □

## A.5 Proof of Theorem 4

*Proof.* It suffices to show that $P_\beta(t)$ is increasing in $\beta$ for all $t \in T$. By differentiating Equation (3.4) with respect to time, we obtain

$$\bar{g}'(t) = -\left(\int_X f(x)^{-1}\mathbf{1}_{\{f(x)\geq\bar{g}(t)\}}dx\right)^{-1}.$$

The Lipschitz term due to the changing domain of integration vanishes because $1 - \frac{\bar{g}(t)}{f(x)} = 0$ on $\{x \in X : f(x) = \bar{g}(t)\}$.

As the leveling function $\bar{g}$ is absolutely continuous, the probability of discovery is absolutely continuous with respect to $\beta$ with derivative

$$\begin{aligned}
\partial_\beta P_\beta(t) &= -t\bar{g}'(\beta t)\int_X \mathbf{1}_{\{f(x)\geq\bar{g}(\beta t)\}}dx - t\bar{g}(\beta t) + (2-\beta)t^2\bar{g}'(\beta t)\\
&= \frac{t}{\int_X f(x)^{-1}\mathbf{1}_{\{f(x)\geq\bar{g}(\beta t)\}}dx}\left(\int_X \mathbf{1}_{f(x)\geq\bar{g}(\beta t)}dx - \int_X \frac{\bar{g}(\beta t)}{f(x)}\mathbf{1}_{\{f(x)\geq\bar{g}(\beta t)\}}dx - (2-\beta)t\right)\\
&= \frac{t}{\int_X f(x)^{-1}\mathbf{1}_{\{f(x)\geq\bar{g}(\beta t)\}}dx}\left(\beta t - (2-\beta)t\right)\\
&= \frac{2(\beta-1)t^2}{\int_X f(x)^{-1}\mathbf{1}_{\{f(x)\geq\bar{g}(\beta t)\}}dx}\\
&\geq 0
\end{aligned}$$

$\beta$-almost everywhere, where the third equality follows from Equation (3.4). □

## A.6 Proof of Theorem 6

*Proof.* Verifying that the strategy profile is a Nash equilibrium is analogous to the arguments in Section 3.3.1 and we shall not replicate the proof here. In any symmetric strategy profile and hence any symmetric Nash equilibrium, the payoff of each player is $\pi/n$ by symmetry.

We first show that, for any strategy $\rho$, the leveling strategy $\bar{\rho}$ guarantees player $i$ payoff $\pi/n$ when all other players employ $\rho$. Since the leveling strategy is $t$-absolutely continuous, the payoff of player $i$ equals to the time integral of flow payoffs by Fubini's theorem

$$\begin{aligned} u\left(\bar{\rho}, \rho_{-i}\right) &= \int_X \int_T f(x)(1-\rho_{-i}(t,x))\partial_t\bar{\rho}(t,x)dtdx \\ &= \int_T \int_X f(x)(1-\rho_{-i}(t,x))\partial_t\bar{\rho}(t,x)dxdt\,. \end{aligned}$$

As the leveling profile gives payoff $\pi/n$ for all players, it suffices to show that, for all $t \in (0,1)$, the leveling strategy $\bar{\rho}(t,\cdot)$ minimizes the flow payoff of $\bar{\rho}$

$$\min_{\rho(t,\cdot)} \int_X f(x)(1-\rho(t,x))^{n-1}\partial_t\bar{\rho}(t,x)dx \tag{A.8}$$

subject to the capacity constraint at $t$. Without the constraint $\rho(t,\cdot) \leq 1$, the relaxed problem must is equivalent to the first-order condition

$$(n-1)f(x)\partial_t\bar{\rho}(t,x)(1-\rho(t,x))^{n-2} = C$$

for $\{x \in X : \rho(t,x) > 0\}$ almost everywhere, together with the complementary slackness condition

$$(n-1)f(x)\partial_t\bar{\rho}(t,x) \leq C$$

for $\{x \in X : \rho(t,x) = 0\}$ almost everywhere, for some Lagrange multiplier $C \geq 0$. It is then straightforward to show that the leveling strategy $\rho(t,\cdot)$ solves the two conditions and hence the minimization problem.

From here, uniqueness can be shown along the idea of Theorem 2 by constructing a modified leveling strategy that yields strictly higher payoff, as in the case of $n = 2$. We provide here a shorter proof that takes advantage of the strict convexity of minimization problem (A.8) when $n > 2$. With duplication among other players, they can no longer achieve the minimum payoff of player $i$ by any other strategies.

We proceed to show that the leveling strategy $\bar{\rho}$ yields payoff strictly above $\pi/n$ when all other players employ $\rho \neq \bar{\rho}$. Since the strategy is not leveling, there exist time interval $(t_0, t_1)$ and positive-measure set $A \subset X$ such that $\rho(t,x) \neq \bar{\rho}(t,x)$ for all $t \in (t_0,t_1)$ and $x \in A$. The strict convexity of minimization problem (A.8) implies that the minimizer $\bar{\rho}(t,\cdot)$ is unique. Therefore, the flow payoff of $\bar{\rho}$ is strictly above the minimum over $(t_0, t_1)$.

For any symmetric strategy profile $(\rho, ..., \rho)$ for $\rho \neq \bar{\rho}$, all players have a profitable deviation to $\bar{\rho}$. Therefore, the strategy profile is not an equilibrium. □

## A.7 Proof of Theorem 7

We first show a technical lemma.

**Lemma 7.** *Let $v_1$ and $v_2$ be continuous, strictly decreasing, and convex functions from $[0,1]$ to $\mathbb{R}_+$ with $v_1(1) = v_2(1) = 0$ and $\int_0^1 v_1(s)ds = \int_0^1 v_2(s)ds$. Denote $v_1^{-1}$ and $v_2^{-1}$ as their respective inverses, with the extension of value $0$ outside of their domains. Then $\int_0^t v_1(s)ds \geq \int_0^t v_2(s)ds$ for all $t \in [0,1]$ if and only if $\int_0^z v_1^{-1}(y)dy \leq \int_0^z v_2^{-1}(y)dy$ for all $z \in \mathbb{R}_+$.*

*Proof.* We prove the inequality $\int_0^t v_1(s)ds \leq \int_0^t v_2(s)ds$ for all $t \in T$ under the hypothesis that $\int_0^z v_1^{-1}(y)dy \geq \int_0^z v_2^{-1}(y)dy$ for all $z \in \mathbb{R}_+$. The converse is analogous.

Since $v_m$, $m \in \{1,2\}$, is strictly decreasing and convex, it has a strictly negative derivative almost everywhere on $(0,1)$. The integration by parts and then a change of variables give

$$\int_0^t v_m(s)ds = tv_m(t) - \int_0^t s dv_m(s) = tv_m(t) + \int_{v_m(t)}^{\infty} v_m^{-1}(y)dy. \tag{A.9}$$

Take $t = 1$ (and hence $v_m(t) = 0$), we obtain

$$\int_0^1 v_m(s)ds = \int_0^{\infty} v_m^{-1}(y)dy. \tag{A.10}$$

Equation (A.9) and Equation (A.10) combines to give,

$$\int_0^t v_m(s)ds = tv_m(t) + \int_0^1 v_m(s)ds - \int_0^{v_m(t)} v_m^{-1}(y)dy. \tag{A.11}$$

Thus the desired inequality holds on the set $S := \{t \in T : v_1(t) = v_2(t)\}$ because it follows from Equation (A.11) that

$$\int_0^t (v_1(s) - v_2(s))ds = \int_0^{v_1(t)} (v_2^{-1}(y) - v_1^{-1}(y))dy \leq 0.$$

The set $S$ is closed by the continuity of $v_1$ and $v_2$, and it contains $t = 1$ by assumption. The inequality holds trivially at $t = 0$. Denote $S^* := S \cup \{0\}$.

For any $t \notin S^*$, define two endpoints $\underline{t} := \max\{s \in S^* : s < t\}$ and $\bar{t} := \min\{s \in S^* : s > t\}$. They are well-defined because $S^*$ is closed. The difference $v_1(s) - v_2(s)$ has the same sign over $(\underline{t}, \bar{t})$ by continuity, so its integral $\int_0^{t'} (v_1(s) - v_2(s))ds$ is monotonic over the same interval. As the desired inequality holds at the endpoints, it actually holds over the entire interval, and at $t$ in particular. □

*Proof of Theorem 7.* Suppose $n' > n$. Let $\bar{g}'$ and $\bar{g}$ be the leveling function associated with $n'$ players and $n$ players respectively. Parallel to the proof of the benchmark case, the probabilities of simultaneous discovery are both zero, and the flow probabilities of discovery are $n'\bar{g}'$ and $n\bar{g}$ respectively. By Lemma 7, it suffices to prove the stochastic order between their inverses $h'(\cdot/n')$ and $h(\cdot/n)$.

By the Fubini theorem, the integral can be written as

$$\begin{aligned}\int_0^y h\left(\frac{z}{n}\right) dz &= \int_0^y \int_X \left(1 - \left(\frac{z}{nf(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{nf(x)\geq z\}}(x,z)dxdz \\ &= \int_X \int_0^y \left(1 - \left(\frac{z}{nf(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{nf(x)\geq z\}}(x,z)dzdx \\ &= \int_X I(n,y,f(x))dx,\end{aligned}$$

where the integrand is given by

$$I(n,y,f(x)) = \begin{cases} y - \frac{n-1}{n}\left(\frac{y}{nf(x)}\right)^{\frac{1}{n-1}} y, & \text{if } n \geq \frac{y}{f(x)}; \\ f(x), & \text{if } n < \frac{y}{f(x)}. \end{cases}$$

We obtain the desired inequality by noting that $I(n,y,f(x))$ is decreasing in $n$. □

## A.8 Proof of Theorem 8

*Proof.* Let $\bar{g}_m$ be the equilibrium leveling functions for prior $f_m$, and its inverse is $h_m$, $m \in \{1,2\}$, where

$$h_m(y) = \int_X \left(1 - \left(\frac{y}{f_m(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{f_m(x)\geq y\}}(x)dx. \tag{A.12}$$

As the probability of simultaneous discovery is zero, the flow probabilities of discovery are $2\bar{g}_1$ and $2\bar{g}_2$ respectively. The desired conclusion is $\int_0^t n\bar{g}_1(s)ds \geq \int_0^t n\bar{g}_2(s)ds$ for all $t \in Y$. Since $\bar{g}_1$ and $\bar{g}_2$ satisfies the assumptions of Lemma 7, it suffices to show their inverses satisfy $\int_0^z h_1(y)dy \leq \int_0^z h_2(y)dy \quad \forall z \in \mathbb{R}_+$. By Equation (A.12), the Fubini theorem, and then a change of variables, the integral can be written as

$$\int_0^z h_m(y)dy = \int_0^z \int_X \left(1 - \left(\frac{y}{f_m(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{f_m(x)\geq y\}}(x,y)dxdy$$

$$
\begin{aligned}
&= \int_X \int_0^z \left(1 - \left(\frac{y}{f_m(x)}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{f_m(x) \geq y\}}(x, y) dy dx \\
&= \int_0^\infty \int_0^z \left(1 - \left(\frac{y}{w}\right)^{\frac{1}{n-1}}\right) \mathbf{1}_{\{w \geq y\}}(w, y) dy d\lambda \circ f_m^{-1} \\
&= \int_0^\infty I(w, z) d\lambda \circ f_m^{-1}
\end{aligned}
$$

where the integrand is given by

$$
I(w, z) = \begin{cases} z - \frac{n-1}{n} \left(\frac{z}{w}\right)^{\frac{1}{n-1}} z, & \text{if } w \geq z; \\ \frac{w}{n}, & \text{if } w < z. \end{cases}
$$

The result follows because $I(\cdot, z)$ is an increasing concave function and $\lambda \circ f_1^{-1}$ is a mean-preserving spread of $\lambda \circ f_2^{-1}$. $\square$