

Investment strategy and selection bias: An equilibrium perspective on overoptimism*

Philippe Jehiel[†]

4th November 2016

Abstract

Investors of new projects consider the returns of implemented projects delivering the same impression, and invest if the empirical mean return exceeds the cost. The steady states of such economies result in suboptimal investment decisions due to the selection bias in the sampling procedure and the dispersion of impressions across investors. Assuming better impressions are associated with higher returns, investors' assessments of their projects are overoptimistic, and there is overinvestment as compared with the rational benchmark. The presence of rational investors aggravates the overoptimism bias of sampling investors, thereby illustrating a negative externality imposed by rational investors.

*Thanks to Roland Bénabou, Milo Bianchi, Ignacio Esponda, David Hirshleifer, Charles Manski, Stephen Morris, Rani Spiegler, anonymous reviewers, and seminar participants at UCL behavioral game theory workshop, PSE, Princeton student workshop, Game theory French symposium, ESSET 2015, Bristol Game Theory workshop 2016, D-tea 2016 where preliminary versions of the ideas expressed here were presented.

[†]PSE, 48 boulevard Jourdan, 75014 Paris, France and University College London ; jehiel@enpc.fr

1 Introduction

A key aspect of entrepreneurial activity consists in deciding whether and when to make investments on the basis of the ex ante impressions decision makers get from the projects that reach them. In a Bayesian framework, the impression would be modelled as a signal received by the decision maker. Together with the knowledge of how signals, returns and costs are jointly distributed, the decision maker would be able to make the optimal investment decision using the standard Bayesian updating machinery. Yet, it is not clear how decision makers would know the joint distribution of signals, returns and costs. Instead, decision makers are likely to look at implemented projects and consider the realized returns observed in those projects so as to build statistics (even informally) about what impressions imply in terms of the distribution of returns.¹

A natural approach consists for the decision maker in aggregating the returns observed in those implemented projects that look similar to the current project of interest in the sense of delivering the same impression to the decision maker from an ex ante viewpoint. The decision maker would invest in a current project associated with some impression a if the returns observed over the implemented projects associated to the same impression a for him are sufficiently high in a statistical sense that may in general depend on the risk tolerance of the decision maker as well as his estimate of the costs. For example, if the decision maker is risk neutral and knows the cost c -an assumption that is maintained for simplicity in the rest of the paper- he would invest if the observed mean return exceeds the cost c in the considered pool, and he would not invest otherwise.

Given a pool of implemented projects, the heuristics just proposed would give rise to a new distribution of implemented projects for the current generation of decision makers, and this new distribution would itself be used for the derivation of the investment strategy of the next generation of investors. Assuming stationarity in the arrival of new projects and decision makers, I am interested in understanding the properties of the steady states generated by such dynamic systems, which I will refer to as equilibria with sampling investors.

When analyzing such equilibria, I will be assuming that the impressions obtained by

¹The very reason why statistics would in general rely exclusively on such data is that data on non-implemented projects are typically not accessible.

investors about a given project need not be the same across investors. More precisely, every investor's impression about a given project will be modelled as an independent signal realization drawn from a distribution assumed to be common to all investors and that typically depends on the return of the project. I will also be assuming that a higher realization of investors' signals is more representative of a higher return in the sense that the joint distribution of the return and investors' signals satisfies the monotone likelihood ratio property. A canonical illustration (routinely considered in finance models, see for example Grossman and Stiglitz, 1976) stipulates that conditional on the return x , the impression of investor i takes the form $a_i = x + \varepsilon_i$ where ε_i is the investor-specific realization of a normal distribution with mean 0 and variance σ . While such a specification will be used to illustrate the main findings, I am allowing for more general specifications of the distribution of impressions and returns.

The main findings of the paper are as follows. In equilibrium, no matter what impression investors get from their project, sampling investors have an overly optimistic assessment of the expected returns of their project. As a result, there is more investment in the equilibrium with sampling investors than what the optimal strategy based on the knowledge of the distribution of signals and returns would dictate. Modelling the impression as a noisy signal about the return as described above, I note that the overoptimism bias is more pronounced for intermediate realizations of the impressions, and that the welfare loss induced by the excessive investment is lowest either when the noise is small or when it is very significant so that the biggest welfare loss is obtained for intermediate levels of informativeness of the impressions. Finally, when investors are either rational (making the optimal investment decisions) or sampling, I note that the overoptimism and overinvestment biases of the sampling investors are all the more severe that the share of rational investors is bigger, thereby illustrating a negative externality that rational investors impose on sampling investors.

The overoptimism and overinvestment biases identified in this paper are the consequences of two ingredients: the selection neglect implicit in the heuristics used by sampling investors and the hypothesis that the impressions given by the same project are not the same across investors.² Each effect is key for the derivation of the biases.

²The monotone likelihood ratio property also plays a role as illustrated later through an example.

Clearly, as a direct implication of the law of large numbers, if decision makers could have access also to non-implemented projects, the heuristics of sampling investors (now allowing for the aggregation of both implemented and non-implemented projects according to the delivered impression) would lead them to have the correct estimate of the expected return for each possible impression. But, the samples considered by sampling investors are biased in that they include only projects that were implemented, and these are not randomly drawn projects. The sampling heuristics can be viewed as reflecting a kind of selection neglect in that it treats the biased samples as if they were not biased. Selection neglect has been documented in a number of psychological studies (Nisbett and Ross, 1980), and even if decision makers were aware of selection bias, it would be very hard to fully adjust for it.³ But, selection neglect alone is not enough to explain the overoptimism and overinvestment biases arising in the equilibrium with sampling investors. If all investors were to get the same impression about any given project, there would be no pro-investment bias in equilibrium. Biases would not arise in this case because assuming investment occurs with positive probability for some impression, the sample of implemented projects associated with that common impression would be unbiased. The dispersion of beliefs even when exposed to the same objective facts has been documented in various fields in particular in experts' surveys about inflation expectations (Mankiw et al., 2004) or in relation to traders' reactions to public announcements (Kandel and Pearson (1995)). Likewise, it is most likely at work regarding investors' impressions about objectively similar projects.

An illustrative example:

To illustrate some of the main findings, consider a setting in which the returns can take two equally likely values \underline{x} , \bar{x} with the cost c lying in between these two values, i.e.

³From a theoretical viewpoint, correcting the bias would require some structural knowledge about how impressions are generated and how other investors process the data. While dispensing from a parametric knowledge, the knowledge that the impressions of different investors are iid conditional on the return realization would still be required and such a knowledge cannot be inferred from the data (see Manski (2004) for related discussions as to why correcting the selection bias requires a lot of knowledge). From an empirical viewpoint, Elton et al. (1996) show that the bias persists in the context of assessing the performance of funds even though everyone there is aware that funds have a tendency to disappear when they perform poorly and thus that the sample consisting of still alive funds is not representative of all funds.

$\underline{x} < c < \bar{x}$. Assume decision makers can get three possible impressions labelled *Good*, *Medium* and *Bad*, and when the return is high (resp. low), the decision maker gets an impression that is either *Good* (resp. *Bad*) or *Medium* each with probability half. Thus, when the impression is *Good*, it is optimal to invest, since a *Good* impression can only come from a high return project. Similarly, when the impression is *Bad*, it is optimal to not invest, since a *Bad* impression can only come from a low return project. Assuming that $c > \frac{\underline{x} + \bar{x}}{2}$, it is optimal to not invest when the impression is *Medium*, since given the symmetry of the problem, Bayesian updating would then tell the decision maker that the two returns \underline{x} and \bar{x} are equally likely.

Consider a sampling investor who would observe in his pool only projects handled by rational investors. Since rational investors invest only when their impression is *Good*, the pool of implemented projects would all have high returns. Thus, a sampling investor would choose to invest when getting impression *Medium*, since half of the implemented projects would give him impression *Medium* and all of them would correspond to a high return. In the equilibrium with sampling investors only, decision makers invest more than in the rational case, but potentially less than a sampling investor would do when surrounded with rational investors only. The reason why the investment decisions of sampling investors may be altered is that the presence of sampling investors results in the presence of low return projects in the pool of implemented projects, and such a compositional effect reduces the pro-investment bias, even if it does not eliminate it, as implied by the main result of the paper. More precisely, within the proposed example, in the equilibrium with sampling investors only, when $c < \frac{2\bar{x} + \underline{x}}{3}$ decision makers invest when they get impressions *Good* or *Medium*, but when $c > \frac{2\bar{x} + \underline{x}}{3}$, only a fraction of projects associated to impression *Medium* is implemented by sampling investors and the perceived expected return associated to that impression coincides exactly with the cost c in equilibrium.⁴

⁴Observe that if the impressions obtained by different investors were always the same, investment when the impression is *Medium* would not be possible. By contradiction, if investment sometimes occurred with the *Medium* impression, a sampling investor would rightly get at the conclusion when the impression is *Medium* that there is an equal chance that the return is high or low (unlike in the dispersed impression scenario in which some projects delivering impression *Medium* had been decided by investors who had got a *Good* impression), and thus he would not invest.

Related literature:

This paper can be viewed as bringing together ideas from the literatures on overconfidence, bounded rationality, and econometrics combining them in a novel way. The econometrics literature has been discussing at length selection bias but essentially from the viewpoint of the analyst, assuming economic agents are perfectly rational (see Heckman, 1979). By contrast, this paper assumes that sampling investors are subject to selection neglect, and it analyzes the consequences this may cause on the efficiency of the decision making. The literature on overconfidence has documented that entrepreneurs tend to be overly optimistic about their projects (see for example Cooper et al. (1988) or Malmendier and Tate (2005)), which has generally been used to justify that investors rely on subjective priors or attach excessive precision to the signals they receive (see for example Xiong (2013) or Daniel and Hirshleifer (2015) for such a use in finance models). By contrast, this paper derives the overoptimism bias from selection neglect, assuming only the data about implemented projects are available to investors. As mentioned above, the equilibrium approach pursued in this paper allows me to relate the degree of overoptimism to the informativeness of the objective signals received by the investors, which can in principle be tested (and would not be implied by the subjective prior approach).⁵ Finally, the literature on bounded rationality has developed various solution concepts allowing for misspecified expectations (see in particular, the analogy-based expectation equilibrium (Jehiel, 2005), and the cursed equilibrium (Eyster and Rabin, 2005)), and it has sometimes connected such equilibrium approaches to selection bias (see in particular the behavioral equilibrium (Esponda, 2008)). This paper adopts an equilibrium perspective in line with the solution concepts just mentioned, but unlike the previous approaches it applies it to pure decision problems and derives the overoptimism bias in such non-strategic contexts.⁶

⁵Theoretical approaches to overconfidence that complement the one discussed in this paper include: 1) Rabin and Schrag (1999) who derive overconfidence from another psychological bias, the confirmatory bias that leads agents to sometimes behave as if they had not made observations that go against their current beliefs, 2) Van den Steen (2004) who defines overconfidence as the subjective belief that one performs better than others, which Van den Steen derives from a revealed preference argument in a subjective prior world - Van de Steen's insight will be further discussed later, and 3) several studies that derive overconfidence from motivated cognition purposes among which Bénabou and Tirole (2002), Köszegi (2006) or Bénabou (2015).

⁶Viewing nature as a player, I suggest later how the equilibrium with sampling investors can be viewed

Relating erroneous judgements to selection neglect appears in other studies. For example, Denrell (2003) discusses how one can wrongly believe that risky projects are associated with high mean performance if data on failed projects are not accessible or less visible (essentially because large negative shocks tend not to be recorded in the accessible data).⁷ However, to the best of my knowledge, this paper is the first to consider an equilibrium approach to this in the sense of viewing the pool of data agents have access to as resulting from the erroneous judgments agents make based on their observations. The endogeneity discussed here is essential for example for the understanding of why the presence of extra rational investors makes the overoptimism bias worse (which is consistent with the empirical finding of Lerner and Malmendier (2013) who observe that a higher share of entrepreneurial peers decreases entrepreneurship, see subsection 4.3 for elaborations).

The rest of the paper is organized as follows. Section 2 presents the investment problem. Section 3 analyzes the overoptimism and overinvestment biases arising in the equilibrium with sampling investors. Section 4 develops further insights in particular allowing for a mix of rational and sampling investors, and adding dynamic considerations in relation to cycling and the convergence to steady state.

2 The investment problem

A large number of investors idealized as a continuum is considered. Each investor assumed to be risk neutral has to decide whether or not to invest in one project that is different for

as an analogy-based expectation equilibrium after appropriately defining the extensive-form game and the required analogy partitions (see the working paper version Jehiel (2015) for more details). I believe such connections are useful in that they allow to develop a unifying theme of the effects of imperfect learning over different applications. In a recent paper, Spiegler (2016) discusses how to interpret the equilibrium with sampling investors using the tool of Bayesian networks that he has recently introduced into economics. It should be mentioned that his interpretation requires that decision makers would have access to the joint distribution of signals and returns, and in this case I would argue decision makers should be able to make the optimal investment decisions.

⁷In a very different context, Streufert (2000) discusses the idea that parents of poor neighborhoods may not consider the data of successful youngsters who would leave the neighborhood, and as a result would have a downward biased perception of the returns to schooling. As Denrell though, he does not discuss how the schooling decisions generated by such erroneous perceptions would affect the pool of data from which agents base their estimates.

each investor. The cost of every project is c . The return of a project is random and can take various possible values x in a set $X \subset \mathbb{R}$ (assumed for simplicity to consist of finitely many values). Before making his decision, an investor knows the cost c but does not know the return realization x of his project. However, he observes a signal realization a for his project. The signal realization a can be thought of as representing the impression that the investor gets from the project. It takes values in (\underline{a}, \bar{a}) with $\underline{a} < \bar{a}$ (where I allow that $\underline{a} = -\infty$ and $\bar{a} = +\infty$).⁸ Based on a , the investor has to decide whether or not to invest. If the project is implemented (i.e., the investor decides to invest), it is accessible by everyone, and the obtained return x is assumed to be freely observable. Non-implemented projects are not accessible. When an investor has access to a project, he can freely generate an impression similarly to how he generates an impression for his own project. That is, he can observe a signal realization a for every implemented project. It should be mentioned that investors are not assumed to be observing the impressions of other investors. They only observe their own impressions.

Returns and impressions are generated similarly for all projects and for all investors. Importantly, I assume that conditional on a return realization, the impressions of two different investors are statistically independent so as to reflect that impressions are influenced by effective returns but investors' impressions (about the same objective project) are heterogeneous. Specifically, for each project, the probability that the return realization be $x \in X$ is $l(x) \geq 0$ with $\sum_{x \in X} l(x) = 1$. Conditional on the realization x of a project, the signal realization a observed by any investor for this project is assumed to be distributed according to the density $f(\cdot | x)$ with support (\underline{a}, \bar{a}) , and two different investors i and i' get two independent draws $a_i, a_{i'}$ from this distribution. Assuming that the distribution of a takes the form of a density will simplify the exposition of the analysis but is not required (both the example in introduction and one example below assume a can take finitely many values). For concreteness, one may think of the situation in which the impression a would take the form $x + \varepsilon$ where x is the return and ε is the realization of a noise term for example distributed according to a normal distribution with mean 0 and variance σ . While such a specification will serve as the leading example to illustrate

⁸In the sequel, for any continuous function $g(\cdot)$, I will refer to $\lim_{a \rightarrow \bar{a}} g(a)$ (resp. $\lim_{a \rightarrow \underline{a}} g(a)$) as $g(\bar{a})$ (resp. $g(\underline{a})$).

the main results, the analysis developed next allows for much more general specifications of the distributions of the impression as a function of the return.

If the investor knew how the signal realization a and the return realization x are jointly distributed, he could adjust the optimal strategy. Given the assumed risk neutrality, the optimal investment strategy requires the investor to invest upon observing a whenever $E(x | a) > c$ and to not invest whenever $E(x | a) < c$ where $E(x | a)$ is derived from $l(\cdot)$ and $f(\cdot | x)$ by Bayes' law, i.e.,

$$E(x | a) = \frac{\sum_{x \in X} l(x) f(a | x) \cdot x}{\sum_{x \in X} l(x) f(a | x)}.$$

In the following, I will assume that the optimal strategy requires that, for some signal realizations a , it is best to invest.

Importantly, I am assuming that neither $l(\cdot)$ nor $f(\cdot | \cdot)$ is *a priori* known to the investor. Instead, I am embedding the above investment environment into a multi-period framework in which in every period, a new cohort of investors has to make investment decisions similar to the ones just described, and investors design their investment strategy by considering the past data accessible to them. According to the above observability assumptions, an investor in a given period can freely access those past projects that were implemented, observe the corresponding returns and derive the associated signals/impressions according to the process described above. I am assuming that in order to decide whether or not to invest, investors adopt the following heuristic. When getting a signal realization (an impression) a for his current project, the decision maker gathers all projects in the past for which he gets the same signal realization (impression) a . Then he computes the empirical mean return in those projects (this only requires averaging the x observed in those projects having the same a signal realization), and he invests whenever the obtained empirical mean return is above the cost c , and he does not invest otherwise. I will consider the steady states of such a dynamic system and refer to the resulting investment strategies as equilibria with sampling investors (in Section 4, I briefly consider whether and when the dynamic system converges to a steady state). In order to rule out trivial situations in which there would be no investment at all, I will also assume that whatever the observed signal there is a tiny probability (assumed to be the same for

all signal realizations) that the decision maker invests.

To define formally an equilibrium with sampling investors, let $q(a)$ denote the (steady state) probability with which an investor observing a would invest, and assume q is bounded away from 0 for some positive measure of signals. The probability of observing an implemented project with return x conditional on the impression a being in $A \subseteq (\underline{a}, \bar{a})$ would be

$$\Pr(x \mid a \in A, \text{implemented}; q) = \frac{l(x) \Pr(a \in A, \text{implemented} \mid x; q)}{\sum_{x' \in X} l(x') \Pr(a \in A, \text{implemented} \mid x'; q)}$$

where

$$\Pr(a \in A, x \mid \text{implemented}; q) = \Pr(a \in A \mid x) \int_{\underline{a}}^{\bar{a}} q(b) f(b \mid x) db$$

given that a randomly drawn project with return realization x would be implemented with probability $\int_{\underline{a}}^{\bar{a}} q(b) f(b \mid x) db$ (the investor in charge of such a project would receive signal b according to the density $f(\cdot \mid x)$ and invests then with probability $q(b)$).

Thus, the empirical mean return of implemented projects for which the investor gets the signal realization a would be

$$\widehat{v}(a; q) = \frac{\sum_{x \in X} l(x) f(a \mid x) \int_{\underline{a}}^{\bar{a}} q(b) f(b \mid x) db \cdot x}{\sum_{x \in X} l(x) f(a \mid x) \int_{\underline{a}}^{\bar{a}} q(b) f(b \mid x) db}$$

as results from the induced proportion of projects with return x in the pool of implemented projects associated to impression a . This leads to the following definition:⁹

Definition 1 *An investment strategy $q(\cdot)$ over (\underline{a}, \bar{a}) that induces some investment with positive probability is an equilibrium with sampling investors if $q(a) > 0$ implies $\widehat{v}(a; q) \geq c$ and $q(a) = 0$ implies $\widehat{v}(a; q) \leq c$.*

⁹To present formally the equilibrium with trembles, one may define the set Q_n of q such that $q(a) \geq \frac{1}{n}$ for all a , define an $\frac{1}{n}$ -equilibrium to be $q_n(\cdot)$ such that $q_n(a) > \frac{1}{n}$ implies $\widehat{v}(a; q_n) \geq c$ and $q_n(a) = \frac{1}{n}$ implies $\widehat{v}(a; q) \leq c$, and define an equilibrium to be the limit as n grows large of q_n such that q_n is an $\frac{1}{n}$ -equilibrium. Our assumption that the optimal investment strategy involves non-trivial investment will imply that all equilibria must result in positive probability of investment.

The strategy just defined is the result of a fixed point. The probabilities $q(b)$ with which other investors choose to invest when getting signal b affect the compositions of return x projects in the pool of implemented projects, which in turn affects the probability with which an investor getting signal a is willing to invest. In equilibrium, these two probability mappings should be the same.

Comments. 1) In the working paper version (see Jehiel, 2015), I envision projects as being described by strings of attribute realizations and each individual investor as observing just one attribute realization. Assuming there are as many attributes as there are investors and that each attribute realization has the same distribution conditional on the return realization, one gets a formulation similar to the one developed above. When there are finitely many attributes, the observations of two different investors would not be independent conditional on the return realization, which would result in some extra complications (see the section on correlation in Jehiel, 2015). 2) One may interpret an equilibrium with sampling investors as defined above as an analogy-based expectation equilibrium (Jehiel, 2005) in which nature would be considered as a player, and the game would let nature first select a project describing the return realization and the vector of signal realization for every investor, then the investor would have to decide whether or not to invest on the basis of the observed signal and in case of investment nature would implement the return realization. The analogy partition of a given investor required to support the above equilibrium with sampling investors consists in bundling all the decision nodes of nature regarding the (second) choice of return that correspond to the same signal realization of the considered investor (see Jehiel (2015) for further details in the multi-attribute specification).

3 Overoptimism as a result of selection bias

I analyze the above investment environment assuming that a higher signal realization is more representative of a higher return. Such a condition is satisfied whenever conditional on x , a is distributed according to a normal distribution with mean x and variance σ , and it is satisfied for many other specifications. In particular, it is without loss of generality whenever there are two possible return realizations $x = \underline{x}, \bar{x}$, since then a can be reordered

so that the likelihood ratio $\frac{f(a|\bar{x})}{f(a|x)}$ is increasing with a . Formally, the following monotone likelihood ratio property is assumed to hold:

Assumption (MLRP): For any $a' > a$ and $x' > x$, it holds that: $\frac{f(a'|x')}{f(a|x')} > \frac{f(a'|x)}{f(a|x)}$.¹⁰

3.1 Equilibrium characterization

Proposition 1 *Under MLRP, there exists a unique equilibrium with sampling investors. The equilibrium is such that for some threshold a^S , a decision-maker chooses to invest if the observed signal realization a is above a^S and to not invest otherwise where a^S is uniquely defined by*

$$\frac{\sum_{x \in X} f(a^S | x)[1 - F(a^S | x)]l(x) \cdot x}{\sum_{x \in X} f(a^S | x)[1 - F(a^S | x)]l(x)} = \begin{cases} \geq c & \text{if } a^S = \underline{a} \\ c & \text{if } a^S \in (\underline{a}, \bar{a}) \end{cases} \quad (1)$$

Proof of Proposition 1:

Suppose that, in equilibrium, investment occurs with probability $q(a)$ when $a \in (\underline{a}, \bar{a})$ is observed. As already highlighted, the perceived expected return of a project with signal realization a would then be:

$$\hat{v}(a; q) = \frac{\sum_{x \in X} l(x)f(a | x) \int_{\underline{a}}^{\bar{a}} q(b)f(b | x)db \cdot x}{\sum_{x \in X} l(x)f(a | x) \int_{\underline{a}}^{\bar{a}} q(b)f(b | x)db}$$

Since $a \rightarrow \hat{v}(a, q)$ is increasing (whatever $q(\cdot)$) by MLRP, one can infer that investors must follow a threshold strategy, i.e. for some z , invest if $a > z$ and do not invest if $a < z$ where z (if interior) is defined by $\hat{v}(z, q) = c$.

Define

$$H(a, z) = \frac{\sum_{x \in X} l(x)f(a | x)[1 - F(z | x)] \cdot x}{\sum_{x \in X} l(x)f(a | x)[1 - F(z | x)]} \quad (2)$$

This is the perceived expected return of an a -project when other investors follow the z -threshold investment strategy. One has:

¹⁰ Assuming $f(\cdot | x)$ is smooth, this can be formulated as requiring that $\frac{\partial f(a|x)/\partial a}{f(a|x)}$ is increasing in x .

Lemma 1 Under MLRP, $H(\cdot, \cdot)$ is increasing in a and z .

Proof of Lemma 1. The monotonicity in a has already been noted. The monotonicity in z follows from the observation that under MLRP, the hazard rate $\frac{f(z|x)}{1-F(z|x)}$ decreases with x (see any textbook or the monotone likelihood ratio entry of wikipedia) and thus $x \rightarrow \frac{\frac{\partial}{\partial z}[1-F(z|x)]}{1-F(z|x)} = \frac{-f(z|x)}{1-F(z|x)}$ increases with x . **Q.E.D.**

Given the assumption that there is some investment in the rational case, it follows that $H(\bar{a}, \underline{a}) > c$. An equilibrium must employ a threshold strategy z as already noted (by the monotonicity of $H(\cdot, \cdot)$ in a) and the threshold z must satisfy

$$H(z, z) = \begin{cases} \geq c & \text{if } z = \underline{a} \\ c & \text{if } z \in (\underline{a}, \bar{a}) \\ \leq c & \text{if } z = \bar{a} \end{cases} \quad (3)$$

Given that $H(\bar{a}, \underline{a}) > c$, the monotonicity of $H(\cdot, \cdot)$ in the second argument implies that $H(\bar{a}, \bar{a}) > c$, and thus the latter case can be ignored. Suppose then that $H(\underline{a}, \underline{a}) < c$. The continuity of H ensures that there exists $z \in (\underline{a}, \bar{a})$ satisfying $H(z, z) = c$. Hence, there must exist $z > \bar{a}$ satisfying (3).

Consider now $\bar{a} \geq z_1 > z_2 \geq \underline{a}$. Clearly, $H(z_1, z_1) > H(z_2, z_2)$ and (3) cannot be simultaneously satisfied for $z = z_1$ and z_2 . One concludes that there is only one equilibrium, and that this equilibrium is a threshold equilibrium a^S where a^S is uniquely defined to satisfy (1). **Q.E.D.**

Given that equilibrium is unique, one can unambiguously speak of the equilibrium subjective value that an investor observing the signal realization a assigns to the expected mean return of the project. It is denoted by $v^S(a)$. Using the definition (2) of $H(\cdot, \cdot)$, one has $v^S(a) = H(a, a^S)$ where a^S is as defined in Proposition 1.

3.2 Comparison to the rational benchmark

A rational investor is defined to be one who makes the optimal investment decision based on the true statistical distributions as defined by the densities $f(\cdot | \cdot)$ and the probabilities

$l(\cdot)$. Accordingly, upon observing the realization a of the signal, a rational investor rightly perceives the expected return of the project to be

$$v^R(a) = E(x | a) = \frac{\sum_{x \in X} l(x) f(a | x) \cdot x}{\sum_{x \in X} l(x) f(a | x)}$$

A rational investor invests if this value is above c and does not otherwise.

Let $a^R \in (\underline{a}, \bar{a})$ be uniquely defined by¹¹

$$v^R(a^R) = \begin{cases} \geq c & \text{if } a^R = \underline{a} \\ c & \text{if } a^R > \underline{a} \end{cases}$$

A rational investor invests whenever $a > a^R$ and he does not when $a < a^R$.

Using the $H(\cdot, \cdot)$ function introduced in (2), it is readily verified that $v^R(a) = H(a, \underline{a})$ (since for all x , $F(\underline{a} | x) = 0$), and thus a^R is uniquely defined by

$$H(a^R, \underline{a}) = \begin{cases} \geq c & \text{if } a^R = \underline{a} \\ c & \text{if } a^R > \underline{a} \end{cases}$$

There are two ways to think of a rational decision maker, as just described. One way is to hold the view that a rational investor knows $l(\cdot)$ and $f(\cdot | \cdot)$ to start with and does the corresponding Bayesian updating when observing a , as already suggested. Another way is to hold the view that a rational investor is an experienced decision maker who has had sufficiently many learning opportunities to find out the investment strategy (defined as a function of the observed signal) that delivers the highest expected payoff.

Whatever the interpretation of the rational investment strategy, I note that the sampling heuristic used by decision makers to assess the expected return of a -projects in the equilibrium shown in Proposition 1 leads them to have an overly optimistic perception as compared with the rational perception, since $H(a, a^S)$ is bigger than $H(a, \underline{a})$ (the rational assessment) due to the monotonicity of $H(\cdot, \cdot)$ in its second argument. This observation and its implication for the volume of investment is summarized in the next Proposition

¹¹Uniqueness comes again from MLRP, which ensures that $v^R(\cdot)$ is increasing in a . The fact that $a^R < \bar{a}$ comes from the assumption that there is investment with positive probability in the optimal solution.

whose complete proof appears in the Appendix.

Proposition 2 *Under MLRP, in equilibrium, sampling investors overvalue the expected returns of a -projects as compared with the rational benchmark whatever the signal observation a , i.e. $v^S(a) \geq v^R(a)$ for all a . There is at least as much investment in the equilibrium with sampling investors as in the rational benchmark. That is, $a^S \leq a^R$.*

It is tempting to think that the overoptimism arising in the equilibrium with sampling investors is the mere consequence of the fact that the subset of projects investors have access to is biased toward having higher returns. However, this property alone would not necessarily give rise to the overoptimism bias without the additional statistical structures assumed in the main model. In particular, assume in contrast to the main model that the impressions that different investors get about a given project are always the same across investors (they may of course differ across projects). This can be modeled by assuming that for every project with return x , every investor observes the same signal a drawn from the distribution $f(\cdot | x)$ (unlike in the main model in which two different investors were assumed to get independent draws from $f(\cdot | x)$). In this case, letting $q(a)$ denote the steady state probability that there is investment whenever the (common) signal realization is a and assuming that $q(a) > 0$ (which would always hold true if some trembling behavior were assumed), the subjective assessment of a sampling investor observing the signal realization a would be:

$$\frac{\sum_{x \in X} l(x) f(a | x) q(a) \cdot x}{\sum_{x \in X} l(x) f(a | x) q(a)}.$$

This expression simplifies into the rational expression $v^R(a)$ shown above, thereby implying that there is no overoptimism bias in the corresponding equilibrium with sampling investors. Note that there is no bias even though the sample investors have access to is biased toward the high return ones (since in equilibrium only those projects for which $a > a^R$ would be voluntarily implemented).

Thus, the overoptimism bias identified in Proposition 2 requires the combination of the sampling heuristic and the assumption that the impressions produced by any project are heterogeneous across investors. As it turns out, the MLRP assumption also plays a

role in the result. While the MLRP assumption is a very natural one (it is without loss of generality if x can take only two values and it holds whenever a is a noisy signal about x for many specifications of the noise distribution), I note that the result of Proposition 2 would not necessarily hold if the MLRP assumption were removed (still assuming as in the main model that two different investors get independent draws from $f(\cdot | x)$ for a project with return x). This is illustrated through the following example.

Example 1 Returns x can take four values $x = -2, -1, 1, 2$ which are equally likely (i.e., $l(x) = 1/4$ for all x); a can take three values $a = a_1, a_2, a_3$; cost c is 0. The distribution of a given x is summarized in the following table in which the number at the intersection of the a_i row and the x column -referred to as $p_i(x)$ - is the probability that signal a_i is drawn conditional on the return realization being x .

		2	1	-1	-2
a_1	0.1	0.4	0.1	0.24	
a_2	0.1	0.31	0.5	0	
a_3	0.8	0.29	0.4	0.76	

The rational investment strategy requires that there is investment when $a = a_1$ or a_2 but not when $a = a_3$.¹² Noting that $\sum_x p_i(x)(p_1(x)+p_2(x)) \cdot x$ is positive for $i = 1$ but negative for $i = 2, 3$ implies that if all other investors were following the rational investment strategy, a sampling investor would choose to invest when observing a_1 but not when observing a_2 or a_3 . The equilibrium with sampling investors would then take the following form: invest when observing a_1 , invest with probability μ such that $\sum_x p_2(x)(p_1(x) + \mu p_2(x)) \cdot x = 0$ when observing a_2 and not invest when observing a_3 . Overall, there would be less investment than in the optimal strategy, and the subjective assessment attached to an a_2 project -which would coincide with $c = 0$ - would be lower than the rational assessment.

¹²This follows because $\sum_x p_i(x) \cdot x > 0$ for $i = 1, 2$ for not for $i = 3$.

3.3 Informativeness, overoptimism, and overinvestment

In order to illustrate the findings obtained so far, it may be useful to consider the following simple scenario. The cost c is normalized to 0. Return x can take two values $\underline{x} = -1$ or $\bar{x} = 1$ with the same probability ($l(x) = 1/2$ for $x = -1$ and 1). Conditional on x , the signal a takes the form $a = x + \varepsilon$ where ε is the realization of a normal distribution with mean 0 and variance σ . In such a symmetric setting, the optimal strategy requires to invest if $a > 0$ and to not invest otherwise. That is, $a^R = 0$ (whatever σ). After simple arrangements, the threshold a^S that arises in the equilibrium with sampling investors is characterized by

$$PDF(a^S + 1)(1 - CDF(a^S + 1)) = PDF(a^S - 1)(1 - CDF(a^S - 1))$$

where PDF and CDF stand respectively for the density and the cumulative of the normal distribution with mean 0 and variance σ . Clearly, $a^S < 0$ given that $PDF(1) = PDF(-1)$ and $CDF(1) > CDF(-1)$, which confirms the general result of Proposition 2. In the next graph, the overoptimism bias $v^S(a) - v^R(a)$ known to be positive by Proposition 2 is depicted as a function of a , assuming the variance σ of the normal distribution is 1.¹³

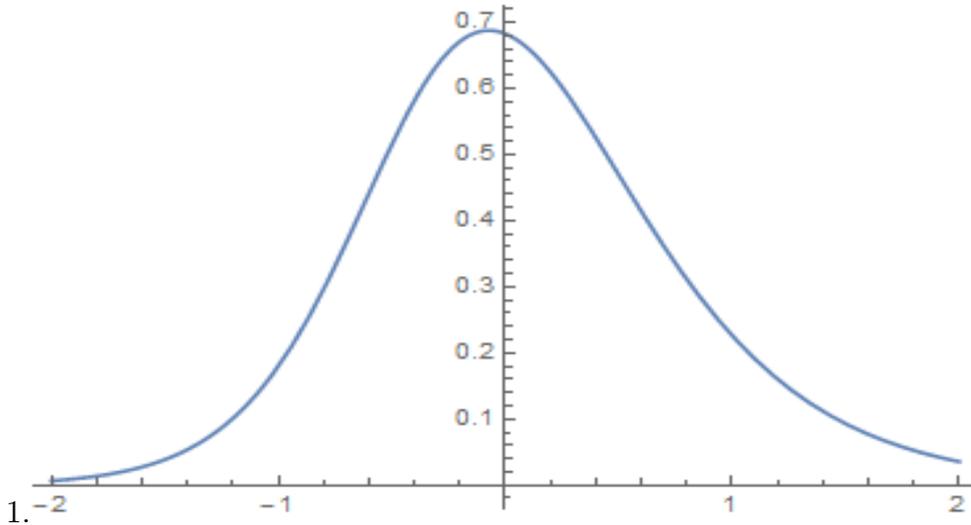


FIGURE 1

¹³By way of comparison, $v^R(0) = 0$ and $v^R(+\infty) = 1$.

Interestingly, one observes that the bias is maximal for intermediate values of a and goes to 0 as a approaches $-\infty$ or $+\infty$. This is no coincidence. When a approaches either $\underline{a} = -\infty$ (respectively, $\bar{a} = +\infty$), the likelihood that the return is $x = \underline{x}$ (resp. \bar{x}) becomes very large.¹⁴ More generally, there would be no bias for extreme realizations of a (for a close to \underline{a} or \bar{a}) whenever x can take two values $x = \underline{x}$ or \bar{x} , $\underline{x} < c < \bar{x}$ and $\frac{f(a|\bar{x})}{f(a|\underline{x})}$ approaches 0 (resp. $+\infty$) as a approaches \underline{a} (resp. \bar{a}). This is because as a mere consequence of extreme signals being very informative, an investor observing such a signal realization would mostly see projects with homogeneous return realizations in his sample, thereby making the selection bias negligible.

Another quantity I consider now is the welfare loss induced in the equilibrium with sampling investors as compared with the rational benchmark. In the context of the example, the welfare loss can be expressed as

$$WL = \frac{1}{2} \int_{a^S}^0 (PDF(a+1) - PDF(a-1)) da$$

representing the aggregate loss due to the suboptimal implementation of projects with signal realizations a falling in $(a^S, 0)$. The following graph depicts how WL varies with the variance σ .¹⁵

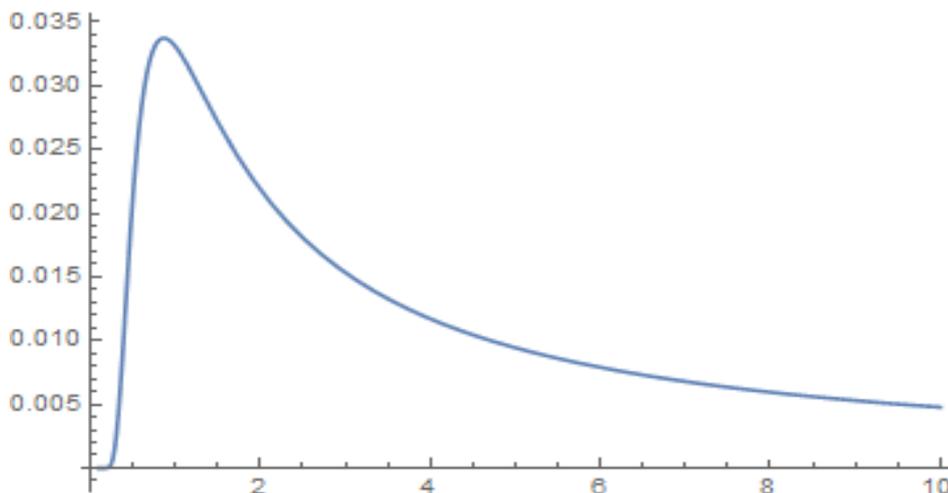


FIGURE 2

¹⁴This is so because $PDF(a^S + 1)/PDF(a^S - 1) \rightarrow +\infty$ (resp. $-\infty$) as $a \rightarrow -\infty$ (resp. $+\infty$).

¹⁵By way of comparison, the value of the expected maximal welfare is 0.5 when $\sigma = 0$, 0.34 when $\sigma = 1$, and 0 when $\sigma = \infty$.

As can be seen, the welfare loss is single-peaked, converges to 0 as σ goes to 0 or $+\infty$, and is thus maximal for intermediate values of the overall informativeness of the signal as parameterized here by σ . Such properties would hold more generally. Indeed, when the signal is very informative, the effect of the selection bias on the performance of the decision made by sampling investors becomes negligible as already noted, and thus when signals are globally informative with very high probability (as is the case when σ is large), the aggregate welfare loss is small. At the other extreme, when the signals are poorly uninformative, in generic situations, either investors would always invest or they would never invest, and when there is investment it would have to be optimal.¹⁶ Hence, beyond the canonical example considered here, the maximal welfare loss would arise for intermediate levels of the informativeness of the signals.

As illustrated above, the degree of overoptimism arising in the equilibrium with sampling investors and the welfare consequences of it depend on the informativeness of signals.¹⁷ Such a dependence would not necessarily arise in the subjective prior approach to overoptimism, which typically puts no structure on how investors from their subjective prior and thus on how overoptimism varies with the primitives of the model. It may be mentioned here that within the subjective prior paradigm, Van den Steen (2004) makes the interesting and simple observation that as a consequence of a revealed preference argument, no matter how subjective priors are modelled, others' decisions always look (weakly) suboptimal from the subjective viewpoint of any agent, thereby leading to the systematic subjective belief that one performs better than others. Such a relative overoptimism bias would also arise in my setting. If one were to ask any given investor at an ex ante stage (i.e., before he receives a signal realization for his project) whether he thinks he would perform better than other investors, he would be affirmative. Relative overoptimism arises here for the very same reason highlighted by Van den Steen that the investor believes (based on his subjective perception of the mapping between his perception and

¹⁶The considered example is non-generic in the sense that if investors receive no signal, they are indifferent as to whether or not to invest, but the conclusion still holds true in this case, since whether or not they invest it is equally good.

¹⁷This can in principle be the subject of empirical tests. In contexts in which market conditions are very clear or very unclear (so that signals are either very informative or very uninformative), one should observe less overoptimism bias than in contexts in which market conditions are partially predictable.

the investment decision) that he can screen projects better than others.¹⁸

4 Further insights

4.1 When rational investors exert negative externalities

Suppose the population of investors is mixed. A share $1 - \lambda$ of investors (referred to the sampling investors) proceeds as described in the main model: They observe a signal realization a for their project, sample all implemented projects in which they get the same signal realization a , and invest if the observed empirical mean return exceeds the cost c . A share λ of investors (referred to as rational investors) makes the optimal investment decision based on the observation of the signal realization. Signals and returns are distributed as in the main model.

It should be noted that in the sample considered by the sampling investors, there are both projects held by sampling investors and by rational investors. Since the decision rule is not the same for sampling and rational investors, the selection bias is typically affected by the heterogeneity of the population of investors. The purpose of the next Proposition is to investigate the effect of cognitive heterogeneity on the performance of sampling investors (rational investors are unaffected by the presence of sampling investors given that they face a decision problem and they behave optimally).

To pave the way toward the main result of this subsection, observe that sampling investors follow in equilibrium a threshold strategy that consists in investing in a project with signal realization a only if a exceeds a^* where a^* is defined by

$$\frac{\sum_{x \in X} f(a^* | x)[(1 - \lambda)(1 - F(a^* | x)) + \lambda(1 - F(a^R | x))]l(x) \cdot x}{\sum_{x \in X} f(a^* | x)[(1 - \lambda)(1 - F(a^* | x)) + \lambda(1 - F(a^R | x))]l(x)} = \begin{cases} \geq c & \text{if } a^* = \underline{a} \\ c & \text{if } a^* \in (\underline{a}, \bar{a}) \end{cases}$$

The left hand-side of this expression represents how a sampling investor subjectively

¹⁸Formally, his perceived ex ante payoff would be given by $E(v^S(a) - c | a > a^S) \Pr(a > a^S)$ whereas his perception of other investors' performance would be correctly given by $E(v^S(a)) - c$ where the density of a is the one arising in the pool of implemented projects, i.e., $f^S(a) = \sum_x l(x)[1 - F(a^S | x)]f(a | x) / \sum_x l(x)[1 - F(a^S | x)]$. Since $v^S(\cdot)$ is increasing and $v^S(a^S) = c$, the former expression is larger than the latter, confirming Van den Steen's relative overoptimism insight in the present setting.

assesses the mean return of a project with signal realization a^* , and it requires in equilibrium that if a^* is interior, this perceived mean return should be equal to the cost c .

The difference with the main model is that when a sampling investor makes an observation of another project, with probability λ she is facing a rational investor who invests only if the signal realization observed by the rational investor is larger than the rational threshold a^R (as defined in subsection 3.2), and with probability $1 - \lambda$ she is facing another sampling investor who invests if the signal realization he observes is larger than a^* . Given that conditional on the return realization, the signals are independently distributed across investors, the above expression follows.

Denote the above threshold a^* by $a^S(\lambda)$. One has previously seen that when there is no rational investor around ($\lambda = 0$), it holds that $a^S(0) \leq a^R$. The effect of λ on $a^S(\lambda)$ is unambiguously given by:

Proposition 3 *Under MLRP, the higher the share λ of rational investors, the more severe the pro-investment bias of sampling investors. That is, $a^S(\lambda)$ is weakly decreasing in λ , and for all λ , $a^S(\lambda) \leq a^S(0) \leq a^R$.*

The intuition behind Proposition 3 whose detailed proof appears in the Appendix is simple. If an investor is surrounded with more rational decision makers, the decisions made by others are better, and thus when sampling from these to form an assessment regarding the profitability of the project it appears to the investor that the project is even more profitable. The selection bias is more severe, which leads the investor to make a poorer decision. In some sense, rational investors exert a negative externality on those investors who follow the sampling heuristic.

It is natural to consider the effect of an increase of λ on welfare. Given Proposition 3, an increase of λ deteriorates the welfare of sampling investors, but at the same it increases the share of rational investors whose welfare is larger. Aggregating these two effects leads to ambiguous comparative statics in general. When the share of rational investors is sufficiently large, an increase of λ always enhances expected welfare (essentially because there are too few sampling investors who suffer from the negative externality imposed by rational investors). When the share of rational investors is sufficiently away from 1, the negative effect on sampling investors of increasing λ may dominate for some distributional

assumptions resulting in an overall negative impact of increasing the share of rational investors on expected welfare.¹⁹

4.2 Convergence to steady state

Embedding the above framework into a dynamic setting in which new cohorts of investors sample from previous cohorts of investors naturally leads to asking when we should expect to see convergence to steady state as considered in the main analysis. Another legitimate concern is whether the overoptimism bias identified in the main analysis would still arise in case there would be no convergence. To model the dynamics most simply, consider within the MLRP scenario discussed above a sequence of time periods $t = 1, 2, \dots$. Assume that in every period $t > 1$ there is a new cohort of investors of the same mass who sample from the implemented projects handled by the cohort of investors living in period $t - 1$, and assume to fix ideas that in the first period investors choose to invest whatever signal they observe.

In such a dynamic setting, investors in period t would adopt a threshold strategy z_t specifying to invest if the observed signal realization a is above z_t and to not invest otherwise where the sequence of z_t would be characterized inductively by $z_1 = \underline{a}$ (since the first generation of investors was assumed to invest always) and for all $t > 1$, the threshold z_{t+1} would be uniquely defined by $H(z_{t+1}, z_t) = c$ (assuming $H(\underline{a}, z) < c < H(\bar{a}, z)$ for all

¹⁹To illustrate these two points, consider a two return \underline{x}, \bar{x} scenario with $\underline{x} < c < \bar{x}$ and $l(\underline{x}) = l(\bar{x}) = 1/2$. Simple calculations yield that the marginal effect of λ on global welfare $\frac{dW}{d\lambda}$ can be written as $\frac{A+B}{2}$ where

$$\begin{aligned} A &= \int_{a^S(\lambda)}^{a^R} (f(a | \underline{x})(c - \underline{x}) - f(a | \bar{x})(\bar{x} - c)) da \\ B &= (1 - \lambda) \frac{da^S(\lambda)}{d\lambda} (f(a^S(\lambda) | \underline{x})(c - \underline{x}) - f(a^S(\lambda) | \bar{x})(\bar{x} - c)) \end{aligned}$$

A (resp B) is shown to be positive (resp. negative) using the MLRP property, $f(a^R | \underline{x})(c - \underline{x}) = f(a^R | \bar{x})(\bar{x} - c)$, $a^S(\lambda) < a^R$ and $\frac{da^S(\lambda)}{d\lambda} < 0$.

When λ is close to 1, B becomes negligible and thus $\frac{dW}{d\lambda} > 0$. When λ is away from 1, A can be made small relative to B by having a sufficiently small probability that signal realizations a fall in $(a^S(\lambda), a^R)$ (this is consistent with MLRP which only requires that $\frac{f(a|\bar{x})}{f(a|\underline{x})}$ is increasing in a , but puts no restriction on how likely the various a are). In such cases, $\frac{dW}{d\lambda} < 0$ holds. In the leading example in which conditional on x , $a = x + \varepsilon$ where ε is drawn from a normal distribution with variance $\sigma = 1$, one can show that $\frac{dW}{d\lambda} > 0$ for all λ .

z). It appears that z_2 coincides with a^R , and using the monotonicity of H , it can be shown by induction that the sequence $(z_{2k+1})_{k \geq 1}$ is weakly decreasing and satisfies $z_{2k+1} \geq a^S$ for all k while the sequence $(z_{2k})_{k \geq 1}$ is weakly increasing and satisfies $z_{2k} \leq a^S$ for all k where a^S is the equilibrium threshold defined in Proposition 1. Thus, $(z_{2k+1})_{k \geq 1}$ converges to z^* and $(z_{2k})_{k \geq 1}$ converges to z_* with $z_* \leq a^S \leq z^*$. If $z_* = z^* = a^S$ the system converges to the steady state described in Proposition 1. If $z_* < a^S < z^*$, the system converges to a limit two-period cycle in which in odd periods there is less activity as dictated by the threshold strategy z^* and in even periods there is more activity as dictated by the threshold strategy z_* . Whether the system converges or cycles depends on how the slope $\frac{\partial H / \partial z}{\partial H / \partial a}$ compares to 1. When it is uniformly lower than 1, (as is the case for the leading example with variance $\sigma = 1$), there is convergence. When it is larger than 1 in the neighborhood of $a = z = a^S$, the two-period limit cycle prevails.²⁰

It should be noted that in the above dynamics whether or not there is convergence, the overoptimism and overinvestment biases hold in every period (this follows from the monotonicity of H and the observation that $H(a^R, \underline{a}) = c$). Moreover, since $z_t \leq a^R$ for all t and $z_2 = a^R$, the monotonicity of H implies that the smallest z_t which corresponds to the most biased investment strategy is obtained in period 3 when the samples considered by the current cohort consist of projects handled by rational investors. In all subsequent periods, because sampled investors adopt suboptimal strategies, the sampling heuristic leads to less severe biases.

4.3 Cycling with heterogeneous investors

It is natural to combine dynamics as just considered with the possibility that investors could vary in their degree of sophistication, some of them being rational and others being subject to selection neglect as proposed in the main model. A full-fledged dynamic model along these lines would aim at endogenizing entry and exit of entrepreneurs, assuming for example entrepreneurs' sophistication vary with their experience. Analyzing such a model is clearly beyond the scope of this paper. Yet, in order to illustrate that some

²⁰If investors were sampling from all previous cohorts rather than just the most recent one, I suspect the convergence scenario would be made more likely (because such a sampling device would smoothen the reaction to previous behaviors), but more work is needed to establish this formally.

rich dynamics can be expected, consider the following stylized setting. In each period $t = 1, 2, \dots$ a new cohort of agents decides whether or not to become entrepreneur. Every entrepreneur faces the same distribution of projects as described above but agents may have different outside options assumed to be drawn independently across agents from a distribution with cumulative G . In every period, the share of rational agents is λ and the share of sampling agents is $1 - \lambda$. Let w^R denote the expected payoff a rational investor gets by becoming an entrepreneur (i.e., $w^R = E(\max v^R(a) - c, 0)$), and let $w^S(\lambda)$ denote the expected payoff a sampling investor subjectively expects to get when facing a share λ (resp. $1 - \lambda$) of rational (resp. sampling) investors.²¹ Rational agents become entrepreneur whenever their outside option falls below w^R , i.e. with probability $G(w^R)$. Sampling agents who would sample from a mix λ of rational investors and $1 - \lambda$ of sampling investors would become entrepreneur with probability $G(w^S(\lambda))$. Thus assuming the cohort of (sampling) agents in period t samples from the implemented projects in period $t - 1$, the share λ_t of rational investors in period t would follow the dynamic:

$$\lambda_t = \frac{\mu G(w^R)}{\mu G(w^R) + (1 - \mu) G(w^S(\lambda_{t-1}))}.$$

As can be inferred from the above analysis, $w^S(\cdot)$ is increasing in λ . Thus, a higher share of rational investors in period t would lead more sampling agents to become entrepreneurs in period $t + 1$, which would result in a lower share of rational investors in period $t + 1$. Depending on the shape of G , such a dynamic system may either converge to a limit share λ^* of rational investors or lead to long term cycling between high and low shares (away and respectively above and below λ^*) of rational investors, corresponding respectively to low and high levels of entrepreneurial activity.²² The prediction that there is more chance to become entrepreneur if one is exposed to fewer (better skilled) entrepreneurial cases should be subject of further empirical investigation, but it seems in agreement with Lerner and Malmendier's (2013) finding that a higher share of entrepreneurial peers decreases

²¹With the notation previously introduced, $w^S(\lambda) = E[\max(H(a, a^S(\lambda)) - c, 0)]$ where the density of a is $f_\lambda(a) = \frac{\sum_{x \in X} f(a|x)[(1-\lambda)(1-F(a^S(\lambda)|x)) + \lambda(1-F(a^R|x))]l(x)}{\sum_{x \in X} [(1-\lambda)(1-F(a^S(\lambda)|x)) + \lambda(1-F(a^R|x))]l(x)}$.

²² λ^* is a solution to $\lambda^* = \frac{\mu G(w^R)}{\mu G(w^R) + (1-\mu)G(w^S(\lambda^*))}$ and if G has sufficient mass around $w^S(\lambda^*)$ one should expect cycling to emerge.

entrepreneurship.²³

²³Lerner and Malmendier also note that the reduction is driven by a reduction in unsuccessful entrepreneurial ventures, which can be related to our finding that when sampling from a more active cohort the investment decisions get closer to the rational ones.

Appendix

Proof of Proposition 2: The first part, $v^S(a) \geq v^R(a)$, is proven noting that $v^S(a) = H(a, a^S)$, $v^R(a) = H(a, \underline{a})$ and using the monotonicity of H in z (see Lemma 1). As for the second part, whenever $a^R < \bar{a}$, one has that $H(a^R, \underline{a}) \geq c$ and thus $H(a^R, a^R) \geq c$ by the monotonicity of H in its second argument. This implies that $a^S \leq a^R$, as desired. **Q.E.D.**

Proof of Proposition 3. Define $H(a, z, \lambda) = \frac{\sum_{x \in X} f(a|x)[(1-\lambda)(1-F(z|x)) + \lambda(1-F(a^R|x))]l(x) \cdot x}{\sum_{x \in X} f(a|x)[(1-\lambda)(1-F(z|x)) + \lambda(1-F(a^R|x))]l(x)}$.

Lemma Under MLRP, H is increasing in a and z . It is decreasing in λ for $z \leq a^R$.

Proof. H increasing in a follows directly from MLRP.

H increasing in z follows from the observation that $\frac{f(z|x)}{(1-\lambda)(1-F(z|x)) + \lambda(1-F(a^R|x))}$ is decreasing in x , which is proven in the same way as the decreasing hazard rate property.²⁴

H decreasing in λ for $z \leq a^R$ follows because $\frac{F(a^R|x) - F(z|x)}{\lambda(F(a^R|x) - F(z|x)) + 1 - F(z|x)}$ is increasing in x for $z \leq a^R$, which follows because $\frac{1 - F(a^R|x)}{1 - F(z|x)}$ is decreasing in x (which follows from the fact MLRP implies the first order stochastic dominance property noting that $\frac{F(a|x)}{1 - F(z|x)}$ is the cumulative of F conditional on x and a being no smaller than z and that MLRP still holds when we truncate the support of a). **Q. E. D.**

Proving that $a^S(\lambda)$ is smaller than a^R follows by noting that $H(a^R, a^R, \lambda) \geq H(a^R, 0, 0)$. Proving that $a^S(\lambda)$ is decreasing follows by noting that for an interior solution $H(a^S(\lambda), a^S(\lambda), \lambda) = c$, and thus if $\lambda' > \lambda$, $H(a^S(\lambda), a^S(\lambda), \lambda') \leq c$ (by the monotonicity of H in λ), which implies that $a^S(\lambda') \leq a^S(\lambda)$ (by the monotonicity of H in a and z). **Q.E.D.**

²⁴Specifically, integrate $f(a_1 | x_1)f(a_0 | x_0) \geq f(a_0 | x_1)f(a_1 | x_0)$ (which holds for all $a_1 \geq a_0$, $x_1 \geq x_0$) in a_1 from a_0 to \bar{a} and multiply by $1 - \lambda$ and integrate in a_1 from a^B to \bar{a} and multiply by λ to obtain that

$$\frac{f(a | x_0)}{(1 - \lambda)(1 - F(a^B | x_0)) + \lambda(1 - F(a | x_0))} \geq \frac{f(a | x_1)}{(1 - \lambda)(1 - F(a^B | x_1)) + \lambda(1 - F(a | x_1))}$$

as required.

References

- [1] Bénabou R. (2015): 'The economics of motivated beliefs,' *Revue d'économie politique* 125, 665-685.
- [2] Bénabou R. and J. Tirole (2002): 'Self-confidence and personal motivation,' *Quarterly Journal of Economics* 117, 871-915.
- [3] Cooper A., C. Woo, and W. Dunkelberg (1988): 'Entrepreneurs' perceived chances of success,' *Journal of Business Venturing* 3, 97-108.
- [4] Daniel, K. and D. Hirshleifer (2015): 'Overconfident investors, predictable returns, and excessive trading,' *Journal of Economic Perspectives* 29, 61-88.
- [5] Denrell, J. (2003): 'Vicarious learning, undersampling of failure, and the myths of management,' *Organization Science* 14, 227-243.
- [6] Elton E. J., M. J. Gruber, and C. R. Blake (1996): 'Survivorship bias and mutual fund performance,' *Review of Financial Studies* 9, 1097-1120.
- [7] Esponda, I. (2008): 'Behavioral equilibrium in economies with adverse selection,' *American Economic Review*, 98, 1269-1291.
- [8] Eyster E. and M. Rabin (2005): 'Cursed equilibrium,' *Econometrica*, 73, 1623-1672.
- [9] Grossman, S. J. and Stiglitz, J. E. (1976): 'Information and competitive price systems,' *American Economic Review* 66(2), 246-253.
- [10] Heckman J. (1979): 'Sample selection bias as a specification error,' *Econometrica* 47, 153-161.
- [11] Jehiel P (2005): 'Analogy-based expectation equilibrium,' *Journal of Economic Theory* 123, 81-104.
- [12] Jehiel P (2015): 'Investment strategy and selection bias: An equilibrium perspective on overconfidence,' discussion paper CEPR.

- [13] Kandel, E. and N. Pearson (1995): 'Differential interpretation of public signals and trade in speculative markets,' *Journal of Political Economy* 103, 831-672.
- [14] Köszegi, B. (2006): 'Ego utility, overconfidence, and task choice,' *Journal of the European Economic Association* 4, 673-707.
- [15] Lerner, J. and U. Malmendier (2013): 'With a little help from my (random) friends: Success and failure in post-business school entrepreneurship,' *Review of Financial Studies* 26, 2411-52.
- [16] Malmendier U. and G. Tate (2005): 'CEO overconfidence and corporate investment,' *Journal of Finance*, 60, 2661-2700.
- [17] Mankiw G., R. Reiss, and J. Wolfers (2004): 'Disagreement about inflation expectations,' *NBER Macroeconomics Annual* 18, 209-248.
- [18] Manski C. (2004): 'Measuring expectations,' *Econometrica* 72, 1329-1376.
- [19] Nisbett R. E. and L. Ross (1980): *Human inference: Strategies and shortcomings of social judgment*, Prentice-Hall, Englewood Cliffs, NJ.
- [20] Rabin, M. and J. L. Schrag (1999): 'First impressions matter: A model of confirmatory bias,' *Quarterly Journal of Economics* 114, 37-82.
- [21] Spiegel R (2016): 'Data monkeys: A model of naive extrapolation from partial statistics,' mimeo.
- [22] Streufert P. (2000): 'The effect of underclass social isolation on schooling choice,' *Journal of Public Economic Theory* 2, 461-482.
- [23] Tversky A. and D. Kahneman (1974): 'Judgment under uncertainty: Heuristics and biases,' *Science* 185, 1124-1131.
- [24] Van den Steen E. (2004): 'Rational overoptimism (and other biases),' *American Economic Review* 94, 1141-1151.
- [25] Xiong, X. (2013): 'Bubbles, crises, and heterogeneous beliefs,' *Handbook for Systemic Risk*, edited by J-P Fouque and J. Langsman, Cambridge University Press, 663-713.