BANDIT PROBLEMS WITH LÉVY PAYOFF PROCESSES

ASAF COHEN AND EILON SOLAN

Abstract. We study one-arm Lévy bandits in continuous time, which have one safe arm that yields a constant payoff s, and one risky arm that can be either of type High or Low; both types yield stochastic payoffs generated by a Lévy process. The expectation of the Lévy process when the arm is High is greater than s, and lower than s if the arm is Low.

The decision maker (DM) has to choose, at any given time t, the fraction of resource over the time interval $[t, t + dt)$ to be allocated to each arm. We show that under proper conditions on the Lévy processes, there is a unique optimal strategy, which is a cut-off strategy, and we provide an explicit formula for the cut-off and the corresponding expected payoff from the data of the problem. We also examine the case where the DM has incorrect prior over the type of the risky arm, and we calculate the expected payoff gained by a DM who plays the optimal strategy that corresponds to the incorrect prior.

In addition, we study some applications of the results: (a) we show how to price information in one arm Lévy bandit problem, and (b) we investigate who fares better in one-arm bandit problems: an optimist who assigns a probability higher than the true probability to High, or a pessimist who assigns a probability lower than the true probability to High.

1. Introduction

Consider a firm that has to determine, on an ongoing basis, how much to invest in the research of new technologies for its next line of products. The firm faces a tradeoff between exploration and exploitation: on the one hand, it can adopt the technology that seems most successful according to the research conducted so far, thereby exploiting its investment in research, and on the other hand, it can continue investing in other technologies, in the hope of finding an even better technology for its products. If the firm decides to stop investing in a given technology, then no information will be obtained on that technology, so even if it is actually better than the finally adapted technology, it will never be adapted.

A similar tradeoff between exploration versus exploitation arises, e.g., in the market of venture capital funds, where each fund has to decide in which start-up companies to invest, and in clinical trials, where pharmaceutical companies have to decide which new drugs or treatments to explore.

To concentrate on the issue of exploration versus exploitation, one assumes that there are no exogenous factors that affect the firm's decision (such as new technologies or drugs that are introduced by competitors). The optimization problem that the firm faces has been modelled in the literature as a multi-arm bandit problem (see, e.g., Bergemann and Valimaki (2006), Keller, Rady and Cripps (2005), Basenko and Wu (2008), Klein and Rady (2008)): a decision maker has finitely many actions, called arms, each one yields a payoff with an unknown distribution, that is taken from a finite set of distributions. Each time the decision maker chooses an arm, he obtains a payoff, and improves his information regarding the correct payoff distribution of the arm that he just chose.

Gittins and Jones (1979) in discrete time, and Kaspi and Mandelbaum (1995) in continuous time, proved that the optimal strategy of the decision maker has a particularly simple form: the decision maker calculates a real number, an index, to each arm, based on the payoffs that the arm obtained in the past; he then has to choose at each period the arm with the highest index. It turns out that to calculate the index of an arm it is sufficient to consider an auxiliary problem with two arms: the arm to which we calculate the index, and an arm that yields a constant payoff. The literature therefore focuses on such problems, called one-arm bandit problems.

The simple structure of the optimal strategy allows one to find an explicit formula for the index when the payoff is one of two distributions that have simple form. Berry and Friestedt (1985) provide the solution to the problem in discrete time, e.g. when the payoff distribution is one of two Bernoulli distributions, and in continuous time, e.g., when the payoff distribution is one of two Brownian motions. Karatzas (1984) characterized the index when the payoff's distribution is a diffusion process. Using Markovian excursion theory, Kaspi and Mandelbaum (1995) characterized the index when the payoff's distribution is a Levy process, and they were able to obtain an explicit form for the index for special distributions. By studying the dynamic programming equation that describes the problem, Keller, Rady and Cripps (2005) and Keller and Rady (2008) provide an explicit form for the index when the payoff's distribution is Poisson.[1]

In practice, payoff processes have a complex form, exhibiting both small random changes that can be modelled by a Brownian motion, and large shocks that can be modelled by a Poisson distribution. In fact, Carr and Wu (2004) argue that almost all economic phenomena can be described by time shifts of Levy processes, and therefore it is desirable to study the bandit problem when the payoff distribution is one of finitely many Levy distributions.

In the present paper we provide an explicit solution to the one-arm bandit problem when the payoff's distribution is one of two Levy processes. We assume that one distribution, called *High*, dominates the other, called *Low*, in a strong sense (see Assumption 2.1 below). To eliminate trivial cases, we assume that the constant payoff that is generated by the safe arm is lower than the expected payoff generated by the High distribution, and higher than the expected payoff generated by the Low distribution.

In such a case in discrete time, the optimal strategy is a cut-off strategy: the decision maker keeps on experimenting as long as the posterior belief that the distribution is High is higher than some cut-off, and, once the posterior probability that the distribution is High falls below the cut-off, the decision maker switches to the safe arm. As is well known, in continuous time the notion of a strategy is not well defined. We prove that when the two payoff distributions are Levy processes that satisfy our requirements, among all strategies that define a unique play, the optimal strategy is a cut-off strategy, and we provide an explicit expression for the cut-off, in terms of the data of the problem. When particularized to the models studied by Kaspi and Mandelbaum (1995), Keller, Rady and Cripps (2005)

[1]These authors also studied the strategic setup, in which several decision makers have the same set of arms, and their arm's payoff distribution is the same (and unknown), and they compared the cooperative solution to the non-cooperative solution.

and Keller and Rady (2008), our expression reduces to the expressions that they obtained.

We then apply our solution to three problems. First, in the context of multi-arm bandit problems in continuous time, we provide an explicit expression for the index of the arm. This characterization allows one to describe the optimal strategy in a multi-arm bandit problem, when each arm's payoff can be one of two Levy processes that satisfy our assumptions.

Second, we use our characterization to find the fair price of additional information. Suppose that the decision maker can purchase additional information, which refines his information regarding the payoff's distribution. Plainly, the more information that the decision maker has, the higher is his optimal payoff. Suppose that the payoff's distribution, without the additional information and with the additional information, is a Levy process. Our technique can be used to provide an explicit expression for the fair price of the additional information.

Third, we use our characterization to study the role of optimism and pessimism in bandit problem. A decision maker is called optimist if his prior probability that the payoff's distribution is High is higher than the true probability, and he is called a pessimist if his prior probability that the payoff's distribution is High is lower than the true probability. Using our characterization we find who fares better: an optimist or a pessimist.

The rest of the paper is arranged as follows. The model and the main results appear in Section 2. All proofs appear in Section 3.

2. THE MODEL AND THE MAIN RESULTS

2.1. **Reminder on Lévy Processes.** Lévy processes are the continuous-time analog of discrete time random walks with i.i.d. increments. A Lévy process $X = (X(t))_{t \geq 0}$ is a continuous-time stochastic process that (a) starts at the origin: $X(0) = 0$, (b) admits càdlàg modification,[2] and (c) has stationary independent increments. Few examples for Lévy processes are a Brownian motion, a Poisson process, and a compound Poisson process. The latter is a continuous time process in which jumps arrive according to a Poisson process and the jumps are i.i.d.[3]

We now present the Lévy-Ito decomposition of Lévy processes. Let $(X(t))$ be a Lévy process. For every Borel measurable set $A \subseteq \mathbb{R}\ \{0\}$, and every $t_0 \in \mathbb{R}$, let $N(t_0, A)$ be the number of jumps of $(X(t))$ in the time interval $[0, t_0]$ with jump size in A:

$$N(t, A) = \sharp\{0 \leq s \leq t | \Delta X(t) := X(t) - X(t-) \in A\} = \sum_{0 \leq s \leq t} \chi_A(\Delta X(s)),$$

[2]That is, it is continuous from the right, and has limits from the left: for every t_0, the limit $X(t_0-) := \lim\limits_{t \nearrow t_0} X(t)$ exists a.s. and $X(t_0) = \lim\limits_{t \searrow t_0} X(t)$.

[3]Formally, let $\lambda > 0$ and let D be a distribution over $\mathbb{R}\backslash\{0\}$. A compound Poisson process with rate λ and jump size distribution D is a continuous-time stochastic process given by $X(t) = \sum\limits_{i=1}^{N(t)} D_i$, where $N(t)$ is a Poisson process with rate λ and D_i are independent and identically distributed random variables, with distribution function D, which are also independent of $(N(t))_{t \geq 0}$.

where χ_A is the indicator of the set A. By Applebaum (2004), for every $t \in \mathbb{R}$ one can define a Borel measure on $\mathcal{B}(\mathbb{R}\backslash\{0\})$ by:

$$\varphi_t(A) = E[N(t, A)] = \int N(t, A)(\omega) dP(\omega),$$

where (Ω, P) is the underlying probability space. The measure $\nu(A) := \varphi_1(A)$ is called the Lévy measure of $(X(t))$, or the intensity measure associated with $(X(t))$.

If the Lévy measure ν is finite, that is, if $\nu(\mathbb{R}\backslash\{0\}) < \infty$, then the expected number of jumps in the time interval $[0, 1]$, and therefore in any compact time interval, is finite a.s. If $\nu(\mathbb{R}\backslash\{0\}) = \infty$, the number of jumps in any compact time interval is infinite a.s.

By the Lévy-Ito decomposition, every Lévy process can be represented as follows:

$$X(t) = \mu t + B_\sigma(t) + \int\limits_{x > |1|} x N(t, dx) + \int\limits_{x \le |1|} x \tilde{N}(t, dx), \tag{2.1}$$

with Lévy measure ν, where $B_\sigma(t)$ is Brownian motion with standard deviation σ, and $\tilde{N}(t, A) = N(t, A) - t\nu(A)$ is the compensated Poisson process, which is independent of $B_\sigma(t)$. When the Lévy measure is finite, the last two terms in (2.1) vanish, and $X(t) = \mu t + B_\sigma(t) + L(t)$, where $L(t)$ is a compound Poisson process independent of $B_\sigma(t)$: jumps arrive at a Poisson rate with expectation $\nu(\mathbb{R}\backslash\{0\})$, and the distribution of each jump is given by the Lévy measure ν.

2.2. **Lévy Bandits with a Finite Lévy Measure.** A decision maker operates a one arm bandit machine in continuous time, with a safe arm that yields a constant payoff s, and a risky arm that yields a stochastic payoff $(X(t))$, which is a Lévy process. The risky arm can be of two types, High or Low. We denote the arm's type by θ: if the type is High (resp. Low) we set $\theta = \theta_1$ (resp. $\theta = \theta_2$). If $\theta = \theta_i$, $i = 1, 2$, the risky arm yields payoff $(X^i(t))$, which is a Lévy process. In this section we assume that the Lévy measures of both $(X^1(t))$ and $(X^2(t))$ are finite, and therefore a.s. there are only finite many jumps in each compact time interval. Denote the Lévy-Ito decomposition of $(X^i(t))$, $i = 1, 2$, by $X^i(t) = \mu_i t + B_\sigma(t) + L^i(t)$, where $L^i(t)$ is a compound Poisson process with measure ν_i, independent of $B_\sigma(t)$.

Set $\bar{\nu}_i := \nu_i(\mathbb{R} \setminus \{0\})$, and denote by $H_i := \int h\nu_i(dh)/\bar{\nu}_i$ the expected jumps size of $(X^i(t))$. The expectation of the risky arm at time $t = 1$ if $\theta = \theta_i$ is $g_i := E[X^i(1)] = \bar{\nu}_i H_i + \mu_i$.

Throughout we make the following assumption, which states that the High type is better than the Low type in a strong sense.

Assumption 2.1.
A1. $g_2 < s < g_1 < \infty$.
A2. $\sigma_1 = \sigma_2 = \sigma$.
A3. $\mu_1 \ge \mu_2$.
A4. *for every* $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$, $\nu_2(A) \le \nu_1(A) < \infty$.

The first part of Assumption 2.1 merely says that a High (resp. Low) type provides higher (resp. lower) expected payoff than the safe arm, and it rules out trivial cases.

The second part states that both the High and Low type Lévy processes admit the same standard deviations of the Brownian motion component. Otherwise, the DM can distinguish between the arms in any infinitesimal time interval.

The third and forth parts of the assumption are less innocuous; they require that the Lévy measure of the High type will dominate the Lévy measure of the Low type in a strong sense, and the drift of the High type will dominate the drift of the Low type.[4]

At each time instance t, the DM chooses the proportion of time to devote to each of the two arms. If he chooses to devote a proportion k of the current time instance to the risky arm, then he receives an instant payoff $dY^k = dY^k_S + dY^k_R$, where $dY^k_S = (1-k)sdt$ is the payoff from the safe arm, and $dY^k_R = dY^k_P + dY^k_B$ is the payoff from the risky arm. Here dY^k_P is the payoff from the compound Poisson process with the Lévy measure $k \cdot \nu_i$, $i = 1, 2$, and $dY^k_B(t) := k\mu dt + \sqrt{k}\sigma dZ(t)$ is the payoff from the Brownian motion with drift, where $Z(t)$ is a standard Brownian motion.

A strategy κ is a (measurable) function, that assigns to each history a number in the interval $[0, 1]$, that is interpreted as the amount of time in the interval $[t, t + dt)$ devoted to the risky arm. In continuous-time, it is usually assumed that a strategy is predictable, that is, to determine the behavior at time t it is sufficient to know the history strictly before time t. Formally, κ_t is $\mathcal{F}^I_{t-}$-measurable, where $\mathcal{F}^I_{t-}$ is the σ-algebra generated by the stochastic process $(I(t))$ of the discounted payoff with discount rate r, up to (excluding) time t: $I(t) := \int_0^t re^{-rt} dY^\kappa(t)$. It is well known that in discrete-time bandits with independent identically distributed payoffs and geometric discounting, the optimal strategy depends only on the posterior belief (see Berry and Fristedt, 1985). Exponential discounting is the continuous-time analogues of geometric discounting. Therefore we will restrict ourselves to the class of Markovian strategies.

Definition 2.2. A *Markovian strategy* is a function $\kappa : [0, 1] \to [0, 1]$.
The interpretation of κ is as follows: if $p_t := P(\theta = \theta_1 | \mathcal{F}^I_{t-})$ is the posterior belief at time t that the risky arm is High, then the proportion of time in the time interval $[t, t + dt)$ that is devoted to the risky arm is $\kappa(p_t)$.

As is well known, the play under a Markovian strategy is not well defined. Indeed, let p_0 be the prior belief of the DM, and consider the following Markovian strategy κ:

$$k(p) = \begin{cases} 0 & \text{if } p = p_0, \\ 1 & \text{if } p \neq p_0. \end{cases} \tag{2.2}$$

That is, the DM plays safe if his belief is p_0, and he plays risky otherwise. If the payoff process is such that when the DM always chooses the risky arm, the belief is different than p_0 for every $t > 0$,[5]there are two histories consistent with κ: one in which the DM always chooses the safe arm, and one in which he chooses the

[4]In fact, our results hold without assuming $\mu_1 \geq \mu_2$; we use this restriction out of convenience.

[5]such a case occurs,e.g., when the risky arm is either a constant zero (Low), or yields payoffs according to a Poisson process (High). Keller, Rady and Cripps (2005) show that the posterior in this case decreases until a jump arrives, and then the posterior is equal to 1.

safe arm at $t = 0$ and the risky arm afterwards. To overcome this difficulty while capturing the intuition that continuous time is an idealized model of discrete time, we restrict ourselves to a sub-class of Markovian strategies. [at each time instance t the DM plays the same action over the time interval $[t, t + \Delta t)$, according to the belief at time $t-$ (the limit from the left always exists due to the càdlàg property of Lévy processes), and taking $\Delta t \to 0$]. In the example above, in this construction the safe arm is always chosen. It is not known if the restriction mentioned is enough. Therefore we also restrict the discussion for strategies that do not have two consistent histories. Various classes of strategies that satisfy the restriction are cut-off strategies: play risky as long as the posterior belief is higher then p^*, and safe otherwise, and piecewise cut-off strategies: any piecewise constant function on the posterior interval $[0, 1]$ taking values in the action interval $[0, 1]$.

2.3. **The Optimal Strategy.** The expected discounted payoff under a strategy κ when the prior is $p_0 = p$ is

$$\begin{aligned} V_\kappa(p) &= E\left[\int_0^\infty re^{-rt} dY^\kappa(t)\right] \\ &= pE\left[\int_0^\infty re^{-rt} dY^\kappa(t)\middle|\theta_1\right] + (1-p)E\left[\int_0^\infty re^{-rt} dY^\kappa(t)\middle|\theta_2\right]. \end{aligned}$$

Let $U(p) = \sup_\kappa V_\kappa(p)$ be the maximal payoff a DM can achieve. As we show below, a DM has an optimal strategy, so in fact the supremum in the definition of $U(p)$ is achieved. The function $V_\kappa(p)$ is linear with respect to p, and therefore $U(p)$, as the supremum of linear functions, is convex. By always choosing the safe arm, the DM can achieve at least s; Since $U(0) = s$, the convexity of $U(p)$ implies that U is non-decreasing.

Proposition 2.3. *$U(p)$ is monotone non-decreasing, convex, and continuous in p.*

It follows from Proposition 2.3 that there is p^* such that $U(p) = s$ if $p \leq p^*$ and $U(p) > s$ otherwise, so that the strategy $\kappa \equiv 0$ that always chooses the safe arm, is optimal for $[0, p^*]$.

Our first theorem states that there is a unique optimal strategy, which is a cut-off strategy. Moreover it provides the exact cut-off and the corresponding expected payoff in terms of the data of the problem. Let α be the unique solution of

$$f(\eta) := \int \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^\eta + \eta(\bar{\nu}_1 - \bar{\nu}_2) - \bar{\nu}_2 + \frac{1}{2}(\eta+1)\eta\left(\frac{\mu_1-\mu_2}{\sigma}\right)^2 - r = 0 \tag{2.3}$$

in $(0, \infty)$. The existence and uniqueness of such a solution are proved in Lemma 3.4 below.

Theorem 2.4. *Denote $p^* := \frac{\alpha(s-g_2)}{(\alpha+1)(g_1-s)+\alpha(s-g_2)}$. Under Assumption 2.1, the unique optimal strategy is*

$$\kappa^* = \begin{cases} 0 & \text{if } p \leq p^*, \\ 1 & \text{if } p > p^*. \end{cases}$$

The expected payoff under κ^* *is*

$$U(p) = V_{\kappa^*}(p) = \begin{cases} s & \text{if } p \le p^*, \\ g_2 + (g_1 - g_2)p + C(1-p)(\frac{1-p}{p})^\alpha & \text{if } p > p^*, \end{cases} \tag{2.4}$$

where $C = \frac{s - g_2 - p^*(g_1 - g_2)}{(1-p^*)(\frac{1-p^*}{p^*})^\alpha}$.

This theorem generalizes the results of Keller, Rady and Cripps (2005) and Bolton and Harris (1999) for the one agent problem. In Keller, Rady and Cripps (2005), the risky arm is either the constant zero (Low, so that $\nu_2 \equiv 0$), or yields a payoff according to a Poisson process (High). If the risky arm is High, the only component in the Lévy-Ito decomposition is the compound Poisson component, which is a standard Poisson process with constant jump size $\bar{h}$, so that $\nu_1(\bar{h}) = \lambda$ and zero otherwise. Therefore, $g_1 = \lambda\bar{h}$, $g_2 = 0$, and $\alpha = r/\lambda$.
In Bolton and Harris (1999), the only component in the risky arm is the Brownian motion with drift. Therefore $\nu_i \equiv 0$, so that $g_1 = \mu_1$, $g_2 = \mu_2$, and $\alpha = (-1 + \sqrt{1 + 8r\sigma^2/(\mu_1 - \mu_2)})/2$.

The cut-off p^* is an increasing function of α. As can be expected, α is increasing in the discount rate r and in $\nu_2(dh)$, and it is decreasing in $\nu_1(dh)$ and in $|\mu_1 - \mu_2|$: the DM switches to the safe arm earlier as the discount rate increases, or as jumps provide less information. Moreover, $\alpha(r = 0) = 0$ and $\alpha(r = \infty) = \infty$. In Sections 2.5 and 2.6 we extend the results to the case that the prior belief of the DM is not the true prior p_0, and to Lévy processes with infinite measure. In Sections 2.4, 2.7 and 3.7 we provide three applications to our results and techniques.

2.4. **Gittins Index.** Suppose a DM faces a multi-arm Lévy bandit, with n independent arms, such that each arm is risky and can be of two types, High or Low, as defined in Section 2.2. For each arm there is a prior probability that the type is High. At each time instance t, the DM must operate one arm. According to Gittins and Jones (1979) the optimal strategy can be formulated as follows: for each k, attach an index to the k-th arm called the Gittins index, that depends only on its data (g_1^k, g_2^k, p_2^k, α^k), and the history of the payoff received from that arm. At each time instance t, the DM chooses the arm with the highest index. The Gittins index is the parameter s such that the DM will be indifferent between the safe arm and the risky arm given his prior p_0.

Corollary 2.5. *The Gittins index for a Lévy bandit, in which all arms satisfy Assumption 2.1 is* $GI(g_1, g_2, p, \alpha) = \frac{\alpha p(g_1 - g_2) + pg_1 + \alpha g_2}{\alpha + p}$.

2.5. **The Payoff with Incorrect Prior.** In decision problems it is usually assumed that the decision maker (DM) either knows the true state of nature, or has some prior distribution over the set of states of nature. Experiments show that the prior distribution that decision makers have is often different than the true distribution. The phenomenon of overconfidence – assigning a too high probability to the good state of nature – has been observed in various areas (Svenson (1981), Baumhart (1968), Larwood and Whittaker (1977), Cross (1977), Weinstein (1980), Camerer and Lovallo (1999)). Babcock and Loewenstein (1997) argue that biases in bargaining may be self serving, and Heifetz, Shannon and Spiegel (2007) show that biases of preferences may be stable.

In every decision problem, a DM who correctly perceives the prior distribution will fare better than a DM who has some bias, and believes that the prior distribution is different than the correct one. Indeed, an optimal strategy of a DM who correctly perceives the prior distribution is a strategy that yields the highest possible gains for this prior distribution, so it yields at least as much as any other strategy, in particular, optimal strategies for incorrect prior distributions.

Denote the initial belief of the DM for the High type by q_0, and suppose that it may be different from the true probability p_0. In this section we give an exact formula for the payoff, assuming the DM plays optimally given his belief. We will also describe the optimal strategy from a different point of view, not as a cut-off strategy. This point of view is arguably closer to the way people perceive the decision problem that the DM faces.

Suppose that until time t, DM chose the risky arm, observed jumps $h_1, ...h_n$ from a compound Poisson process, and $Y_B^1(t)$ is the Brownian motion with drift payoff of the risky arm at to time t. The posterior belief $q_t := P_t(\theta_1 | h_1, ...h_n; Y_B^1(t); q_0)$ of a DM who holds the prior q_0 is:[6]

$$q_t = \frac{q_0 \frac{1}{\sqrt{2\pi t}\sigma} e^{-\frac{(Y_B^1(t)-\mu_1 t)^2}{2\sigma^2 t}} e^{-\bar{\nu}_1 t} \prod_t \nu_1(dh_j)}{q_0 \frac{1}{\sqrt{2\pi t}\sigma} e^{-\frac{(Y_B^1(t)-\mu_1 t)^2}{2\sigma^2 t}} e^{-\bar{\nu}_1 t} \prod_t \nu_1(dh_j) + (1-q_0) \frac{1}{\sqrt{2\pi t}\sigma} e^{-\frac{(Y_B^1(t)-\mu_2 t)^2}{2\sigma^2 t}} e^{-\bar{\nu}_2 t} \prod_t \nu_2(dh_j)}$$
$$= \frac{q_0 e^{\mu_1 Y_B^1(t)/\sigma^2 - \mu_1^2 t/2\sigma^2} e^{-\bar{\nu}_1 t} \prod_t \nu_1(dh_j)}{q_0 e^{\mu_1 Y_B^1(t)/\sigma^2 - \mu_1^2 t/2\sigma^2} e^{-\bar{\nu}_1 t} \prod_t \nu_1(dh_j) + (1-q_0) e^{\mu_2 Y_B^1(t)/\sigma^2 - \mu_2^2 t/2\sigma^2} e^{-\bar{\nu}_2 t} \prod_t \nu_2(dh_j)}. \tag{2.5}$$

Indeed, $\frac{1}{\sqrt{2\pi t}\sigma} e^{-\frac{(Y_B^1(t)-\mu_i t)^2}{2\sigma^2 t}}$ is the probability of receiving the payoff $Y_B^1(t)$, given the type θ_i, and $e^{-\bar{\nu}_i t} \frac{(\bar{\nu}_i t)^n}{n!} \prod_t \frac{\nu_i(dh_j)}{\bar{\nu}_i}$ is the probability of receiving the jumps that occurred until time t, given the type θ_i. The first equality in 2.5 is the Bayesian belief updating, using the independence of the components in the Lévy-Ito decomposition, given the type of the risky arm, and the second equality is obtained by eliminating common components.

Suppose the DM follows a cut-off strategy κ' with cut-off p',
$$\kappa' = \begin{cases} 0 & \text{if } p \le p', \\ 1 & \text{if } p > p'. \end{cases}$$

If $q_0 \le p'$, the DM will always choose the safe arm. If $q_0 > p'$ the DM will initially choose the risky arm. Then DM plays as long as $q_t > p'$, which, by Eq. (2.5), is equivalent to:

[6] $\prod_t \nu_i(dh_j)$ is the product of $\nu_i(dh_j)$ over all jumps h_j that occur up to time $t-$. Similarly, we use the notation $\sum_t$ as the sum over all jumps up to time $t-$.

$$\frac{q_0 e^{\mu_1 Y_B^1(t)/\sigma^2 - \mu_1^2 t/2\sigma^2 - \bar{\nu}_1 t}}{(1-q_0) e^{\mu_2 Y_B^1(t)/\sigma^2 - \mu_2^2 t/2\sigma^2 - \bar{\nu}_2 t}} \prod_t \frac{\nu_1(dh)}{\nu_2(dh)} > \frac{p'}{1-p'}. \tag{2.6}$$

By taking the natural logarithm, and rearranging the resulting terms, we obtain that this inequality is equivalent to:

(2.7)
$$\frac{1}{\sigma} Y_B^1(t) > \left[\left(\frac{\mu_1+\mu_2}{2\sigma}\right) + \frac{\sigma(\bar{\nu}_1 - \bar{\nu}_2)}{\mu_1 - \mu_2}\right] t - \frac{\sigma}{\mu_1-\mu_2} \times \left[\ln\left(\frac{q_0}{1-q_0}\right) - \ln\left(\frac{p'}{1-p'}\right)\right] - \frac{\sigma}{\mu_1-\mu_2} \sum_t \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right).$$

The right hand-side in (2.7) is a piecewise linear function $F \cdot t - E - G_t$ of t, where the slope $F := \left(\frac{\mu_1+\mu_2}{2\sigma}\right) + \frac{\sigma(\bar{\nu}_1-\bar{\nu}_2)}{\mu_1-\mu_2}$ is independent of t, the intercept at $t = 0$ is $E := \frac{\sigma}{\mu_1-\mu_2} \times \left[\ln\left(\frac{q_0}{1-q_0}\right) - \ln\left(\frac{p'}{1-p'}\right)\right]$, and $G_t := \frac{\sigma}{\mu_1-\mu_2} \sum_t \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right)$. Denote by $G_h := \frac{\sigma}{\mu_1-\mu_2} \ln\left(\frac{\nu_1(dh)}{\nu_2(dh)}\right)$ the contribution of a jump of size h to the intercept.

From Eq. (2.7) we obtain the following alternative description of the optimal strategy: The DM has a time dependent cut-off which is piecewise linear. The slope of the cut-off function is always F, and whenever there is a jump of size h, the cut-off decreases by G_h (see figure –). The DM chooses the risky arm as long as his current payoff from the continuous part of the Lévy process, $Y_B^1(t)$, exceeds the cut-off.

Therefore, at first, the DM plays until the payoff from the continuous part divided by the standard deviation satisfies: $\frac{1}{\sigma} Y_B^1(t) \le F \cdot t - E$; if a jump of size h occurs before he stops, then E decreases by G_h and this behavior repeats itself. In case there is no Brownian motion component, the DM chooses the risky arm for a fixed amount of time, and then switches to the safe arm, unless a jump occurs, and then the amount of time to choose the risky arm increases, as a function of $\frac{\nu_1(dh)}{\nu_2(dh)}$.

Theorem 2.6. *Under Assumption 2.1, for every $p_0 \in [0,1]$, the payoff of a DM using the cut-off strategy with cut-off p', denoted $U(p_0, q_0)$, is as follows:*[7] *if $q_0 \ge p'$,*

(2.8)
$$U(p_0,q_0) = (s-g_1)p_0 \left(\frac{1-q_0}{q_0}\right)^{\alpha+1} \left(\frac{p'}{1-p'}\right)^{\alpha+1} + (s-g_2)(1-p_0)\left(\frac{1-q_0}{q_0}\right)^{\alpha} \left(\frac{p'}{1-p'}\right)^{\alpha} + g_1 p_0 + g_2(1-p_0),$$

while if $q_0 < p'$, it is given by $U(p_0, q_0) = s$.

One can verify that when the DM holds the correct prior, and plays according to the optimal strategy κ^*, Eq. (2.8) coincides with (2.4): $U(p_0, p_0) = U(p_0)$. Note that $U(p_0, q_0)$ is continuous in all its parameters.

[7] We omit the dependency of U on p', s, g_1, g_2.

2.6. **Lévy Bandits with an Infinite Lévy Measure.** In this section we deal with Lévy processes with infinite Lévy measures ν_i, $i = 1, 2$. We make the following assumption.

Assumption 2.7.
AA5. $g_2 < s < g_2$.
AA6. $\sigma_1 = \sigma_2 = \sigma$.
AA7. $\mu_1 \geq \mu_2$.
AA8. $\nu_1(R \setminus \{0\}) = \nu_2(R \setminus \{0\}) = \infty$.
AA9. $\nu_1(R \setminus \{0\}) - \nu_2(R \setminus \{0\}) < \infty$.
AA10. $\nu_1(A) \geq \nu_2(A)$ *for every* $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$.

As in Assumption 2.1, the first, second and last parts of Assumption 2.7 state that the High type is better than the Low type in a strong sense. The second part states that both the High and Low type Lévy processes admit the same standard deviations of the Brownian motion components. The forth part of the assumption states that both types have infinite Lévy measures, while the third part states that $\nu_1 - \nu_2$ is a finite measure. In case $\nu_1 - \nu_2$ is an infinite measure the DM can distinguish between the types in any infinitesimal time interval.

Let $B_0 \in \mathcal{B}(\mathbb{R} \setminus \{0\})$ be a maximal set (up to measure zero), such that $\nu_1(B_0) \geq 0 = \nu_2(B_0)$; occurrence of a jump from B_0 indicates that the risky arm is High. The Radon-Nikodym derivative $g(h) = \frac{\nu_1(dh)}{\nu_2(dh)}$ exists on $\mathbb{R} \setminus (\{0\} \cup B_0)$. The assumption that $\nu_1 - \nu_2$ is a finite measure implies that $\nu_1(B_0) < \infty$. In this section we limit the analysis only to cut-off strategies. The expected discounted payoff $V_\kappa(p)$ that a strategy yields, as well as the optimal payoff $U(p)$, are defined as in Section 2.2.

It is well known that the Lévy measure ν_i satisfies $\nu_i([-\varepsilon, \varepsilon]^c) < \infty$[8] for every $\varepsilon > 0$ (see Applebaum (2004)). We will therefore analyze the problem with infinite Lévy measure as the limit of models with finite Lévy measures: we will study the case in which the DM does not observe, nor receives, the payoff of jumps with absolute value less then $1/n$. Denote this problem by D_n, and denote the optimal payoff in this problem by $U_n(p)$. The problem D_n falls into the class studied in Sections 2.2, and 2.3, and therefore the optimal strategy κ_n is a cut-off strategy. We will solve the original problem D by considering the limit of the solutions for D_n, in Proposition (2.10) we prove that $\lim_{n\to\infty} U_n(p)$ exists.

Denote $A_n = (-1/n, 1/n)^c$. For every cut-off strategy κ, let $(Y_n^\kappa(t))_t$ be the process generated by the payoff process $(Y^\kappa(t))_t$, ignoring jumps with absolute value less then $1/n$ (jumps in A_n^c). Let $\mathcal{F}_t^n := \sigma\left((Y_n^\kappa(s))_{s\leq t}\right)$ (resp. $\mathcal{F}_t := \sigma\left((Y^\kappa(s))_{s\leq t}\right)$) be the σ-algebra generated by the process $(Y_n^\kappa(s))_s$ (resp. $(Y^\kappa(s))_s$), up to time t.
Let $\left(Y_{n,B_0}^\kappa(t)\right)_t$ (resp. $\left(Y_{B_0}^\kappa(t)\right)_t$) be the process generated by the payoff process $(Y^\kappa(t))_t$, ignoring jumps in $A_n^c \cup B_0$ (resp. B_0). Let $\mathcal{G}_t^n := \sigma\left(\left(Y_{n,B_0}^\kappa(s)\right)_{s\leq t}\right)$ (resp. $\mathcal{G}_t := \sigma\left((Y^\kappa(s))_{s\leq t}\right)$) be the σ-algebra generated by the process $\left(Y_{n,B_0}^\kappa(s)\right)_s$ (resp. $\left(Y_{B_0}^\kappa(s)\right)_s$), up to time t.

[8]For every set $X \subseteq \mathbb{R}$, X^c is the complement of X.

We will now construct the posterior process related to the general Lévy case. By Assumption 2.7, $\nu_2(A_n) \leq \nu_1(A_n) < \infty$, so that Assumption 2.1 is satisfied for the problem D_n, and therefore we can use the posterior process (2.5). Suppose the DM always operates the risky arm. Then the posterior process using the filtaration $(\mathcal{G}_t^n)_t$ satisfies:

(2.9)

$$\ln\left(\frac{p(\theta_1|\mathcal{G}_t^n)}{1-p(\theta_1|\mathcal{G}_t^n)}\right) = \ln\left(\frac{p_0}{1-p_0}\right) + \frac{Y_B^1(t)}{\sigma^2}(\mu_1-\mu_2) - \frac{t}{2\sigma^2}(\mu_1^2-\mu_2^2) - (\bar{\nu}_1^n - \bar{\nu}_2^n)t + \sum_{\substack{t \\ h_j \in A_n \setminus B_0}} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right),$$

where $\bar{\nu}_i^n = \int_{A_n} \nu_i(dh)$ is the rate of jumps with absolute value at least $1/n$ given the type θ_i. From this process we can see that the contribution of each jump h is by a factor $\frac{\nu_1(dh)}{\nu_2(dh)}$.

Since the processes $\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right)$, and $\left(\ln\left(\frac{p(\theta_1|\mathcal{G}_t^n)}{1-p(\theta_1|\mathcal{G}_t^n)}\right)\right)_n$ have stationary independent increments, one can verify that both are Lévy processes. Moreover, by Assumption 2.7(AA10) it follows that these Lévy processes have only positive jumps a.s. The next Lemma states that the Lévy process $\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right)$ is the limit of the Lévy processes $\ln\left(\frac{p(\theta_1|\mathcal{G}_t^n)}{1-p(\theta_1|\mathcal{G}_t^n)}\right)$. The following lemma describes the evolution of the posterior process when ignoring jumps in B_0 (that reveal that the risky arm is high). It also says that this posterior process is the limit, as n goes to ∞, of the posterior processes when ignoring jumps in $A_n^c \cup B_0$.

Lemma 2.8. *Suppose the DM always operates the risky arm. Then the posterior process using the filtration $(\mathcal{G}_t)_t$ satisfies:*

(2.10)

$$\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right) = \ln\left(\frac{p_0}{1-p_0}\right) + \frac{Y_B^1(t)}{\sigma^2}(\mu_1-\mu_2) - \frac{t}{2\sigma^2}(\mu_1^2-\mu_2^2) - (\bar{\nu}_1 - \bar{\nu}_2)t + \sum_{\substack{t \\ h_j \in \mathbb{R}\setminus(\{0\}\cup B_0)}} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right).$$

Moreover, $P\left(p(\theta_1|\mathcal{G}_t^n)(\omega) \to p(\theta_1|\mathcal{G}_t)(\omega)\ \forall t\right) = 1$.

The next proposition states that a DM who operates only the risky arm, and ignores jumps that belong to B_0, will never be sure whether the arm is High or Low. The intuition is that since jumps with size that is not in B_0 may occur under both X^1 and X^2, the conditional probability of the High type given such jumps cannot be 0 or 1. In particular, if all jumps are in $\mathbb{R} \setminus B_0$, that is, if $\nu_1(B_0) = 0$, then the DM will never know the true type of the risky arm. Since in continuous time the amount of information that is revealed in a compact time interval may be unbounded, it is not straight forward that such a result should hold.

Proposition 2.9. *Under Assumption 2.7, $0 < p_t(\theta_1|\mathcal{G}) < 1$ with probability 1, for every $t \geq 0$.*

By Proposition 2.9, and by the discussion above, it follows that the posterior probability 1 is achieved if and only if a jump of size in B_0 occurs. Formally:

$$\ln\left(\frac{p(\theta_1|\mathcal{F}_t^n)(\omega)}{1-p(\theta_1|\mathcal{F}_t^n)(\omega)}\right) = \ln\left(\frac{p(\theta_1|\mathcal{G}_t^n)(\omega)}{1-p(\theta_1|\mathcal{G}_t^n)(\omega)}\right) + \delta_{B_0\cap A_n}(\omega) \tag{2.11}$$

convergence a.s to

$$\ln\left(\frac{p(\theta_1|\mathcal{F}_t)(\omega)}{1-p(\theta_1|\mathcal{F}_t)(\omega)}\right) = \ln\left(\frac{p(\theta_1|\mathcal{G}_t)(\omega)}{1-p(\theta_1|\mathcal{G}_t)(\omega)}\right) + \delta_{B_0}(\omega), \tag{2.12}$$

where $\delta_D(\omega) = \infty$ if a jump that belongs to D occurred along the path ω, and $\delta_D(\omega) = 0$ otherwise. The statement follows since $p(\theta_1|\mathcal{F}_t)(\omega) = 1$ if and only if $\ln\left(\frac{p(\theta_1|\mathcal{F}_t)(\omega)}{1-p(\theta_1|\mathcal{F}_t)(\omega)}\right) = \infty$.

Consider the problem D_n. For $i = 1, 2$, let $\bar{\nu}_i^n = \int_{A_n} \nu_i(dh)$ be the intensity of jumps with size in A_n, let $H_i^n = \int_{A_n} h\nu_i(dh)$ be the expected jump size, given that a jump from A_n occurred, let $g_i^n = \bar{\nu}_i^n H_i^n + \mu_i$ be the expected payoff ignoring jumps in A_n^c, and let α_n be the solution of $\int_{A_n} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha_n} + \alpha_n(\bar{\nu}_1^n - \bar{\nu}_2^n) - \bar{\nu}_2^n + \frac{1}{2}(\alpha_n+1)\alpha_n(\tilde{\mu_1} - \tilde{\mu_2})^2 = r$. The following proposition lists the continuity properties needed to analyze the case of general Lévy processes. It states that as the size of the interval which the DM ignores vanishes, the solution converges to the optimal solution of the limit problem. We will then obtain that the optimal payoff the DM obtains, ignoring jumps outside A_n, converges to the optimal payoff using all the information available.

Proposition 2.10. *Under Assumption 2.7,* $\lim_{n\to\infty} U_n(p) = U(p)$.

We are now ready to find the optimal strategy for Lévy processes with infinite measures. Let $g_i' = \lim_{n\to\infty} g_i^n$ be the limit of the expected payoffs, let $\alpha' = \lim_{n\to\infty} \alpha_n$ be the limit of the α_n,[9] let $p' = \lim_{n\to\infty} p_n^* = \lim_{n\to\infty} \frac{\alpha_n(s-g_2^n)}{(1+\alpha_n)(g_1^n-s)+\alpha_n(s-g_2^n)}$ be the limit of the optimal cut-offs in D_n, let $C' = \lim_{n\to\infty} C^n = \lim_{n\to\infty} \frac{s-g_2^n-p^n(g_1^n-g_2^n)}{(1-p^n)(\frac{1-p^n}{p^n})^{\alpha_n}}$ be the limit of the constants C_n of Theorem 2.4 in D_n, and let $\kappa' = \begin{cases} 0 & \text{if } p \le p', \\ 1 & \text{if } p > p', \end{cases}$ be a cut-off strategy defined by the limit of the optimal cut-off strategies in D_n.

Theorem 2.11. *Under Assumption 2.7, κ' is an optimal strategy in the problem D among all cut-off strategies, and it is the only optimal cut-off strategy. The value function is:*

$$U(p) = V_{\kappa'}(p) = \begin{cases} s & \text{if } p \le p', \\ g_2 + (g_1 - g_2)p + C'(1-p)(\frac{1-p}{p})^{\alpha'} & \text{if } p > p'. \end{cases}$$

We now present the equivalent of Theorem 2.6 for Lévy processes with infinite Lévy measures. For every $q_0 \in [0,1]$, let $\kappa(q_0)$ be the strategy that plays as κ, assuming the prior belief is q_0 rather then p_0. Define $M_\kappa := \int_0^\infty re^{-rt}dY^\kappa(t)$, the discounted payoff under the strategy κ. Set $N_\kappa^n := \int_0^\infty re^{-rt}dY_n^\kappa(t)$. Recall that $Y_n^\kappa(t)$ is the process generated by the payoff process $Y^\kappa(t)$, ignoring jumps in A_n^c.

[9] Existence of the limit follows by monotonicity, see Lemma 3.7 below.

This is the discounted payoff under κ, ignoring jumps in A_n^c. Let $\mathcal{H}_t$ be a σ-algebra ($\mathcal{H}_t \in \{\mathcal{F}_t^n, \mathcal{F}_t\}$), denote $E[M_\kappa|\mathcal{H}]$ (resp. $E[N_\kappa^n|\mathcal{H}]$) the expected payoff (resp. the expected payoff ignoring jumps in A_n^c) for a DM using the cut-off strategy κ, and updating his belief according to $(\mathcal{H}_t)_t$.

Theorem 2.12. *Under Assumption 2.7, the expected payoff of the DM, if he plays according to the optimal cut-off strategy κ', when the true prior is p_0 and the DM's subjective belief is $q_0 \geq p_0$, is:*

(2.13)

$$U(p_0, q_0) = E\left[M_{\kappa'(q_0)}|\mathcal{F}\right] = (s - g_1')p_0 \left(\frac{1-q_0}{q_0}\right)^{\alpha'+1} \left(\frac{p'}{1-p'}\right)^{\alpha'+1} + (s - g_2')(1-p_0)\left(\frac{1-q_0}{q_0}\right)^{\alpha'} \left(\frac{p'}{1-p'}\right)^{\alpha'} + g_1' p_0 + g_2'(1-p_0).$$

2.7. **Optimism vs Pessimism.** As mentioned before, the phenomenon of over confidence is common in many decision problems. In this section we apply the result of Section 2.6, and investigate who will fare better in one-arm bandit problems, an optimist who assigns a probability higher than the true probability to the High type, or a pessimist who assigns a probability lower than the true probability to the High type.

Suppose that there are two decision makers, DM1 and DM2, who face independent identical copies of the decision problem. DM1 is an *optimist*, and believes that the probability of High is $p_0 + \rho$, where $\rho > 0$, while DM2 is a *pessimist*, and believes that the probability of High is $p_0 - \rho$, and both play optimally given their beliefs. If $p_0 - \rho \leq p^*$, the pessimist will always choose the safe arm, since according to his subjective belief the prior is at most the cut-off. Assume then that $p_0 - \rho > p^*$. For every $\epsilon \in [p^* - p_0, 1 - p_0]$ denote by $V_{p_0}(\epsilon)$ the expected payoff for a DM playing optimally according to the incorrect prior $p_0 + \epsilon$, where ϵ may be negative. It turns out that the answer regarding who will fare better, an optimist or a pessimist, depends on α.

Theorem 2.13. *Assume that Assumption 2.7 holds.*
1. If $\alpha > 1$, then for every $\epsilon > 0$ such that $p^* < p_0 \pm \epsilon \leq 1$ we have $V_{p_0}(\epsilon) > V_{p_0}(-\epsilon)$: an optimist will fare better.
2. If $0 < \alpha < 1$, then for every $\epsilon > 0$ such that $p^* < p_0 \pm \epsilon \leq \frac{\alpha+2}{3}$ we have $V_{p_0}(\epsilon) > V_{p_0}(-\epsilon)$: an optimist will fare better; for every $\epsilon > 0$ such that $\frac{\alpha+2}{3} < p_0 \pm \epsilon < 1$ we have $V_{p_0}(\epsilon) < V_{p_0}(-\epsilon)$: a pessimist will fare better.

Thus, optimism is better than pessimism, unless α is low and $p_0 - \varepsilon$ is high. That is, pessimism is better only if the following two conditions are satisfied,
1) The pessimist assigns high probability to the High type.
2) The two DM's are sufficiently patient, or it is difficult to distinguish between the two types of the risky arm.
Otherwise, an optimist will fare better.

Since $p^* = \frac{\alpha(s-g_2)}{(\alpha+1)(g_1-s)+\alpha(s-g_2)}$, the condition $p^* < \frac{\alpha+2}{3}$ is not always satisfied. If $3g_1 + g_2 \geq 4s$, then $p^* < \frac{\alpha+2}{3}$. If $3g_1 + g_2 < 4s$, then $p^* < \frac{\alpha+2}{3}$ if and only if $\alpha < \frac{4s-3g_1-g_2-\sqrt{(4s-3g_1-g_2)^2-8(g_1-g_2)(g_1-s)}}{2(g_1-g_2)}$ or $\alpha > \frac{4s-3g_1-g_2+\sqrt{(4s-3g_1-g_2)^2-8(g_1-g_2)(g_1-s)}}{2(g_1-g_2)}$.

2.8. **Information Pricing.** Often the DM can purchase additional information, that improves his information regarding the state of nature. In the present section we illustrate the use of our technique to the following problem: suppose the DM faces a one-arm bandit problem with Lévy payoffs, and suppose that the DM observes the continuous part of the Lévy process as well as jumps with absolute value at least $\frac{1}{m}$, though his payoff is affected by jumps with absolute value at least $\frac{1}{n}$. If $n < m$, the DM's payoff is affected by small jumps that he does not observe, while if $n > m$ the DM observes small jumps that do not affect his payoff. How much the DM would pay to be able to observe small jumps (that may or may not affect his payoff, yet they increase the rate of learning)?

In the first part of Theorem 2.14 below we characterize the value of the problem when the DM does not observe small jumps that do affect his payoff, while in the second part of Theorem 2.14 we characterize the value of the problem when the DM observes small jumps that do not affect his payoff. The difference between two such values is the amount that the DM will be willing to pay for additional information. The fair price of additional information in other cases can be similarly derived.

Let β_1 be the unique solution of

(2.14)
$$f_1(\eta) := \int_{A_m} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\eta} + \eta(\bar{\nu}_1^n - \bar{\nu}_2^n) + \bar{\nu}_1^n - \bar{\nu}_1^m - \bar{\nu}_2^n + \frac{1}{2}(\eta+1)\eta\left(\frac{\mu_1-\mu_2}{\sigma}\right)^2 - r = 0$$

in $(-1,\infty)$. Let β_2 be the unique solution of

(2.15)
$$f_2(\eta) := \int_{A_m} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\eta} + \eta(\bar{\nu}_1^n - \bar{\nu}_2^n) - \bar{\nu}_2^m + \frac{1}{2}(\eta+1)\eta\left(\frac{\mu_1-\mu_2}{\sigma}\right)^2 - r = 0$$

in $(0,\infty)$. The existence and uniqueness of such solutions are proved in Lemma 3.8 below.

Theorem 2.14. *1.Let $n < m$. The expected payoff to a decision maker who uses a cut-off strategy κ' with cut-off p', observes the continuous part of the process as well as jumps larger than $\frac{1}{m}$, and whose payoff is determined by the continuous part of the process as well as jumps larger than $\frac{1}{n}$, is*

$$V_{p'}(p) := E[N^n_{\kappa'}|\mathcal{F}^m](p) = g_2^n + (g_1^n - g_2^n)p + (s-g_1^n)(1-p)\left(\frac{1-p}{p}\right)^{\beta_1}\left(\frac{p'}{1-p'}\right)^{(\beta_1+1)}$$
$$+ (s-g_2^n)(1-p)\left(\frac{1-p}{p}\right)^{\beta_2}\left(\frac{p'}{1-p'}\right)^{\beta_2}$$

for every $p > p'$, and $V_{p'}(p) = s$ otherwise.

2. Let $n < m$. The expected payoff to a decision maker who uses a cut-off strategy κ' with cut-off p', who observes the continuous part of the process as well as jumps *larger than* $\frac{1}{m}$*, and whose payoff is determined by the continuous part of the process as well as jumps larger than* $\frac{1}{n}$*, is*

$$V_{p'}(p) := E[N^n_{\kappa'}|\mathcal{F}^m](p) = g^n_2 + (g^n_1 - g^n_2) + C_{n,p',\alpha_m}(1-p)\left(\frac{1-p}{p}\right)^{\alpha_m}$$

for every $p > p'$*, and* $V_{p'}(p) = s$ *otherwise, where* $C_{n,p',\alpha_m} := \frac{s-g^n_2-p'(g^n_1-g^n_2)}{(1-p')(\frac{1-p'}{p'})^{\alpha_m}}$.

In both cases, using the same method that we use to prove Theorem 2.4, one can find the optimal strategy.

3. Appendix

3.1. **Notations and Formulas on General Lévy Processes.** Throughout the proofs we use formulas that are derived using Borodin (1996) p.197 - 223. Hereafter, $B^\mu(t) = Z(t) + \left[\frac{2\tilde{\mu} - \tilde{\mu_1} - \tilde{\mu_2}}{2} - \frac{\bar{\nu}_1 - \bar{\nu}_2}{\tilde{\mu_1} - \tilde{\mu_2}}\right] t$, is a standard Brownian motion with drift $\left[\frac{2\tilde{\mu} - \tilde{\mu_1} - \tilde{\mu_2}}{2} - \frac{\bar{\nu}_1 - \bar{\nu}_2}{\tilde{\mu_1} - \tilde{\mu_2}}\right] t$. Recall that, $\tilde{\mu} = \frac{\mu}{\sigma}$ is determined by θ. Following the notations of Section 3.5, denote by $F_\mu := \frac{2\tilde{\mu} - \tilde{\mu_1} - \tilde{\mu_2}}{2} - \frac{\bar{\nu}_1 - \bar{\nu}_2}{\tilde{\mu_1} - \tilde{\mu_2}}$ the drift of $B^\mu(t)$, and denote

(3.1)

$$\begin{aligned}
&p_{t,h,x} := P(\theta_1|\tau < T, \tau = t, B^\mu(\tau) \in dx, h) \\
&= \frac{P(\tau < T, \tau = t, B^\mu(\tau) \in dx, h|\theta_1)P(\theta_1)}{P(\tau < T, \tau = t, B^\mu(\tau) \in dx, h|\theta_1)P(\theta_1) + P(\tau < T, \tau = t, B^\mu(\tau) \in dx, h|\theta_2)P(\theta_2)} \\
&= \frac{p_0 P_{\theta_1}(\tau < T) f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\nu_1(dh)/\bar{\nu}_1}{p_0 P_{\theta_1}(\tau < T) f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\nu_1(dh)/\bar{\nu}_1 + (1-p_0)P_{\theta_2}(\tau < T) f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\nu_2(dh)/\bar{\nu}_2}.
\end{aligned}$$

This is the posterior belief that the type is θ, given that the first jump that occurred in the time interval $[t, t+dt)$ has size h, it occurred before the DM switched to the safe arm, and the Brownian motion $B^\mu(t)$ is in the interval $[x, x+dx)$.

The probability that the DM switches to the safe arm before the first jump appeared is:

$$P_\theta(\tau > T) = P_\theta(\inf_{0<s<\tau} B^\mu(s) \le -E) = e^{-E(F_\mu + \sqrt{2\bar{\nu} + F^2_\mu})}. \tag{3.2}$$

The expected discounted payoff from the continuous part of the risky arm, until the switching time to the safe arm, given the DM switched before the first jump, is

(3.3)

$$\begin{aligned}
E_\theta\left[\int_0^T re^{-rt}dY^1_B(t)\middle|\tau > T\right] &= \mu E_\theta\left[\int_0^T re^{-rt}dt\middle|\tau > T\right] + \sigma E_\theta\left[\int_0^T re^{-rt}dZ_t\middle|\tau > T\right] \\
&= \mu E_\theta[1 - e^{-rT}|\tau > T] = \mu(1 - E_\theta[e^{-rT}|\tau > T]).
\end{aligned}$$

The expected discounted payoff from the safe arm, after the switching time to the safe arm, given the DM switched before the first jump, is:

$$E_\theta\left[\int_T^\infty re^{-rt}sdt|\tau > T\right] = sE_\theta\left[e^{-rT}|\tau > T\right]. \tag{3.4}$$

The expected discounted payoff from the continuous partof the risky arm, until the first jump occurs, given the first jump occurred before the switching time, is:

$$E_\theta\left[\int_0^\tau re^{-rt}dY_B^1(t)|\tau < T\right] = \mu(1 - E_\theta\left[e^{-r\tau}|\tau < T\right]). \tag{3.5}$$

The expressions on the right hand side of (3.4), and (3.5), can be re-written using the following two lists of equalities:

(3.6)

$$\begin{aligned}
E_\theta[e^{-rT}|\tau > T] &= \frac{E_\theta[e^{-rT}, \tau > T]}{P_\theta(\tau > T)} = \frac{\int e^{-rt_1}P_\theta(T \in dt_1, t_1 < \tau)dt_1}{P_\theta(\tau > T)} \\
&= \frac{\int e^{-rt_1}P_\theta(T \in dt_1)P_\theta(t_1 < \tau)dt_1}{P_\theta(\tau > T)} = \frac{\int e^{-rt_1}P_\theta(T \in dt_1)e^{-\bar{\nu}t_1}dt_1}{P_\theta(\tau > T)} \\
&= \frac{\int e^{-(r+\bar{\nu})t_1}P_\theta(T \in dt_1)dt_1}{P_\theta(\tau > T)} = \frac{E_\theta[e^{-(r+\bar{\nu})T}]}{P_\theta(\tau > T)} \\
&= \frac{e^{-E(F_\mu+\sqrt{2(r+\bar{\nu})+F_\mu^2})}}{P_\theta(\tau > T)} = \frac{P_\theta(\tau^r > T)}{P_\theta(\tau > T)},
\end{aligned}$$

where $P_\theta(\tau^r > t) = e^{-(r+\bar{\nu})t}$, and $P_\theta(\tau^r > T) = e^{-E(F_\mu+\sqrt{2(\bar{\nu}+r)+F_\mu^2})}$.

(3.7)

$$\begin{aligned}
E_\theta[e^{-r\tau}|\tau < T] &= \int e^{-rt_1}P_\theta(\tau \in dt_1|\tau < T)dt_1 = \int \frac{e^{-rt_1}P_\theta(\tau \in dt_1, t_1 < T)}{P_\theta(\tau < T)}dt_1 \\
&= \int \frac{e^{-rt_1}P_\theta(\tau \in dt_1)P_\theta(t_1 < T)}{P_\theta(\tau < T)}dt_1 = \int \frac{e^{-rt_1}\bar{\nu}e^{-\bar{\nu}t_1}P_\theta(t_1 < T)}{P_\theta(\tau < T)}dt_1 \\
&= \frac{\bar{\nu}}{\bar{\nu}+r}\int (\bar{\nu}+r)\frac{e^{-(r+\bar{\nu})t_1}P_\theta(t_1 < T)}{P_\theta(\tau < T)}dt_1 \\
&= \frac{\bar{\nu}}{\bar{\nu}+r}\int \frac{P_\theta(\tau^r \in dt_1)P_\theta(t_1 < T)}{P_\theta(\tau < T)}dt_1 = \frac{\bar{\nu}}{\bar{\nu}+r}\frac{P_\theta(\tau^r < T)}{P_\theta(\tau < T)}.
\end{aligned}$$

In the proof of Theorem 2.6 we use the following identities:

(3.8)

$$\int_0^\infty f^\theta_{(B^\mu(\tau),\tau)|\tau<T}(x,t)e^{-\gamma t}dt = \frac{e^{-\gamma t}P_\theta\left(\inf_{0<s<t} B^\mu(s) \geq -E, B^\mu(t) \in dx, \tau \in dt\right)}{P_\theta(\tau < T)} =$$

$$= \frac{\int e^{-\gamma t}P_\theta\left(\inf_{0<s<t} B^\mu(s) \geq -E, B^\mu(t) \in dx\right) P_\theta(\tau \in dt)dt}{P_\theta(\tau < T)} =$$

$$= \frac{\int e^{-\gamma t}P_\theta\left(\inf_{0<s<t} B^\mu(s) \geq -E, B^\mu(t) \in dx\right) \bar{\nu}e^{-\bar{\nu}t}dt}{P_\theta(\tau < T)} =$$

$$= \frac{\bar{\nu}}{\bar{\nu}+\gamma}\frac{\int (\bar{\nu}+\gamma)e^{-(\gamma+\bar{\nu})t}P_\theta\left(\inf_{0<s<t} B^\mu(s) \geq -E, B^\mu(t) \in dx\right) dt}{P_\theta(\tau < T)} =$$

$$= \frac{\bar{\nu}}{\bar{\nu}+\gamma}\frac{\int P_\theta\left(\inf_{0<s<t} B^\mu(s) \geq -E, B^\mu(t) \in dx\right) P_\theta(\tau^\gamma \in dt)dt}{P_\theta(\tau < T)} =$$

$$= \frac{\bar{\nu}}{\bar{\nu}+\gamma}\frac{\int P_\theta\left(\inf_{0<s<t} B^\mu(s) \geq -E, B^\mu(t) \in dx, \tau^\gamma \in dt\right) dt}{P_\theta(\tau < T)} =$$

$$= \frac{\bar{\nu}}{\bar{\nu}+\gamma}\frac{P_\theta(B^\mu(\tau^\gamma) \in dx, \tau^\gamma < T)}{P_\theta(\tau < T)}.$$

(3.9)

$$\int_{-E}^{\infty}\int_{0}^{\infty} f^{\theta}_{(B^{\mu}(\tau),\tau)|\tau<T}(x,t)e^{-\gamma t}e^{-\delta x}dtdx = \frac{\bar{\nu}}{\bar{\nu}+\gamma}\frac{\int_{-E}^{\infty} P_{\theta}(B^{\mu}(\tau^{\gamma})\in dx, \tau^{\gamma}<T)e^{-\delta x}dx}{P_{\theta}(\tau<T)} =$$

$$= \frac{\bar{\nu}}{(\bar{\nu}+\gamma)P_{\theta}(\tau<T)}\int_{-E}^{\infty} P_{\theta}\left(\inf_{0<s<\tau^{\gamma}} B^{\mu}(s) > -E, B^{\mu}(\tau^{\gamma})\in dx\right)e^{-\delta x}dx =$$

$$= \frac{\bar{\nu}}{(\bar{\nu}+\gamma)P_{\theta}(\tau<T)}\int_{-E}^{\infty}\left[P_{\theta}(B^{\mu}(\tau^{\gamma})\in dx)e^{-\delta x} - P_{\theta}\left(\inf_{0<s<\tau^{\gamma}} B^{\mu}(s)\le -E, B^{\mu}(\tau^{\gamma})\in dx\right)e^{-\delta x}\right]dx =$$

$$= \frac{\bar{\nu}+\gamma}{\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}}\frac{\bar{\nu}}{(\bar{\nu}+\gamma)P_{\theta}(\tau<T)}\cdot$$

$$\cdot\int_{-E}^{\infty}\left[e^{(F_{\mu}-\delta)x-|x|\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}} - e^{-x(\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}+\delta-F_{\mu})}e^{-2E\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}}\right]dx =$$

$$= \frac{\bar{\nu}}{\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}P_{\theta}(\tau<T)}\cdot\left[\int_{0}^{\infty} e^{-x(\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}+\delta-F_{\mu})} + \int_{-E}^{0} e^{x(\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}-\delta+F_{\mu})}\right.$$

$$\left. -e^{-2E\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}}\int_{-E}^{\infty} e^{-x(\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}+\delta-F_{\mu})}\right] =$$

$$= \frac{\bar{\nu}(1-e^{-E(\sqrt{2(\bar{\nu}+\gamma)+F_{\mu}^2}+F_{\mu}-\delta)})}{(\bar{\nu}+\gamma+\delta F_{\mu}-\delta^2/2)P_{\theta}(\tau<T)}.$$

Where the first equality follows by (3.8).

3.2. **Proof of Proposition 2.3.** It is left to prove that U is continuous. Since $U(p)$ is convex, it is continuous on $(0,1)$. If the DM had known the true type of the arm, his optimal payoff would have been $pg_1+(1-p)s$. Since this information is not available, $U(p)\le pg_1+(1-p)s$, which implies $\lim_{p\to 0}U(p)=s$. Since the DM can follow the strategy that always selects the risky arm, $U(p)\ge pg_1+(1-p)g_2$. This implies that $U(1)=g_1=\lim_{p\to 1}U(p)$.

3.3. **The Control Equation.** In this section we express the optimal payoff $U(p)$ using the dynamic programming principle. This representation extends that in Bolton and Harris (1999) to the more general setup of Lévy processes. To simplify notations, it is convenient to divide various expressions by the standard deviation σ: set $\tilde{\mu}=\frac{\mu}{\sigma}$, and denote $d\tilde{Y_B^k} := \frac{1}{\sigma}dY_B^k = \sqrt{k}\tilde{\mu}dt + dZ(t)$. Then $P(d\tilde{Y_B^k}|\tilde{\mu}) = P(d\tilde{Y_B^k}|\mu) = \frac{1}{\sqrt{2\pi dt}}e^{-\frac{(d\tilde{Y_B^k}(t)-\sqrt{k}\tilde{\mu}dt)^2}{2dt}} = C\cdot e^{\sqrt{k}\tilde{\mu}d\tilde{Y_B^k}-\frac{k\tilde{\mu}^2dt}{2}}$, where C is independent of μ. Denote $\tilde{P}(d\tilde{Y_B^k}|\tilde{\mu}) := 1+\sqrt{k}\tilde{\mu}d\tilde{Y_B^k}$, and note that $e^{\sqrt{k}\tilde{\mu}d\tilde{Y_B^k}-\frac{k\tilde{\mu}^2dt}{2}} = 1+\sqrt{k}\tilde{\mu}d\tilde{Y_B^k}+o(dt^2)$.

Let $p = P_t(\theta_1)$ be the belief at time t that the risky arm is High. Then:

$$\begin{aligned} P_{t+dt}(\theta_1) &= \frac{P(dY_B^k|\theta_1)P(dY_P^k|\theta_1)p}{P(dY_B^k|\theta_1)P(dY_P^k|\theta_1)p + P(dY_B^k|\theta_2)P(dY_P^k|\theta_2)(1-p)} \\ &= \frac{P(d\tilde{Y_B^k}|\theta_1)P(dY_P^k|\theta_1)p}{P(d\tilde{Y_B^k}|\theta_1)P(dY_P^k|\theta_1)p + P(d\tilde{Y_B^k}|\theta_2)P(dY_P^k|\theta_2)(1-p)} \\ &= \frac{\tilde{P}(d\tilde{Y_B^k}|\theta_1)P(dY_P^k|\theta_1)p}{\tilde{P}(d\tilde{Y_B^k}|\theta_1)P(dY_P^k|\theta_1)p + \tilde{P}(d\tilde{Y_B^k}|\theta_2)P(dY_P^k|\theta_2)(1-p)}. \end{aligned}$$

This is the Bayesian belief updating given the risky arm type θ using the independence of the components in the Lévy-Ito decomposition. The next Lemma expresses the change in the posterior belief over time. The Lemma handles separately the case where there are no jumps in the time interval $[t, t+dt)$, and the case where there is a jump in this interval. Let $\check{dp} := P_{t+dt}(\theta_1) - P_t(\theta_1)$, be the change in the posterior, given that there are no jumps in the time interval $[t, t+dt)$.

Lemma 3.1.
1. *Suppose that there are no jumps during the time interval $[t, t+dt)$. Then:*

$$\check{dp} = p(1-p)\sqrt{k}(\tilde{\mu_1} - \tilde{\mu_2})d\tilde{Z} - p(1-p)k(\bar{\nu}_1 - \bar{\nu}_2)dt,$$

where $d\tilde{Z} = d\tilde{Y_B^k} - \sqrt{k}(p\tilde{\mu_1} + (1-p)\tilde{\mu_2})dt$ is a standard Brownian motion.
2. *Suppose that during the interval $[t, t+dt)$ a jump of size h occurred. Then:*

$$P_{t+dt}(\theta_1) = P_h + \sqrt{k}(\tilde{\mu_1} - \tilde{\mu_2})P_h(1-P_h)d\tilde{Z}_2,$$

where $P_h := \frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}$,[10] *and $d\tilde{Z}_2 := d\tilde{Y_B^k} - \sqrt{k}(\tilde{\mu_1}P_h + \tilde{\mu_2}(1-P_h))dt$.*

[10]Here $\frac{\nu_2(dh)}{\nu_1(dh)}$ is the Radon-Nikodym derivative, which exists by Assumption 2.1(A4).

Proof of Lemma 3.1. The proof of the lemma is standard and non-inspiring. The first statement follows from a long chain of equalities.

$$
\begin{aligned}
\check{dp} := dP_t =& \frac{p(1-p)[\tilde{P}(d\tilde{Y_B^k}|\theta_1)P(dY_P^k|\theta_1) - \tilde{P}(d\tilde{Y_B^k}|\theta_2)P(dY_P^k|\theta_2)]}{\tilde{P}(d\tilde{Y_B^k}|\theta_1)P(dY_P^k|\theta_1)p + \tilde{P}(d\tilde{Y_B^k}|\theta_2)P(dY_P^k|\theta_2)(1-p)} \\
=& \frac{p(1-p)[(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})(e^{-\bar{\nu}_1 kdt}) - (1+\sqrt{k}\tilde{\mu_2}d\tilde{Y_B^k})(e^{-\bar{\nu}_2 kdt})]}{(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})(e^{-\bar{\nu}_1 kdt})p + (1+\sqrt{k}\tilde{\mu_2}d\tilde{Y_B^k})(e^{-\bar{\nu}_2 kdt})(1-p)} \\
=& \frac{p(1-p)[(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})(1-\bar{\nu}_1 kdt) - (1+\sqrt{k}\tilde{\mu_2}d\tilde{Y_B^k})(1-\bar{\nu}_2 kdt)]}{(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})(1-\bar{\nu}_1 kdt)p + (1+\sqrt{k}\tilde{\mu_2}d\tilde{Y_B^k})(1-\bar{\nu}_2 kdt)(1-p)} \\
=& \frac{p(1-p)[\sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})d\tilde{Y_B^k} - (\bar{\nu}_1-\bar{\nu}_2)kdt]}{1+\sqrt{k}(p\tilde{\mu_1}+(1-p)\tilde{\mu_2})d\tilde{Y_B^k} - k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)dt} \\
=& p(1-p)[\sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})d\tilde{Y_B^k} - (\bar{\nu}_1-\bar{\nu}_2)kdt]\cdot \\
& \cdot[1-\sqrt{k}(p\tilde{\mu_1}+(1-p)\tilde{\mu_2})d\tilde{Y_B^k} + k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)dt + k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)^2 dt] \\
=& p(1-p)[\sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})(d\tilde{Y_B^k} - \sqrt{k}(p\tilde{\mu_1}+(1-p)\tilde{\mu_2})dt) - k(\bar{\nu}_1-\bar{\nu}_2)dt] \\
=& p(1-p)\sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})d\tilde{Z} - p(1-p)k(\bar{\nu}_1-\bar{\nu}_2)dt.
\end{aligned}
$$

In the calculations we used the fact that $d\tilde{Z} = d\tilde{Y_B^k} - \sqrt{k}(p\tilde{\mu_1}+(1-p)\tilde{\mu_2})dt$ is a standard Brownian motion (see Bolton and Harris (1999)), and the Brownian motion properties: $dZ^2 = dt$, and $dZdt = 0$. We also ignored terms of order $dt^{3/2}$ and above.

We now prove the second statement.

$$
\begin{aligned}
P_{t+dt}(\theta_1) =& \frac{p\nu_1(dh)(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})}{[\nu_1(dh)p+\nu_2(dh)(1-p)] + \sqrt{k}[\nu_1(dh)\tilde{\mu_1}p+\nu_2(dh)\tilde{\mu_2}(1-p)]d\tilde{Y_B^k}} \\
=& \frac{P_h(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})}{1+\sqrt{k}[\tilde{\mu_1}P_h+\tilde{\mu_2}(1-P_h)]d\tilde{Y_B^k}} \\
=& P_h(1+\sqrt{k}\tilde{\mu_1}d\tilde{Y_B^k})[1-\sqrt{k}(P_h\tilde{\mu_1}+(1-P_h)\tilde{\mu_2})d\tilde{Y_B^k} + k(P_h\tilde{\mu_1}+(1-P_h)\tilde{\mu_2})^2 dt] \\
=& P_h[1+\sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})(1-P_h)(d\tilde{Y_B^k} - \sqrt{k}(\tilde{\mu_1}P_h+\tilde{\mu_2}(1-P_h))dt)] \\
=& P_h + \sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})P_h(1-P_h)d\tilde{Z}_2,
\end{aligned}
$$

□

Thus, if there are no jumps during the time interval $[t, t+dt)$, the Brownian motion's contribution to the posterior belief is $p(1-p)\sqrt{k}(\tilde{\mu_1}-\tilde{\mu_2})d\tilde{Z}$, and the compound Poisson process' contribution to the posterior belief is $-p(1-p)k(\bar{\nu}_1-\bar{\nu}_2)dt$, which is negative due to Assumption 2.1(A4).
If a jump of size h occurs during the interval $[t, t+dt)$, then the Brownian motion's contribution is $d\tilde{Y_B^k} - \sqrt{k}(\tilde{\mu_1}P_h+\tilde{\mu_2}(1-P_h))dt$, and the compound Poisson process'

contribution is $P_h := \frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}$. The latter is the Bayesian update of the probability that the risky arm is High given that a jump of size h occurred. Note that $p < P_h$.
We now formulate the control equation that describes the optimal payoff.

(CE)
$$U(p) = \max_{k\in[0,1]} \{[(1-k)s+k(p(\bar{\nu}_1 H_1+\mu_1)+(1-p)(\bar{\nu}_2 H_2+\mu_2))]rdt+e^{-rdt}E[U(p+dp)]\},$$

where k is the control variable. The first term within the maximization is the expected instantaneous payoff, and the last term is the discounted expected continuation payoff. The following lemma provides a more convenient form to the control equation, in terms of the second derivative of U.

Lemma 3.2. *The following equality holds:*
(CE2)
$$\begin{aligned} U(p) = \max_{k\in[0,1]} \{&(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_1)) \\ &+ \frac{1}{r}\left[k\int (p\nu_1(dh) + (1-p)\nu_2(dh))U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)\right. \\ &- kp(1-p)(\bar{\nu}_1 - \bar{\nu}_2)U'(p) - k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)U(p) \\ &+ \left.\left.\frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu_1} - \tilde{\mu_2})^2\right]\right\} \quad p-a.s. \end{aligned}$$

Proof of Lemma 3.2. Since $U(p)$ is convex, $U(p)$ is twice differentiable p-a.s. (in the sense of the Lebesgue measure). With probability $[p(1-k\bar{\nu}_1 dt) + (1-p)(1-k\bar{\nu}_2 dt)]$ there are no jumps in the interval $[t, t+dt)$. Using the Taylor expansion of U, and ignoring terms of order $dt^{3/2}$ and higher, we obtain that the optimal payoff is $U(d+\check{d}p) = U(p) + U'(p)\check{d}p + \frac{1}{2}U''(p)(\check{d}p)^2$ a.s.[11]

With probability $[pk\nu_1(dh)dt + (1-p)k\nu_2(dh)dt]$ there is a jump of size h, and the optimal payoff is $U(P_h + \sqrt{k}(\tilde{\mu_1} - \tilde{\mu_2})P_h(1-P_h)d\tilde{Z}_2) = U(P_h) + U'(P_h)\hat{d}P_h + \frac{1}{2}U''(P_h)\hat{d}P_h$, where $\hat{d}P_h := \sqrt{k}(\tilde{\mu_1} - \tilde{\mu_2})P_h(1-P_h)d\tilde{Z}_2$.

During the subsequent calculations we use the following Eqs. (3.10), (3.11), (3.12) and (3.13), that can be derived from Lemma 3.1:

$$E[\check{d}p] = -kp(1-p)(\bar{\nu}_1 - \bar{\nu}_2)dt. \tag{3.10}$$

This is the expectation of the change in the belief, given that no jump occurred during the interval $[t, t+dt)$.

$$E[\check{d}p^2] = kp^2(1-p)^2(\tilde{\mu_1} - \tilde{\mu_2})^2 dt. \tag{3.11}$$

This is the second moment of the change in the belief, given that no jump occurred during the interval $[t, t+dt)$.

$$E[\hat{d}P_h] = C_1 \cdot dt. \tag{3.12}$$

[11]Since we can ignore terms or order $dt^{3/2}$ and higher, it is sufficient to consider the Taylor expansion up to the second derivative.

This is the expected contribution of the Brownian motion part to the posterior belief, given that a jump of size h occurred during the interval $[t, t+dt)$.

$$E[\hat{d}P_h^2] = C_2 \cdot dt. \tag{3.13}$$

This is the second moment of the contribution of the Brownian motion part to the posterior belief, given that a jump of size h occurred during the interval $[t, t+dt)$. In Eqs. (3.12) and (3.13), C_1 and C_2 are constants. Using the above notations we obtain from (CE):

(3.14)

$$\begin{aligned}
U(p) = \max_{k\in[0,1]} &\{[(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2))]rdt \\
&+ (1-rdt)\left[kdt\left[\int\left[U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)\right.\right.\right. \\
&+U'\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)C_1 dt \\
&+ \left.U''\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)C_2 dt\right](p\nu_1(dh)+(1-p)\nu_2(dh))\Big] \\
&+ [1-k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)dt]\left.\left[U(p)+U'(p)E(\check{d}p)+\frac{1}{2}U''(p)E(\check{d}p^2)\right]\right]\Big\} \quad p-a.s.
\end{aligned}$$

Using $dt^2 = 0$ we obtain:

(3.15)

$$\begin{aligned}
U(p) = \max_{k\in[0,1]} &\{[(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2))]rdt \\
&+ kdt\int U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)(p\nu_1(dh)+(1-p)\nu_2(dh)) \\
&+ U(p) - kdtp(1-p)(\bar{\nu}_1-\bar{\nu}_2)U'(p)dt - k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)U(p)dt - rU(p)dt \\
&+ \frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu_1}-\tilde{\mu_2})^2 dt\Big\} \quad p-a.s.
\end{aligned}$$

Eliminating $U(p)$ from both sides, and dividing by dt, we obtain after simple algebraic manipulations:

$$\begin{aligned}
U(p) = \max_{k\in[0,1]} &\{(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2)) \\
&+ \frac{1}{r}\left[k\int(p\nu_1(dh)+(1-p)\nu_2(dh))U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)\right. \\
&- kp(1-p)(\bar{\nu}_1-\bar{\nu}_2)U'(p) - k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)U(p) \\
&+\left.\frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu_1}-\tilde{\mu_2})^2\right]\Big\} \quad p-a.s.,
\end{aligned}$$

as desired. □

From Eq. (3.14) it follows that the contribution of the continuation payoff given that a jump of size h occurred during the time interval $[t, t+dt)$ is

(3.16)
$$kdt\left[\int\left[U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right) + U'\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)C_1dt\right.\right.$$
$$\left.\left.+\ U''\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)C_2dt\right](p\nu_1(dh)+(1-p)\nu_2(dh))\right].$$

Since only C_1 and C_2 depend on the parameters of the Brownian motion, and since C_1 and C_2 do not appear in Eq. (3.15), it follows that if a jump occurs during the time interval $[t, t+dt)$, the information from the compound Poisson process has more impact than the information of the Brownian motion.

According to (CE2), the payoff is the maximum over the control payoff k of the expectation of the current flow payoff $[(1-k)s + k(p(\bar{\nu}_1 H_1 + a) + (1-p)(\bar{\nu}_2 H_2 + b))]$ plus the discounted value of the continuation payoff

$$\frac{1}{r}\left[k\int(p\nu_1(dh)+(1-p)\nu_2(dh))U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)\right.$$
$$\left.-\ kp(1-p)(\bar{\nu}_1-\bar{\nu}_2)U'(p) + \frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu_1}-\tilde{\mu_2})^2 - k(p\bar{\nu}_1+(1-p)\bar{\nu}_2)U(p)\right].$$

A solution to this maximization problem is :

$$k = \begin{cases} 0 & \text{if } b(p,U) < s - [pg_1 + (1-p)g_2], \\ \in [0,1] & \text{if } b(p,U) = s - [pg_1 + (1-p)g_2], \\ 1 & \text{if } b(p,U) > s - [pg_1 + (1-p)g_2], \end{cases} \tag{3.17}$$

where

$$b(p,U) = \frac{1}{r}\left[\int(p\nu_1(dh)+(1-p)\nu_2(dh))U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)\right.$$
$$-\ p(1-p)(\bar{\nu}_1-\bar{\nu}_2)U'(p) - (p\bar{\nu}_1+(1-p)\bar{\nu}_2)U(p)$$
$$\left.+\frac{1}{2}U''(p)p^2(1-p)^2(\tilde{\mu_1}-\tilde{\mu_2})^2\right].$$

The function within the maximization in (CE2) is linear in k. Therefore it achieves it's maximum at $k=1$ or $k=0$ p-a.s. From Proposition 2.3, $U(p)$ is non-decreasing and continuous, and therefore there is p^* such that $U(p) = s$ for every $p \le p^*$. Thus $k=0$ is optimal for $p < p^*$. For every $p > p^*$, we have $U(p) > s$, so that in this case $k=1$ is optimal p-a.s.[12]

[12] Recall that Eq. (CE2) is satisfied p-a.s., since $U'(p)$ and $U''(p)$ exist p-a.s., and in Section 3.4 we show that the optimal strategy is really a cut-off strategy.

3.4. **Characterizing the optimal strategy and the value.** When it is optimal to play safe, that is, when the optimal solution of (CE) is $k^*=0$, we have $U(p)=s$. When it is optimal to play risky, that is, when the optimal solution of (CE) is $k^*=1$, it follows from Lemma 3.2 that $U(p)$ solves the following functional differential equation:

$$\begin{aligned}\text{(FDE)}\qquad U(p) &= p(\bar\nu_1 H_1+\mu_1)+(1-p)(\bar\nu_2 H_2+\mu_2)\\ &+\frac{1}{r}\Bigg[\int (p\nu_1(dh)+(1-p)\nu_2(dh))U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)\\ &-p(1-p)(\bar\nu_1-\bar\nu_2)U'(p)-(p\bar\nu_1+(1-p)\bar\nu_2)U(p)\\ &+\frac{1}{2}U''(p)p^2(1-p)^2(\tilde{\mu_1}-\tilde{\mu_2})^2\Bigg]\ \textit{a.s. in } (p^*,1).\end{aligned}$$

A solution $U(p)$ for this equation must be smooth (Friedman (1969), p.56).[13] Therefore, $U(p)$ satisfies Eq. (FDE) in $(p^*,1)$ always, and $k=1$ is optimal in $(p^*,1)$. Hence, there is an optimal cut-off strategy κ^* with cut-off p^*.

The next lemma suggests one solution to Eq. (FDE). To prove it, substitute the expression for $U(p)$ defined in Eq. (3.18) below into Eq. (FDE).

Lemma 3.3. *One smooth solution to Eq. (FDE) is*

$$U(p)=g_2+(g_1-g_2)p+C(1-p)\left(\frac{1-p}{p}\right)^{\alpha}, \tag{3.18}$$

where $\alpha\in(0,\infty)$ *solves the equation* $f(\eta)=\int\nu_2(dh)(\frac{\nu_2(dh)}{\nu_1(dh)})^{\eta}+\eta(\bar\nu_1-\bar\nu_2)-\bar\nu_2+\frac{1}{2}(\eta+1)\eta(\tilde{\mu_1}-\tilde{\mu_2})^2-r=0$.

In fact, one can verify that

$$U(p)=g_2+(g_1-g_2)p+C(1-p)\left(\frac{1-p}{p}\right)^{\alpha}+D(1-p)\left(\frac{1-p}{p}\right)^{\beta}$$

solves Eq. (FDE), where α is as in the statement of Lemma 3.3, and β is the unique solution of $f(\eta)=0$ in $(-\infty,0)$.[14] The following Lemma assures that α is well defined.

Lemma 3.4. *Define a function* $f:[0,\infty]\to\mathbb{R}\cup\{\infty\}$ *by*

$$f(\eta):=\int\nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\eta}+\eta\left(\bar\nu_1-\bar\nu_2\right)-\bar\nu_2+\frac{1}{2}\left(\eta+1\right)\eta(\tilde{\mu_1}-\tilde{\mu_2})^2-r. \tag{3.19}$$

The equation $f(\eta)=0$ *admits a unique solution in the interval* $(0,\infty)$.

Proof.
The function f is a continuous function that satisfies $f(0)<0$ and $f(\infty)=\infty$. To show that $f(\eta)=0$ has a unique solution, it is therefore sufficient to prove that f is

[13] To see that the conditions of Friedman (1969) are satisfied, substitute $f(p)=\int(p\nu_1(dh)+(1-p)\nu_2(dh))U\left(\frac{p\nu_1(dh)}{p\nu_1(dh)+(1-p)\nu_2(dh)}\right)-U(p)(r+(p\bar\nu_1+(1-p)\bar\nu_2))$. Since $U(p)$ is continuous, so is f, and we get $U\in C^2$, as Friedman (1969) requires.

[14] In effect, such a solution β must be smaller then -1.

increasing in η. Note that $\frac{1}{2}(\eta+1)\eta(\tilde{\mu_1}-\tilde{\mu_2})^2-r-\bar{\nu}_2$ is increasing in η. It remains to prove that $\int \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^\eta+\eta(\bar{\nu}_1-\bar{\nu}_2)$ is increasing in η as well. Since

$$\int \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^\eta+\eta(\bar{\nu}_1-\bar{\nu}_2)=\int\left[\nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^\eta+\eta(\nu_1(dh)-\nu_2(dh))\right],$$

it is sufficient to prove that $g(\eta)=\nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^\eta+\eta(\nu_1(dh)-\nu_2(dh))$ is increasing in η. Now,

$$\begin{aligned}g'(\eta)&=-\nu_2(dh)\ln\left(\frac{\nu_1(dh)}{\nu_2(dh)}\right)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^\eta+(\nu_1(dh)-\nu_2(dh))\\&>-\nu_2(dh)\ln\left(\frac{\nu_1(dh)}{\nu_2(dh)}\right)+(\nu_1(dh)-\nu_2(dh))>0,\end{aligned}$$

and since $-\ln(x)+x-1>0$ for every $x\neq 1$, it follows that $g(\eta)$ is increasing, as desired. ☐

We now prove that Eq. (FDE) has a unique solution.

Lemma 3.5. *For every $p_1<p_2$, and every $u_1,u_2\in\mathbb{R}$, there is a unique solution $U(p)$ satisfying Eq. (FDE) in the interval (p_1,p_2) with the boundary conditions $U(p_1)=u_1$, $U(p_2)=u_2$.*

Proof.

Since Eq. (FDE) is a non-homogenous linear equation in U, if U and V are two solutions of Eq. (FDE), then $U-V$ is a solution of the homogenous version of Eq. (FDE). To prove the lemma it is therefore sufficient to fix a solution U of the homogenous version of Eq. (FDE) that satisfies $U(p_1)=U(p_2)=0$ and to prove that $U\equiv 0$.

Suppose that U achieves its maximum at $\hat{p}$. Then $U'(\hat{p})=0$, and $U''(\hat{p})\leq 0$. It follows that

$$\begin{aligned}U(\hat{p})=\frac{1}{r}&\left[\int(\hat{p}\nu_1(dh)+(1-\hat{p})\nu_2(dh))U\left(\frac{\hat{p}\nu_1(dh)}{\hat{p}\nu_1(dh)+(1-\hat{p})\nu_2(dh)}\right)\right.\\&\left.+\frac{1}{2}U''(\hat{p})\hat{p}^2(1-\hat{p})^2(\tilde{\mu_1}-\tilde{\mu_2})^2-(\hat{p}\bar{\nu}_1+(1-\hat{p})\bar{\nu}_2)U(\hat{p})\right].\end{aligned}$$

Simple algebraic manipulations imply that:

$$\begin{aligned}(r+\hat{p}\bar{\nu}_1+(1-\hat{p})\bar{\nu}_2)U(\hat{p})&=\int(\hat{p}\nu_1(dh)+(1-\hat{p})\nu_2(dh))U\left(\frac{\hat{p}\nu_1(dh)}{\hat{p}\nu_1(dh)+(1-\hat{p})\nu_2(dh)}\right)\\&+\frac{1}{2}U''(\hat{p})\hat{p}^2(1-\hat{p})^2(\tilde{\mu_1}-\tilde{\mu_2})^2\\&\leq U(\hat{p})\int(\hat{p}\nu_1(dh)+(1-\hat{p})\nu_2(dh))+0\\&=(\hat{p}\bar{\nu}_1+(1-\hat{p})\bar{\nu}_2)U(\hat{p}),\end{aligned}$$

where the inequality holds since $U''(\hat{p}) \leq 0$. Since $r > 0$ we conclude that $U(\hat{p}) = 0$. A similar argument shows that the minimum of U in (p_1, p_2) is 0, so that $U(p) \equiv 0$ on (p_1, p_2) as desired. □

As mentioned before, there is an optimal cut-off strategy with corresponding payoff U. Lemmas 3.3, 3.4 and 3.5 prove that U is the unique solution of Eq. (FDE). We now prove Theorem 2.4, which provides an explicit form to the optimal strategy and to the payoff function.

Proof of Theorem 2.4.
Recall that κ^* is the optimal cut-off strategy (with cut-off p^*). To complete the proof of the theorem, we provide an explicit expression to p^* and to U. To this end, we first prove that the derivative from the right of U at p^* is 0.

Let $U'_R(p^*) = \lim\limits_{\varepsilon \searrow 0} \frac{U(p^*+\varepsilon)-U(p^*)}{\varepsilon}$ be the right derivative of U at the cut-off p^*. Since U is convex, $U'_R(p^*)$ is well-defined. Since U is non-decreasing, $U'_R(p^*) \geq 0$. We now prove that $U'_R(p^*) \leq 0$. As defined in Section 2.6, $M_\kappa := \int_0^\infty re^{-rt}dY^\kappa(t)$. Recall that for every $q_0 \in [0,1]$, $\kappa(q_0)$ is the strategy that plays as κ, assuming the prior belief is q_0 rather then p_0. Then we note that

$$\lim_{\epsilon \to 0} E\left[M_{\kappa^*(p^*+\epsilon)} \middle| \, \theta = s\right] \quad \forall \theta \in \{\theta_1, \theta_2\}.$$

Indeed, κ^* stops when the posterior is p^*, and therefore $E\left[M_{\kappa^*(p^*+\epsilon)} \middle| \, \theta = s\right]$. If ϵ is small, then with probability close to 1 the posterior will decrease beneath the threshold p^* "quite fast": if there is a Brownian motion component, the fluctuations will cause fast stopping, since in any infinitesimal time interval the Brownian motion decreases beneath its initial value. If there is no Brownian motion component, there are no jumps in infinitesimal time interval, and therefore the compound Poisson process will cause the posterior to decrease until a jump arrives, since the Lévy measure is finite.[15]

Since κ^* is the optimal cut-off strategy, and since it is independent of the prior belief p_0, we deduce that for every $p \in [0,1]$, $U(p) = V_{\kappa^*(p)}(p)$. Therefore:

$$
\begin{aligned}
(3.20) \qquad U'_R(p^*) &= \lim_{\epsilon \to 0} \frac{U(p^*+\epsilon) - U(p^*)}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{V_{\kappa^*(p^*+\epsilon)}(p^*+\epsilon) - V_{\kappa^*(p^*)}(p^*)}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[(p^*+\epsilon)E[M_{\kappa^*(p^*+\epsilon)}|\theta_1] \; + (1-p^*-\epsilon)E[M_{\kappa^*(p^*+\epsilon)}|\theta_2] \right] \\
&\qquad -p^*E[M_{\kappa^*(p^*)}|\theta_1] - (1-p^*)E[M_{\kappa^*(p^*)}|\theta_2]] \\
&= \lim_{\epsilon \to 0} \left\{ \frac{1}{\epsilon} \left[p^*E[M_{\kappa^*(p^*+\epsilon)}|\theta_1] \; + (1-p^*)E[M_{\kappa^*(p^*+\epsilon)}|\theta_2] \right. \right. \\
&\qquad -p^*E[M_{\kappa*(p^*)}|\theta_1] - (1-p^*)E[M_{\kappa^k(p^*)}|\theta_2]] \\
&\qquad \left. +E[M_{\kappa^*(p^*+\epsilon)}|\theta_1] - E[M_{\kappa^*(p^*+\epsilon)}|\theta_2] \right\},
\end{aligned}
$$

[15]By the same argument we get that $V_{\tilde{\kappa}}(p)$ is continuous in $\tilde{p}$, where $\tilde{\kappa}$ is a cut-off strategy with cut-off $\tilde{p}$.

By the optimality of $U(p)$, $U(p^*) \geq V_{\kappa^*(p^*+\epsilon)}(p^*)$, and therefore,

(3.21)
$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[p^*E[M_{\kappa^*(p^*+\epsilon)}|\theta_1]+(1-p^*)E[M_{\kappa^*(p^*+\epsilon)}|\theta_2]-p^*E[M_{\kappa^*(p^*)}|\theta_1]-(1-p^*)E[M_{\kappa^*(p^*)}|\theta_2]\right]$$
$$=\lim_{\epsilon\to 0}\frac{1}{\epsilon}[V_{\kappa^*(p^*+\epsilon)}(p^*)-V_{\kappa^*(p^*)}(p^*)]=\lim_{\epsilon\to 0}\frac{1}{\epsilon}[V_{\kappa^*(p^*+\epsilon)}(p^*)-U(p^*)]\leq 0,$$
and

(3.22)
$$\lim_{\epsilon\to 0}\left(E[M_{\kappa^*(p^*+\epsilon)}|\theta_1]-E[M_{\kappa^*(p^*+\epsilon)}|\theta_2]\right)=0,$$

Substituting (3.21) and (3.22) in (3.20) we deduce that $U'_R(p^*)\leq 0$.

As mentioned in Lemma 2.3, $U(p^*)=s$, and $U(1)=g_1$. By Lemma 3.5 the unique solution to Eq. (FDE) on $(p^*,1]$ is $U(p)=g_2+(g_1-g_2)p+C(1-p)\left(\frac{1-p}{p}\right)^{\alpha}$. By imposing $U(p^*)=s$ and $U'_R(p^*)=0$ we get $p^*=\frac{\alpha(s-g_2)}{(\alpha+1)(g_1-s)+\alpha(s-g_2)}$, and $C=\frac{s-g_2-p^*(g_1-g_2)}{(1-p^*)(\frac{1-p^*}{p^*})^{\alpha}}$. Uniqueness follows by (3.17).

□

3.5. **Incorrect Prior.** To find the expected discounted payoff for a DM who plays according to an incorrect prior we present here a condition which is equivalent to (2.7). Substituting $Y^1_B(t)=\mu t+\sigma Z(t)$ and $\tilde{\mu}=\frac{\mu}{\sigma}$ in (2.7), we get:

$$Z(t)+\left[\frac{2\tilde{\mu}-\tilde{\mu_1}-\tilde{\mu_2}}{2}-\frac{\bar{\nu}_1-\bar{\nu}_2}{\tilde{\mu_1}-\tilde{\mu_2}}\right]t>-\frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\left[\ln\left(\frac{q_0}{1-q_0}\right)-\ln\left(\frac{p'}{1-p'}\right)\right]-\frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\sum_t\ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right).$$
It follows that the DM selects the risky arm until the first time t that satisfied:

(3.23)
$$\begin{aligned}B^{\mu}(t):=Z(t)+&\left[\frac{2\tilde{\mu}-\tilde{\mu_1}-\tilde{\mu_2}}{2}-\frac{\bar{\nu}_1-\bar{\nu}_2}{\tilde{\mu_1}-\tilde{\mu_2}}\right]t\\ &\leq-\frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\left[\ln\left(\frac{q_0}{1-q_0}\right)-\ln\left(\frac{p'}{1-p'}\right)\right]-\frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\sum_t\ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right).\end{aligned}$$

Proof of Theorem 2.6. If the prior belief of the DM, q_0, satisfies $q_0\geq p'$ then there is a bijection relation between E and q_0, where $E=\frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\times\left[\ln\left(\frac{q_0}{1-q_0}\right)-\ln\left(\frac{p'}{1-p'}\right)\right]$ is the intercept in Eq. (2.7) at $t=0$. If $q_0\leq p'$ then the DM always chooses the risky arm, which is equivalent to $E=0$. Therefore, we will use the notation $U(p_0,E)$ instead of $U(p_0,q_0)$ when the former is more convenient. We now prove that under Assumption 2.1, for every $p_0\in[0,1]$ and every $E\in[0,\infty)$, the payoff of the DM is $U(p_0,E)=g_1p_0+g_2(1-p_0)+(s-g_1)p_0e^{-(\tilde{\mu_1}-\tilde{\mu_2})(\alpha+1)E}+(s-g_2)(1-p_0)e^{-(\tilde{\mu_1}-\tilde{\mu_2})\alpha E}$.

Using Eq. (3.23), we construct an integral equation, to find the utility function for a DM who has a prior belief q_0 while the actual probability of High is p_0.
Let τ be the stopping time of the first jump. Let T be the first time t satisfying (3.23). The DM chooses the risky arm until the stopping time T, and then he

switches to the safe arm. The calculations use dynamic programming in which the continuation payoff is determined by the time of the first jump, τ, and the value of the continuous part of the payoff at that time.

Constructing the integral equation: The DM chooses the risky arm until the minimum between the stopping time of the first jump τ, and the stopping time T. We distinguish between two case. In case the DM stops before the time of the first jump τ, we calculate the expected discounted payoff from the risky arm until the stopping time T, and the expected discounted payoff from the safe arm after the stopping time T. In case the first jump occurrs before the stopping time T, we calculate the expected discounted payoff received from the continuous part $Y_B^1(t)$ until time τ. We add the expected discounted payoff from the first jump, and the expected discounted continuation payoff, updating both the posterior $p_{t,h,x}$, and the intercept $E+G_h+x$, according to the time of the first jump, the first jump's size, and the value of the continuous part of the payoff.

The notations used are as follows: $P_\theta(\tau > T)$ is the probability the DM switches to the safe arm before a jump occurs. If $\tau > T$ then the expected payoff from the risky arm is $E_{\theta_1}[\int_0^T re^{-rt}dY_B^1(t)|\tau > T]$, and the expected payoff from the safe arm is $E_{\theta_1}[\int_T^\infty re^{-rt}sdt|\tau > T]$. $P_\theta(\tau < T)$ is the probability a jump occurs before the DM switches to the safe arm. If $\tau < T$ then the expected payoff until the first jump is $E_{\theta_1}[\int_0^\tau re^{-rt}dY_B^1(t)|\tau < T]$. $f^\theta_{(B^\mu(\tau),\tau)|\tau<T}(x,t)$ is the probability that the first jump occurs in the interval $[t, t+dt)$, and B^μ belongs to $[x, x+dx)$, given a jump occurs before the DM switches to the safe arm. $re^{-rt}\frac{1}{\bar\nu}\int h\nu(dh)$ is the expected discounted payoff from the first jump, and $\frac{1}{\bar\nu_1}\int \nu_1(dh)e^{-rt}U(p_{t,h,x}, E+G_h+x))$ is the expected discounted continuation payoff, updating both the posterior $p_{t,h,x}$, and the intercept $E+G_h+x$ at time t. With these notations, the expected payoff is as follows. In the equalities that follow we write below each equality the equalities from Section 3.1 that are used in order to derive it.

(3.24)

$$
\begin{aligned}
&U(p_0,E) = p_0P_{\theta_1}(\tau<T)\left[E_{\theta_1}[\textstyle\int_0^\tau re^{-rt}dY_B^1(t)|\tau<T]\right.\\
&\left.+\textstyle\int_{-E}^\infty\int_0^\infty f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\left(re^{-rt}\frac{1}{\bar\nu_1}\int\nu_1(dh)h+\frac{1}{\bar\nu_1}\int\nu_1(dh)e^{-rt}U(p_{t,h,x},E+G_h+x)\right)\right]\\
&+p_0P_{\theta_1}(\tau>T)\left[E_{\theta_1}[\textstyle\int_0^T re^{-rt}dY_B^1(t)|\tau>T]+E_{\theta_1}[\int_T^\infty re^{-rt}sdt|\tau>T]\right]\\
&+(1-p_0)P_{\theta_2}(\tau<T)\left[E_{\theta_2}[\textstyle\int_0^\tau re^{-rt}dY_B^1(t)|\tau<T]\right.\\
&\left.+\textstyle\int_{-E}^\infty\int_0^\infty f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\left(re^{-rt}\frac{1}{\bar\nu_2}\int\nu_2(dh)h+\frac{1}{\bar\nu_2}\int\nu_2(dh)e^{-rt}U(p_{t,h,x},E+G_h+x)\right)\right]\\
&+(1-p_0)P_{\theta_2}(\tau>T)\left[E_{\theta_2}[\textstyle\int_0^T re^{-rt}dY_B^1(t)|\tau>T]+E_{\theta_2}[\int_T^\infty re^{-rt}sdt|\tau>T]\right]
\end{aligned}
$$

$$\underset{\{3,4,5\}}{=} \mu_1 p_0(1-P_{\theta_1}(\tau>T))(1-E_{\theta_1}[e^{-r\tau}|\tau<T])$$

$$+p_0P_{\theta_1}(\tau<T)H_1\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)re^{-rt}$$

$$+p_0P_{\theta_1}(\tau<T)\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\tfrac{1}{\bar\nu_1}\int e^{-rt}U(p_{t,h,x},E+G_h+x)\nu_1(dh)$$

$$+\mu_1p_0P_{\theta_1}(\tau>T)(1-E_{\theta_1}[e^{-rT}|\tau>T])+sp_0P_{\theta_1}(\tau>T)E_{\theta_1}[e^{-rT}|\tau>T]$$

$$+\mu_2(1-p_0)P_{\theta_2}(\tau>T))(1-E_{\theta_2}[e^{-r\tau}|\tau<T])$$

$$+(1-p_0)P_{\theta_2}(\tau<T)H_2\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)re^{-rt}$$

$$+(1-p_0)P_{\theta_2}(\tau<T)\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\tfrac{1}{\bar\nu_2}\int e^{-rt}U(p_{t,h,x},E+G_h+x)\nu_2(dh)$$

$$+\mu_2(1-p_0)P_{\theta_2}(\tau>T)(1-E_{\theta_2}[e^{-rT}|\tau>T]+s(1-p_0)P_{\theta_2}(\tau>T)E_{\theta_2}[e^{-rT}|\tau>T]$$

$$\underset{\{6,7,8\}}{=} \mu_1p_0P_{\theta_1}(\tau>T)\left(1-\tfrac{P_{\theta_1}(\tau^r>T)}{P_{\theta_1}(\tau>T)}\right)+sp_0P_{\theta_1}(\tau>T)\tfrac{P_{\theta_1}(\tau^r>T)}{P_{\theta_1}(\tau>T)}$$

$$+p_0(1-P_{\theta_1}(\tau>T))H_1r\cdot\tfrac{\bar\nu_1}{\bar\nu_1+r}\cdot\tfrac{1-P_{\theta_1}(\tau^r>T)}{1-P_{\theta_1}(\tau>T)}$$

$$+\mu_1p_0(1-P_{\theta_1}(\tau>T))\left(1-\tfrac{\bar\nu_1}{\bar\nu_1+r}\cdot\tfrac{1-P_{\theta_1}(\tau^r>T)}{1-P_{\theta_1}(\tau>T)}\right)$$

$$+\mu_2(1-p_0)P_{\theta_2}(\tau>T)\left(1-\tfrac{P_{\theta_2}(\tau^r>T)}{P_{\theta_2}(\tau>T)}\right)+s(1-p_0)P_{\theta_2}(\tau>T)\tfrac{P_{\theta_2}(\tau^r>T)}{P_{\theta_2}(\tau>T)}$$

$$+(1-p_0)(1-P_{\theta_2}(\tau>T))H_2r\cdot\tfrac{\bar\nu_2}{\bar\nu_2+r}\cdot\tfrac{1-P_{\theta_2}(\tau^r>T)}{1-P_{\theta_2}(\tau>T)}$$

$$+\mu_2(1-p_0)(1-P_{\theta_2}(\tau>T))\left(1-\tfrac{\bar\nu_1}{\bar\nu_1+r}\cdot\tfrac{1-P_{\theta_2}(\tau^r>T)}{1-P_{\theta_2}(\tau>T)}\right)$$

$$+\int_{-E}^{\infty}\int_0^{\infty}\int e^{-rt}U(p_{t,h,x},E+G_h+x)\cdot$$

$$\cdot[p_0P_{\theta_1}(\tau<T)f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\nu_1(dh)/\bar\nu_1+(1-p_0)P_{\theta_2}(\tau<T)f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\nu_2(dh)/\bar\nu_2]$$

$$=p_0\left(\tfrac{\mu_1r}{\bar\nu_1+r}+\tfrac{\bar\nu_1H_1r}{\bar\nu_1+r}\right)+(1-p_0)\left(\tfrac{\mu_2r}{\bar\nu_2+r}+\tfrac{\bar\nu_2H_2r}{\bar\nu_2+r}\right)$$

$$+p_0P_{\theta_1}(\tau^r>T)\left(s-\tfrac{\mu_1r}{\bar\nu_1+r}-\tfrac{\bar\nu_1H_1r}{\bar\nu_1+r}\right)+(1-p_0)P_{\theta_2}(\tau^r>T)\left(s-\tfrac{\mu_2r}{\bar\nu_2+r}-\tfrac{\bar\nu_2H_2r}{\bar\nu_2+r}\right)$$

$$+\int_{-E}^{\infty}\int_0^{\infty}\int e^{-rt}U(p_{t,h,x},E+G_h+x)\cdot$$

$$\cdot[p_0P_{\theta_1}(\tau<T)f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\nu_1(dh)/\bar\nu_1+(1-p_0)P_{\theta_2}(\tau<T)f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\nu_2(dh)/\bar\nu_2]$$

$$=p_0\left(\tfrac{\mu_1r}{\bar\nu_1+r}+\tfrac{\bar\nu_1H_1r}{\bar\nu_1+r}\right)+(1-p_0)\left(\tfrac{\mu_2r}{\bar\nu_2+r}+\tfrac{\bar\nu_2H_2r}{\bar\nu_2+r}\right)$$

$$+p_0e^{-E(F_{\mu_1}+\sqrt{2(\bar\nu_1+r)+F^2_{\mu_1}})}\left(s-\tfrac{\mu_1r}{\bar\nu_1+r}-\tfrac{\bar\nu_1H_1r}{\bar\nu_1+r}\right)$$

$$+(1-p_0)e^{-E(F_{\mu_2}+\sqrt{2(\bar\nu_2+r)+F^2_{\mu_2}})}\left(s-\tfrac{\mu_2r}{\bar\nu_2+r}-\tfrac{\bar\nu_2H_2r}{\bar\nu_2+r}\right)$$

$+\int_{-E}^{\infty}\int_0^{\infty}\int e^{-rt}U(p_{t,h,x},E+G_h+x)\cdot$

$\cdot[p_0P_{\theta_1}(\tau<T)f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\nu_1(dh)/\bar{\nu}_1+(1-p_0)P_{\theta_2}(\tau<T)f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\nu_2(dh)/\bar{\nu}_2].$

Simplifying the last expression, the integral equation is:

(IE)

$$U(p,E)=Ap+B(1-p)+Cpe^{-m_1E}+D(1-p)e^{-m_2E}+\int_{-E}^{\infty}\int_0^{\infty}\int e^{-rt}U(p_{t,h,x},E+G_h+x)g(x,t,h),$$

where A,B,C, and D are constants, and

$$g(x,t,h)=[p_0P_{\theta_1}(\tau<T)f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\nu_1(dh)/\bar{\nu}_1 + (1-p_0)P_{\theta_2}(\tau<T)f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\nu_2(dh)/\bar{\nu}_2]. \tag{3.25}$$

We show now that (IE) admits a unique solution U in the region $[0,1]\times[0,\infty)$.

Boundary values
First, we find the values of $U(p,E)$ on the boundary of the region $[0,1]\times[0,\infty]$. Note that $U(p,0)=s$, and $U(p,\infty)=pg_1+(1-p)g_2$ for every $p\in[0,1]$, .
We now argue that there is a unique solution for (IE) when $p=0$. $U(0,E)$ is a function of E with two boundary conditions, at $E=0$ and at $E=\infty$. Suppose that $U(0,E),V(0,E)$ solve (IE). Then $W(0,E):=U(0,E)-V(0,E)$ satisfies: $W(0,E)=\int_{-E}^{\infty}\int_0^{\infty}\int e^{-rt}W(0,E+G_h+x)g(x,t,h)$, and $W(0,0)=W(0,\infty)=0$. Let $\hat{E}$ be a critical point, where W achieves its maximum. Assume to the contrary that $W(0,\hat{E})>0$. By (3.8) it follows that $\int_{-\hat{E}}^{\infty}\int_0^{\infty}\int e^{-rt}g(x,t,h)<1$. Therefore,

$$\begin{aligned}W(0,\hat{E})&=\int_{-\hat{E}}^{\infty}\int_0^{\infty}\int e^{-rt}W(0,\hat{E}+G_h+x)g(x,t,h)\\&\le W(0,\hat{E})\int_{-\hat{E}}^{\infty}\int_0^{\infty}\int e^{-rt}g(x,t,h)<W(0,\hat{E}),\end{aligned}$$

which implies that $W(0,E)\le 0$. Similarly, one can obtain that $W(0,E)\ge 0$, so that $W(0,E)\equiv 0$, and the solution is unique.
Similar arguments show that (IE) admits a unique solution on $[0,\infty)$ when $p=1$.

Since $U(p,E)$ is uniquely determined on the boundary of the region $[0,1]\times[0,\infty]$, similar arguments show the uniqueness of the solution $[0,1]\times[0,\infty)$. Using Eqs. (3.1) - (ref9) one can verify that the solution for (IE) is

$$U(p_0,E)=g_1p_0+g_2(1-p_0)+(s-g_1)p_0e^{-(\tilde{\mu_1}-\tilde{\mu_2})(\alpha+1)E}+(s-g_2)(1-p_0)e^{-(\tilde{\mu_1}-\tilde{\mu_2})\alpha E}.$$
□

3.6. Lévy Bandits with an Infinite Lévy Measure.

Proof of Lemma 2.8. Fix $s \in \mathbb{Q}$. By the martingale convergence theorem it follows that $p(\theta_1|\mathcal{G}_s^n)(\omega) \underset{n\to\infty}{\to} p(\theta_1|\mathcal{G}_s)(\omega)$ a.s. Therefore,

$$P\left(p(\theta_1|\mathcal{G}_s^n)(\omega) \underset{n\to\infty}{\to} p(\theta_1|\mathcal{G}_s)(\omega) \quad \forall s \in \mathbb{Q}\right) = 1. \tag{3.26}$$

By Eq. (2.9) and (3.26), it follows that for every $s \in \mathbb{Q}$

$$\ln\left(\frac{p(\theta_1|\mathcal{G}_s)}{1-p(\theta_1|\mathcal{G}_s)}\right) = \ln\left(\frac{p_0}{1-p_0}\right) + \frac{Y_B^1(s)}{\sigma^2}(\mu_1-\mu_2) - \frac{s}{2\sigma^2}(\mu_1^2-\mu_2^2) - (\bar{\nu}_1-\bar{\nu}_2)s + \sum_{h_j \in \mathbb{R}\setminus(\{0\}\cup B_0)}^{s} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right),$$

Since the Lévy process $\left(\ln\left(\frac{p(\theta_1|\mathcal{G}_s)}{1-p(\theta_1|\mathcal{G}_s)}\right)\right)$ is continuous on the left[16] it follows that for every $t \geq 0$

$$\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right) = \ln\left(\frac{p_0}{1-p_0}\right) + \frac{Y_B^1(t)}{\sigma^2}(\mu_1-\mu_2) - \frac{t}{2\sigma^2}(\mu_1^2-\mu_2^2) - (\bar{\nu}_1-\bar{\nu}_2)t + \sum_{h_j \in \mathbb{R}\setminus(\{0\}\cup B_0)}^{t} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right),$$

and the first statement of the Lemma is proved.

Define the processes $R(t) := \left(\sum^t_{h_j\in\mathbb{R}\setminus(\{0\}\cup B_0)} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right)\right)$, and for every n, $R_n(t) := \left(\sum^t_{h_j\in A_n\setminus B_0} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right)\right)$. Since $\lim_{n\to\infty}(\bar{\nu}_1^n - \bar{\nu}_2^n) = (\bar{\nu}_1-\bar{\nu}_2)$, by Eq. (2.9) and (2.10), it is sufficient to prove that $P\left(R_n(t)(\omega) \underset{n\to\infty}{\to} R(t)(\omega) \quad \forall t \in \mathbb{R}\setminus\mathbb{Q}\right) = 1$. Fix ω such that $p(\theta_1|\mathcal{G}_s^n)(\omega) \underset{n\to\infty}{\to} p(\theta_1|\mathcal{G}_s)(\omega)$ for every $s \in \mathbb{Q}$, and that $\ln\left(\frac{p(\theta_1|\mathcal{G}_s)}{1-p(\theta_1|\mathcal{G}_s)}\right)(\omega)$ is left-continuous. Fix $t \in \mathbb{R}\setminus\mathbb{Q}$. Note that by Assumption 2.7(AA10), $R_n(t)(\omega)$ is non-decreasing with n, and bounded by $R(t)(\omega)$. Therefore, the limit $\lim_{n\to\infty} R_n(t)(\omega)$ exists, and satisfies $\lim_{n\to\infty} R_n(t)(\omega) \leq R(t)(\omega)$. By Assumption 2.7(AA10), $(R(t))_t$, and for every n, $(R_n(t))_t$ are non-decreasing processes. Therefore, for every $s \in \mathbb{Q}$ satisfies $s < t$, and every n, $R_n(s)(\omega) \leq R_n(t)(\omega)$. Fix s, and take the limit $n \to \infty$, it follows that $R(s)(\omega) = \lim_{n\to\infty} R_n(s)(\omega) \leq \lim_{n\to\infty} R_n(t)(\omega)$, where the first equality follows by the choice of ω. Since $R(t)(\omega)$ is left-continuous, by taking the limit $s \nearrow t$, we get $R(t)(\omega) \leq \lim_{n\to\infty} R_n(t)(\omega)$.

□

Proof of Proposition 2.9.
The Radon-Nikodym derivatives $g(h) = \frac{\nu_1(dh)}{\nu_2(dh)}$ and $l(h) = \frac{\nu_1(dh)-\nu_2(dh)}{\nu_1(dh)}$ exist on

[16]Recall that the process $(p(\theta_1|\mathcal{G}_t))_t \geq 0$ is $\mathcal{G}_{t-}$-adapted, i.e. the posterior belief at time t is updated according to the payoff received up to time $t-$.

$\mathbb{R} \setminus (\{0\} \cup B_0)$, and $0 \le l(h) < 1$, $1 \le g(h) < \infty$, ν_1-a.s. For each $i \ge 1$ define the sets: $B_i = \{h | \frac{1}{i+1} \le l(h) < \frac{1}{i}\} = \{h | \frac{i+1}{i} \le g(h) < \frac{i}{i-1}\}$, (we use the convention $\frac{1}{1-1} = \infty$). We need the following two inequalities:
For every $i \ge 1$,

$$\frac{1}{i+1}\nu_1(B_i) \le \int_{B_i} l(h)\nu_1(dh) = \nu_1(B_i) - \nu_2(B_i) < \infty, \tag{3.27}$$

and therefore $\nu_1(B_i) < \infty$. It follows that

$$\sum_{i=1}^{\infty} \frac{1}{i+1}\nu_1(B_i) \le \nu_1(\mathbb{R} \setminus (\{0\} \cup B_0)) - \nu_2(\mathbb{R} \setminus (\{0\} \cup B_0)) < \infty, \tag{3.28}$$

since $\nu_1 - \nu_2$ is a finite measure. Notice that $\sum\limits_{i \ge 2} \nu_1(B_i) \ln\left(\frac{i}{i-1}\right) < \infty$ if and only if $\sum\limits_{i \ge 1}{}_t \nu_1(B_i)\frac{1}{i+1} < \infty$, so in particular

$$\sum_{i=2}^{\infty} \ln\left(\frac{i}{i-1}\right)\nu_1(B_i) < \infty. \tag{3.29}$$

By Eq. (2.10),[17], when $g(h) = 1$, a jump of size h does not change the posterior belief. Therefore, we treat only jumps with size in $\bigcup\limits_{i \ge 1} B_i$. Note that $P\left(p(\theta_1|\mathcal{G}_t) < 1\right) = 1$ if and only if $P\left(\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right) < \infty\right) = 1$, which, by Eq. (2.10), is equivalent to $P\left(\sum\limits_{h_j}{}_t \ln(g(h_j)) < \infty\right) = 1$. Since

$$P\left(\sum_{\substack{t\\h_j}} \ln(g(h_j)) < \infty\right) = p_0 P\left(\sum_{\substack{t\\h_j}} \ln(g(h_j)) < \infty \middle| \theta_1\right) + (1-p_0) P\left(\sum_{\substack{t\\h_j}} \ln(g(h_j)) < \infty \middle| \theta_2\right),$$

it is sufficient to show that $P\left(\sum\limits_{h_j}{}_t \ln(g(h_j)) < \infty \middle| \theta_1\right) = 1$. Since $\nu_1(B_1) < \infty$, up to time t there are only finitely many jumps with size in B_1. Therefore, the contribution of these jumps to $\sum\limits_{h_j}{}_t \ln(g(h_j))$ is finite, and can be ignored. From now on in the sum $\sum\limits_{h_j}{}_t \ln(g(h_j))$ we ignore jumps with size in B_1.

Observe that $\sum\limits_{h_j}{}_t \ln(g(h_j)) \le \sum\limits_{i \ge 2} J_i^t \ln\left(\frac{i}{i-1}\right)$, where J_i^t denotes the number of jumps with size in B_i that occur up to time t. By the properties of compound Poisson processes, and by (3.27), we get $J_i^t \sim Poisson(\nu_1(B_i)t)$ for every $i \ge 2$. Hence $E\left[\sum\limits_{h_j}{}_t \ln(g(h_j)) \middle| \theta_1\right] \le E\left[\sum\limits_{i \ge 2} J_i^t \ln\left(\frac{i}{i-1}\right) \middle| \theta_1\right] \underset{i \to \infty}{\to} \sum\limits_{i \ge 2} t\nu_1(B_i) \ln\left(\frac{i}{i-1}\right) < \infty$, where the last inequality follows by (3.29). $\square$

Recall that $M_\kappa = \int_0^\infty re^{-rt}dY^\kappa(t)$, is the discounted payoff under the strategy κ, and $N_\kappa^n = \int_0^\infty re^{-rt}dY_n^\kappa(t)$ is the discounted payoff under κ, ignoring jumps

[17] In fact, it is sufficient to use Eq. (3.26) to prove the proposition.

in A_n^c, where $Y_n^\kappa(t)$ is the process generated by the payoff process $Y^\kappa(t)$, ignoring jumps in A_n^c. Set $L_\kappa^n := M_\kappa^n - N_\kappa^n$; this is the discounted payoff under κ from jumps in A_n.

Lemma 3.6 below states that if a DM observes only jumps with absolute values higher than $1/m$, and plays according to the strategy κ (that is, he chooses the risky arm as long as his belief, which is updated according to $\mathcal{F}_t^m$, is higher than the cut-off of κ), then the expected discounted payoff converges to the expected payoff using the same strategy κ when updating the belief according to all information available.

Lemma 3.6. *For every cut-off strategies κ, and $(\kappa_n)_{n=1}^\infty$ one has:*
L1. $\lim_{n\to\infty} E[L_{\kappa_n}^n|\mathcal{F}](p) = 0$, $\lim_{n\to\infty} E[L_\kappa^n|\mathcal{F}^n](p) = 0$.
L2. $\lim_{m\to\infty} E[N_\kappa^n|\mathcal{F}^m](p) = E[N_\kappa^n|\mathcal{F}](p)$.
L3. $\lim_{m\to\infty} E[M_\kappa|\mathcal{F}^m](p) = E[M_\kappa|\mathcal{F}](p)$.

Proof of Lemma 3.6.
We start by proving L1. $E(L_\kappa^n|\mathcal{F})(p)$ is the expected payoff generated only by jumps in $(-1/n, 1/n)$, under the cut-off strategy κ. For a sufficiently large n the expected payoff will be small.

We now prove L3. The proof of L2 is similar and therefore omitted. $E(M_\kappa|\mathcal{F}^m)(p) = E[\int_0^\infty re^{-rt}dY^\kappa(t)](p)$ is the expected discounted payoff under the cut-off strategy κ, when the posterior probability is updated according to $\mathcal{F}^m$. Let s be sufficiently large, such that the expected payoff from time s and on is small enough.

We use the convention $\inf \phi = 0$. Define:

$$T_n := \inf_{0\le t\le s} \{t\,|p(\theta_1|\mathcal{F}_t^n) \le p'\} = \inf_{0\le t\le s} \left\{ t \,\middle|\, \ln\left(\frac{p(\theta_1|\mathcal{F}_t^n)}{1-p(\theta_1|\mathcal{F}_t^n)}\right) \le \ln\left(\frac{p'}{1-p'}\right) \right\};$$

this is the first time up to time s, in which the posterior belief reaches the value p', when the belief is updated using $\mathcal{F}_t^n$.

Set

$$T := \inf_{0\le t\le s} \{t|p(\theta_1|\mathcal{F}_t) \le p'\} = \inf_{0\le t\le s} \left\{ t \,\middle|\, \ln\left(\frac{p(\theta_1|\mathcal{F}_t)}{1-p(\theta_1|\mathcal{F}_t)}\right) \le \ln\left(\frac{p'}{1-p'}\right) \right\}; \tag{3.30}$$

this is the first time up to time s, in which the posterior belief reaches the value p', when the belief is updated using $\mathcal{F}_t$. By the dominated convergence theorem, it is sufficient to show that T_n convergence to T a.s. We prove it in two steps.

Step 1: $\lim_{n\to\infty} T_n \ge T$

By Proposition 2.9, $p(\theta_1|\mathcal{F}_t) = 1$, or equivalently, a jump from B_0 occurrs, if and only if there exists n such that $p(\theta_1|\mathcal{F}_t^m) = 1$ for every $m \ge n$. In this case the DM never stops according to neither $\mathcal{F}^m$ nor $\mathcal{F}$.

Let ω be a path with no jumps that have size in B_0 during the time interval $(0, s)$. Since $P\left(p(\theta_1|\mathcal{G}_t^n)(\omega) \to p(\theta_1|\mathcal{G}_t)(\omega)\ \forall t\right) = 1$, we can assume that $p(\theta_1|\mathcal{G}_t^n)(\omega) \to$

$p(\theta_1|\mathcal{G}_t)(\omega)$ for all $0<t<s$ a.s. For every $n\geq 1$, define the process

$$Q(\theta_1|\mathcal{G}^n_t) := \ln\left(\frac{p(\theta_1|\mathcal{G}^n_t)}{1-p(\theta_1|\mathcal{G}^n_t)}\right) - (\bar\nu_1-\bar\nu_2)t + (\bar\nu^n_1-\bar\nu^n_2)t$$
$$= \ln\left(\frac{p_0}{1-p_0}\right) + \frac{Y^1_B(t)}{\sigma^2}(\mu_1-\mu_2) - \frac{t}{2\sigma^2}(\mu_1^2-\mu_2^2) - (\bar\nu_1-\bar\nu_2)t + \sum_{\substack{t\\ h_i\in A_n}} \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right). \tag{3.31}$$

As mentioned in Section 2.6, for every $n\geq 1$ the process $\ln\left(\frac{p(\theta_1|\mathcal{G}^n_t)}{1-p(\theta_1|\mathcal{G}^n_t)}\right)$ is a Lévy process with only positive jumps a.s., and therefore so is $Q(\theta_1|\mathcal{G}^n_t)$. Note that, $Q(\theta_1|\mathcal{G}^n_t)(\omega)\to\ln\left(\frac{p(\theta_1|\mathcal{G}_t)(\omega)}{1-p(\theta_1|\mathcal{G}_t)(\omega)}\right)$ for every t ω-a.s., in the sense that

$$P\left(Q(\theta_1|\mathcal{G}^n_t)(\omega)\to\ln\left(\frac{p(\theta_1|\mathcal{G}_t)(\omega)}{1-p(\theta_1|\mathcal{G}_t)(\omega)}\right)\ \ \forall t\right)=1. \tag{3.32}$$

Define

$$S_n := \inf_{0\leq t\leq s}\left\{t\,\middle|\,Q(\theta_1|\mathcal{G}^n_t)\leq\ln\left(\frac{p'}{1-p'}\right)\right\}. \tag{3.33}$$

By Eq. (2.10), (3.31), and Assumption 2.7(AA10), it follows that for every n, and every time t, $Q(\theta_1|\mathcal{G}^n_t)\leq\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right)$, and therefore, for every $n\geq 1$, $S_n\leq T$. By Eq. (3.31), and Assumption 2.7(AA10), it follows that $Q(\theta_1|\mathcal{G}^n_t)$ is non-decreasing with n for every $t\in(0,s)$. Therefore, there is a stopping time T' s.t. $S_n\nearrow T'\leq T$. Moreover, for every $n<m$, $\quad Q(\theta_1|\mathcal{G}^n_{S_m})\leq Q(\theta_1|\mathcal{G}^m_{S_m})=\ln\left(\frac{p'}{1-p'}\right)$. Fixing n, and taking the limit $m\to\infty$, we get $Q(\theta_1|\mathcal{G}^n_{T'-})\leq\ln\left(\frac{p'}{1-p'}\right)$. Taking now the limit $n\to\infty$, by Lemma 2.8, and Eq. (3.32), we get

$$\ln\left(\frac{p(\theta_1|\mathcal{G}_{T'-})}{1-p(\theta_1|\mathcal{G}_{T'-})}\right) = \lim_{n\to\infty}\ln\left(\frac{p(\theta_1|\mathcal{G}_{S_n})}{1-p(\theta_1|\mathcal{G}_{S_n})}\right) = Q(\theta_1|\mathcal{G}^n_{T'-})\leq\ln\left(\frac{p'}{1-p'}\right). \tag{3.34}$$

Since the process $p(\theta_1|\mathcal{G}_t)$ is left continuous, one has $p(\theta_1|\mathcal{G}_{T'})=p(\theta_1|\mathcal{G}_{T'-})\leq p'$, and so $T\leq T'$. We found that $T'\leq T$, and $T\leq T'$, therefore $T=T'$. By Eq. (2.9), (3.31), and Assumption 2.7(AA10), it follows that $Q(\theta_1|\mathcal{G}^n_t)\leq\ln\left(\frac{p(\theta_1|\mathcal{G}^n_t)}{1-p(\theta_1|\mathcal{G}^n_t)}\right)$, and therefore $S_n\leq T_n$. By taking the limit $n\to\infty$, we get $\lim_{n\to\infty}T_n\geq\lim_{n\to\infty}S_n=T$.

Step 2: $\lim_{n\to\infty}T_n\leq T$

The Lévy process $\left(\ln\left(\frac{p(\theta_1|\mathcal{G}_t)}{1-p(\theta_1|\mathcal{G}_t)}\right)\right)$, has only positive jumps, and therefore, by Bertoin (1996) (Theorem 1, p.189), once the process reaches the level $\ln\left(\frac{p'}{1-p'}\right)$, it must go below this level in any small time interval. Therefore for every $\delta>0$ there exists $0<\epsilon(\omega)<\delta$ such that $p(\theta|\mathcal{G}_{T+\epsilon})(\omega)<p'$ ω-a.s. By Lemma 2.8, $\lim_{n\to\infty}p(\theta|\mathcal{G}^n_{T+\epsilon})(\omega)=p(\theta|\mathcal{G}_{T+\epsilon})(\omega)<p'$ ω-a.s. It follows that for every $\delta>0$, $\lim_{n\to\infty}T_n\leq T+\delta$. Since δ is arbitrary, $\lim_{n\to\infty}T_n\leq T$, as desired. $\square$

The next Lemma states that, fixing a cut-off strategy, additional information is more profitable.[18] Let κ_n be the optimal strategy of D_n.

Lemma 3.7.
1. If $n < m$ then $\alpha_m \le \alpha_n$.
2. If $n \le l < m$ then $E[N^n_{\kappa_n}|\mathcal{F}^l](p) \le E[N^n_{\kappa_n}|\mathcal{F}^m](p)$.

Proof of Lemma 3.7.
1. Fix $n < m$. We show that $\alpha_m \le \alpha_n$. In the proof of Lemma 3.4 we showed that (3.35)

$$f_m(\alpha) := \int_{A_m} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha} + \alpha(\bar{\nu}_1^m - \bar{\nu}_2^m) - \bar{\nu}_2^m + \frac{1}{2}(\alpha+1)\alpha(\tilde{\mu_1} - \tilde{\mu_2})^2 - r.$$

is increasing in $[0,\infty)$.
Since by definition $f_m(\alpha_m) = 0$, it is left to prove that $f_m(\alpha_n) \ge 0$. Simple algebraic manipulations show that:

$$\begin{aligned}
f_m(\alpha_n) =& \int_{A_m} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha_n} + \alpha_n(\bar{\nu}_1^m - \bar{\nu}_2^m) - \bar{\nu}_2^m + \frac{1}{2}(\alpha_n+1)\alpha_n(\tilde{\mu_1} - \tilde{\mu_2})^2 - r \\
=& \int_{A_n} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha_n} + \alpha_n(\bar{\nu}_1^n - \bar{\nu}_2^n) - \bar{\nu}_2^n + \frac{1}{2}(\alpha_n+1)\alpha_n(\tilde{\mu_1} - \tilde{\mu_2})^2 - r \\
+& \int_{A_m\setminus A_n} \nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha_n} + \alpha_n[\bar{\nu}_1^m - \bar{\nu}_1^n - (\bar{\nu}_2^m - \bar{\nu}_2^n)] - (\bar{\nu}_2^m - \bar{\nu}_2^n) \\
=& f_n(\alpha_n) + \int_{A_m\setminus A_n} \left[\nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha_n} + \alpha_n[\nu_1(dh) - \nu_2(dh)] - (\nu_2(dh))\right].
\end{aligned}$$

Since, $f_n(\alpha_n) = 0$, it is left to prove that

$$\int_{A_m\setminus A_n} \left[\nu_2(dh)\left(\frac{\nu_2(dh)}{\nu_1(dh)}\right)^{\alpha_n} + \alpha_n[\nu_1(dh) - \nu_2(dh)] - (\nu_2(dh))\right] \ge 0.$$

This inequality follows by the same arguments as in the proof of Lemma 3.4.

2. Fix $n < m$. By Theorem 2.4, for every $p > p_n^*$, one has $E[N^n_{\kappa_n}|\mathcal{F}^n_s](p) = g_2^n + (g_1^n - g_2^n)p + C_n(1-p)(\frac{1-p}{p})^{\alpha_n}$, with $C_n = \frac{s - g_2^n - p_n^*(g_1^n - g_2^n)}{(1-p^*)(\frac{1-p^*}{p^*})^{\alpha_n}}$ and the cut-off, $p_n^* = \frac{\alpha_n(s-g_2^n)}{(\alpha_n+1)(g_1^n - s) + \alpha_n(s - g_2^n)}$. By constructing an integral equation similar to the one that appears in the proof of Theorem 2.6, we prove that for every $p > p_n^*$,

$$E[N^n_{\kappa_n}|\mathcal{F}^m](p) = g_2^n + (g_1^n - g_2^n)p + C_{n,\alpha_m}(1-p)\left(\frac{1-p}{p}\right)^{\alpha_m}, \tag{3.36}$$

where $C_{n,\alpha_m} := \frac{s - g_2^n - p_n^*(g_1^n - g_2^n)}{(1-p_n^*)(\frac{1-p_n^*}{p_n^*})^{\alpha_m}}$. Since $\alpha_m \le \alpha_l$, we get $E[N^n_{\kappa_n}|\mathcal{F}^l](p) \le E[N^n_{\kappa_n}|\mathcal{F}^m](p)$.

Proof of Eq. (3.36). Let τ be the stopping time of the first jump with size that belongs to A_m. Let T be the stopping time of the first time t that satisfied:

[18] It is well known that in one-player optimization problems, additional information cannot hurt the DM, since he can always ignore it. However, a Markovian strategy does not allow a player to forget additional information, and therefore the statement is not trivial.

(3.37)

$$\begin{aligned}B^{\mu}(t) &:= Z(t) + \left[\frac{2\tilde{\mu} - \tilde{\mu_1} - \tilde{\mu_2}}{2} - \frac{\bar{\nu}_1^m - \bar{\nu}_2^m}{\tilde{\mu_1} - \tilde{\mu_2}}\right] t \\ &\leq -\frac{1}{\tilde{\mu_1} - \tilde{\mu_2}}\left[\ln\left(\frac{p_0}{1-p_0}\right) - \ln\left(\frac{p_n^*}{1-p_n^*}\right)\right] - \frac{1}{\tilde{\mu_1} - \tilde{\mu_2}}\sum_t \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right).\end{aligned}$$

In the notations of Section 2.5, $F := \left[\frac{2\tilde{\mu} - \tilde{\mu_1} - \tilde{\mu_2}}{2} - \frac{\bar{\nu}_1^m - \bar{\nu}_2^m}{\tilde{\mu_1} - \tilde{\mu_2}}\right]$, $E := \frac{1}{\tilde{\mu_1} - \tilde{\mu_2}}\left[\ln\left(\frac{p_0}{1-p_0}\right) - \ln\left(\frac{p_n^*}{1-p_n^*}\right)\right]$, and $G_h := \frac{1}{\tilde{\mu_1} - \tilde{\mu_2}}\sum_t \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right)$. The DM chooses the risky arm until the stopping time T, and then he switches to the safe arm. The calculations below use dynamic programming in which the continuation payoff is determined by the time τ of the first jump with size in A_m, and the value of the continuous part of the payoff at that time.

Constructing the integral equation: The construction of the integral equation is similar to the one in the proof of Theorem 2.6, with one modification: the first jump with size in A_m is added to the expected payoff if and only if its size is in A_n. The integral equation is as follows:[19]

(3.38)

$$\begin{aligned}&U(p_0, E) = p_0 P_{\theta_1}(\tau < T)\left[E_{\theta_1}\left[\int_0^{\tau} re^{-rt} dY_B^1(t)\middle|\tau < T\right]\right. \\ &\left.+\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\left(re^{-rt}\frac{1}{\bar{\nu}_1^m}\int_{A_n}\nu_1(dh)h + \frac{1}{\bar{\nu}_1^m}\int_{A_m}\nu_1(dh)e^{-rt}U(p_{t,h,x}, E+G_h+x)\right)\right] \\ &+ p_0 P_{\theta_1}(\tau > T)\left[E_{\theta_1}\left[\int_0^{T} re^{-rt} dY_B^1(t)\middle|\tau > T\right] + E_{\theta_1}\left[\int_T^{\infty} re^{-rt} s dt\middle|\tau > T\right]\right] \\ &+ (1-p_0) P_{\theta_2}(\tau < T)\left[E_{\theta_2}\left[\int_0^{\tau} re^{-rt} dY_B^1(t)\middle|\tau < T\right]\right. \\ &\left.+\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\left(re^{-rt}\frac{1}{\bar{\nu}_2^m}\int_{A_n}\nu_2(dh)h + \frac{1}{\bar{\nu}_2^m}\int_{A_m}\nu_2(dh)e^{-rt}U(p_{t,h,x}, E+G_h+x)\right)\right] \\ &+ (1-p_0) P_{\theta_2}(\tau > T)\left[E_{\theta_2}\left[\int_0^{T} re^{-rt} dY_B^1(t)\middle|\tau > T\right] + E_{\theta_2}\left[\int_T^{\infty} re^{-rt} s dt\middle|\tau > T\right]\right].\end{aligned}$$

[19] The different terms, relative to Eq. (3.24) are: $re^{-rt}\frac{1}{\bar{\nu}_1^m}\int_{A_n}\nu_1(dh)h$, and $re^{-rt}\frac{1}{\bar{\nu}_2^m}\int_{A_n}\nu_2(dh)h$.

Using similar calculations to those in Section 3.5, we get:

$$
\begin{aligned}
(3.39)\qquad U(p_0, E) &= p_0 \left(\frac{\mu_1 r}{\bar{\nu}_1^m + r} + \frac{\bar{\nu}_1^n H_1^n r}{\bar{\nu}_1^m + r} \right) + (1 - p_0) \left(\frac{\mu_2 r}{\bar{\nu}_2^m + r} + \frac{\bar{\nu}_2^n H_2^n r}{\bar{\nu}_2^m + r} \right) \\
&+ p_0 e^{-E(F_{\mu_1} + \sqrt{2(\bar{\nu_1}+r)+F_{\mu_1}^2})} \left(s - \frac{\mu_1 r}{\bar{\nu}_1^m + r} - \frac{\bar{\nu}_1^n H_1^n r}{\bar{\nu}_1^m + r} \right) \\
&+ (1-p_0) e^{-E(F_{\mu_2} + \sqrt{2(\bar{\nu_2}+r)+F_{\mu_2}^2})} \left(s - \frac{\mu_2 r}{\bar{\nu}_2^m + r} - \frac{\bar{\nu}_2^n H_2^n r}{\bar{\nu}_2^m + r} \right) \\
&+ \int_{-E}^{\infty} \int_0^{\infty} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) \cdot \\
&\quad \cdot \Big[p_0 P_{\theta_1}(\tau < T) f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t) \nu_1(dh)/\bar{\nu}_1 \\
&\quad + (1-p_0) P_{\theta_2}(\tau < T) f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t) \nu_2(dh)/\bar{\nu}_2 \Big].
\end{aligned}
$$

One can verify that for every $p > p_n^*$, the unique solution of Eq. (3.39) is

$U(p, E) = g_2^n + (g_1^n - g_2^n)p + (s - g_1^n)pe^{-(\alpha_m+1)(\tilde{\mu_1}-\tilde{\mu_2})E} + (s - g_1^n)(1 - p)e^{-\alpha_m(\tilde{\mu_1}-\tilde{\mu_2})E}$.

Therefore, for every $p > p_n^*$, the unique solution is

$$
(3.40)\qquad E[N^n_{\kappa_n}|\mathcal{F}^m](p) = g_2^n + (g_1^n - g_2^n)p + C_{n,\alpha_m}(1 - p) \left(\frac{1-p}{p} \right)^{\alpha_m}.
$$

□

Proof of Proposition 2.10.
Let κ_n be the optimal strategy under $\mathcal{F}^n$, and let κ_ϵ be a cut-off strategy such that $E[M_{\kappa_\epsilon}|\mathcal{F}](p) \geq U(p) - \epsilon$. We prove $\lim_{n\to\infty} U_n(p) = U(p)$ in two steps; we first prove that $\lim_{n\to\infty} U_n(p) \geq U(p) - \epsilon$ for every $\epsilon >$, and subsequently we prove that $\lim_{n\to\infty} U_n(p) \leq U(p)$.
By the definition of U_n, and the optimality of κ_n, we get by Lemma 3.6

(3.41)
$$
U_n(p) = E[N^n_{\kappa_n}|\mathcal{F}^n_s](p) \geq E[N^n_{\kappa_\epsilon}|\mathcal{F}^n_s](p) = E[M_{\kappa_\epsilon}|\mathcal{F}^n_s](p) - E[L^n_{\kappa_\epsilon}|\mathcal{F}^n_s](p) \underset{n\to\infty}{\to} E[M_{\kappa_\epsilon}|\mathcal{F}](p).
$$

Since κ_ϵ is ϵ-optimal we obtain $\lim_{n\to\infty} U_n(p) \geq U(p) - \epsilon$.

By Lemmas 3.6(L2) and 3.7, it follows that $E[N^n_{\kappa_n}|\mathcal{F}^n](p) \leq E[N^n_{\kappa_n}|\mathcal{F}](p)$ for every n. By the definition of U, Eq. (3.41), and the optimality of κ_n, we deduce

$$
\begin{aligned}
(3.42)\qquad U(p) &\geq E[M_{\kappa_n}|\mathcal{F}](p) = E[N^n_{\kappa_n}|\mathcal{F}](p) + E[L^n_{\kappa_n}|\mathcal{F}](p) \\
&\geq E[N^n_{\kappa_n}|\mathcal{F}_n](p) + E[L^n_{\kappa_n}|\mathcal{F}](p) = U_n(p) + E[L^n_{\kappa_n}|\mathcal{F}](p).
\end{aligned}
$$

By Lemma 3.6 (L1), $\lim_{n\to\infty} E[L^n_{\kappa_n}|\mathcal{F}](p) = 0$. Therefore, $U(p) \geq \lim_{n\to\infty} U_n(p)$. □

Proof of Theorem 2.11.
1. **Optimality** . Fix $p > p'$. By Lemma 3.6(L3),

$$
\begin{aligned}
(3.43)\quad E[M_{\kappa'}|\mathcal{F}](p) &= \lim_{n\to\infty} [E[M_{\kappa'}|\mathcal{F}_n](p)] = \lim_{n\to\infty} [E[N^n_{\kappa'}|\mathcal{F}_n](p) + E[L^n_{\kappa'}|\mathcal{F}_n](p)] \\
&= \lim_{n\to\infty} E[N^n_{\kappa'}|\mathcal{F}_n](p).
\end{aligned}
$$

By Theorem 2.6, we obtain for $p > p'$:

$$
\begin{aligned}
\lim_{n\to\infty} E[N^n_{\kappa'}|\mathcal{F}_n](p) &= \lim_{n\to\infty} \Bigg[g_1^n p + g_2^n(1-p) + (s-g_1^n)p\left(\frac{1-p}{p}\right)^{\alpha_n+1}\left(\frac{p'}{1-p'}\right)^{\alpha_n+1} \\
&\quad + (s-g_2^n)(1-p)\left(\frac{1-p}{p}\right)^{\alpha_n}\left(\frac{p'}{1-p'}\right)^{\alpha_n}\Bigg] \\
&= g_1' p + g_2'(1-p) + (s-g_1')p\left(\frac{1-p}{p}\right)^{\alpha'+1}\left(\frac{p'}{1-p'}\right)^{\alpha'+1} \\
&\quad + (s-g_2')(1-p)\left(\frac{1-p}{p}\right)^{\alpha'}\left(\frac{p'}{1-p'}\right)^{\alpha'} \\
&= g_2' + p(g_1'-g_2') + C'(1-p)\left(\frac{1-p}{p}\right)^{\alpha'} \\
&= \lim_{n\to\infty}\left[g_2^n + p(g_1^n - g_2^n) + C^n(1-p)\left(\frac{1-p}{p}\right)^{\alpha_n}\right] \\
&= \lim_{n\to\infty}\left[E[N^n_{\kappa_n}|\mathcal{F}_n](p)\right] = \lim_{n\to\infty} U_n(p) = U(p),
\end{aligned} \tag{3.44}
$$

where the last equality follows from Proposition 2.10. Observe that for every $p < p'$:

$$E[M_{\kappa'}|\mathcal{F}](p) = s = U(p). \tag{3.45}$$

By Eqs. 3.43, 3.45 and 3.44 we obtain that $E[M_{\kappa'}|\mathcal{F}](p) = U(p)$, so that κ' is optimal.

2. **Uniqueness**. Let $\bar\kappa$ be a cut-off strategy, different from the optimal cut-off strategy κ'. As in Eq. (3.43),and the optimality of (κ_n), we obtain,

$$E[M_{\kappa'}|\mathcal{F}](p) = \lim_{n\to\infty} E[N^n_{\kappa_n}|\mathcal{F}_n](p) \geq \lim_{n\to\infty} E[N^n_{\bar\kappa}|\mathcal{F}_n](p) = E[M_{\bar\kappa}|\mathcal{F}](p). \tag{3.46}$$

We show now that for any cut-off strategy $\bar\kappa$, with cut-off $\bar p$, κ' is strictly better than $\bar\kappa$. As in Eq. (3.44), for every $p > \bar p$,

$$
\begin{aligned}
E[M_{\bar\kappa}|\mathcal{F}](p) &= \lim_{n\to\infty} E[N^n_{\bar\kappa}|\mathcal{F}_n](p) \\
&= \lim_{n\to\infty} \Bigg[g_1^n p + g_2^n(1-p) + (s-g_1^n)p\left(\frac{1-p}{p}\right)^{\alpha_n+1}\left(\frac{\bar p}{1-\bar p}\right)^{\alpha_n+1} \\
&\quad + (s-g_2^n)(1-p)\left(\frac{1-p}{p}\right)^{\alpha_n}\left(\frac{\bar p}{1-\bar p}\right)^{\alpha_n}\Bigg] \\
&= g_1' p + g_2'(1-p) + (s-g_1')p\left(\frac{1-p}{p}\right)^{\alpha'+1}\left(\frac{\bar p}{1-\bar p}\right)^{\alpha'+1} \\
&\quad + (s-g_2')(1-p)\left(\frac{1-p}{p}\right)^{\alpha'}\left(\frac{\bar p}{1-\bar p}\right)^{\alpha'}.
\end{aligned} \tag{3.47}
$$

If $p' < \bar{p}$, then $E[M_{\kappa'}|\mathcal{F}](\bar{p}) > s = E[M_{\bar{\kappa}}|\mathcal{F}](\bar{p})$,where the inequality follows by Eq. (3.44). Therefore, κ' is strictly better then $\bar{\kappa}$. If $\bar{p} < p'$, then for every $\bar{p} < p < p'$, $E[M_{\bar{\kappa}}|\mathcal{F}](p) \le E[M_{\kappa'}|\mathcal{F}](p) = s$. Since $E[M_{\bar{\kappa}}|\mathcal{F}](\bar{p}) = s$, and since the function $E[M_{\bar{\kappa}}|\mathcal{F}](p) = g_1'p + g_2'(1-p) + (s - g_1')p \left(\frac{1-p}{p}\right)^{\alpha'+1} \left(\frac{\bar{p}}{1-\bar{p}}\right)^{\alpha'+1} + (s - g_2')(1-p) \left(\frac{1-p}{p}\right)^{\alpha'} \left(\frac{\bar{p}}{1-\bar{p}}\right)^{\alpha'}$ is not constant on any interval, it follows that there exists $\epsilon > 0$ sufficiently small such that $E[M_{\bar{\kappa}}|\mathcal{F}](p+\epsilon) < s$ and therefore $\bar{\kappa}$ is not optimal.

□

Proof of Theorem 2.12.
Let $E[M_{\kappa'(q)}|\mathcal{F}](p)$, be the expected discounted payoff for a DM with subjective belief $q > p'$. As in the proof of Theorem 2.11, and by Lemma 3.6,

(3.48)
$$\begin{aligned} E[M_{\kappa'(q)}|\mathcal{F}](p) &= \lim_{n\to\infty} \left[E[M_{\kappa'(q)}|\mathcal{F}_n](p)\right] = \lim_{n\to\infty} \left[E[N^n_{\kappa'(q)}|\mathcal{F}_n](p) + E[L^n_{\kappa'(q)}|\mathcal{F}_n](p)\right] \\ &= \lim_{n\to\infty} E[N^n_{\kappa'(q)}|\mathcal{F}_n](p). \end{aligned}$$

By Theorem 2.6,

$$\begin{aligned} \lim_{n\to\infty} E[N^n_{\kappa'(q)}|\mathcal{F}_n](p) \quad &= \lim_{n\to\infty} \quad \Bigg[g_1^n p + g_2^n(1-p) + (s - g_1^n)p\left(\frac{1-q}{q}\right)^{\alpha_n+1}\left(\frac{p'}{1-p'}\right)^{\alpha_n+1} \\ &\qquad +(s-g_2^n)(1-p)\left(\frac{1-q}{q}\right)^{\alpha_n}\left(\frac{p'}{1-p'}\right)^{\alpha_n}\Bigg] \\ &= \qquad g_1'p + g_2'(1-p) + (s-g_1')p\left(\frac{1-q}{q}\right)^{\alpha'+1}\left(\frac{p'}{1-p'}\right)^{\alpha'+1} \\ &\qquad +(s-g_2')(1-p)\left(\frac{1-q}{q}\right)^{\alpha'}\left(\frac{p'}{1-p'}\right)^{\alpha'}. \end{aligned}$$

□

Proof of Theorem 2.13. Substituting $q_0 = p_0 + \epsilon$ in Eq. (2.8) (or in Eq. (2.13)), yields

$$V_{p_0}(\epsilon) = p_0(s-g_1)\left(\frac{p^*}{1-p^*}\right)^{\alpha+1}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha+1} + (1-p_0)(s-g_2)\left(\frac{p^*}{1-p^*}\right)^{\alpha}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha},$$

where α is the unique solution of (2.3) in $(0,\infty)$.

Simple algebraic manipulation yield:

$$
\begin{aligned}
V_{p_0}(\epsilon) &= g_1 p_0 + g_2(1-p_0) + p_0(s-g_1)\left(\frac{p^*}{1-p^*}\right)^{\alpha+1}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha+1} \\
&\quad +(1-p_0)(s-g_2)\left(\frac{p^*}{1-p^*}\right)^{\alpha}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha} \\
&= [g_1 p_0 + g_2(1-p_0)] + \left(\frac{p^*}{1-p^*}\right)^{\alpha} \cdot \\
&\quad \cdot \left[p_0(s-g_1)\frac{p^*}{1-p^*}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha+1} + (1-p_0)(s-g_2)\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha}\right] \\
&= [g_1 p_0 + g_2(1-p_0)] + \left(\frac{\alpha}{\alpha+1}\right)^{\alpha}\left(\frac{s-g_2}{g_1-s}\right)^{\alpha} \cdot \\
&\quad \cdot \left[-p_0(s-g_2)\frac{\alpha}{\alpha+1}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha+1} + (1-p_0)(s-g_2)\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha}\right] \\
&= [g_1 p_0 + g_2(1-p_0)] + \left(\frac{\alpha}{\alpha+1}\right)^{\alpha}\left(\frac{s-g_2}{g_1-s}\right)^{\alpha}\frac{1}{p_0(s-g_2)} \cdot \\
&\quad \cdot \left[-\frac{\alpha}{\alpha+1}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha+1} + \frac{1-p_0}{p_0}\left(\frac{1-p_0-\epsilon}{p_0+\epsilon}\right)^{\alpha}\right].
\end{aligned}
$$

For every $x \geq 0$ define: $W(x) = V_p(x) - V_p(-x)$. This is the difference between the payoff of an optimist and the payoff of a pessimist. Straightforward calculations show that
$W'(x) = \frac{\alpha x}{p}\left[-\left(\frac{1-(p+x)}{p+x}\right)^{\alpha}\frac{1}{(p+x)^2(1-(p+x))} + \left(\frac{1-(p-x)}{p-x}\right)^{\alpha}\frac{1}{(p-x)^2(1-(p-x))}\right]$
so that $W(0) = W'(0) = 0$. Suppose $\alpha > 1$. Since $\frac{1}{(p+x)^3} < \frac{1}{(p-x)^3}$, and $\left(\frac{1-(p+x)}{p+x}\right)^{\alpha-1} < \left(\frac{1-(p-x)}{p-x}\right)^{\alpha-1}$ we get $W'(x) > 0$ for every $x > 0$ such that $p^* < p \pm x \leq 1$, and so $W(x) > 0$: an optimist will fare better.

If $0 < \alpha < 1$, it is easy to verify that $p^* < \frac{\alpha+2}{3}$. Since the function $\left(\frac{1-y}{y}\right)^{\alpha}\frac{1}{y^2(1-y)}$ decreases for $0 < y < \frac{\alpha+2}{3}$ and increases for $\frac{\alpha+2}{3} < y < 1$. For every $x > 0$ such that $p^* < p \pm x \leq \frac{\alpha+2}{3}$ we get $W'(x) > 0$: an optimist will fare better. For every $x > 0$ such that $\frac{\alpha+2}{3} < p \pm x$ we get $W'(x) < 0$, and so $W(x) < 0$: a pessimist will fare better. □

3.7. **Information Pricing.**

Lemma 3.8. *Let $f_1(\eta)$ and $f_1(\eta)$ be the functions defined in (2.14) and (2.15). The equation $f_1(\eta) = 0$ admits a unique solution in the interval $(-1, \infty)$, and the equation $f_2(\eta) = 0$ admits a unique solution in the interval $(0, \infty)$.*

Proof of Lemma 3.8.
The function f_1 is a continuous function that satisfies $f_1(-1) < 0$ and $f_1(\infty) = \infty$. To show that the equation $f_1(\eta) = 0$ in the interval $[-1, \infty)$ has a unique solution it is therefore sufficient to prove that f_1 is convex in η in $(-1, 0)$ and increasing in η in $(0, \infty)$. The function f_2 is a continuous function that satisfies $f_2(0) < 0$ and $f_2(\infty) = \infty$. To show that $f_2(\eta) = 0$ has a unique solution it is therefore sufficient

to prove that f_2 is increasing in η in $(0,\infty)$. The verification that these properties hold is done as in the proof of Lemma 3.4 $\square$

Proof of Theorem 2.14. We construct an integral equation similar to the one in the proofs of Theorem 2.6 and Lemma 3.7.

Let τ be the stopping time of the first jump with size that belongs to A_n. Let the stopping time T be the first time t that satisfied:

$$\begin{aligned}B^{\mu}(t) &:= Z(t) + \left[\frac{2\tilde{\mu} - \tilde{\mu_1} - \tilde{\mu_2}}{2} - \frac{\bar{\nu}_1^n - \bar{\nu}_2^n}{\tilde{\mu_1} - \tilde{\mu_2}}\right] t \\ &\leq -\frac{1}{\tilde{\mu_1} - \tilde{\mu_2}}\left[\ln\left(\frac{p_0}{1-p_0}\right) - \ln\left(\frac{p'}{1-p'}\right)\right] - \frac{1}{\tilde{\mu_1} - \tilde{\mu_2}}\sum_t \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right).\end{aligned} \tag{3.49}$$

In the notations of Section 2.5, $F := \left[\frac{2\tilde{\mu}-\tilde{\mu_1}-\tilde{\mu_2}}{2} - \frac{\bar{\nu}_1^n-\bar{\nu}_2^n}{\tilde{\mu_1}-\tilde{\mu_2}}\right]$, $E := \frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\left[\ln\left(\frac{p_0}{1-p_0}\right) - \ln\left(\frac{p'}{1-p'}\right)\right]$, and $G_h := \frac{1}{\tilde{\mu_1}-\tilde{\mu_2}}\sum_t \ln\left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)}\right)$. The DM chooses the risky arm until the stopping time T, then switches to the safe arm. The calculations below use dynamic programming. There are two ways to construct the integral equation, both are similar to the proof of Theorem 2.6.

(a) The continuation payoff is determined by the stopping time τ of the first jump with size in A_m, and the value of the continuous part of the payoff at that time. The modification here is that if the first jump with size in A_m occurs before time T, then the expected payoff from jumps in $A_n \setminus A_m$ up to time τ is being added to the expected payoff.

(b) The continuation payoff is determined by the stopping time τ of the first jump with size in A_n, and the value of the continuous part of the payoff at that time. The modification here is that we distinguish between two cases. If the first jump with size in A_n occurs before the stopping time T and if it belongs to A_m, then the DM updates both his posterior $p_{t,h,x}$ and the intercept $E + G_h + x$, according to the time τ, the size h of the jump, and the value of the continuous part of the payoff at the stopping time τ. The jump's size is added to the expected payoff. If the first jump with size in A_n occurs before the stopping time T and if it belongs to $A_n \setminus A_m$, then the DM updates both his posterior $p_{t,x}$, and the intercept $E + x$ according to the value of the continuous part of the payoff at the stopping time τ. The jump's size is not added to his expected payoff.

The posterior $p_{t,x}$ is updated as follows

$$\begin{aligned}p_{t,x} &:= P(\theta_1 | \tau < T, \tau = t, B^{\mu}(\tau) \in dx) \\ &= \frac{P(\tau < T, \tau = t, B^{\mu}(\tau) \in dx | \theta_1) P(\theta_1)}{P(\tau < T, \tau = t, B^{\mu}(\tau) \in dx | \theta_1) P(\theta_1) + P(\tau < T, \tau = t, B^{\mu}(\tau) \in dx | \theta_2) P(\theta_2)} \\ &= \frac{p_0 P_{\theta_1}(\tau < T) f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)}{p_0 P_{\theta_1}(\tau < T) f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t) + (1-p_0) P_{\theta_2}(\tau < T) f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)},\end{aligned}$$

It is more convenient to use the second way to construct the integral equation:
(3.50)

$$\begin{aligned}
U(p_0,E) &= p_0 P_{\theta_1}(\tau<T)\left[E_{\theta_1}\left[\int_0^{\tau} re^{-rt}dY_B^1(t)\Big|\,\tau<T\right]\right.\\
&+\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_1}_{(B^{\mu_1}(\tau),\tau)|\tau<T}(x,t)\left(re^{-rt}\frac{1}{\bar\nu_1^n}\int_{A_n}\nu_1(dh)h+\frac{1}{\bar\nu_1^n}\int_{A_n\setminus A_m}\nu_1(dh)e^{-rt}U(p_{t,x},E+x)\right.\\
&\left.\left.+\frac{1}{\bar\nu_1^n}\int_{A_m}\nu_1(dh)e^{-rt}U(p_{t,h,x},E+G_h+x)\right)\right]\\
&+p_0P_{\theta_1}(\tau>T)\left[E_{\theta_1}\left[\int_0^{T} re^{-rt}dY_B^1(t)\Big|\,\tau>T\right]+E_{\theta_1}\left[\int_T^{\infty} re^{-rt}sdt\Big|\,\tau>T\right]\right]\\
&+(1-p_0)P_{\theta_2}(\tau<T)\left[E_{\theta_2}\left[\int_0^{\tau} re^{-rt}dY_B^1(t)\Big|\,\tau<T\right]\right.\\
&+\int_{-E}^{\infty}\int_0^{\infty} f^{\theta_2}_{(B^{\mu_2}(\tau),\tau)|\tau<T}(x,t)\left(re^{-rt}\frac{1}{\bar\nu_2^n}\int_{A_n}\nu_2(dh)h+\frac{1}{\bar\nu_2^n}\int_{A_n\setminus A_m}\nu_2(dh)e^{-rt}U(p_{t,x},E+x)\right.\\
&\left.\left.+\frac{1}{\bar\nu_2^m}\int_{A_m}\nu_2(dh)e^{-rt}U(p_{t,h,x},E+G_h+x)\right)\right]\\
&+(1-p_0)P_{\theta_2}(\tau>T)\left[E_{\theta_2}\left[\int_0^{T} re^{-rt}dY_B^1(t)\Big|\,\tau>T\right]+E_{\theta_2}\left[\int_T^{\infty} re^{-rt}sdt\Big|\,\tau>T\right]\right].
\end{aligned}$$

The different terms, relative to Eq. (3.24), here are:

$$\begin{aligned}
&re^{-rt}\frac{1}{\bar\nu_1^n}\int_{A_n}\nu_1(dh)h+\frac{1}{\bar\nu_1^n}\int_{A_n\setminus A_m}\nu_1(dh)e^{-rt}U(p_{t,x},E+x)\\
&+\frac{1}{\bar\nu_1^n}\int_{A_m}\nu_1(dh)e^{-rt}U(p_{t,h,x},E+G_h+x),
\end{aligned}$$

and

$$\begin{aligned}
&re^{-rt}\frac{1}{\bar\nu_2^n}\int_{A_n}\nu_2(dh)h+\frac{1}{\bar\nu_2^n}\int_{A_n\setminus A_m}\nu_2(dh)e^{-rt}U(p_{t,x},E+x)\\
&+\frac{1}{\bar\nu_2^m}\int_{A_m}\nu_2(dh)e^{-rt}U(p_{t,h,x},E+G_h+x).
\end{aligned}$$

One can verify that for every $p>p_m^*$, the unique solution of Eq. (3.50) is
$U(p,E)=g_2^n+(g_1^n-g_2^n)p+(s-g_1^n)pe^{-(1+\beta_1)(\tilde{\mu_1}-\tilde{\mu_2})E}+(s-g_2^n)(1-p)e^{-\beta_2(\tilde{\mu_1}-\tilde{\mu_2})E}$.
Therefore, for every $p>p_m^*$, the unique expected discounted payoff is:

$$\begin{aligned}
E[N^n_{\kappa'}|\mathcal{F}^m](p) &= g_2^n+(g_1^n-g_2^n)p+(s-g_1^n)(1-p)\left(\frac{1-p}{p}\right)^{\beta_1}\left(\frac{p'}{1-p'}\right)^{(\beta_1+1)}\\
&+(s-g_2^n)(1-p)\left(\frac{1-p}{p}\right)^{\beta_2}\left(\frac{p'}{1-p'}\right)^{\beta_2},
\end{aligned}$$

as desired.

The proof of the second part of the theorem is similar to the proof of Eq. (3.36), and is therefore omitted. □

4. Future directions

Our results call for further research.

- We studied the case that the High distribution dominates the Low distribution in a strong sense (see Assumptions 2.1). These assumptions ensure that the discontinuities of the process of posterior belief are always to one direction. It is interesting to know whether our characterization holds under only assumptions A1 and A2.
- We assumed that the payoff is distributed according to a Levy process. It will be interested to solve the model when the payoff is distributed according to a geometric Levy process.
- Bolton and Harris (1999), Keller, Rady and Cripps (2004), and Keller and Rady (2008) study a strategic version of the model, in which several decision makers face identical unknown arms. Klein and Rady (2008) study a strategic version of the model, in which the arms of two decision makers are negatively correlated. It will be interesting to solve these models, when the payoff's distribution are general Levy processes.

References

1. Applebaum D.(2004) Lévy Processes and Stochastic Calculus. Cambridge.
2. Babcock L. and Loewenstein G. (1997) Explaining Bargaining Impasse: The Role of Self-Serving Biases. *Journal of Economic Perspectives*, **11**, 109-126.
3. Baumhart R. (1968) *An Honest Profit*. Prentice Hall.
4. Bergemann D. and Valimaki J. (2006). Bandit Problems. HECER discussion paper number 93.
5. Berry D. A. and Fristedt, B. (1985). Bandit Problems: Sequential Allocation of Experiments. Chapman and Hall, New York.
6. Bertoin, J. (1996) Lévy Processes. Cambridge University Press. Cambridge.
7. Besanko D. and Wu J. (2008) The Impact of Market Structure on R&D Competition with Learning. Preprint.
8. Bolton P. and C. Harris (1999) Strategic Experimentation. *Econometrica*, **67**, 349374.
9. Borodin A.N. (1996) Handbook of Brownian motion - facts and formulae. Basel, Birkhauser.
10. Camerer C.F. and Lovallo D. (1999) Overconfidence and Excess Entry: An Experimental Approach. *The American Economic Review*, **89**, 306-318.
11. Carr.P and Wu.L (2004) Time-Changed Lévy Processes and Option Pricing. *Journal of Financial Economics*, **71**, 113141.
12. Cross P. (1977) Not can, But will College Teaching be Improved? *New Directions for Higher Education*, **17**, 1-15.
13. Friedman A. (1969) Partial Differential Equations, Holt, Rinehart & Winston.
14. Geman, H., Madan, D., and Yor M. (2001) Time Changes for Lévy Processes. *Mathematical Finance*, **17**, 7996.
15. Gittins J.C. and Jones D.M (1979) A Dynamic Allocation Index for the Discounted Multi-armed Bandit Problem. *Biometrika*, **66**, 561-565
16. Heifetz A., Shannon C. and Spiegel Y. (2007) The Dynamic Evolution of Preferences. *Economic Theory*, **32**, 251-286.
17. Karatzas, I. (1984) Gittins indices in the dynamic allocation problem for diffusion processes. *Ann. of Probab*, **12**, 173192.
18. Kaspi.H and Mandelbaum.A (1995) Lévy Bandits: Multi-Armed Bandits Driven by Lévy Processes. Ann. Appl. Probab., **5**, 541-565.
19. Keller, G., Rady.S and Cripps M. (2005) Strategic Experimentation with Exponential Bandits. *Econometrica*, **73**, 39-68.
20. Klein.N and Rady.S (2008) Negatively Correlated Bandits, under revision.
21. Larwood L. and Whittaker W. (1977) Managerial Myopia: Self-Serving Biases in Organizational Planning. *Journal of Applied Psychology*, **62**, 194-198.

22. Matacz A. (2000) Financial Modeling and Option Theory with the Truncated Lévy Process. *International Journal of Theoretical and Applied Finance.*, **3**, 143.
23. McLennan A. (1984) Price Dipersion and Incomplete Learning in the Long Run. *Journal of Economic Dynamic and control*, **7**, 331-347.
24. Moscarini G. and Squintani F. (2004) Competitive Experimentation with Private Information. *Cowles Foundation DP 1489.*
25. Robbins H. (1952) A Two-Armed Some Aspects of the Sequential Design of Expirements. *Bull.Amer.Math.Soc*, **58**, 527-535.
26. Roberts and Weitzman (1981) Funding Criteria for Research, Development, and Exploration Projects. *Econometrica*, **49**, 1261-1288.
27. Rothschild M. (1974) A Two-Armed Bandit Theory of Market Pricing. *Journal of Economic Theory*, **9**, 185-202.
28. Rustichini A. and Wolinsky A. (1995) Learning About Variable Demand in the Long Run. *Journal of Economic Dynamics and Control*, **19**, 1283-1292.
29. Svenson O. (1981) Are We all Less Risky and More Skillful than our Fellow Drivers? *Acta Psychologica*, **47**, 143-148.
30. Weinstein N.D. (1980) Unrealistic Optimism about Future Life Events. *Journal of Personality and Social Psychology*, **39**, 806-820.
31. Weitzman M.L.. (1979) Optimal Search for the Best Alternative. *Econometrica*, **47**, 641-654.