Information acquisition and choice under uncertainty *

Isabelle Brocas
University of Southern California and CEPR

Juan D. Carrillo
University of Southern California and CEPR

Address for correspondence:

Isabelle Brocas and Juan D. Carrillo

Department of Economics, University of Southern California

3620 S. Vermont Ave., Los Angeles, CA 90089-0253, USA

email: <brocas@usc.edu> and <juandc@usc.edu>

Running head: INFORMATION ACQUISITION AND CHOICE UNDER UNCERTAINTY

Abstract

This paper presents a model where individuals have imperfect information and there is an opportunity cost of learning. It shows that the endogenous decision to collect costly information before taking an action has a systematic effect on choices. More precisely, consider two alternatives with ex ante identical expected payoff but different variances. The model predicts that, after the learning process is stopped, a majority of individuals will select the alternative with largest payoff-variance. The result persists when agents have multiple sources of information. Applications to entrepreneurial investments, composition of advisory committees, and judicial decision-making are discussed.

1 Motivation

Empirical studies show a high rate of failure in new businesses (for data, see e.g. Camerer and Lovallo (1999) and the references there-in). Explanations based on hit-and-run strategies or a skewed distribution of profits with positive expected returns can rationalize the willingness of entrepreneurs to engage in these high-risk, low-probability activities. The literatures in psychology and behavioral finance argue, on the contrary, that a rational cost-benefit analysis fails short to explain these choices. These theories claim that an "irrational" tendency to optimism and overconfidence (loosely defined as an individual holding an excessively positive belief in his capabilities or chances of success) provides a more accurate account for this behavioral tendency.[1] It is this same irrational belief that pushes researchers to pursue adventurous innovation strategies.

The present paper discusses a different and possibly complementary force for this observed tendency to engage in high risk enterprises. We consider individuals with imperfect knowledge about the environment (or about themselves) who choose between alternatives with ex ante identical expected payoffs but different risks. We argue that if learning is feasible but sequential and costly, then the *endogenous* decision to collect information generates in a population of *rational* individuals a systematic and testable tendency to favor the alternatives characterized by highest risk. Stated differently, the paper shows that, in settings where the collection of information is dynamic and endogenous, a population of rational individuals display an aggregate form of behavior which may look like driven by irrational beliefs. Naturally, we do not argue that imperfect knowledge and endogenous information acquisition provide an explanation for all the choices documented above. In that respect, the paper just adds one new element to the discussion: risky decisions may be favored not because of irrational beliefs and cognitive limitations but because of rational learning and an option value argument.

To illustrate our theory, consider the following stylized example. Two risk-neutral entrepreneurs must decide between two investment strategies. The preferences of these entrepreneurs are identical in most respects. In particular, for any given belief about the relative chances of success of these investments, not only they both prefer to undertake the same one, but they also incur the same utility loss if the other investment is selected. There is, however, one subtle difference: the first investment strategy is more risky for one entrepreneur whereas the second strategy is more risky for the other. This difference in risks may reflect, for example, the fact that entrepreneurs start with different core activities. Pursuing a strategy that builds on existing technology or competence is intrinsically less risky than giving up the current technology or competence in order to pursue a radically new strategy. Entrepreneurs initially share the same belief regarding the relative value of both strategies but they can independently acquire extra evidence at the expense of postponing the investment decision. Finally, we assume that delay is costly: the project may become obsolete or less valuable, and the profits postponed are discounted at a positive rate. Given the same starting belief and the identical behavior and utility loss of both entrepreneurs for any given belief, one could think that their choices would be indistinguishable in a stochastic sense. However, this intuition is incorrect: after the information acquisition process, each entrepreneur will choose his more risky strategy with higher probability than his less risky one, both when it is ex-post revealed to be the best alternative and when it is ex-post revealed to be the worst one.

The key for the result lies in the opportunity cost of learning. Suppose that the preliminary evidence points towards one of the investments. The opportunity cost of sampling is greatest for the entrepreneur who derives highest payoff if that investment is chosen and turns out to be successful, that is, for the entrepreneur with highest payoff variance under this investment. This individual is then more tempted than the other to stop the information acquisition process, and enjoy the high expected payoff of his (hopefully correct) decision. Overall, these two entrepreneurs would behave identically if the amount of information collected were exogenously fixed. However, the asymmetry in the *total* payoff of making the right decision combined with the costly endogenous choice of learning implies that, in expectation, they will end up choosing different actions and therefore committing different investment errors. The reader may find obvious that each entrepreneur favors the investment that has the potential to yield highest payoff. However, one should realize that by adopting such strategy, entrepreneurs are also committing more often the mistakes that are most costly.

The result has two immediate consequences for the design of advisory committees. Suppose that a firm requests the opinion of several employees regarding the optimal investment strategy and aggregates the information. If, for some reason (related to profit maximization or not), the firm has a preference for a particular investment, it can increase the probability that this investment is proposed simply by choosing advisors whose payoff variance is greatest for that investment. Perhaps more surprisingly, a firm concerned with maximizing the probability of choosing the correct investment, will optimally select all advisors of the same type. Thus, the systematic differences in choices (and errors) should persist even when multiple sources of information are available.

Note that, because all agents are rational in our model, the amount of information collected is always optimal. Costly learning implies that entrepreneurs decide without being fully informed, and therefore make wrong choices with positive probability. Thus, the systematic differences in choices and in the type of mistakes the entrepreneurs make relates to their different likelihood of choosing (rightly or wrongly) one investment or the other, and not on whether they sample optimally. Also, the cost of acquiring information is a delayed (and therefore discounted) payoff and/or a probability of the project becoming obsolete. In either case, it is proportional to the expected payoff if sampling is stopped and the action with highest expected payoff undertaken. This is crucial as it implies that the project with highest payoff variance has also the highest opportunity cost of sampling. If, instead, we assumed a fixed sampling cost, all entrepreneurs would choose the different investments with identical probabilities and the effect highlighted in the paper would disappear. Last, the paper discusses other applications such as research strategies of firms, court judgments under civil law, and career choices.

Related literature The paper is related to two strands of the literature. First, to the individual choice models developed independently by Zabojnik (2004), Van den Steen (2004), Santos-Pinto and Sobel (2005) and more recently Benoît and Dubra (2007). These works concentrate on a single activity that requires ability and show that agents may perceive themselves as "better" than their objective ranking. The argument in Zabojnik (2004) is based on an opportunity cost of learning (as in our paper) and an exogenous utility function convex in ability. Under appropriate initial conditions on the discount factor, the initial ability and the degree of convexity of the utility, only individuals with an expected ability below a certain threshold experiment, generating the bias. In Van den Steen (2004) and Santos-Pinto and Sobel (2005) agents evaluate situations using different criteria: they are endowed with heterogeneous beliefs and heterogeneous preferences about which skills are valuable, respectively. Agents can invest in an action or in improving these skills. The key issue is that agents evaluate the skills of others according to their own criteria rather than the criteria of others. This, again, generates a bias in self-assessment. Benoît and Dubra (2007) demonstrate that a prior distribution of beliefs and a private signal impose very little statistical restrictions on a summary of the posterior beliefs held in the population (e.g., whether their belief is above or below the x-percentile). In particular, the authors find an upper bound on the fraction of individuals who can rate themselves above $x\%$ which is strictly greater than x for all x. Overall, these papers explain why a majority of individuals may hold above average or even above median beliefs concerning a certain positive trait. Our setting is different in that our agents choose between several alternatives. Although we share with these papers the result that one option will be systematically favored, our main goal is to explore the behavioral consequences. In particular, we emphasize the role of the payoff-variance of the different alternatives in determining the propensity of individuals to take different actions and therefore commit different types of errors. We also argue the existence of a testable relationship between delay and type of action undertaken. Finally, we show how this systematic tendency to favor certain choices can be exploited by third parties.

Second, since we build a model of costly learning with an optimal stopping rule, the paper can be seen as a particular application of optimal experimentation (see e.g. the statistical literature on multi-armed bandits summarized in Berry and Fristedt (1985)). There are two features that make our model different from the main economic applications studied in this literature. First, unlike in Bolton and Harris (1999) or Keller and Rady (1999) for example, the agent does not decide at each date on which arm he experiments. Instead, the decision to keep accumulating evidence produces a signal about the relative likelihood of each state. Second, most of this literature "highlights the fundamental trade-off between the conflicting objectives of learning and obtaining high current payoffs" (Aghion et al. (1991, p. 623)). More precisely, by experimenting with one arm, the agent obtains the payoff associated with that alternative. Thus, he may choose a highly informative arm with low expected payoff in order to learn how to behave in the future. In our paper, experimenting has a different implicit cost: the discount factor applied to the action eventually taken. It thus depends on the current belief about which action is optimal and it is only borne when the experimentation process is stopped.

The plan of the paper is the following. We first present a model in which a decision maker has imperfect information about the state of nature and chooses between two (risky) actions. We are particularly interested in the behavior of agents with "seemingly the same" motivations. Two agents have the same motivations if, for any given belief, they share the same difference in expected utility between the actions (section 2). We show that their different incentives to acquire information affects their behavior in systematically different ways (section 3). We also determine the effect on third parties when actions generate externalities (section 4). Last, we provide some concluding remarks (section 5).

2 The model

2.1 States, actions and utilities

We consider the following model. There are I types of agents in the economy ($i \in I$) and two states of the world A and B denoted by s. Agents choose among a finite set of irreversible actions $\gamma \in \Gamma$. The ex-post utility of a type-i agent is a function $u_i(\gamma, s)$ of the action and the true state. For each state, there is one action that provides the highest utility. Naturally, this action is selected if the state is known. However, agents initially have imperfect knowledge about the state. More precisely, they share a common prior p that the true state is A. The expected payoff of taking action γ is:

$$u_i(\gamma) = p\, u_i(\gamma, A) + (1-p)\, u_i(\gamma, B)$$

For expositional purposes, we will study a simpler version with only two actions $\Gamma = \{a, b\}$. As we will develop in the discussion of our results, this restriction is made with little loss of generality. Action a is optimal if the state is A and action b is optimal if the state is B. Last and foremost, the variance in payoffs is different across actions. To capture this property, we assume that the utility representation for a type-i agent is:

$$u_i(a, \cdot) = \begin{cases} x_i & \text{if} \quad s = A \\ -x_i & \text{if} \quad s = B \end{cases} \qquad \text{and} \qquad u_i(b, \cdot) = \begin{cases} -y_i & \text{if} \quad s = A \\ y_i & \text{if} \quad s = B \end{cases}, \tag{1}$$

with $x_i > 0$ and $y_i > 0$. This representation allows us to restrict the attention to the most interesting cases where it is possible to compare the variances of the actions and to have clear-cut results. Indeed, it is easy to see that when $x_i > y_i$, then action a has the highest variance in payoffs.

2.2 Information

Before making a decision, each agent can learn about the likelihood of the states. We denote by $\tau_{i,t}$ the decision of agent i at a given date $t \in \{0, 1, ..., T-1\}$, where T is finite but arbitrarily large. At each date, his options are either to take the optimal action conditional on his current information ($\tau_{i,t} = \gamma \in \{a, b\}$) or wait until the following period ($\tau_{i,t} = w$). If the agent undertakes an (irreversible) action, then payoffs are realized and the game ends. Waiting has costs and benefits. On the one hand, the delay implied by the decision to wait one more period before acting is costly. We denote by δ (< 1) the discount factor. Alternatively, $1 - \delta$ can be interpreted as the probability that all options vanish, in which case the agent obtains no payoff. On the other hand, the agent obtains between dates t and $t+1$ one signal $\sigma \in \{\alpha, \beta\}$ imperfectly correlated with the true state. Information improves the quality of the decision made by the agent. As long as the agent waits, he keeps the option of undertaking action a or b in a future period, except at date T where waiting is not possible anymore, so the agent's options are reduced to $\tau_{i,T} \in \{a, b\}$.[2] The relation between signal and state is the following:

$$\Pr[\,\alpha \mid A\,] = \Pr[\,\beta \mid B\,] = \theta \quad \text{and} \quad \Pr[\,\alpha \mid B\,] = \Pr[\,\beta \mid A\,] = 1 - \theta,$$

where $\theta \in (1/2, 1)$ captures the accuracy of information: as θ increases, the informational content of a signal σ increases (when $\theta \to 1/2$ signals are uninformative, and when $\theta \to 1$ one signal perfectly informs the agent about the true state).[3]

Suppose that a number n_α of signals α and a number n_β of signals β are revealed during the $n_\alpha + n_\beta$ periods in which the agent waits. Using standard statistical techniques, it is possible to compute the agent's posterior belief about the state:

$$\begin{aligned}\Pr(A \mid n_\alpha, n_\beta) &= \frac{\Pr(n_\alpha, n_\beta \mid A)\Pr(A)}{\Pr(n_\alpha, n_\beta \mid A)\Pr(A) + \Pr(n_\alpha, n_\beta \mid B)\Pr(B)} \\ &= \frac{\theta^{n_\alpha - n_\beta} \cdot p}{\theta^{n_\alpha - n_\beta} \cdot p + (1-\theta)^{n_\alpha - n_\beta} \cdot (1-p)}\end{aligned}$$

It is interesting to notice that the posterior depends exclusively on the difference between the number of signals α and the number of signals β. So, roughly speaking, two different signals "cancel each other out" for the purpose of computing the expected belief. The relevant variable which will be used from now on is $n \equiv n_\alpha - n_\beta \in \mathbb{Z}$. We define the posterior probability $\mu(n) \equiv \Pr(A \mid n_a, n_b)$. Rearranging terms, we have:[4]

$$\mu(n) = \frac{1}{1 + \left(\frac{1-\theta}{\theta}\right)^n \frac{1-p}{p}}$$

Last, when solving the model, we will treat n as a real number (instead of an integer as we should in order to be rigorous). This mathematical abuse is made for technical convenience.

2.3 Types

Different types of agents have different preferences, which translate into different cardinal representations of their utility. From a general perspective, there are two cases. In some situations, agents with the same belief simply disagree on the optimal action. They will end up making different choices both when they learn and when they choose between actions. In some other situations, agents with the same belief agree on the action to take. One objective of this paper is to show that they still might end up making different learning decisions and taking different actions subsequently.

To focus on these second type of situations (see the next section for some examples), we assume that for any given belief, all types of agents have the same difference in expected utility between every pair of actions. This means not only that they have the same preferred action when confronted to the same evidence, but also that they have the same willingness to pay to make the decision. We will say that these different types of agents "for Identical Beliefs are Identical in Behavior and Utility Difference" (ibibud). The property is summarized as follows.

Definition *Agents are* IBIBUD *if and only if:*

$$u_i(\gamma) - u_i(\gamma') = u_{i'}(\gamma) - u_{i'}(\gamma') \qquad \forall \;\; i, i' \in I, \;\; \gamma, \gamma' \in \Gamma, \;\; p$$

which, in particular, implies that $\arg\max_{\gamma} u_i(\gamma) \equiv \arg\max_{\gamma} u_{i'}(\gamma)$ *for all* $i, i' \in I, \; p$.

Given our simplified two-action model, it is sufficient to restrict to two types of agents: action a has the highest variance in payoffs for type-1 agents and action b has the highest variance in payoffs for type-2 agents. Let $x_1 = h$ and $y_1 = l$ with $h > l$, then, it is sufficient to restrict to the case where $x_2 = l$ and $y_2 = h$. The IBIBUD property translates into:

$$u_i(a) - u_i(b) = (h + l)(2p - 1) \;\; \forall \; i \quad \Rightarrow \quad \gamma_i = a \;\text{ if }\; p > 1/2 \;\text{ and }\; \gamma_i = b \;\text{ if }\; p < 1/2 \;\; \forall \; i.$$

Figure 1 provides a graphical representation of these utilities.[5]

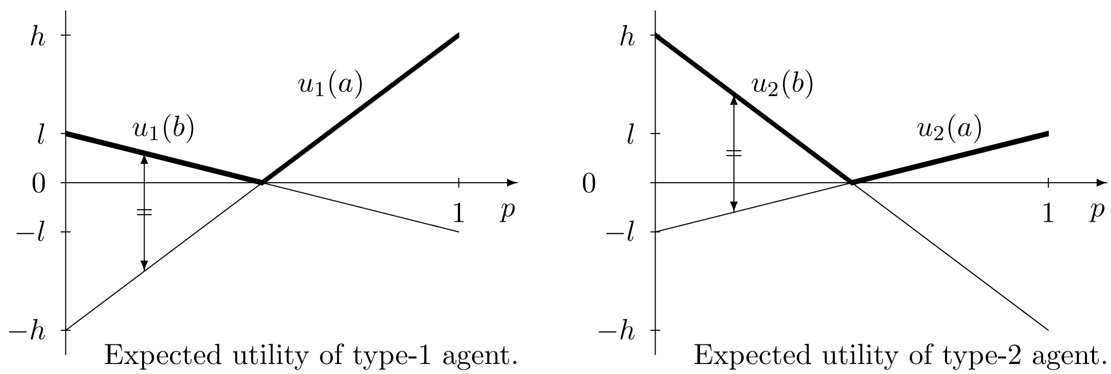


Figure 1. Utility representations for type-1 and type-2 agents.

2.4 Examples

In our theory, an individual must eventually take an irreversible decision that is ex post optimal only in one (ex ante unknown) state of the world. Information can be obtained before making a choice at a cost. We briefly review a series of situations in which those ingredients are present.

Entrepreneurial investments under uncertainty Our leading example is about choice between different risky investments. A firm must decide which investment strategy to follow, a or b: the development of a product based on current know-how or one that radically departs from it; an R&D strategy that builds on existing technology or one that requires the development of a new technology; an investment that consolidates the existing costumer base or one that expands to a different population; a product that exploits complementarities with the existing portfolio or one that opens a new niche for the firm. The ex ante unknown state of the economy, market conditions, and consumer preferences, A or B, determine which investment will be relatively more successful. Finally, a choice that involves diversification is intrinsically more risky for a firm than one that builds on core competence (existing vs. new knowledge, technology, customers or product). Because firms in the same market have different backgrounds, what is considered high risk for one firm may be low risk for another and vice versa; the difference across types (1 or 2) captures this heterogeneity.

Court judgements under civil law A judge must choose whether to release (action a) or convict (action b) an offender who is innocent (state A) with probability p and guilty (state B) with probability $1 - p$. The judge can acquire information about the culpability of the accused at the cost of delaying the sentence. Letting the prisoner free is the riskiest choice for a type-1 judge (payoff $u_1(a, \cdot) \in \{-h, h\}$) whereas convicting him is the riskiest choice for a type-2 judge (payoff $u_2(b, \cdot) \in \{-h, h\}$). However, for any belief p, the differential in utility between convicting and releasing the offender is the same for both judges (IBIBUD property). An alternative interpretation is that there is only one judge and i represents the type of offense (robbery, murder, etc.). For these different offenses, conviction and acquittal involve different objective risks.

Career choices under imperfect self knowledge An adolescent chooses whether to pursue a career in sports (a) or to continue his intellectual education (b). Success in sports depends largely on "talent" (physical strength, coordination, performance under pressure). States A and B denote respectively a person with high talent and low talent for sport relative to his talent for intellectual activities. Training and repeated exposure to the activity provides information at a cost. Indeed, each year of non-exclusive attention decreases the long-run expected return in either domain. Last, earnings have a higher variance in sports (h or $-h$) than in intellectual endeavors (l or $-l$). Thus, there is only one relevant type in this application.

This example can be relabeled as an individual who decides whether to become an entrepreneur and open his own business (the high risk activity) or accept a job as an employee in a firm (low risk activity). Entrepreneurial talent is most valuable in new business ventures whereas discipline and team spirit is most important when working in a firm.

3 Information acquisition and optimal decision-making

A first goal of our study is to determine how a type-i agent acquires information before making a decision (section 3.1). Another objective is to compare the behavior of individuals with apparently similar motivations, that is, individuals who satisfy the IBIBUD property (section 3.2). We want to determine whether they exhibit different patterns of information acquisition and, if so, why. We also want to analyze how these different sampling strategies affect posterior beliefs (which measure the ex post confidence in the state) and actions. Then, we want to find out which type of mistakes are eventually made: how often action a is undertaken under state B, and action b under state A. The next objective is to determine whether the preferences of agents can be inferred from choices, the only observable variable (section 3.3). We also study what happens when we consider a different cost of information acquisition (section 3.4). Finally, we discuss the importance of the main ingredients of the model (section 3.5).

3.1 Option value of waiting and optimal stopping rule

Given the information revelation structure presented in section 2.2, agents face a trade-off between delay and information. This trade-off has been analyzed in a related setting in the literature on investment under uncertainty (see e.g. Dixit and Pindyck (1994) for a summary). In these models however, time is continuous and there is only one risky action to take. Our model can thus be seen as an extension of this literature to the case where two risky options are available. In this new setting and conditional on making a choice now, the opportunity cost of taking one action is not fixed anymore. This, in turn, also affects the option value of waiting.

In order to find the optimal stopping rule, we first determine the value function V_i^t that a type-i agent maximizes at date t. It can be written as:

$$V_i^t(n)=\begin{cases}\max\Big\{x_i(2\mu(n)-1),\delta\Big[\nu(n)V_i^{t+1}(n+1)+(1-\nu(n))V_i^{t+1}(n-1)\Big]\Big\} & \text{if } \mu(n)\geqslant\frac{1}{2}\\ \max\Big\{y_i(1-2\mu(n)),\delta\Big[\nu(n)V_i^{t+1}(n+1)+(1-\nu(n))V_i^{t+1}(n-1)\Big]\Big\} & \text{if } \mu(n)<\frac{1}{2}\end{cases} \quad (2)$$

where $\nu(n)=\mu(n)\theta+(1-\mu(n))(1-\theta)$. In words, at date t and given a difference of signals n that implies a posterior $\mu(n)>1/2$, type-i agent chooses between taking action a with expected payoff $x_i\mu-x_i(1-\mu)$ or waiting. In the latter case, signal α (respectively β) is received with probability ν (respectively $1-\nu$) and the value function in the following period $t+1$ becomes $V_i^{t+1}(n+1)$ (respectively $V_i^{t+1}(n-1)$), discounted at the rate δ. For $\mu(n)<1/2$, the argument is the same, except that the optimal action if the agent does not wait is b with payoff $-y_i\mu+y_i(1-\mu)$. Given (2), we can determine the optimal strategy for each type. This technical result is key for the subsequent analysis.

Lemma 1 *For all $\delta<1$, there exist $(n_{i,t}^{*},n_{i,t}^{**})$ at each date t s.t.:*

$$\tau_{i,t}=b \;\; \textit{if} \;\; n\leqslant n_{i,t}^{*},\;\; \tau_{i,t}=a \;\; \textit{if} \;\; n\geqslant n_{i,t}^{**} \;\; \textit{and} \;\; \tau_{i,t}=w \;\; \textit{if} \;\; n\in(n_{i,t}^{*},n_{i,t}^{**}).$$

Besides, we have $\mu(n^*_{i,t}) < 1/2 < \mu(n^{**}_{i,t})$.

Proof. See Appendix A1. □

The idea is simple. Agents trade-off the costs of delaying their choice between actions a and b with the benefits of acquiring more accurate information. When $\mu(n) > 1/2$, waiting becomes more costly as n increases, because delaying the action one extra period reduces the expected payoff by an amount proportional to $2\mu(n) - 1$. Conversely, when $\mu(n) < 1/2$, waiting becomes more costly as n decreases, because delaying the action reduces the expected payoff by an amount proportional to $1 - 2\mu(n)$. In other words, at each date t, there are two cutoffs $\mu(n^{**}_{i,t}) > 1/2$ and $\mu(n^*_{i,t}) < 1/2$ for a type-i agent. When $\mu \geqslant \mu(n^{**}_{i,t})$, the individual is "reasonably confident" that the true state is A, and when $\mu \leqslant \mu(n^*_{i,t})$, he is "reasonably confident" that the true state is B. In either case, the marginal gain of improving the information about the true state is offset by the marginal cost of a reduction in the expected payoff due to the delay it implies. As a result, he strictly prefers to stop learning and take his optimal action. For intermediate beliefs, that is when $\mu(n) \in (\mu(n^*_{i,t}), \mu(n^{**}_{i,t}))$, a type-$i$ agent prefers to keep accumulating evidence.

3.2 Different decisions by agents with the same motivations

In this section, we want to compare the behavior of IBIBUD agents. We consider the two types we already introduced, that is, $(x_1 = h, y_1 = l)$ and $(x_2 = l, y_2 = h)$, with $h > l$. Action a has the highest variance in payoffs for type-1 agents, whereas action b has the highest variance in payoffs for type-2 agents. Our next result is the following.

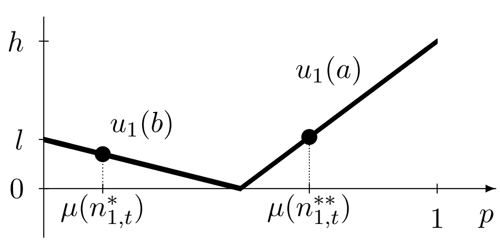
Proposition 1 *For all* $\delta < 1$ *and for all* t*, type-1 agents require less evidence in favor of* A *to take action* a *and more evidence in favor of* B *to take action* b *than type-2 agents. Formally,* $\mu(n^*_{1,t}) < \mu(n^*_{2,t}) < 1/2 < \mu(n^{**}_{1,t}) < \mu(n^{**}_{2,t})$.

Proof. See Appendix A1. □

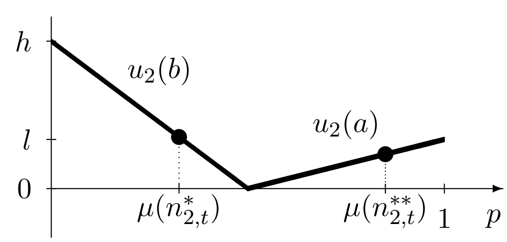
First of all, note that by the symmetry of types 1 and 2, $\mu(n^{**}_{1,t}) = 1 - \mu(n^{*}_{2,t})$ and $\mu(n^{*}_{1,t}) = 1 - \mu(n^{**}_{2,t})$. It immediately implies that:

$$\mu(n^{**}_{1,t}) - 1/2 < 1/2 - \mu(n^{*}_{1,t}) \qquad \text{and} \qquad \mu(n^{**}_{2,t}) - 1/2 > 1/2 - \mu(n^{*}_{2,t}).$$

These inequalities state that the confidence of a type-1 agent on the true state being A when he chooses to take action a is smaller than his confidence on the true state being B when he chooses to take action b. By symmetry, the opposite is true for a type-2 agent. Comparing the two agents, it means that a type-1 agent will need fewer evidence in favor of A in order to decide to stop collecting news and take action a and more evidence in favor of B in order to stop collecting news and take action b than a type-2 agent. The intuition for this result is simply that, given the delay associated with the accumulation of evidence, the marginal cost of learning is proportional to the agent's expected payoff of taking an action. Formally, for a type-1 individual, it is proportional to $h(1-\delta)$ when $\mu > 1/2$ (action a) and to $l(1-\delta)$ when $\mu < 1/2$ (action b). As a result and other things being equal, it is relatively less interesting to keep experimenting when the action currently optimal is a rather than b. The argument for a type-2 agent is symmetric. The shape of these cutoffs is graphically represented in Figure 2.



Stopping rule for a type-1 agent.



Stopping rule for a type-2 agent.

Figure 2. Stopping rules for type-1 and type-2 agents.

When $T \to +\infty$, then $n^*_{i,t} \to n^*_i$ and $n^{**}_{i,t} \to n^{**}_i$ for all t. Denote by $\Pr(\tau_i = \gamma_i \mid s)$ the probability that a type-i individual eventually undertakes action γ_i ($\in \{a, b\}$) when the true state is s ($\in \{A, B\}$). Also, let $\mu^{**}_i \equiv \mu(n^{**}_i)$ and $\mu^*_i \equiv \mu(n^*_i)$. Then, the posterior beliefs at the stopping rules are μ^*_1 and μ^{**}_1 for type-1 agents and μ^*_2 and μ^{**}_2 for type-2 agents. Given agents are symmetric, we have $\mu^{**}_2 = 1 - \mu^*_1$ and $\mu^*_2 = 1 - \mu^{**}_1$. To simplify notations, let $\mu^{**} \equiv \mu^{**}_1$ and $\mu^* \equiv \mu^*_1$ (then $\mu^{**}_2 = 1 - \mu^*$ and $\mu^*_2 = 1 - \mu^{**}$). Suppose that type-1 and type-2 agents start with the same prior belief $p \in (1 - \mu^{**}, \mu^{**})$. Each agent chooses the amount of information he collects before undertaking an action and the signals obtained by the agents are independent. Their optimal stopping rule is given by Lemma 1. We can compare the relative probabilities that each agent undertakes action a and action b.

Proposition 2 *For all $p \in (1 - \mu^{**}, \mu^{**})$, $\delta < 1$, $h > l > 0$ and when $T \to \infty$, type-1 agents take action a wrongly more often than type-2 agents. Similarly, type-1 agents take action b wrongly less often than type-2 agents. Moreover, as the difference in payoffs between actions increases, the difference in behavior between types 1 and 2 increases.*

Proof. The first part of Proposition 2 is a direct consequence of $\mu(n^{**}_2) > \mu(n^{**}_1)$ and $\mu(n^*_2) > \mu(n^*_1)$. These inequalities imply that $\Pr(\tau_1 = a \mid B) > \Pr(\tau_2 = a \mid B)$ and $\Pr(\tau_1 = b \mid A) < \Pr(\tau_2 = b \mid A)$; The second part results from the fact that, also by Lemma 1, $\frac{\partial n^*_1}{\partial h} < 0$, $\frac{\partial n^{**}_1}{\partial h} < 0$, $\frac{\partial n^*_1}{\partial l} > 0$, $\frac{\partial n^{**}_1}{\partial l} > 0$ and by symmetry $\frac{\partial n^*_2}{\partial h} > 0$, $\frac{\partial n^{**}_2}{\partial h} > 0$, $\frac{\partial n^*_2}{\partial l} < 0$, $\frac{\partial n^{**}_2}{\partial l} < 0$. Then $\frac{\partial \Pr(\tau_1 = a \mid s)}{\partial h} > 0 > \frac{\partial \Pr(\tau_2 = a \mid s)}{\partial h}$ and $\frac{\partial \Pr(\tau_1 = a \mid s)}{\partial l} < 0 < \frac{\partial \Pr(\gamma_2 = a \mid s)}{\partial l}$ for all s. These comparative statics fulfill the purpose of our analysis. However, for the reader interested, the analytical expressions of the probabilities $\Pr(\tau_i \mid s)$ are derived in Brocas and Carrillo (2007, Lemma 1) for an initial prior p and exogenous stopping posteriors μ^* and μ^{**}.[6] These are given by: $\Pr(\tau_1 = a \mid A) = \frac{p - \mu^*}{\mu^{**} - \mu^*} \frac{\mu^{**}}{p}$, $\Pr(\tau_1 = a \mid B) = \frac{p - \mu^*}{\mu^{**} - \mu^*} \frac{1 - \mu^{**}}{1 - p}$, $\Pr(\tau_2 = a \mid A) = \frac{p - (1 - \mu^{**})}{\mu^{**} - \mu^*} \frac{1 - \mu^*}{p}$, $\Pr(\tau_2 = a \mid B) = \frac{p - (1 - \mu^{**})}{\mu^{**} - \mu^*} \frac{\mu^*}{1 - p}$. □

Proposition 2 shows that, even if type-1 and type-2 agents are IBIBUD –and therefore have intrinsically the same motivations– they will make systematically different choices, at least in a stochastic sense. As shown in Lemma 1, a type-1 agent is relatively more likely to stop collecting news when the preliminary evidence points towards the optimality of action a than when it points towards the optimality of action b (i.e., when the first few signals are mainly α rather than β). Stated differently, the evidence in favor of A needed to induce a type-1 agent to take action a is smaller than the evidence in favor of B needed to induce him to take action b. The opposite is true for a type-2 agent. As a result, in equilibrium, a type-1 agent is more likely to take action a by mistake (i.e., when the true state is B) and less likely to take action b by mistake (i.e., when the true state is A) than a type-2 agent. Note that the endogenous choice to acquire information is crucial for this result: by definition of IBIBUD, the two types of agents would take action a with the same expected probability if the number of signals they receive were externally or exogenously imposed. Also, as the difference in the variance of payoffs $(h - l)$ increases, the likelihood that the two agents behave differently also increases: type-1 takes more often action a by mistake and less often action b by mistake whereas the opposite is true for type-2. Last, the fact that type-1 agents are less likely to take action b when the state is A automatically implies that they are more likely to take action a when the state is A. Thus, Proposition 2 can be best stated as "type-1 agents are more likely to take action a and less likely to take action b, *both rightly and wrongly*, than type-2 agents."

We now provide a simple numerical example to give an idea of the propensity of agents to make different types of mistakes. Consider the extreme situation in which $h > 0$ and $l \to 0$.[7] From the proof of Proposition 2, the probability that a type-i agent makes the wrong decision is:

$$\Pr(\tau_1 = a \mid B) = \frac{p}{1-p} \times \frac{1-\mu^{**}}{\mu^{**}} \quad \text{and} \quad \Pr(\tau_1 = b \mid A) \to 0$$

$$\Pr(\tau_2 = a \mid B) \to 0 \quad \text{and} \quad \Pr(\tau_2 = b \mid A) = \frac{1-p}{p} \times \frac{1-\mu^{**}}{\mu^{**}}$$

A type-1 agent will never take action b mistakenly, and a type-2 agent will never take action a mistakenly. Simple comparative statics about the likelihood of taking the wrong action given a prior probability p and a stopping posterior μ^{**} are illustrated in Figure 3.

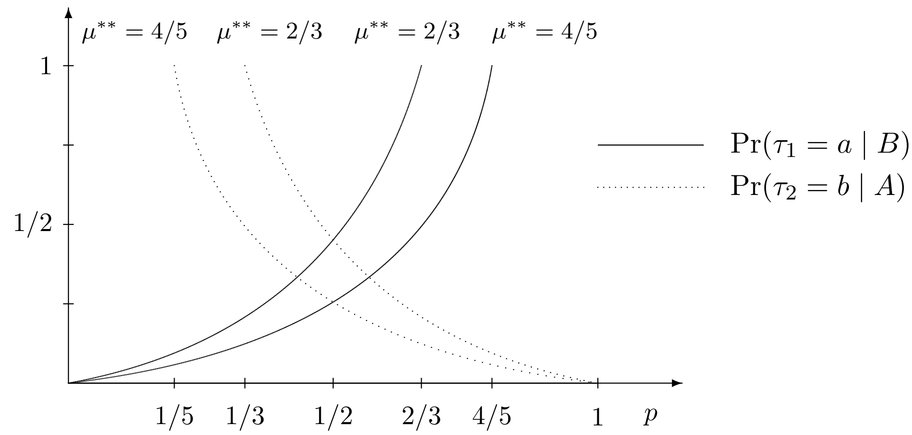


Figure 3. Frequency of mistakes by type-1 and type-2 agents.

Last, note that μ^{**} is increasing in δ, and $\lim_{\delta \to 1} \mu^{**} = 1$. As individuals become more patient, they acquire more information and make fewer mistakes. If they are infinitely patient, the cost of waiting vanishes. It then becomes optimal for both types to be (almost) perfectly informed before choosing any action, and there are (almost) no mistakes in equilibrium.

3.3 Revealed preferences

Suppose only choices are observable. Do choices convey any information about the preferences of agents? In principle, agents might end up making decisions for many different reasons and it might be difficult to identify a clear relationship between preferences and choices. Agents who often take action a might simply prefer that action. But, as our theory suggests, a tendency to favor a certain action can also arise in the absence of such strict preference. Overall, behavior is not a good indicator of preferences and a systematic tendency to behave in a certain way does not necessarily result from a bias in perceptions or preferences.

Our analysis suggests however that observed choices can be sometimes informative about the preferences of decision-makers. To be more precise, suppose that the individual starts with a prior belief $p = 1/2$ and his preferences are known up to the true type. We have the following result.

Proposition 3 *Agents' types can be partly inferred from (i) the decisions they reach; (ii) the delay in making decisions; and (iii) the frequency of their mistakes.*

First, we have shown in the previous section that the alternative that can potentially yield highest payoff (that is, the one with highest payoff-variance) will be adopted more often. In that case, it is possible to infer the preferences by observing the decisions of agents. Type-1 agents will take action a more often than type-2 agents. Conversely, type-2 agents will take action b more often than type-1 agents.

Second, given the optimal learning strategy and compared to type-2 agents, type-1 agents will reach more quickly the stopping rule commanding to take action a than the stopping rule commanding to take action b. In other words, the alternative that can potentially yield highest payoff will be adopted not only more often but also more rapidly. This positive relation between delay and type of decision can, in principle, be tested empirically.

Finally, if the state is observable ex post, it is possible to determine whether a mistake was made or not. The frequency of the mistakes can then be used to infer the type of the agent. In our case, a type-1 agent is more often wrong than a type-2 agent when he takes action a and less often wrong when he takes action b. Again, this prediction can be empirically tested.

3.4 Robustness

A crucial ingredient for the results presented so far is our specific way of modelling the cost of information acquisition. Because, delayed payoffs are discounted at a positive rate, the opportunity cost of waiting is greater the bigger the expected payoff of choosing an action. So, for example, starting from $p = 1/2$ it is more costly to wait after one signal in favor of the high-variance alternative than after one signal in favor of the low variance alternative. Formally, the opportunity cost of stopping the sampling process is $(1-\delta)\, h(2\mu(1)-1)$ in the former case and $(1-\delta)\, l(1-2\mu(-1)) \equiv (1-\delta)\, l(2\mu(1)-1)$ in the latter. Naturally, identical results would be obtained if, instead of a discount factor, we assumed that the possibility of acting vanishes between two dates with probability $p = 1-\delta$.

By contrast, the results *would not hold* if the only cost of sampling was a fixed per unit fee (and no delay). Formally, the value function of a type-i agent at date t, $\tilde{V}_i^t$, would be:

$$\tilde{V}_i^t(n) = \begin{cases} \max\Big\{x_i(2\mu(n)-1), \nu(n)\tilde{V}_i^{t+1}(n+1) + (1-\nu(n))\tilde{V}_i^{t+1}(n-1) - c\Big\} & \text{if } \mu(n) \geqslant \frac{1}{2} \\ \max\Big\{y_i(1-2\mu(n)), \nu(n)\tilde{V}_i^{t+1}(n+1) + (1-\nu(n))\tilde{V}_i^{t+1}(n-1) - c\Big\} & \text{if } \mu(n) < \frac{1}{2} \end{cases} \quad (3)$$

with c (> 0) denoting the cost per unit of sampling. Indeed, in Appendix A2 we show that under this alternative specification, the willingness to experiment is identical in the low and high variance alternatives, so it is also identical for a type-1 and a type-2 agent. The reason is simple. Sampling has a benefit and a cost. The benefit is the possibility of finding enough information in support of the currently unfavored alternative that would lead to a switch of action weighted by the incremental expected gain of implementing this action reversal. Because of the IBIBUD property, this incremental gain is identical for both types of agents. The cost is simply the amount to pay for extra information. With a fixed per unit fee c, this cost is also identical for both types of agents. If the cost and the benefit are the same, the optimal stopping rule is also the same.[8]

Since a fixed cost affects the total amount of sampling but not the relative propensity to experiment on each alternative, all our results survive when we combine the fixed cost described in this section with the opportunity cost developed in the main body of the paper. For most applications, the cost of experimentation is likely to be a combination of delay, probability of not being able to act in the future and per unit fee. The relative importance of each of them will depend on the specific case. For example, in the investment application acquiring information has two major costs: the project may become obsolete or may be undertaken by a rival entrepreneur, and the profits are delayed and therefore discounted. By contrast, for judicial decision-making the most important cost is the time, effort and money spent in the collection of each piece of evidence. Finally, the cost of delaying the choice between an intellectual and a sport career is a decrease in the probability of success in either domain.

3.5 Discussion

To conclude this section, we briefly discuss the importance of some other ingredients of the model and some possible interpretations of the results.

It should be clear by now that agents in our model *are not* fooled, deceived or misled. Contrary to the behavioral literature on optimism or overconfidence (see the references in the introduction), our agents have no cognitive limitations that would lead to systematic biases in their beliefs. Instead, they are rational; they accumulate and interpret signals in a bayesian way, and choose optimally given their information. Differences in choices between the different types of agents (and therefore in outcomes and in the type of mistakes incurred) are solely due to differences in their marginal incentives to learn about the state of the economy. In other words, in our paper the tendency to favor risky alternatives in entrepreneurial endeavors after a small amount of evidence is a profit-maximizing strategy: the high risk and low chances of success are recognized, but the opportunity cost of accumulating more evidence is too important. Technically, the point is very simple. The endogenous decision to acquire information does not affect the first-order moment of beliefs. That is, the average belief in the population always coincides with the true average. However, it may influence the higher-order moments. In particular, it can affect the *skewness* in the distribution of beliefs. Given a limited set of actions, two populations whose distribution of beliefs have the same average but different skewness will exhibit different aggregate behaviors.[9]

The model relies on irreversibility of actions or no learning after the decision is made. Irreversibility is quite natural in the judicial example, but either assumption can be too extreme in investment choices for example. Nevertheless, one should realize that partial irreversibility is enough to generate a short-run tendency to favor the riskiest alternative. Moreover, if the environment changes stochastically, information becomes obsolete over time, preventing the agent from learning the state with certainty. In that case, the willingness to favor risky choices will persist also in the long run, even under partial reversibility.

Geometrically, the utility of a type-2 agent is just a rotation of the utility of a type-1 agent (see Figure 1). It is then easy to see that the effect of payoff-variance in the delay and likelihood of taking certain alternatives will hold if, keeping IBIBUD, we increase the action space. From a theoretical viewpoint, it would be interesting to study a more general version of this two armed bandit problem, as it could provide novel insights about the relationship between the value of information and the "curvature" of the utility function.

4 Micro motives and macro consequences

We have argued in the previous sections that agents with the same motivations can end up making different choices, resulting in different types of mistakes. In many contexts, the decision might affect other agents in the economy, and those agents might be more or less sensitive to a given type of mistake. In the next subsections, we assess the mistakes from the perspective of third parties when externalities are present.

4.1 Preferences over IBIBUD agents

Note that, in our model, agents select a stopping rule that increases the probability of taking the action with highest payoff. The other side of the coin is that, with this strategy, agents are also increasing the probability of making the mistakes that are most costly. Because the types of mistakes incurred are systematically different, the parties involved will invariably have preferences over which type of agent they prefer to face.

To analyze this point in more detail, let us consider a third party with preferences summarized by the utility function $v(\gamma, s)$. Also, let $\Pr(A) \equiv p \in (\mu_i^*, \mu_i^{**})$ and assume that the third party does not pay the cost of learning ($\delta = 1$). Given the stopping rule used by a type-i agent, the expected utility of the third party is:

$$\begin{aligned}\hat{v}(\mu_i^*, \mu_i^{**}) \quad = \quad & p \Big[\Pr(\tau_1 = a \mid A)\, v(a, A) + \Pr(\tau_1 = b \mid A)\, v(b, A) \Big] \\ & + (1-p) \Big[\Pr(\tau_1 = a \mid B)\, v(a, B) + \Pr(\tau_1 = b \mid B)\, v(b, B) \Big]\end{aligned}$$

which simplifies as:

$$\hat{v}(\mu_i^*, \mu_i^{**}) = \frac{p - \mu_i^*}{\mu_i^{**} - \mu_i^*} \Big[\mu_i^{**} v(a, A) + (1 - \mu_i^{**}) v(a, B) \Big] + \frac{\mu_i^{**} - p}{\mu_i^{**} - \mu_i^*} \Big[\mu_i^* v(b, A) + (1 - \mu_i^*) v(b, B) \Big]$$

A simple inspection of this function yields the following result.

Proposition 4 *Different types of agents have different effects on the welfare of third parties. In particular, (i) an* IBIBUD *third party with preferences of the form $v(\cdot, \cdot) = u_1(\cdot, \cdot)$ or $v(\cdot, \cdot) = u_2(\cdot, \cdot)$ strictly prefers type-1 agents when $p < 1/2$ and type-2 agents when $p \geqslant 1/2$; and (ii) third parties with a strict preference for an action can appoint a priori unbiased agents to manipulate collective decision-making.*

Proof. We have $\frac{\partial v}{\partial \mu_i^*} \propto \mu_i^{**} \Big[v(b, A) - v(a, A) \Big] + (1 - \mu_i^{**}) \Big[v(b, B) - v(a, B) \Big]$. Similarly, $\frac{\partial v}{\partial \mu_i^{**}} \propto \mu_i^* \Big[v(b, A) - v(a, A) \Big] + (1 - \mu_i^*) \Big[v(b, B) - v(a, B) \Big]$. Then, if $v(b, A) < v(a, A)$ and $v(b, B) < v(a, B)$, both derivatives are negative. Now let $v(a, A) = x$, $v(a, B) = -x$,

$v(b,A) = -y$ and $v(b,B) = y$, then $\frac{\partial v}{\partial \mu_i^*} \propto (x+y)(1-2\mu_i^{**}) < 0$ and $\frac{\partial v}{\partial \mu_i^{**}} \propto (x+y)(1-2\mu_i^*) > 0$. Also,

$$\hat{v}(\mu_1^*, \mu_1^{**}) \equiv \hat{V}^1 = x(2\mu^{**}-1)\frac{p-\mu^*}{\mu^{**}-\mu^*} + y(1-2\mu^*)\frac{\mu^{**}-p}{\mu^{**}-\mu^*}$$

$$\hat{v}(\mu_2^*, \mu_2^{**}) \equiv \hat{V}^2 = x(1-2\mu^*)\frac{p-1+\mu^{**}}{\mu^{**}-\mu^*} + y(2\mu^{**}-1)\frac{1-\mu^*-p}{\mu^{**}-\mu^*}$$

and $\hat{V}^1 - \hat{V}^2 \propto (1-2p)[1-\mu^{**}-\mu^*]$. For all $p \in (1-\mu^{**}, \mu^{**})$, we have $1-\mu^{**}-\mu^* > 0$ and therefore $\hat{V}^1 - \hat{V}^2 \gtrless 0$ if $p \lesseqgtr 1/2$. □

The result is intuitive. Given that IBIBUD agents end up making different choices and therefore commit different types of mistakes, they affect third parties differently. Because they do not have to pay the cost of learning, all third parties who care about taking the correct action (a under A and b under B) want to learn the true state and therefore prefer an agent who acquires as much information as possible. If third parties are also affected by the delay, an interior stopping rule becomes optimal also from their perspective. However, even the extreme case where maximum information is optimal has an interesting property: all third parties who care about taking the correct action and would take the same decision as both types of agents for a given prior, have strict preferences over types. If the initial belief suggests to take action a ($p > 1/2$), they all want to delegate the decision to a type-2 agent. The reason is simply that they anticipate that a type-1 agent will stop with little evidence towards state A and therefore take action a "too often." The best chance to discover action a is incorrect is to appoint a type-2 agent who will continue learning until there is substantial evidence in favor of A. The same argument applies when the initial belief suggests b is optimal. The implications for the investment example described in section 2.4 are simple but interesting. For instance, consider a manager whose preferences are represented by the utility function $u_k(\gamma, s)$ with $k \in \{1,2\}$. Suppose that he must delegate both the information acquisition and the investment decision to one of his employees whose preferences are represented by $u_i(\gamma, s)$. If the manager must compensate the employee for the sampling process, then it is trivially optimal to select an individual with identical interests, $i = k$. This conclusion does not necessarily hold when the manager does not compensate the employee for the information cost nor suffers if there is a delay. In that case, the manager tries to maximize the information obtained before acting. This is achieved by selecting an employee who is reluctant to stop in the state favored by the prior. Summing up, a type-1 manager who initially believes that state A is more likely than B finds it optimal to appoint an employee who stops with little evidence towards A (that is, a type-1 agent) if the manager pays for the sampling cost. However, if the manager does not pay for it, then he prefers to appoint an individual who samples relatively more given the prior belief and therefore takes more often the optimal action (that is, a type-2 agent).

The second part of the proposition states that decisions can be manipulated if third parties can choose the agents' type. It is a direct consequence of Propositions 1 and 2. If a third party wants action a to be taken independently of the state, it will optimally delegate the decision to an agent who is most likely to take action a, that is, a type-1 agent. The implications are, again, immediate. A manager with a vested interest in one particular action can impose his preferences with relatively high probability and, at the same time, not be considered partisan: he simply needs to delegate the decision to an employee whose payoff variance is very high under the action preferred by the manager and very low under the other action.

4.2 Committees

Another natural question is whether aggregating the information that IBIBUD individuals can collect would alleviate mistakes. For instance, suppose that a welfare maximizing principal can ask several type-1 and type-2 agents their opinion about which action a or b should be taken. For simplicity, assume that agents care only about providing the correct appraisal (whether their suggestion is followed by the principal or not) and that their utility is captured with the functions $u_i(\gamma, s)$ described in section 2. This behavior is rational if, for example, appraisal and state are ex-post revealed and agents have career-concerns: their payoff is then a function of the quality of their suggestion, and not a function of the final action undertaken. In this setting, each agent's optimal rule for the acquisition of information coincides with the rule described in Lemma 1, so increasing the number of agents can only decrease the probability of an incorrect decision.[10] We assume that the number of agents is fixed but the principal can choose the proportion of type-1 and type-2 agents. Given that the two types of agents commit systematically different errors, we want to determine whether it is optimal to select all agents of the same type or to have appraisals from agents of both types. In other words, we are interested in studying the optimal composition of an advisory committee, and we ask the following question: is it better to be surrounded by individuals who tend to favor the same or opposite actions?

To address this issue, we consider the simplest version of our model. We denote by γ_i^j the recommendation made by the j^{th} type-i agent. We suppose that $l \to 0$, so that $\Pr(\gamma_1^j = b \mid A) = 0$ and $\Pr(\gamma_2^j = a \mid B) = 0$ for all j. The total number of agents is fixed and equal to n. The principal chooses x, the number of type-1 agents, $n - x$ being the number of type-2 agents. Last, in order to avoid any exogenous reason to prefer one type of agent over another, we assume that the principal's sole concern is to minimize the probability of a mistake, i.e., $v(a, A) = v(b, B) > v(a, B) = v(b, A)$. If we denote by $\gamma_P \in \{a, b\}$ the action taken eventually by the principal, we have the following result.

Proposition 5 *If* $p < 1/2$*, then* $x = n$*. The principal chooses* $\gamma_P = a$ *if* $\gamma_1^j = a \;\forall j$ *and* $\gamma_P = b$ *otherwise. Also,* $\Pr(\gamma_P = b \mid A) = 0$ *and* $\Pr(\gamma_P = a \mid B) = \left(\frac{p}{1-p} \times \frac{1-\mu^{**}}{\mu^{**}}\right)^n$.

If $p > 1/2$*, then* $x = 0$*. The principal chooses* $\gamma_P = b$ *if* $\gamma_2^j = b \;\forall j$ *and* $\gamma_P = a$ *otherwise. Also,* $\Pr(\gamma_P = b \mid A) = \left(\frac{1-p}{p} \times \frac{1-\mu^{**}}{\mu^{**}}\right)^n$ *and* $\Pr(\gamma_P = a \mid B) = 0$.

<u>Proof</u>. Fix x. Given $l \to 0$, we have $\Pr(\gamma_1 = b \mid A) = 0$ and $\Pr(\gamma_2 = a \mid B) = 0$, so the only possible error arises when all type-1 agents announce $\gamma_1^j = a$ ($j \in \{1, ..., x\}$) and all

type-2 agents announce $\gamma_2^k = b$ ($k \in \{1, ..., n-x\}$). The remaining question is whether, if this happens, the principal will take action a or action b.

Suppose that the principal minimizes costs with $\gamma_P = a$. The expected loss is then:

$$L_A(x) = \Pr(B) \cdot \prod_{j=1}^{x} \Pr(\gamma_1^j = a \mid B) \cdot \prod_{k=1}^{n-x} \Pr(\gamma_2^k = b \mid B) = (1-p)\left(\frac{p}{1-p} \times \frac{1-\mu^{**}}{\mu^{**}}\right)^x$$

So, conditional on taking $\gamma_P = a$, the principal optimally sets $x = n$, and the loss is:

$$L_A(n) = (1-p)\left(\frac{p}{1-p} \times \frac{1-\mu^{**}}{\mu^{**}}\right)^n \tag{4}$$

Suppose that the principal minimizes costs with $\gamma_P = b$. The expected loss is then:

$$L_B(x) = \Pr(A) \cdot \prod_{j=1}^{x} \Pr(\gamma_1^j = a \mid A) \cdot \prod_{k=1}^{n-x} \Pr(\gamma_2^k = b \mid A) = p\left(\frac{1-p}{p} \times \frac{1-\mu^{**}}{\mu^{**}}\right)^{n-x}$$

So, conditional on taking $\gamma_P = b$, the principal optimally sets $x = 0$, and the loss is:

$$L_B(0) = p\left(\frac{1-p}{p} \times \frac{1-\mu^{**}}{\mu^{**}}\right)^n \tag{5}$$

Last, from (4) and (5): $L_A(n) \lesseqgtr L_B(0) \Leftrightarrow (1-p)\left(\frac{p}{1-p}\right)^n \lesseqgtr p\left(\frac{1-p}{p}\right)^n \Leftrightarrow p \lesseqgtr 1/2$. □

Proposition 5 states that a principal who can choose the source of information will not select a combination of the two types of agents in order to compensate for the different type of errors they are likely to make. Instead, it will be optimal to choose all agents of the same type. As a result, the systematic tendency to favor one action over others is still present with a committee of advisors. The type of mistakes incurred will be identical in nature to the single agent case developed before, but quantitatively smaller due to the greater total amount of information collected. The idea is simple. Since the principal dislikes equally both types of errors, he selects agents so as to minimize their likelihood of committing a mistake, *independently of the nature*. We know from Proposition 2 that the likelihood of providing an incorrect appraisal is inversely proportional to the distance between the prior belief and the posterior at which the agent decides to stop collecting evidence and recommends an action (formally, $\mu^{**} - p$ for a type-1 agent and $p - (1 - \mu^{**})$

for a type-2 agent). Hence, if $p < 1/2$, type-1 agents are relatively less likely to mislead the principal than type-2 agents ($|\mu^{**} - p| > |p - (1 - \mu^{**})|$), so it is optimal to pick only type-1 agents. The opposite is true when $p > 1/2$. Overall, fewer mistakes occur as we increase the number of agents who provide an appraisal. However, the systematic tendency to favor one decision persists. Note that the result is based on the idea that, in order to minimize errors, the principal must encourage the acquisition of information. This is achieved by choosing agents with highest incentives to experiment given the initial prior. In that respect, the conclusion is similar to the one obtained in Proposition 4(i).

Again, the result has interesting implications for the examples presented before. Consider a manager who can appoint a committee of agents in charge of providing independent advice on which investment strategy to follow. Proposition 5 shows that in order to reduce the number of mistakes, all members of the committee should have the same tastes (i.e., similar preferences that result in similar tendencies).[11] Similarly, suppose that a judge has to form a jury and assume that, for a given belief, all members agree on whether the suspect should be convicted or released. The composition of the jury that minimizes mistakes will require all members to be of the same type, and therefore it will still exhibit a systematic tendency to favor one alternative. This, in turn, implies that impartial verdicts (that is, a verdict that errs on both sides with equal probability) are difficult to render even when all members want to minimize mistakes.

5 Concluding remarks

The paper has explored a general distinction between (irrational) systematically biased beliefs and (rational) systematically favored choices that result from the endogenous and costly decision to acquire information. We have pointed out as our major conclusion that actions with highest variance in payoffs across states will generally be favored, at the expense of actions with lowest variance in payoffs across states. In some applications (e.g., R&D strategies by different firms or career choices), the payoffs of the different alternatives are likely to be endogenously determined and inversely related to the fraction of agents who choose the same option. Adding this general equilibrium element and studying whether this possibility increases or decreases the tendency to favor certain alternatives is an interesting extension left for future work.

The conclusion can be of interest for the debate on rationality in decision-making. Consider an individual who chooses between opening a business and working in a firm. The paper argues that a rational individual will be satisfied with little information in favor of high entrepreneurial skills before deciding to open his own business. By contrast, he will need substantial evidence of high team spirit and little entrepreneurial ability in order to decide to work in a firm. As a result, we will observe many more low ability entrepreneurs who start businesses (and thus fail) than high ability ones who work for others. Since ability is not observable (only choices are), this asymmetry in choices and failures may incorrectly lead to the conclusion that a majority of individuals have "excessive" confidence in their entrepreneurial skills.

At the same time, it would be absurd to pretend that our explanation can account for all the evidence of overconfidence and optimism documented in psychology and behavioral economics. First, because the ingredients of our model are not relevant in all settings.[12] Second, because some aggregate beliefs are impossible to reconcile with statistical inference. And third, because the behavioral explanations reviewed in the introduction seem to do a good job in many situations. Yet, we feel that adding this extra element to the discussion can be very useful if we want to improve our understanding of the reasons and situations in which individuals distort their choices.

Notes

* We thank Joel Sobel, Jano Zabojnik, an editor, two anonymous referees and the audiences at various seminars for helpful comments.

[1]See e.g. DeBondt and Thaler (1995) for evidence of managerial optimism and Camerer and Lovallo (1999) for support of this hypothesis in a controlled laboratory environment. Studies also show that optimists can drive realists out of the market (Manove, 1999), that their presence may be socially desirable (Bernardo and Welch, 2001), and that optimistic beliefs can maximize felicity (Brunnermeier and Parker, 2005).

[2]A finite horizon game ensures the existence of a unique stopping rule at each period that can be computed by backward induction. By setting T arbitrarily large we can determine the limiting properties of this optimal stopping rule.

[3]It is equivalent to increase the correlation between signal and state or to increase the number of signals between two dates; both can be captured with the parameter θ.

[4]Given $\theta \in (1/2, 1)$, the following properties of $\mu(n)$ are immediate: (i) $\lim_{n \to -\infty} \mu(n) = 0$, (ii) $\lim_{n \to +\infty} \mu(n) = 1$, and (iii) $\mu(n+1) > \mu(n) \;\; \forall n$.

[5]Because some payoffs are negative, an individual with negative expected utility would prefer to delay the outcome. This counter-intuitive possibility does not arise in our model since, under the optimal action (a if $p > 1/2$ and b if $p < 1/2$), the expected payoff is always non-negative. In any case, all the results and proofs immediately extend if we add a constant k ($> h$) to all utilities, making every payoff positive (that is, $u_1(a, A) = k + h, u_1(a, B) = k - h, u_1(b, B) = k + l, u_1(b, A) = k - l$ and similarly for $u_2(\cdot)$).

[6]The paper uses related techniques to study a different issue. It analyzes a principal/agent model with incomplete contracting and determines the rents obtained by the former due to his ability to control the flow of public information.

[7]This means that $n_1^* \to -\infty$, $n_2^{**} \to +\infty$ and therefore $\mu^* \to 0$. The assumption is by no means necessary. However, it allows us to make clear-cut comparative statics with

only two parameters (p and μ^{**}).

[8]With asymmetric payoffs, one of the alternatives would have an exogenous advantage. In particular, the belief where the individual is indifferent between actions a and b would be $\hat{p} \neq 1/2$. The stopping rule with a cost per unit of sampling would still be symmetric, but only with respect to the belief $\hat{p}$.

[9]This point was first made by Carrillo and Mariotti (2000) in a model with hyperbolic discounting agents and a costless learning technology. It has been recently exploited by Benoît and Dubra (2007) in a different context.

[10]By contrast, if individuals were rewarded as a function of the quality of the final decision, then they would integrate the behavior of other agents in their choice to acquire information (and, possibly, free-ride). The optimal stopping rule would then be modified and it would not be always true that increasing the number of agents improves the quality of the final decision.

[11]The result however should not be overemphasized because the analysis neglects many important issues in the selection of committee members. For example, diversity may be optimal when different opinions in agents with common goals are due to different sources of information.

[12]Among other things, stakes have to be sufficiently small, otherwise the incentives of individuals to become perfectly informed before choosing their optimal action will crowd-out all other motivations (think for example of a patient deciding whether to learn from the doctor his health state concerning a curable disease). Also, incomplete information and costly learning have to be crucial elements at play.

References

1. Aghion, P., P. Bolton, C. Harris and B. Jullien, 1991, "Optimal Learning by Experimentation," *Review of Economic Studies*, 58(4), 621-654.

2. Benoît, J.P. and J. Dubra, 2007, "Overconfidence?" *mimeo*, London Business School.

3. Bernardo, A. and I. Welch, 2001, "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics and Management Strategy*, 10, 301-30.

4. Berry, D.A. and B. Fristedt, 1985, *Bandit Problems: Sequential Allocation of Experiments*, London, UK: Chapman and Hall.

5. Bolton, P. and C. Harris, 1999, "Strategic Experimentation," *Econometrica*, 67(2), 349-374.

6. Brocas, I. and J.D. Carrillo, 2007, "Influence through Ignorance," *RAND Journal of Economics*, 38(4), 931-947.

7. Brunnermeier, M. and J. Parker, 2005, "Optimal Expectations," *American Economic Review*, 95(4), 1092-1118.

8. Camerer, C. and D. Lovallo, 1999, "Overconfidence and Excess Entry: an Experimental Approach," *American Economic Review*, 89, 306-318.

9. Carrillo, J.D. and T. Mariotti, 2000, "Strategic Ignorance as a Self-Disciplining Device," *Review of Economic Studies*, 67(3), 529-544.

10. DeBondt, W.F. and R.H. Thaler, 1995, "Financial Decision-Making in Markets and Firms: a Behavioral Perspective," in R. Jarrow et al. eds., *Handbook in Operation Research and Management*, vol.9, ch.13. Amsterdam: Elsevier Science.

11. Dixit, A.K. and R.S. Pindyck, 1994, *Investment Under Uncertainty*, Princeton, NJ: Princeton University Press.

12. Keller, G. and S. Rady, 1999, "Optimal Experimentation in a Changing Environment," *Review of Economic Studies*, 66(3), 475-507.

13. Manove, M., 1999, "Entrepreneurs, Optimism, and the Competitive Edge," *mimeo*, Boston University.

14. Santos-Pinto, L. and J. Sobel, 2005, "A Model of Positive Self-Image in Subjective Assessments," *American Economic Review*, 95(5), 1386-1402.

15. Van den Steen, E., 2004, "Rational Overoptimism (and Other Biases)," *American Economic Review*, 94, 1141-1151.

16. Zabojnik, J., 2004, "A Model of Rational Bias in Self-Assessment," *Economic Theory*, 23, 259-282.

Appendix A1: Proof of Lemma 1 and Proposition 1

Type-i agent.

Date T. Denote $V_i^T(n) = \max\{x_i(2\mu(n)-1); y_i(1-2\mu(n))\}$ and let:

$$Y_i^t(n) = V_i^t(n) - x_i(2\mu(n)-1) \quad \text{and} \quad W_i^t(n) = V_i^t(n) - y_i(1-2\mu(n)).$$

For $t = T$, we have $Y_i^T(n) = \max\{0; (x_i + y_i)(1-2\mu(n))\}$ and $W_i^T(n) = \max\{0; (x_i + y_i)(2\mu(n)-1)\}$. Since $\mu(n)$ is increasing in n, $W_i^T(n)$ is non-decreasing and $Y_i^T(n)$ is non-increasing in n. Besides, $\lim_{n\to+\infty} \mu(n) = 1$ and $\lim_{n\to-\infty} \mu(n) = 0$, so there exists $\overline{n}$ defined by $\mu(\overline{n}) = 1/2$ such that for all $n > \overline{n}$ then $\tau_{i,T} = a$, and for all $n < \overline{n}$ then $\tau_{i,T} = b$.

Date $T-1$.
Case-1: $n \geqslant \overline{n}$. $V_i^{T-1}(n) = \max\{x_i(2\mu(n)-1); \delta\nu(n)V_i^T(n+1) + \delta(1-\nu(n))V_i^T(n-1)\}$ and

$$Y_i^{T-1}(n) = \max\{0, -(1-\delta)x_i(2\mu(n)-1) + \delta\nu(n)Y_i^T(n+1) + \delta(1-\nu(n))Y_i^T(n-1)\}$$

where $Y_i^{T-1}(n)$ is defined on $(\overline{n}, +\infty)$. Since $\nu(n)$ is increasing in n and $Y_i^T(n)$ is non-increasing in n, we can check that the right-hand side (r.h.s.) of $Y_i^{T-1}(n)$ is decreasing in n, and therefore there exists a cutoff $n_{i,T-1}^{**}$ such that for all $n > n_{i,T-1}^{**}$ then $\tau_{i,T-1} = a$, and for all $n \in [\overline{n}, n_{i,T-1}^{**})$ then $\tau_{i,T-1} = w$. To solve the previous equation, the cutoff has to be such that $n_{i,T-1}^{**} + 1 \geqslant \overline{n}$ and $n_{i,T-1}^{**} - 1 < \overline{n}$, and therefore it is the solution of:

$$0 = x_i \cdot f(n_{i,T-1}^{**}, \delta) - y_i \cdot g(n_{i,T-1}^{**}, \delta)$$

where $f(n_{i,T-1}^{**}, \delta) \equiv 2\mu(n_{i,T-1}^{**}) - 1 - \delta\nu(n_{i,T-1}^{**})(2\mu(n_{i,T-1}^{**}+1)-1)$ and $g(n_{i,T-1}^{**}, \delta) = \delta(1-\nu(n_{i,T-1}^{**}))(1-2\mu(n_{i,T-1}^{**}-1))$. Differentiating with respect to x_i, y_i and δ we have:[13]

$$\begin{aligned}
\frac{\partial n_{i,T-1}^{**}}{\partial x_i}\left[y_i \cdot g_n(n_{i,T-1}^{**}, \delta) - x_i \cdot f_n(n_{i,T-1}^{**}, \delta)\right] &= f(n_{i,T-1}^{**}, \delta) \\
\frac{\partial n_{i,T-1}^{**}}{\partial y_i}\left[x_i \cdot f_n(n_{i,T-1}^{**}, \delta) - y_i \cdot g_n(n_{i,T-1}^{**}, \delta)\right] &= g(n_{i,T-1}^{**}, \delta) \\
\frac{\partial n_{i,T-1}^{**}}{\partial \delta}\left[y_i \cdot g_n(n_{i,T-1}^{**}, \delta) - x_i \cdot f_n(n_{i,T-1}^{**}, \delta)\right] &= x_i \cdot f_\delta(n_{i,T-1}^{**}, \delta) - y_i \cdot g_\delta(n_{i,T-1}^{**}, \delta)
\end{aligned}$$

Given $f(n^{**}_{i,T-1},\delta) > 0$, $g(n^{**}_{i,T-1},\delta) > 0$,[14] $y_i \cdot g_n(n^{**}_{i,T-1},\delta) - x_i \cdot f_n(n^{**}_{i,T-1},\delta) < 0$, $x_i \cdot f_\delta(n^{**}_{i,T-1},\delta) - y_i \cdot g_\delta(n^{**}_{i,T-1},\delta) < 0$, we finally have:

$$\frac{\partial n^{**}_{i,T-1}}{\partial x_i} < 0, \quad \frac{\partial n^{**}_{i,T-1}}{\partial y_i} > 0, \quad \frac{\partial n^{**}_{i,T-1}}{\partial \delta} > 0.$$

Case-2: $n \leqslant \overline{n}$. $V_i^{T-1}(n) = \max\{y_i(1-2\mu(n)); \delta\nu(n)V_i^T(n+1) + \delta(1-\nu(n))V_i^T(n-1)\}$ and

$$W_i^{T-1}(n) = \max\{0, -(1-\delta)y_i(1-2\mu(n)) + \delta\nu(n)W_i^T(n+1) + \delta(1-\nu(n))W_i^T(n-1)\}$$

where $W_i^{T-1}(n)$ is defined on $(-\infty, \overline{n})$. Since $\nu(n)$ is increasing in n and $W_i^T(n)$ is non-decreasing in n, we can check that the r.h.s. of $W_i^{T-1}(n)$ is increasing in n, and therefore there exists a cutoff $n^*_{i,T-1}$ such that for all $n \in (n^*_{i,T-1}, \overline{n}]$ then $\tau_{i,T-1} = w$, and for all $n < n^*_{i,T-1}$ then $\tau_{i,T-1} = b$. This cutoff has to be such that $n^*_{i,T-1} + 1 > \overline{n}$ and $n^*_{i,T-1} - 1 \leq \overline{n}$, so it is solution of:

$$0 = y_i \cdot r(n^*_{i,T-1},\delta) - x_i \cdot s(n^*_{i,T-1},\delta)$$

where $r(n^*_{i,T-1},\delta) = 1-2\mu(n^*_{i,T-1}) - \delta(1-\nu(n^*_{i,T-1}))(1-2\mu(n^*_{i,T-1}-1))$ and $s(n^*_{i,T-1},\delta) = \delta\nu(n^*_{i,T-1})(2\mu(n^*_{i,T-1}+1)-1)$. Again, differentiating with respect to x_i, y_i and δ we have:

$$\begin{aligned}
\frac{\partial n^*_{i,T-1}}{\partial x_i}\left[y_i \cdot r_n(n^*_{i,T-1},\delta) - x_i \cdot s_n(n^*_{i,T-1},\delta)\right] &= s(n^*_{i,T-1},\delta) \\
\frac{\partial n^*_{i,T-1}}{\partial y_i}\left[x_i \cdot s_n(n^*_{i,T-1},\delta) - y_i \cdot r_n(n^*_{i,T-1},\delta)\right] &= r(n^*_{i,T-1},\delta) \\
\frac{\partial n^*_{i,T-1}}{\partial \delta}\left[y_i \cdot r_n(n^*_{i,T-1},\delta) - x_i \cdot s_n(n^*_{i,T-1},\delta)\right] &= x_i \cdot s_\delta(n^*_{i,T-1},\delta) - y_i \cdot r_\delta(n^*_{i,T-1},\delta)
\end{aligned}$$

Given $s(n^*_{i,T-1},\delta) > 0$, $r(n^*_{i,T-1},\delta) > 0$, $y_i \cdot r_n(n^*_{i,T-1},\delta) - x_i \cdot s_n(n^*_{i,T-1},\delta) < 0$, $x_i \cdot s_\delta(n^*_{i,T-1},\delta) - y_i \cdot r_\delta(n^*_{i,T-1},\delta) > 0$, we finally have:

$$\frac{\partial n^*_{i,T-1}}{\partial x_i} < 0, \quad \frac{\partial n^*_{i,T-1}}{\partial y_i} > 0, \quad \frac{\partial n^*_{i,T-1}}{\partial \delta} < 0.$$

The proof is completed using a simple recursive method.[15]

Case-1: $n \geqslant \overline{n}$. $V_i^{t-1}(n) = \max\{x_i(2\mu(n)-1); \delta\nu(n)V_i^t(n+1) + \delta(1-\nu(n))V_i^t(n-1)\}$ and

$$\begin{aligned}
Y_i^t(n) &= \max\{0, -(1-\delta)x_i(2\mu(n)-1) + \delta\nu(n)Y_i^{t+1}(n+1) + \delta(1-\nu(n))Y_i^{t+1}(n-1)\} \\
Y_i^{t-1}(n) &= \max\{0, -(1-\delta)x_i(2\mu(n)-1) + \delta\nu(n)Y_i^t(n+1) + \delta(1-\nu(n))Y_i^t(n-1)\}
\end{aligned}$$

Suppose that the following assumptions **(A1)-(A5)** hold.

(A1): $Y_i^t(n)$ is non-increasing in n and there exists $n_{i,t}^{**}$ such that $\tau_{i,t} = a$ if $n > n_{i,t}^{**}$ and $\tau_{i,t} = w$ if $n \in [\overline{n}, n_{i,t}^{**})$.
(A2): $Y_i^t(n) \geqslant Y_i^{t+1}(n)$ and therefore $n_{i,t}^{**} > n_{i,t+1}^{**}$.
(A3): $Y_i^t(n, x_i) \leqslant Y_i^t(n, x_i')$ if $x_i > x_i'$ (and therefore $\partial n_{i,t}^{**}/\partial x_i < 0$).
(A4): $Y_i^t(n, y_i) \geqslant Y_i^t(n, y_i')$ if $y_i > y_i'$ (and therefore $\partial n_{i,t}^{**}/\partial y_i > 0$).
(A5): $Y_i^t(n, \delta) \geqslant Y_i^t(n, \delta')$ if $\delta > \delta'$ (and therefore $\partial n_{i,t}^{**}/\partial \delta > 0$).

Given **(A1)**, the r.h.s. of $Y_i^{t-1}(n)$ is decreasing in n, so $Y_i^{t-1}(n)$ is non-increasing in n. Therefore, there exists a unique cutoff $n_{i,t-1}^{**}$ such that for all $n > n_{i,t-1}^{**}$ then $\tau_{i,t-1} = a$, and for all $n \in [\overline{n}, n_{i,t-1}^{**})$ then $\tau_{i,t-1} = w$. Also, given **(A2)**, the r.h.s. of $Y_i^{t-1}(n)$ is greater or equal than the r.h.s. of $Y_i^t(n)$ and therefore $Y_i^{t-1}(n) \geqslant Y_i^t(n)$. Overall, both **(A1)** and **(A2)** hold at date $t-1$. Furthermore, $n_{i,t-1}^{**} > n_{i,t}^{**}$. Now, denote:

$$Y_i^{t-1}(n, x_i) = \max\{0, -(1-\delta)x_i(2\mu(n)-1) + \delta\nu(n)Y_i^t(n+1, x_i) + \delta(1-\nu(n))Y_i^t(n-1, x_i)\}$$

$$Y_i^{t-1}(n, x_i') = \max\{0, -(1-\delta)x_i'(2\mu(n)-1) + \delta\nu(n)Y_i^t(n+1, x_i') + \delta(1-\nu(n))Y_i^t(n-1, x_i')\}$$

By **(A3)**, if $x_i > x_i'$ then $Y_i^t(n+1, x_i) \leqslant Y_i^t(n+1, x_i')$ and $Y_i^t(n-1, h) \leqslant Y_i^t(n-1, x_i')$. Therefore, $Y_i^{t-1}(n, h) \leqslant Y_i^{t-1}(n, x_i')$. This means that **(A3)** holds at date $t-1$ and, as a consequence, that $\partial n_{i,t-1}^{**}/\partial x_i < 0$. Using a similar reasoning, it is immediate that **(A4)** and **(A5)** also hold at $t-1$ and therefore that $\partial n_{i,t-1}^{**}/\partial y_i > 0$ and $\partial n_{i,t-1}^{**}/\partial \delta > 0$.

Case-2: $n \leqslant \overline{n}$. $V_i^{t-1}(n) = \max\{y_i(1-2\mu(n)); \delta\nu(n)V_i^t(n+1) + \delta(1-\nu(n))V_i^t(n-1)\}$ and

$$W_i^t(n) = \max\{0, -(1-\delta)y_i(1-2\mu(n)) + \delta\nu(n)W_i^{t+1}(n+1) + \delta(1-\nu(n))W_i^{t+1}(n-1)\}$$
$$W_i^{t-1}(n) = \max\{0, -(1-\delta)y_i(1-2\mu(n)) + \delta\nu(n)W_i^t(n+1) + \delta(1-\nu(n))W_i^t(n-1)\}$$

Suppose that the following assumptions **(A1')-(A5')** hold.

(A1'): $W_i^t(n)$ is non-decreasing in n and there exists $n_{i,t}^*$ such that $\tau_{i,t} = b$ if $n < n_{i,t}^*$ and $\tau_{i,t} = w$ if $n \in (n_{i,t}^*, \overline{n}]$.
(A2'): $W_i^t(n) \geqslant W_i^{t+1}(n)$ and therefore $n_{i,t}^* < n_{i,t+1}^*$.
(A3'): $W_i^t(n, x_i) \leqslant W_i^t(n, x_i')$ if $x_i > x_i'$ (and therefore $\partial n_{i,t}^*/\partial x_i < 0$).
(A4'): $W_i^t(n, y_i) \geqslant W_i^t(n, y_i')$ if $y_i > y_i'$ (and therefore $\partial n_{i,t}^*/\partial y_i > 0$).

(A5'): $W_i^t(n,\delta) \geqslant W_i^t(n,\delta')$ if $\delta > \delta'$ (and therefore $\partial n_{i,t}^*/\partial\delta < 0$).

Given **(A1')**, the r.h.s. of $W_i^{t-1}(n)$ is increasing in n, so $W_i^{t-1}(n)$ is non-decreasing in n. Therefore, there exists a unique cutoff $n_{i,t-1}^*$ such that for all $n < n_{i,t-1}^*$ then $\tau_{i,t-1} = b$, and for all $n \in (n_{i,t-1}^*, \overline{n}]$ then $\tau_{i,t-1} = w$. Also, given **(A2')**, the r.h.s. of $W_i^{t-1}(n)$ is greater or equal than the r.h.s. of $W_i^t(n)$ and therefore $W_i^{t-1}(n) \geqslant W_i^t(n)$. Overall, both **(A1')** and **(A2')** hold at date $t-1$. Furthermore, $n_{i,t-1}^* < n_{i,t}^*$. Now, denote:

$$W_i^{t-1}(n,x_i) = \max\{0, -(1-\delta)y_i(1-2\mu(n)) + \delta\nu(n)W_i^t(n+1,x_i) + \delta(1-\nu(n))W_i^t(n-1,x_i)\}$$

$$W_i^{t-1}(n,x_i') = \max\{0, -(1-\delta)y_i(1-2\mu(n)) + \delta\nu(n)W_i^t(n+1,x_i') + \delta(1-\nu(n))W_i^t(n-1,x_i')\}$$

By **(A3')**, if $x_i > x_i'$ then $W_i^t(n+1,x_i) \leqslant W_i^t(n+1,x_i')$ and $W_i^t(n-1,x_i) \leqslant W_i^t(n-1,x_i')$. Therefore, $W_i^{t-1}(n,x_i) \leqslant W_i^{t-1}(n,x_i')$. This means that **(A3')** holds at date $t-1$ and, as a consequence, that $\partial n_{i,t-1}^*/\partial x_i < 0$. Using a similar reasoning, it is immediate that **(A4')** and **(A5')** also hold at $t-1$ and therefore that $\partial n_{i,t-1}^*/\partial y_i > 0$ and $\partial n_{i,t-1}^*/\partial\delta < 0$.

Type-1 and Type-2 agents.

Type-1 and type-2 agents are fully symmetric. At date t, there exists $n_{1,t}^{**}$ s.t. $\tau_{1,t} = a$ if $n > n_{1,t}^{**}$ and $\tau_{1,t} = w$ if $n \in [\overline{n}, n_{1,t}^{**})$. There also exists $n_{2,t}^*$ s.t. $\tau_{2,t} = b$ if $n < n_{2,t}^*$ and $\tau_{2,t} = w$ if $n \in (n_{2,t}^*, \overline{n}]$. Furthermore, by symmetry, $n_{2,t}^*$ is such that $\overline{n} - n_{2,t}^* = n_{1,t}^{**} - \overline{n}$, that is $\mu(n_{1,t}^{**}) = 1 - \mu(n_{2,t}^*)$. Similarly, if at date t there exists $n_{1,t}^*$ s.t. $\tau_{1,t} = b$ if $n < n_{1,t}^*$ and $\tau_{1,t} = w$ if $n \in (n_{1,t}^*, \overline{n}]$, then there also exists $n_{2,t}^{**}$ s.t. $\tau_{2,t} = a$ if $n > n_{2,t}^{**}$ and $\tau_{2,t} = w$ if $n \in [\overline{n}, n_{2,t}^*)$. Furthermore, $n_{2,t}^{**}$ is such that $n_{2,t}^{**} - \overline{n} = \overline{n} - n_{1,t}^*$, that is $\mu(n_{1,t}^*) = 1 - \mu(n_{2,t}^{**})$.

Note that if $h = l$, then for all t we have $\mu(n_{1,t}^*) = 1 - \mu(n_{1,t}^{**})$ and $\mu(n_{2,t}^*) = 1 - \mu(n_{2,t}^{**})$. As a result, $n_{2,t}^* = n_{1,t}^* < \overline{n}$ and $n_{2,t}^{**} = n_{1,t}^{**} > \overline{n}$. Also, we know that $\frac{\partial n_{1,t}^{**}}{\partial h} < 0$ and $\frac{\partial n_{1,t}^*}{\partial h} < 0$ (which, again by symmetry, implies that $\frac{\partial n_{2,t}^*}{\partial h} > 0$ and $\frac{\partial n_{2,t}^{**}}{\partial h} > 0$). Therefore, for all $h > l$ we have $n_{1,t}^* < n_{2,t}^* < \overline{n} < n_{1,t}^{**} < n_{2,t}^{**}$.

Summing up, when $\delta < 1$, $h > l > 0$ and $T \to +\infty$, we have $n_1^* < n_2^* < \overline{n} < n_1^{**} < n_2^{**}$ where $\mu(n_1^{**}) = 1 - \mu(n_2^*)$ and $\mu(n_1^*) = 1 - \mu(n_2^{**})$. Moreover, $\frac{\partial n_1^{**}}{\partial h} < 0$, $\frac{\partial n_1^{**}}{\partial l} > 0$, $\frac{\partial n_1^{**}}{\partial \delta} > 0$, $\frac{\partial n_1^*}{\partial h} < 0$, $\frac{\partial n_1^*}{\partial l} > 0$, $\frac{\partial n_1^*}{\partial \delta} < 0$ and $\frac{\partial n_2^*}{\partial h} > 0$, $\frac{\partial n_2^*}{\partial l} < 0$, $\frac{\partial n_2^*}{\partial \delta} < 0$, $\frac{\partial n_2^{**}}{\partial h} > 0$, $\frac{\partial n_2^{**}}{\partial l} < 0$, $\frac{\partial n_2^{**}}{\partial \delta} > 0$.

Appendix A2: Fixed per-unit cost of sampling

The decision at date T is the same as in Appendix A1. The rest of the proof follows similar steps as in Appendix A1. We present only a sketch. At date $T-1$, there are two cases.

Case-1: $n \geqslant \overline{n}$. $\tilde{V}_i^{T-1}(n) = \max\{x_i(2\mu(n)-1); \nu(n)\tilde{V}_i^T(n+1) + (1-\nu(n))\tilde{V}_i^T(n-1) - c\}$ and

$$\tilde{Y}_i^{T-1}(n) = \max\{0, \nu(n)\tilde{Y}_i^T(n+1) + (1-\nu(n))\tilde{Y}_i^T(n-1) - c\}$$

where $\tilde{Y}_i^{T-1}(n)$ is defined on $(\overline{n}, +\infty)$. The right-hand side (r.h.s.) of $\tilde{Y}_i^{T-1}(n)$ is decreasing in n, and therefore there exists a cutoff $\tilde{n}_{i,T-1}^{**}$ such that for all $n > \tilde{n}_{i,T-1}^{**}$ then $\tau_{i,T-1} = a$, and for all $n \in [\overline{n}, \tilde{n}_{i,T-1}^{**})$ then $\tau_{i,T-1} = w$. The cutoff is the solution of:

$$c = [x_i + y_i] \cdot \tilde{g}(\tilde{n}_{i,T-1}^{**})$$

where $\tilde{g}(\tilde{n}_{i,T-1}^{**}) = (1-\nu(\tilde{n}_{i,T-1}^{**}))(1-2\mu(\tilde{n}_{i,T-1}^{**}-1))$. Note that $\tilde{g}(n)$ is decreasing in n for all n. Differentiating with respect to x_i, y_i and c we have:

$$\frac{\partial \tilde{n}_{i,T-1}^{**}}{\partial x_i} > 0, \quad \frac{\partial \tilde{n}_{i,T-1}^{**}}{\partial y_i} > 0, \quad \frac{\partial \tilde{n}_{i,T-1}^{**}}{\partial c} < 0.$$

Suppose $x_1 = h$, $y_1 = l$, $x_2 = l$ and $y_2 = h$, then $\frac{\partial \tilde{n}_{1,T-1}^{**}}{\partial h} = \frac{\partial \tilde{n}_{2,T-1}^{**}}{\partial h}$. Then, for all $h > l > 0$, we have $\tilde{n}_{1,T-1}^{**} = \tilde{n}_{2,T-1}^{**}$.

Case-2: $n \leqslant \overline{n}$. $\tilde{V}_i^{T-1}(n) = \max\{y_i(1-2\mu(n)); \nu(n)\tilde{V}_i^T(n+1) + (1-\nu(n))\tilde{V}_i^T(n-1) - c\}$ and

$$\tilde{W}_i^{T-1}(n) = \max\{0, \nu(n)\tilde{W}_i^T(n+1) + (1-\nu(n))\tilde{W}_i^T(n-1) - c\}$$

where $\tilde{W}_i^{T-1}(n)$ is defined on $(-\infty, \overline{n})$. The r.h.s. of $\tilde{W}_i^{T-1}(n)$ is increasing in n, and therefore there exists a cutoff $\tilde{n}_{i,T-1}^{*}$ such that for all $n \in (\tilde{n}_{i,T-1}^{*}, \overline{n}]$ then $\tau_{i,T-1} = w$, and for all $n < \tilde{n}_{i,T-1}^{*}$ then $\tau_{i,T-1} = b$. The cutoff is solution of:

$$c = [x_i + y_i] \cdot \tilde{s}(\tilde{n}_{i,T-1}^{*}, \delta)$$

where $\tilde{s}(\tilde{n}_{i,T-1}^{*}) = \nu(\tilde{n}_{i,T-1}^{*})(2\mu(\tilde{n}_{i,T-1}^{*}+1)-1)$. Again, differentiating with respect to x_i, y_i and c we have:

$$\frac{\partial \tilde{n}_{i,T-1}^{*}}{\partial x_i} < 0, \quad \frac{\partial \tilde{n}_{i,T-1}^{*}}{\partial y_i} < 0, \quad \frac{\partial \tilde{n}_{i,T-1}^{*}}{\partial c} > 0.$$

Suppose $x_1 = h$, $y_1 = l$, $x_2 = l$ and $y_2 = h$, then $\frac{\partial \tilde{n}^*_{1,T-1}}{\partial h} = \frac{\partial \tilde{n}^*_{2,T-1}}{\partial h}$. Then, for all $h > l > 0$, we have $\tilde{n}^*_{1,T-1} = \tilde{n}^*_{2,T-1}$.

The proof is completed using a similar recursive method as in Appendix A1.

Case-1: $n \geqslant \overline{n}$. $\tilde{V}_i^{t-1}(n) = \max\{x_i(2\mu(n) - 1); \nu(n)\tilde{V}_i^t(n+1) + (1-\nu(n))\tilde{V}_i^t(n-1) - c\}$ and

$$\begin{aligned}
\tilde{Y}_i^t(n) &= \max\{0, \nu(n)\tilde{Y}_i^{t+1}(n+1) + (1-\nu(n))\tilde{Y}_i^{t+1}(n-1) - c\} \\
\tilde{Y}_i^{t-1}(n) &= \max\{0, \nu(n)\tilde{Y}_i^{t}(n+1) + (1-\nu(n))\tilde{Y}_i^{t}(n-1) - c\}
\end{aligned}$$

Suppose that the following assumptions **(A1)-(A6)** hold.

(A1): $\tilde{Y}_i^t(n)$ is non-increasing in n and there exists $\tilde{n}^{**}_{i,t}$ such that $\tau_{i,t} = a$ if $n > \tilde{n}^{**}_{i,t}$ and $\tau_{i,t} = w$ if $n \in [\overline{n}, \tilde{n}^{**}_{i,t})$.
(A2): $\tilde{Y}_i^t(n) \geqslant \tilde{Y}_i^{t+1}(n)$ and therefore $\tilde{n}^{**}_{i,t} > \tilde{n}^{**}_{i,t+1}$.
(A3): $\tilde{Y}_i^t(n, x_i) \geqslant \tilde{Y}_i^t(n, x_i')$ if $x_i > x_i'$ (and therefore $\partial \tilde{n}^{**}_{i,t}/\partial x_i > 0$).
(A4): $\tilde{Y}_i^t(n, y_i) \geqslant \tilde{Y}_i^t(n, y_i')$ if $y_i > y_i'$ (and therefore $\partial \tilde{n}^{**}_{i,t}/\partial y_i > 0$).
(A5): $\tilde{Y}_i^t(n, c) \leqslant \tilde{Y}_i^t(n, c')$ if $c > c'$ (and therefore $\partial \tilde{n}^{**}_{i,t}/\partial c < 0$).
(A6): $\tilde{Y}_1^t(n) = \tilde{Y}_2^t(n)$ when $x_1 = h$, $y_1 = l$, $x_2 = l$ and $y_2 = h$ (and therefore $\tilde{n}^{**}_{1,t} = \tilde{n}^{**}_{2,t}$).

Given **(A1)**, the r.h.s. of $\tilde{Y}_i^{t-1}(n)$ is decreasing in n, so $\tilde{Y}_i^{t-1}(n)$ is non-increasing in n. Therefore, there exists a unique cutoff $\tilde{n}^{**}_{i,t-1}$ such that for all $n > \tilde{n}^{**}_{i,t-1}$ then $\tau_{i,t-1} = a$, and for all $n \in [\overline{n}, \tilde{n}^{**}_{i,t-1})$ then $\tau_{i,t-1} = w$. Also, given **(A2)**, the r.h.s. of $Y_i^{t-1}(n)$ is greater or equal than the r.h.s. of $\tilde{Y}_i^t(n)$ and therefore $\tilde{Y}_i^{t-1}(n) \geqslant \tilde{Y}_i^t(n)$. Overall, both **(A1)** and **(A2)** hold at date $t-1$. Furthermore, $\tilde{n}^{**}_{i,t-1} > \tilde{n}^{**}_{i,t}$. Now, denote:

$$\tilde{Y}_i^{t-1}(n, x_i) = \max\{0, \nu(n)\tilde{Y}_i^t(n+1, x_i) + (1-\nu(n))\tilde{Y}_i^t(n-1, x_i) - c\}$$

$$\tilde{Y}_i^{t-1}(n, x_i') = \max\{0, \nu(n)\tilde{Y}_i^t(n+1, x_i') + (1-\nu(n))\tilde{Y}_i^t(n-1, x_i') - c\}$$

By **(A3)**, if $x_i > x_i'$ then $\tilde{Y}_i^t(n+1, x_i) \geqslant \tilde{Y}_i^t(n+1, x_i')$ and $\tilde{Y}_i^t(n-1, h) \geqslant \tilde{Y}_i^t(n-1, x_i')$. Therefore, $\tilde{Y}_i^{t-1}(n, h) \geqslant \tilde{Y}_i^{t-1}(n, x_i')$. This means that **(A3)** holds at date $t-1$ and, as a consequence, that $\partial \tilde{n}^{**}_{i,t-1}/\partial x_i < 0$. Using a similar reasoning, it is immediate that **(A4)** and **(A5)** also hold at $t-1$ and therefore that $\partial \tilde{n}^{**}_{i,t-1}/\partial y_i > 0$ and $\partial \tilde{n}^{**}_{i,t-1}/\partial c < 0$. Last, given **(A6)**, $\tilde{Y}_1^t(n) = \tilde{Y}_2^t(n)$ and $\tilde{n}^{**}_{1,t-1} = \tilde{n}^{**}_{2,t-1}$.

Case-2: $n \leqslant \overline{n}$. $\tilde{V}_i^{t-1}(n) = \max\{y_i(1-2\mu(n)); \nu(n)\tilde{V}_i^t(n+1) + (1-\nu(n))\tilde{V}_i^t(n-1) - c\}$ and

$$\tilde{W}_i^t(n) = \max\{0, \nu(n)\tilde{W}_i^{t+1}(n+1) + (1-\nu(n))\tilde{W}_i^{t+1}(n-1) - c\}$$
$$\tilde{W}_i^{t-1}(n) = \max\{0, \nu(n)\tilde{W}_i^{t}(n+1) + (1-\nu(n))\tilde{W}_i^{t}(n-1) - c\}$$

Suppose that the following assumptions **(A1')-(A6')** hold.

(A1'): $\tilde{W}_i^t(n)$ is non-decreasing in n and there exists $\tilde{n}_{i,t}^*$ such that $\tau_{i,t} = b$ if $n < \tilde{n}_{i,t}^*$ and $\tau_{i,t} = w$ if $n \in (\tilde{n}_{i,t}^*, \overline{n}]$.
(A2'): $\tilde{W}_i^t(n) \geqslant \tilde{W}_i^{t+1}(n)$ and therefore $\tilde{n}_{i,t}^* < \tilde{n}_{i,t+1}^*$.
(A3'): $\tilde{W}_i^t(n, x_i) \leqslant \tilde{W}_i^t(n, x_i')$ if $x_i > x_i'$ (and therefore $\partial \tilde{n}_{i,t}^* / \partial x_i < 0$).
(A4'): $\tilde{W}_i^t(n, y_i) \leqslant \tilde{W}_i^t(n, y_i')$ if $y_i > y_i'$ (and therefore $\partial \tilde{n}_{i,t}^* / \partial y_i < 0$).
(A5'): $\tilde{W}_i^t(n, c) \geqslant \tilde{W}_i^t(n, c')$ if $c > c'$ (and therefore $\partial \tilde{n}_{i,t}^* / \partial c > 0$).
(A6'): $\tilde{W}_1^t(n) = \tilde{W}_2^t(n)$ when $x_1 = h$, $y_1 = l$, $x_2 = l$ and $y_2 = h$ (and therefore $\tilde{n}_{1,t}^* = \tilde{n}_{2,t}^*$).

Given **(A1')**, the r.h.s. of $\tilde{W}_i^{t-1}(n)$ is increasing in n, so $\tilde{W}_i^{t-1}(n)$ is non-decreasing in n. Therefore, there exists a unique cutoff $\tilde{n}_{i,t-1}^*$ such that for all $n < \tilde{n}_{i,t-1}^*$ then $\tau_{i,t-1} = b$, and for all $n \in (\tilde{n}_{i,t-1}^*, \overline{n}]$ then $\tau_{i,t-1} = w$. Also, given **(A2')**, the r.h.s. of $\tilde{W}_i^{t-1}(n)$ is greater or equal than the r.h.s. of $\tilde{W}_i^t(n)$ and therefore $\tilde{W}_i^{t-1}(n) \geqslant \tilde{W}_i^t(n)$. Overall, both **(A1')** and **(A2')** hold at date $t-1$. Furthermore, $\tilde{n}_{i,t-1}^* < \tilde{n}_{i,t}^*$. Now, denote:

$$\tilde{W}_i^{t-1}(n, x_i) = \max\{0, \nu(n)\tilde{W}_i^t(n+1, x_i) + (1-\nu(n))\tilde{W}_i^t(n-1, x_i) - c\}$$

$$\tilde{W}_i^{t-1}(n, x_i') = \max\{0, \nu(n)\tilde{W}_i^t(n+1, x_i') + (1-\nu(n))\tilde{W}_i^t(n-1, x_i') - c\}$$

By **(A3')**, if $x_i > x_i'$ then $\tilde{W}_i^t(n+1, x_i) \leqslant \tilde{W}_i^t(n+1, x_i')$ and $\tilde{W}_i^t(n-1, x_i) \leqslant \tilde{W}_i^t(n-1, x_i')$. Therefore, $\tilde{W}_i^{t-1}(n, x_i) \leqslant \tilde{W}_i^{t-1}(n, x_i')$. This means that **(A3')** holds at date $t-1$ and, as a consequence, that $\partial \tilde{n}_{i,t-1}^* / \partial x_i < 0$. Using a similar reasoning, it is immediate that **(A4')** and **(A5')** also hold at $t-1$ and therefore that $\partial \tilde{n}_{i,t-1}^* / \partial y_i < 0$ and $\partial \tilde{n}_{i,t-1}^* / \partial c > 0$. Last, given **(A6')**, $\tilde{W}_1^t(n) = \tilde{W}_2^t(n)$ and $\tilde{n}_{1,t-1}^* = \tilde{n}_{2,t-1}^*$.