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

Philippe Jehiel†

10th December 2015

Abstract

Prospective investors of new projects consider the returns of implemented projects with similar (observed) attributes 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. Assuming higher attributes are associated with higher returns, there is systematic overinvestment as compared with the Bayesian benchmark, thereby illustrating that selection bias may explain entrepreneurial overconfidence. Extensions are discussed allowing some share of investors to be rational or for correlation between the attributes considered by various investors.

*Thanks to Roland Bénabou, Milo Bianchi, Ignacio Esponda, David Hirshleifer, Charles Manski, Stephen Morris, and seminar participants at UCL behavioral game theory workshop, PSE, Princeton student workshop, Game theory French symposium, ESSET 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

Entrepreneurial activity consists in large part in deciding whether and when to make investments on the basis of the observation of a subset of the characteristics (attributes) of projects. Standard economics assumes that decision makers can relate the (observed) attributes to the distribution of returns and costs of the projects so that optimal (Bayesian) investment decisions can be made. Yet, it is not clear how decision makers can form such rational expectations, since mathematical representations of how attributes, returns and costs are jointly distributed are rarely accessible to decision makers.

This paper considers the steady states of economies in which decision makers sample all those implemented projects with the same set of attributes as those observed in their own project and consider the obtained empirical distribution of returns in the sample as being representative of the return they would obtain if they were to implement their own project (costs are assumed to be known in the model so that there is no need to form expectations about those).

Such a procedure to form expectations is quite commonly used in applied economics (see, for example, Freeman (1971) for the study of the returns to schooling among male students), essentially because it is in general impossible to have access to the returns that would have been obtained in projects that were not implemented (counterfactuals).[1] But it typically suffers from a bias insofar as the sample of those projects (with the same attributes) that were implemented need not be representative of all projects (with the considered attributes). To be more explicit, the fact that another project was implemented means that the decision maker in charge of that project observed attribute realizations that led him to make the investment decision, but the procedure does not take this extra information into account. It should also be stressed that to make the correct inference from the fact that the observed project was implemented, it would be required to know (possibly in a statistical sense) the attributes on which the decision maker of the observed project based his investment decision, but this knowledge too need not be accessible to

[1] Even if it were possible to have access to counterfactuals, one may argue that data about implemented projects are more vivid and more memorable so that they may have more impact on investors' judgments. Such a view is in line with the representativeness heuristic developed in psychology by Tversky and Kahneman (1974) which is sometimes used to explain the neglect of selection bias (see Han and Hirshleifer (2015) for a recent paper in this vein).

investors given that various investors would typically observe different attributes. Such a difficulty makes it more plausible that at least some share of decision makers would rely on the simple sampling heuristic suggested above.

The bias induced by the sampling heuristic is well identified in the econometrics literature and it is referred to as selection bias (see Heckman (1979), Angrist and Imbens (1991) or Manski (1993) for representative contributions on selection bias). Most of the econometrics literature either suggests methods to correct the bias (Heckman (1979) or Angrist and Imbens (1991))[2] or it suggests to elicit beliefs so as to avoid making assumptions on how decision makers form expectations (Manski (2004b)).

By contrast, this paper analyzes the implications of selection bias on the efficiency of investment decisions (as resulting from the sampling heuristic described above), and it explicitly allows different decision makers to make their decisions based on the observations of different attributes, which as will be observed is a key ingredient of the analysis. Specifically, in the basic model, it is assumed that, in addition to what every decision make may commonly observe, every investor observes one single (extra) attribute;[3] the chance that two investors consider the same extra attribute is negligible; conditional on the return, attributes are distributed identically and independently of each other; and a higher realization of attribute is more representative of a higher return in the sense that the joint distributions of returns and attributes satisfy the monotone likelihood ratio property.

In such a setting, it is shown that there is a unique steady state in the sampling heuristic economy. The steady state or equilibrium satisfies a threshold property. That is, a decision maker invests whenever the observed attribute realization a is above a threshold a^*. In the rational (or Bayesian) benchmark, optimal investment decisions also have a threshold property, but the threshold a^B above which investments decisions are made is typically different and larger, $a^B > a^*$. That is, there is more investment in the sampling heuristic world than in the rational world. Moreover, a decision maker observing any attribute realization has always an overly optimistic assessment of the resulting expected

[2]Some of the literature also suggests working with bounds when correcting fully the selection bias is not feasible (see, for example, Manski (2004a)).

[3]Per se, this is without loss of generality given that if a package of attributes is observed, one can always define such a package as a composite attribute.

return as compared with the rational benchmark (with the same observation). That is, investors are overconfident.[4] The intuition for these results is as follows. The sample on which decision makers base their estimates of the expected return is biased toward higher returns as compared with the rational Bayesian benchmark given that it consists of projects in which the observed attribute realization of the corresponding investor was above the threshold a^*, and higher attribute realizations are more representative of higher returns. The threshold is determined as a fixed point given that the threshold used by other investors is a determinant of the threshold used by a given investor (through the compositional effect on the sample) and these thresholds must coincide in steady state (since all attributes have the same informativeness).

While the above insight considers the optimistic biased assessment of investors in absolute terms (i.e., relative to the rational benchmark), I also note that an investor in my setting erroneously believes that on average his own projects when implemented would perform better that others' implemented projects. This is so because investors' views on others' implemented projects are correct (as they are observed by every investor), and investors are overly optimistic about the profitability of their own projects, as just highlighted. Thus, the findings of this paper are in agreement with a number of stylized facts about managerial overconfidence, suggesting that investors tend to be overly optimistic about their own projects but less so about others' projects (see, in particular, Cooper et al. (1988)).

Building on the basic model, I consider various extensions. First, I allow for the possibility that a positive share of decision makers would be rational while others would follow the sampling heuristic. I observe that the presence of rational investors aggravates the selection bias in equilibrium: There is even more investment on the part of the less sophisticated decision makers as one increases the share of rational investors (and overconfidence is also aggravated). The intuition for this result is as follows. As one gets surrounded with more rational investors, decisions are better, and thus sampling from better decisions -which lead to more investment when returns are higher- induces even more bias toward feeding the feeling that it is a good idea to invest no matter what at-

[4] I use in this paper overconfidence interchangeably with optimism. I do not consider overconfidence in relation to biased perceived precision of signals. Biased perceived precision would not affect behaviors in my model given that investors are assumed to be risk-neutral.

tribute realization is observed. Such a negative externality imposed by rational investors on sampling investors is a consequence of the equilibrium approach pursued in this paper, and it would typically be absent if the overconfidence bias were modelled in an exogenous fashion (say through well chosen subjective priors) to start with.

Second, I allow for the extra attributes considered by two investors to be sometimes (with positive probability) the same so as to reflect that, in localized economies, two neighbors may more likely consider the same extra attributes. I first observe that if the correlation is perfect in the sense that all investors consider the same extra attribute, then the investment strategy resulting from the sampling heuristic is rational, thereby illustrating the key role of the heterogeneity of the observations in deriving the above reported bias. I next observe allowing for partial correlation that the steady state no longer takes a threshold form, and there is a range of attribute realizations for which the subjective assessment of the expected return must coincide with the cost. Such a range of indifference would typically not arise in the rational benchmark, thereby illustrating another effect of the equilibrium approach pursued here.

Third, I extend the model to cover simple trading environments with sellers and buyers, thereby illustrating that there is more trade in the sampling heuristic world than in the rational benchmark. In particular, with the sampling heuristic, trade may arise for purely speculative reasons (whereas the no trade theorem would hold in the rational benchmark, see Milgrom and Stokey (1982)), and the heuristic may induce welfare superior outcomes in Akerlof-type environments. The overconfidence that arises due to selection bias in my setting can then potentially be used to shed a new light on the seemingly excessive volumes of trade observed in financial markets.[5]

In the final Section, I briefly discuss how a sophisticated investor or an econometrician would approach the selection bias causing overconfidence in the basic model. I also briefly consider, within the sampling heuristic framework, situations in which various types of attributes with different distributions would be available, and I note that a more informative attribute (in the sense of Blackwell) need not dominate (from a welfare viewpoint) a

[5] Such large volumes of trade are sometimes explained based on subjective prior assumptions which are interpreted through the lens of overconfidence (see Daniel and Hirshleifer (2015) for a recent exposition of this line of research). The equilibrium approach of my paper can be viewed as providing some elements to endogenize the subjective priors considered in that line of research.

less informative attribute due to selection bias. I also suggest in situations in which more informative attributes would be available to decision makers as the volume of investment increases that the sampling heuristic may be beneficial in that it may correct for the resulting externality, thereby better aligning the investment criterion with what would be optimal from a welfare viewpoint.

Related literature:

The above basic insight can be viewed as connecting the econometrics literature on selection bias to the idea that entrepreneurs may be overconfident in the sense of being too optimistic regarding the returns of their projects as compared with the rational Bayesian benchmark (see Cooper et al. (1988), Odean (1998), Heaton (2002), Malmandier and Tate (2005) or Koellinger et al. (2007) for field reports on managerial overconfidence, see also Camerer and Lovallo (1999) for an experiment on overconfidence).[6] The proposed explanation for overconfidence differs from previously proposed explanations (sometimes referring to other notions of overconfidence).[7] For example, Rabin and Schrag (1999) while considering a similar definition of overconfidence, 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. Van den Steen (2004) 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.[8] Finally, Bénabou and Tirole (2002) or Köszegi (2006) refer to overconfidence in a third sense. In their setting, beliefs are perfectly rational (Bayesian) based on the observed information, in contrast to the approaches of this paper or of Rabin and Schrag (1999).

[6] It should be mentioned that in Camerer-Lovallo's treatment in which subjects self-select to participate in the entry game experiment, there is more overconfidence due to subjects missing that other participants are also good at solving puzzles, another form of selection bias.

[7] Denrell (2003) and Hogarth and Kareliaia (2012) consider links between selection bias and managerial erroneous views (yet of a different nature than explored in this paper, see below), but they implicitly attribute the errors to outside observers rather than to investors, thereby not embedding their approach into an equilibrium approach as in this paper. Specifically, Denrell (2003) notes that if an observer only sees successful enterprises as data on failures are less visible, it may create the illusion that riskier projects perform better (for well chosen distributions of risks), and Hogarth and and Kareliaia (2012) note that if managerial assessments are noisy at the time of the investment decision, ex post it may be that implemented projects are not profitable, which may incorrectly be interpreted through the lens of overconfidence.

[8] Observe that such a relative dimension of overconfidence emerges also in my setting.

But information management (acquisition or deletion) is asymmetric following good or bad signal realization either due to ego-utility in Köszegi or time inconsistency in Bénabou-Tirole, thereby leading to a biased distribution of posteriors (at the decision time) that these authors identify with overconfidence.

The expectations formed by investors in the main model are erroneous and yet they are the results of equilibrium forces. As such, this paper relates to various recent equilibrium approaches with non-rational expectations, in particular the analogy-based expectation equilibrium that was first developed for extensive form games in Jehiel (2005) and later applied to games with incomplete information in Jehiel and Koessler (2008) and Ettinger and Jehiel (2010).[9] As it turns out, viewing Nature as a player determining the return of a project after the decision maker has made his investment decision allows me to connect the equilibrium approach developed in this paper with the analogy-based expectation equilibrium (see the end of Section 2).[10] The paper is also related to the behavioral equilibrium developed in Esponda (2008) which has been applied to voting games with asymmetric information and interdependent preferences in Esponda and Pouzo (2014). While the idea of selection bias is present in Esponda (2008) and Esponda and Pouzo (2014) (see also Esponda and Vespa (2015) for a related recent experiment), a key difference is that here selection bias arises in decision problems as a result of learning from others' experiences. By contrast, Esponda's work makes an essential use of multi-player environments in which payoffs are affected by others' actions and it assumes learning is based on one own's experience. Finally, Spiegler (2015) has recently proposed to import the tool of Bayesian network to model bounded rationality in decision problems. Yet, in his interpretation of his solution concept as a steady state, Spiegler implicitly assumes that all decision makers

[9] See also Piccione and Rubinstein (2003) for developing related ideas in the context of pattern recognition and Mullainathan et al. (2008), Jehiel and Samuelson (2012) or Eyster and Piccione (2013) for more recent publications relying on coarse reasoning of the sort covered by the analogy-based expectation equilibrium.

The sampling equilibrium of Osborne and Rubinstein (1998) is less closely related to the present paper since it relies on the bias induced by extrapolations from small samples. Also, less directly related is the valuation equilibrium of Jehiel and Samet (2007) since here every decision maker has only two possible moves (to invest or not to invest in a single investment problem).

[10] The approach in this paper has also some link to the cursed equilibrium given that cursed equilibrium and ABEE are connected (see the discussion in Jehiel and Koessler (2008)). Yet, given that the underlying interaction is one with a single decison maker (and not one with multiple players and interdependent preferences), the link to the cursed equilibrium is less tight.

rely on the same Directed Acyclic Graph to explain the probabilistic process, which does not fit with the present application that makes an essential use of the heterogeneity in how various decision makers form their beliefs.[11]

2 The investment problem

Each project is defined by $\theta = (x, a)$ where $x \in \mathbb{R}$ is the expected return and $a = (a_n)_{n=1}^N$ is the vector of N attributes of the project where N will typically be thought of as being large. Return realizations x are assumed to take finitely many values in X and $l(x)$ denotes the proportion of projects with return x. Attribute realizations a_i would also typically be assumed to take finitely many values in applications. To simplify the exposition of equilibrium, I will assume that they take values in the continuum. Specifically, each attribute a_i is distributed according to a distribution that varies with $x \in X$. Conditional on x, each a_i is distributed independently of each other, according to the same density $f(\cdot \mid x)$ with full support on $[0, 1]$.

I assume that there is a continuum of projects and one decision-maker for each project. Moreover, I assume that the empirical distribution of the ith attribute a_i of projects with return x has a density given by $f(\cdot \mid x)$. That is, some form of the law of large number is assumed to hold in this economy with a continuum of projects.[12]

Each decision maker observes one attribute. This attribute can be interpreted as the extra attribute observed by the decision maker in addition to the common set of attributes observed by everyone (see below comment 2). Having in mind that decision makers consider attributes independently of each other, the probability that two different decision makers consider the same attribute is $\frac{1}{N}$ and taking the limit as N goes to ∞, I will assume it is 0 in the analysis.[13] I will later on consider the case in which two different

[11] As it turns out, the personal equilibrium considered by Spiegler in his context can be viewed as an analogy-based expectation equilibrium of an extensive form game in which the various variables of interest are decided by different players. Analogy partitions should be defined so that they respect the causality graph in the Directed Acyclic Graph describing the Bayesian network.

[12] Such an assumption can be justified with a frequentist approach when a_i takes finitely many values. It is assumed to hold in the setup with a continuum of attribute realizations.

[13] In line with the interpretation of the model as the limit of one with finitely many attribute realizations, one should assume that N is large relative to the number of attribute realizations, and both this number and N should be large.

decision makers may consider the same attribute with positive probability, thereby also covering the case of finite N.

The cost of each project is c. The decision maker of project θ has to decide based on the observation of one attribute picked at random, say attribute realization a_i, whether or not to invest in the project. Decision makers are assumed to be risk-neutral.

I will later on discuss the Bayesian benchmark case in which there is investment whenever $E(x \mid a_i) \geq c$ where the expectation operator E is derived from $l(\cdot)$ and $f(\cdot \mid x)$ by simple composition. But, the main hypothesis is that the decision maker does not know $l(\cdot)$ and $f(\cdot \mid \cdot)$ to start with. Instead, he samples all projects from previous cohorts of decision makers in which an investment decision was made and the same attribute realization a_i as in the current project was observed. Importantly, I am assuming that only implemented projects are observed ex post. That is, θ is revealed when there is investment, but not otherwise. Based on the observation of past implemented projects with same (observed) attribute realization, the decision maker computes the empirical mean x from those projects and invests whenever the empirical mean does not fall short of the cost c. I am interested in analyzing the steady states of such a dynamic sampling process and how they compare to the rational benchmark in which decision makers would make the optimal investment decisions based on their observation.

Considering the empirical mean x of past projects conditional on those projects having the same ith attribute realization a_i as the current project of interest looks like a natural idea, and it would yield the correct (Bayesian) estimate if the sample were representative of all projects having the ith attribute realization a_i. The problem is that only those projects in which there is investment are observed (it seems legitimate to assume that it would be rather difficult to have access to the detailed characteristics of projects that were not implemented). This creates a selection bias in the sampled pool because the decision of other decision makers whether or not to invest cannot be considered to be independent of the variables of interest: Investments by others occur only when these observed attribute realizations that led them to think the expected return of their project would exceed the cost. A key objective of this paper is to understand the exact form taken by this bias.

Formally, a steady state is defined as follows. Let $q(a)$ denote the probability of investing when observing an ith attribute realization a, and assume that $q(a)$ is bounded away from 0 for a positive measure of a.[14] Let

$$\widehat{v}(a) = \frac{\sum_{x \in X} l(x) f(a \mid x) \int_0^1 q(b) f(b \mid x) db \cdot x}{\sum_{x \in X} l(x) f(a \mid x) \int_0^1 q(b) f(b \mid x) db}.$$

Given the assumed investment strategy $q(\cdot)$, $\widehat{v}(a)$ is the steady state mean return of projects with ith attribute realization a in which there was investment. To see this, observe that i) $l(x)f(a \mid x)$ accounts (up to an x-independent scalar) for the density of x-projects with ith attribute a, ii) $\int_0^1 q(b)f(b \mid x)db$ accounts for the steady state probability that an investor looking at an attribute other than the ith would invest in a project with return x, and iii) $\widehat{v}(a)$ is a weighted average of x where the weight of x is proportional to $l(x)f(a \mid x) \int_0^1 q(b)f(b \mid x)db$, the product of the density and probability defined in i) and ii) respectively.

Based on the above definition of $\widehat{v}(a)$, steady state requires that for each a, if $q(a) > 0$, one should have $\widehat{v}(a) \geq c$, and if $q(a) = 0$, one should have $\widehat{v}(a) \leq c$. That is, an investor observing an attribute realization a should decide to invest with positive probability only if the perceived expected return $\widehat{v}(a)$ exceeds the cost c.

Comments. 1) Instead of referring to attributes, one could instead assume that investors receive signals that are informative about returns. Such signals can be interpreted as the overall impressions investors get from their project. More precisely, one could assume that conditional on the return x, investors receive signals that are independently distributed across investors according to the conditional density $f(\cdot \mid x)$. Everything would work similarly, assuming that investors can generate for past projects the signals/impressions they would have gotten at the investment decision time had they been in charge of the past project (and the considered sample would consist of past projects with obtained overall impressions coinciding with the overall impression made by the current project). 2) Coming back to the attribute framework, it may be more realistic in a number of cases to assume that all investors commonly observe a subset of the attributes, say the first m attributes, of all projects, and in addition observe one extra randomly determined attribute of their own project. The same model as above would apply restricting

[14] If this is not so, the notion of steady state can be strengthened by imposing some trembling behavior.

attention to the pool of projects having the same realizations of the first m attributes as the current project of interest (one should redefine accordingly the values of $l(x)$ for this pool of projects).

Link to the analogy-based expectation equilibrium.

The above steady-state can be described using the framework of the analogy-based expectation equilibrium (Jehiel (2005)). Define the investment problem as follows. In stage 1, Nature chooses θ according to the joint distribution as defined by $l(\cdot)$ and $f(\cdot \mid \cdot)$, and Nature also determines according to a uniform distribution which attribute i the decision maker in charge of θ looks at. Let me denote a history at the end of stage 1 by a pair (θ, i). Then, in stage 2, for each (θ, i), the decision-maker chooses whether to invest or not based on the observation of $a_i(\theta)$, the realization of the i-th attribute in θ. If there is no investment, the interaction ends (with a 0 payoff for the decision maker). If there is investment, in stage 3, nature chooses/implements the return x as defined in θ and the final payoff of the investor is $x - c$. In order to form an expectation about the distribution of x if the decision maker chooses to invest at θ, the decision maker bundles all the nodes $((\theta', j), Inv)$ such that the i-th attribute of θ', $a_i(\theta')$ coincides with the observed attribute $a_i(\theta)$ in θ, and the decision maker at θ' chooses to invest. That is, the set of histories $Cl_{a_i} = \{((\theta', j), Inv) \mid a_i(\theta') = a_i\}$ is put into one analogy class for each a_i.

Steady states of the above dynamic sampling process correspond to the analogy-based expectation equilibria in this setting. Given this link, in the sequel I sometimes refer to such steady states as analogy-based expectation equilibria (ABEE) or more compactly as equilibria when there is no possible confusion.

3 Overconfidence as a result of selection bias

I analyze the above investment environment assuming that a higher realization of attribute is more representative of a higher return. More precisely, the following monotone likelihood ratio property is assumed to hold:

Assumption (MLRP): For any $a' > a$ and $x' > x$, it holds that: $\frac{f(a'|x')}{f(a|x')} > \frac{f(a'|x)}{f(a|x)}$.[15]

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

3.1 Equilibrium characterization

Proposition 1 *Under MLRP, there exists a unique steady state of the above investment environment. The equilibrium is such that for some threshold a^*, a decision-maker chooses to invest if the observed attribute realization a is above a^* and not to invest otherwise where a^* is uniquely defined by*

$$\frac{\sum_{x\in X} f(a^* \mid x)[1-F(a^* \mid x)]l(x)\cdot x}{\sum_{x\in X} f(a^* \mid x)[1-F(a^* \mid x)]l(x)} = \begin{cases} \geq c \text{ } \textit{if } a^*=0 \\ c \text{ } \textit{if } a^* \in (0,1) \\ \leq c \text{ } \textit{if } a^*=1 \end{cases} \tag{1}$$

Proof of Proposition 1:

Checking for equilibrium

Given the above strategy, the perceived value of investing for a given realization a of any attribute is:

$$\frac{\sum_{x\in X} l(x)f(a \mid x)[1-F(a^* \mid x)]\cdot x}{\sum_{x\in X} l(x)f(a \mid x)[1-F(a^* \mid x)]} - c$$

Putting aside the cost c of investment, the perceived expected return of the investment is a weighted average of the possible returns x where the weight of x is proportional to $l(x)f(a \mid x)[1-F(a^* \mid x)]$:

1. $l(x)f(a \mid x)$ is the usual Bayesian term accounting for the density of projects having return x and the realization a of the corresponding attribute.

2. $1-F(a^* \mid x)$ amounts for the proportion of investors in projects with returns x who made the investment decision. (Given the assumed equilibrium, it requires that those decision makers in charge of the project got a realization a above a^* in their observed attribute, hence the term $1-F(a^* \mid x)$.)

Thanks to MLRP, it is the case that $a \to \frac{\sum_{x\in X} l(x)f(a|x)[1-F(a^*|x)]\cdot x}{\sum_{x\in X} l(x)f(a|x)[1-F(a^*|x)]}$ is increasing,[16] which

[16]To see that the presence of $1-F(a^* \mid x)$ does not alter the monotonicity property, let $w(x) = p(x)[1-F(a^* \mid x)]$ and apply the standard textbook argument assuming the probability of x relative to that of x' is $w(x)/w(x')$.

then ensures that the above threshold investment strategy with a^*satisfying (1) defines an equilibrium.

Uniqueness

Suppose more generally, that in equilibrium, investment occurs with probability $q(a)$ when $a \in [0,1]$ occurs. Then the perceived expected return of an a project would be:

$$\widehat{v}(a) = \frac{\sum_{x\in X} l(x) f(a \mid x) \int_0^1 q(b) f(b \mid x) db \cdot x}{\sum_{x\in X} l(x) f(a \mid x) \int_0^1 q(b) f(b \mid x) db}$$

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

Define

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

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

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 $x \to \frac{\frac{\partial}{\partial z}[1-F(z|x)]}{1-F(z|x)} = \frac{-f(z|x)}{1-F(z|x)}$ increases with x. **Q.E.D.**

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 in equilibrium

$$H(z,z) = \begin{cases} \geq c \text{ if } z = 0 \\ c \text{ if } z \in (0,1) \\ \leq c \text{ if } z = 1 \end{cases} \tag{3}$$

Clearly, if $1 \geq z_1 > z_2 \geq 0$, then $H(z_1,z_1) > H(z_2,z_2)$ 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 z where z is uniquely defined to satisfy (3).[17]

Q.E.D.

3.2 Comparison to the Bayesian benchmark

A Bayesian (or rational) decision maker who would observe the realization a of an attribute would perceive that the expected return of the project is given by

$$v(a) = \frac{\sum_{x \in X} l(x) f(a \mid x) \cdot x}{\sum_{x \in X} l(x) f(a \mid x)}$$

He would invest if this value is above c and he would not otherwise.

Let $a^B \in [0, 1]$ be uniquely defined by[18]

$$v(a^B) = \begin{cases} \geq c \text{ if } a^B = 0 \\ c \text{ if } a^B \in (0, 1) \\ \leq c \text{ if } a^B = 1 \end{cases}$$

A rational investor invests whenever $a > a^B$ and he does not otherwise.

Using the $H(\cdot, \cdot)$ function introduced in the proof of Proposition 1 (see expression (2)), it is readily verified that $v(a) = H(a, 0)$, and thus a^B is uniquely defined by

$$H(a^B, 0) = \begin{cases} \geq c \text{ if } a^B = 0 \\ c \text{ if } a^B \in (0, 1) \\ \leq c \text{ if } a^B = 1 \end{cases}$$

There are two ways to think of a rational decision maker, as just described. One way is to hold the view that such a rational investor knows $l(\cdot)$ and $f(\cdot \mid \cdot)$ to start with and does the corresponding Bayesian updating when observing a. Another way is to hold the view that such 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

[17] Existence of such a z is guaranteed by the continutity of $z \to H(z, z)$ which ensures (by the theorem of intermediate values) that if $H(0, 0) < c < H(1, 1)$ there must exist a $z \in (0, 1)$ satisfying $H(z, z) = c$.

[18] Uniqueness comes again from MLRP, which ensures that $v(\cdot)$ is increasing in a.

the observed attribute) that delivers the highest expected payoff.[19] Whatever the interpretation of the rational investment strategy, comparing the analogy-based expectation equilibrium threshold a^* to a^B yields:

Proposition 2 *ABEE investors overvalue the expected returns of a-projects as compared with the rational benchmark, i.e. $\widehat{v}(a) \geq v(a)$ where $\widehat{v}(a) = H(a, a^*)$ and $v(a) = H(a, 0)$. There is at least as much investment in the steady state of the ABEE investment scenario as in the rational benchmark. That is, $a^* \leq a^B$.*

Proof of Proposition 2: The first part, $\widehat{v}(a) \geq v(a)$, is proven using the monotonicity of H in z (see Lemma 1). As for the second part, assuming $a^B < 1$ (if $a^B = 1$, the result trivially holds as there is no investment in the Bayesian benchmark case), one has that $H(a^B, 0) \geq c$ and thus $H(a^B, a^B) \geq c$ by the monotonicity of H in its second argument. This implies that $a^* \leq a^B$, as desired. **Q.E.D.**

The heuristic used by decision makers to assess the expected return of a-projects leads them to have an overly optimistic perception, since $H(a, a^*)$ is bigger than $H(a, 0)$ (the rational assessment). This in turn leads to excessive investment as compared to the rational benchmark.

It should be stressed that the excessive optimism of decision makers in our setup was not assumed in the first place. It is rather the result of the selection bias implicit in the heuristic procedure followed by decision makers, which leads them to consider that the observed mean return among adopted projects with the a characteristic is representative of the returns of all a-projects. This view held by decision makers is not correct because the mere fact that other decision makers decided to invest in their project is informative as to what kind of information they observed in their own project, which the heuristic procedure considered here does not correct for.[20] One implication of the approach pursued here is that the degree of overconfidence would endogenously depend on how attributes

[19] Assuming that such an investor focuses on threshold strategies, this only requires finding out the best cut-off value.

[20] I suspect that the overconfidence derived here would continue to hold beyond the MLRP setting, even if then equilibrium would not necessarily take a threshold form and multiplicity of equilibria could not be excluded.

and returns are jointly distributed. For example, when attributes are poorly informative of returns, the model would predict little overconfidence.

Comment. Some authors have reported evidence that entrepreneurs making several investment decisions may be more overconfident than completely novice entrepreneurs making their very first investment decisions (see Ucbasaran et al. (2011)). At first, this seems at odds with the content of Proposition 2 if one views serial entrepreneurs as being rational decision makers. Yet, an alternative view on serial but not excessively experienced entrepreneurs is that such investors would be exposed to more cases than novice entrepreneurs and as such would be more confident in the reliability of their observation, leading them to follow more the sampling heuristic (as compared with novice investors who may find too small datasets a bit too unreliable). With this view, one may reconcile such evidence on serial entrepreneurs with the approach of this paper (admittedly, further elaborations are required to model how investors would adjust their behaviors as a function of the size of the dataset).

3.3 Relative overconfidence

Another well documented bias is that while investors are overly optimistic regarding their own business, they tend to be not so when assessing others' businesses. For example, Cooper et al. (1988) report that while 81% of new business owners thought their business had a success rate above 70%, only 39% thought other businesses were this likely to succeed.

At first glance, it might seem that the selection bias considered in this paper would apply equally when considering one own's business or others' businesses. Yet, this view is incorrect because, on the one hand, investors have access to all implemented projects and as such they can easily have a correct assessment of the profitability of others' implemented projects, and, on the other hand, investors have an incorrect, overly optimistic, assessment of the profitability of their own projects (as shown in Proposition 2). Formally,

Proposition 3 *Assume all investors follow the sampling heuristic. Investors feel on average (before they know the attribute realization of their project) that they perform better than others.*

Proof of Proposition 3: The perceived expected payoff of an ABEE investor is given by $E(\widehat{v}(a) - c \mid a > a^*) \Pr(a > a^*)$ where the density of a is given by $w(a) = \sum_{x \in X} l(x) f(a \mid x)$ and $\widehat{v}(a^*) = c$. The average per-project performance of implemented projects is $E(\widehat{v}(a)) - c$. Given that $a \to \widehat{v}(a)$ is increasing and $\widehat{v}(a^*) = c$, we get the announced result. **Q. E. D.**

The relative overconfidence result of Proposition 3 arises because investors have the feeling they can screen (in an expected return enhancing way) among the pool of projects they consider as representative of all projects and this pool precisely coincides with the pool of implemented projects. To put it differently, in the relative comparison reported in Proposition 3, investors make only the mistake of not realizing that implemented projects are not representative of all projects when making their own investment decision, but they need not attribute to others the same selection bias insofar as to assess the profitability of implemented projects they can safely rely on the direct statistics about implemented projects that are available to them.

Comment. While the result of Proposition 3 applies in ex ante terms (before the investor knows the attribute realization of his own project),[21] it should be mentioned that a similar result would hold at the interim stage (after the investor has made his observations) for a wide range of attribute realizations, if one were to assume that investors observe more about their own project than about others' (which sounds plausible). For example, assume that investors observe no attribute about others' projects. If implemented projects make losses in aggregate (see below for an example of this), then an investor deciding to invest will always hold the view that his project will perform better than the average implemented project of others whatever the observed attribute realization $a > a^*$. This is because to implement a project, it must be that the investor (subjectively) finds his project profitable, and by assumption, implemented projects are not profitable on average (and this aggregate property is rightly perceived by investors).

[21] Such an ex ante comparison may be more compelling for serial (but not too experienced) entrepreneurs than for novice entrepreneurs (for whom the averaging over several decision problems is less relevant).

4 Further insights

4.1 An illustrative example

In order to illustrate more explicitly the above construction, consider the following setting: x can take two values $\underline{x}$ and $\overline{x}$ where $\underline{x} < c < \overline{x}$ so that ex post projects can be categorized as success (when $x = \overline{x}$) or failure (when $x = \underline{x}$). Both values of x are equally likely $l(\underline{x}) = l(\overline{x}) = \frac{1}{2}$. The cumulative of a conditional on x is given by $F(a \mid \underline{x}) = a$ and $F(a \mid \overline{x}) = a^2$ (so that MLRP is satisfied).

Routine calculations yield:

$$a^* = \min(1, \sqrt{\frac{1}{4} + \frac{c - \underline{x}}{2(\overline{x} - c)}} - \frac{1}{2})$$

and $a^B = \min(1, \frac{c-\underline{x}}{2(\overline{x}-c)})$. It is readily verified that $a^* < a^B$ whenever $\frac{c-\underline{x}}{\overline{x}-c} < 4$ (i.e., whenever $a^* < 1$).

In the equilibrium construction, nothing prevents investors from making losses in expectation given that the subjective assessments do not represent the correct expected return conditional on the observation. In the proposed example, ABEE decision makers make losses in aggregate whenever $4 > \frac{c-\underline{x}}{\overline{x}-c} > \frac{3}{2}$ and not otherwise. Whenever there are losses in aggregate, an ABEE investor observing $a > a^*$ would invest and consider his expected net profit $\widehat{v}(a) - c$ to be above the expected profit of implemented projects, which is negative.

4.2 When rational investors exert negative externalities

Suppose the population of investors is mixed. A share $1 - \lambda$ of investors (referred to as ABEE investors) proceeds as described in the main model: They observe one attribute of their project, sample all projects with the same attribute realization which were implemented, and invest if the observed empirical mean return exceeds the cost c. A share λ of investors makes the rational investment decision based on the observation of one attribute realization. Different investors observe different attributes.

It should be noted that in the sample considered by ABEE investors, there are both projects held by ABEE investors and by rational investors. Since the decision rule is not the same for ABEE 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 the cognitive heterogeneity on the performance of ABEE investors (rational investors are unaffected by the presence of ABEE investors given that they face a decision problem and they behave optimally).

To pave the way toward the main result of this section, observe that ABEE investors follow in equilibrium a threshold strategy that consists in investing in a project with attribute realization a only if a exceed a^{**} defined by

$$\frac{\sum_{x\in X} f(a^{**} \mid x)[(1-\lambda)(1-F(a^{**} \mid x)) + \lambda(1-F(a^B \mid x))]l(x)\cdot x}{\sum_{x\in X} f(a^{**} \mid x)[(1-\lambda)(1-F(a^{**} \mid x)) + \lambda(1-F(a^B \mid x))]l(x)} = \begin{cases} \geq c \text{ if } a^{**}=0 \\ c \text{ if } a^{**}\in(0,1) \\ \leq c \text{ if } a^{**}=1 \end{cases}$$

The left hand-side of this expression represents how an ABEE investor subjectively assesses the mean return of a project with attribute 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 an ABEE investor makes an observation of another project, with probability λ she is facing a rational investor who invests only if the observed attribute realization (a different one) is larger than the Bayesian threshold a^B (as defined at the end of subsection 3.2), and with probability $1-\lambda$ she is facing another ABEE investor who invests if the observed attribute realization (again a different one) is larger than a^{**}.

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

Proposition 4 *Under MLRP, the higher the share λ of rational investors, the more severe the pro-investment bias of ABEE investors. That is, $a^*(\lambda)$ is weakly decreasing in λ, and for all λ, $a^*(\lambda) \leq a^*(0) \leq a^B$.*

The proof of this Proposition as well as all subsequent propositions appears in the appendix. But, the intuition behind Proposition 4 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.

To illustrate the idea in an extreme way, suppose all investors around are not only rational but fully informed of the return x of their project (this is not so in the context of Proposition 4 - this is only assumed to make the point clearer). In this case, only projects with returns exceeding the cost c are implemented by other investors and the sampling procedure followed by the ABEE investor leads her to conclude that it is worth investing irrespective of the realization a of the attribute (because whatever the subsample considered all observed returns exceed the cost c). In this case, the ABEE investor would be so overconfident that she would invest always irrespective of the attribute realization she observes. As Proposition 4 shows, the bias gets more and more pronounced as there are more rational investors around even if rational decision makers are not fully informed of the return x of their project and get to observe an attribute of similar informativeness as ABEE investors.

Comment. While Proposition 4 shows that as λ increases, there is more investment among ABEE investors, the overall effect of an increase of λ on total investment (for both ABEE and rational investors) is ambiguous and depends on the assumed densities.

4.3 When the observations of decision makers are correlated

In the above analysis, I have assumed that decision makers observe different attributes of the project. If instead all decision makers observe the same attribute, then for every attribute realization a such that there is investment with positive probability $q(a) > 0$, the sampling heuristic leads to assess the expected return of an a-project according to

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

given that other investors would invest with probability $q(a)$ irrespective of x in a project with the same attribute realization a. This expression simplifies into the rational expression $v(a)$ derived in subsection 3.2, thereby leading to:[22]

Proposition 5 *When all decision makers consider the same attribute and follow the ABEE heuristic, when there is investment, this is the rational decision.*

The next step is to introduce some form of partial correlation between the attributes considered by two different decision makers. Intuitively, as one increases the correlation, one would expect to get closer to the rational benchmark considered in subsection 3.2, and as one decreases the correlation, one would expect to get closer to the solution analyzed in the baseline model above. While this is probably true in some sense, another more subtle insight will emerge from the analysis.

Consider the sampling scenario in which with probability μ the sampled decision maker relies on the same attribute and with probability $1 - \mu$ the sampled decision maker relies on a different attribute.[23] A rationale for this may be that decision makers are divided into a large number of communities, in each community a given and specific attribute is considered, and when sampling, a decision maker either samples from his own community (with probability μ) or he samples randomly (and uniformly) from the other communities (with probability $1 - \mu$).

In such a sampling scenario, it can be shown that under MLRP there is no interior threshold equilibrium. To see this, assume by contradiction that for some $a^* \in (0, 1)$, investment occurs whenever $a \geq a^*$.

For $a < a^*$, the perceived expected return of an a-project would be

$$\widehat{v}^L(a) = \frac{\sum_{x \in X} l(x) f(a \mid x)[1 - F(a^* \mid x)] \cdot x}{\sum_{x \in X} l(x) f(a \mid x)[1 - F(a^* \mid x)]} \tag{4}$$

since a sampled decision maker from the same community (i.e., considering the same attribute $a < a^*$) would not invest (and the corresponding project would not contribute to

[22] If one allows for trembles when there is no investment, then there is also no investment whenever it is suboptimal to invest.

[23] In the scenario with N different attributes, one would have $\mu = 1/N$.

the observed data), thereby explaining that the same expression of subjective assessment as in the main model holds.

For $a > a^*$, the perceived expected return of an a-project would be

$$\widehat{v}^H(a) = \frac{\sum_{x \in X} l(x) f(a \mid x)[(1-\mu)(1-F(a^* \mid x)) + \mu] \cdot x}{\sum_{x \in X} l(x) f(a \mid x)[(1-\mu)(1-F(a^* \mid x)) + \mu]} \tag{5}$$

where the term $(1-\mu)(1-F(a^* \mid x)) + \mu$ corresponds to the fact that an investor from the same community (who would look at the same attribute) would invest for sure (such an investor is sampled with probability μ) and an investor from another community (who looks at a different attribute) would invest only if the attribute realization is above the threshold a^*, that is, with frequency $1 - F(a^* \mid x)$ (such an investor is sampled with probability $1-\mu$).[24]

Yet, under MLRP, it is readily verified that $\widehat{v}^L(a) > \widehat{v}^H(a)$ for any $\mu > 0$ (this is because $\widehat{v}^H(a)$ is in between $\widehat{v}^L(a)$ and the rational evaluation $v(a)$ and $\widehat{v}^L(a) \geq v(a)$ as shown in Proposition 2). This in particular implies that $\widehat{v}^L(a^*) > \widehat{v}^H(a^*)$, which together with the continuity of the $a \to \widehat{v}^L(a)$ and $a \to \widehat{v}^H(a)$ functions makes it impossible that $\widehat{v}^H(a) \geq c$ for all $a > a^*$ and $\widehat{v}^L(a) \leq c$ for all $a < a^*$, as required for the threshold investment strategy to define an equilibrium.

The shape of the equilibrium in this sampling scenario should be modified as follows. For some $\underline{a}$, $\overline{a}$ where $0 \leq \underline{a} \leq \overline{a} \leq 1$, a decision-maker of an a-project does not invest whenever $a < \underline{a}$, always invests whenever $a \geq \overline{a}$ and invests with probability $q(a) > 0$ and does not invest with probability $1 - q(a) > 0$ whenever $\underline{a} < a < \overline{a}$ where for an interior equilibrium it should be that for all $a \in [0,1]$, the subjective expected return attached to an a-project satisfies

$$\widehat{v}(a) = \begin{cases} \geq c \text{ if } a > \overline{a} \\ c \text{ if } a \in (\underline{a}, \overline{a}) \\ \leq c \text{ if } a < \underline{a} \end{cases}$$

with

$$\widehat{v}(a) = \frac{\sum_{x \in X} l(x) f(a \mid x)[\mu q(a) + (1-\mu) K(x)] \cdot x}{\sum_{x \in X} l(x) f(a \mid x)[\mu q(a) + (1-\mu) K(x)]} \tag{6}$$

[24] It should be noted that when $\mu = 1$, $\widehat{v}^H(a)$ corresponds to the rational Bayesian assessment of the expected return of an a-project.

$$K(x)=\int_{\underline{a}}^{1}q(b)f(b\mid x)db \tag{7}$$

using the convention that $q(b)=1$ for $b\geq\overline{a}$.

Note that if investment decisions are governed by $q(\cdot)$, expression (6) corresponds to the subjective assessment of an a-investment given that an investor observing the same attribute a in the same community invests with probability $q(a)$ (such an investor is sampled with probability μ) and an investor from another community invests with probability $q(b)$ when her attribute realization is b which for a given x occurs according to the density $f(b\mid x)$.

It can be shown that such an equilibrium always exists.[25] Moreover, for an interior equilibrium, $\underline{a}$ (which is defined by $\widehat{v}(a)=c$ for $q(a)=0$) is such that $\underline{a}\leq a^{*}$ where a^{*} is the threshold defined in Proposition 1, and $\overline{a}$ (which is defined by $\widehat{v}(a)=c$ for $q(a)=1$) is such that $\overline{a}<a^{B}$ where a^{B} is the rational threshold as defined above.[26]

An interesting feature of the equilibrium is that for a range of attribute realizations with positive measure, i.e. for all $a\in(\underline{a},\overline{a})$, the subjective assessment of an a-project coincides exactly with the cost c. Such a property distinguishes the prediction of this model from that of the Bayesian model (for generic joint distribution of attributes and returns). Moreover, the range of indifference is maximal for interior values of μ (since as μ gets close to 0 or 1, the range shrinks respectively to $\{a^{*}\}$ or $\left\{a^{B}\right\}$). These observations can potentially be used for applied (empirical) purposes. For example, if it turns out

[25] To get a sense of this, define for every increasing function K

$$\widetilde{q}(a)=\frac{1-\lambda}{\lambda}\frac{\sum_{x\in X}p(x)f(a\mid x)K(x)\cdot(c-x)}{\sum_{x\in X}p(x)f(a\mid x)K(x)\cdot(x-c)} \tag{8}$$

and define

$$q(a)=\begin{cases}0\text{ if }\widetilde{\sigma}(a)<0\\ \widetilde{q}(a)\text{ if }0<\widetilde{\sigma}(a)<1\\ 1\text{ if }\widetilde{\sigma}(a)>1\end{cases}$$

Existence follows by noting that that when K is increasing, q as defined above is (weakly) increasing. Moreover for any increasing q, K as defined in (7) is an increasing function of x.

An equilibrium is a fixed point of the compound mapping induced by $K\rightarrow_{1}q$ (according to the (8) formula) and $q\rightarrow_{2}K$ (according to the (7) expression). Such a fixed point is shown to exist using the monotonicity properties just highlighted.

[26] These comparisons are established using the monotonicities mentioned in the previous footnote and the definition of a^{*} and a^{B}.

that for a number of different characteristics, investors find it just profitable to invest, this may be considered to be inconsistent with a Bayesian formulation but not with the ABEE model just proposed.[27]

5 Application to trading

In this Section, I extend the above setup to cover trading applications in which both sellers and buyers may possess private information regarding the value of the goods for sale. I wish to investigate for various prices of the good how the market works if sellers and buyers when forming their expectations about the quality of the goods may be subject to selection bias as considered above.

Specifically, the economy consists of a continuum of sellers and buyers of equal mass. Each seller is matched with one (different) buyer. Sellers and buyers jointly decide whether they wish to transact the good at price p (same price for all transactions). If the seller and the buyer both agree, there is transaction. If either the seller or the buyer refuses, there is no transaction. Prior to the transaction decision, the seller privately observes one among many attribute realizations $b \in [0,1]$ and the buyer privately observes one among many attribute realizations $a \in [0,1]$ of the good for sale.

A good may be of two qualities $\underline{x}$ or $\overline{x}$ assumed to be overall equally likely, $l(\underline{x}) = l(\overline{x}) = \frac{1}{2}$. The value of a quality x good is x for the seller and $h(x) \geq x$ for the buyer. I assume that $h(\overline{x}) \geq \overline{x} > p > h(\underline{x}) \geq \underline{x}$. That is, zero-sum trading environments are covered whenever $h(x) = x$, and so are situations in which the buyer values the good more than the seller for sure as in Akerlof's lemon market model.

The density of any attribute realization a given x is denoted by $f(a \mid x)$. Similarly, the density of b given x is denoted by $g(b \mid x)$. Cumulatives are denoted by $F(a \mid x)$ and $G(a \mid x)$, respectively.

I assume as in the main model that a buyer observing an attribute realization a (resp. a seller observing an attribute realization b) considers, from the previous cohort,

[27] Alternative possible approaches to rationalize indifferences would allow the decision makers to entertain multiple beliefs regarding the joint distribution of attributes and returns. Yet, typical criteria of ambiguity models would not deliver the indifference property except for non-generic specifications. Besides, such approaches would not *a priori* identify a role for the correlation of the attributes considered by two different investors as in this paper.

all transactions that were made and such that the same attribute realization a (resp b) was observed to compute the empirical mean value x in the sample. A buyer accepts the transaction whenever the subjective induced valuation exceeds the price. A seller accepts the transaction whenever the subjective induced valuation is below the price. One interpretation of the restriction to fixed price p mechanisms is that the price is an attribute observed by sellers and buyers, and thus sampling would only occur within those transactions that took place at price p.

Similarly to the benchmark model, the selection bias comes from the fact that only goods which were traded are inspected. I will also compare this scenario to the one in which both sellers and buyers are fully rational, and to the scenario in which sellers are fully rational while buyers follow the sampling heuristic just described. As in the main model, I will assume that both f and g satisfy the monotone likelihood ratio property (MLRP) where higher realizations of a and b are more representative of higher qualities x. I will also assume that $\frac{f(a|\overline{x})}{f(a|\underline{x})}$ converges to 0 as a gets close to 0 ($a \approx 0$ makes it very likely that $x = \underline{x}$) and $\frac{g(b|\overline{x})}{g(b|\underline{x})}$ converges to ∞ as b gets close to 1 ($b \approx 1$ makes it very likely that $x = \overline{x}$). These latter conditions ensure that if there is an equilibrium with positive trade (whatever the cognitive assumptions made on sellers and buyers), it must be interior. In the rest of this Section, when I refer to MLRP, I assume that all these statistical assumptions hold.

Proposition 6 *Under MLRP, assuming both sellers and buyers follow the ABEE heuristic, there is at most one equilibrium with positive trade. When it exists, such an equilibrium takes the following form. Buyers accept to transact whenever their observed attribute realization a exceeds a^*, sellers accept to transact whenever their observed attribute realization b is below b^* where the thresholds a^* and b^* are defined by:*

$$[1 - F(a^* \mid \overline{x})]G(b^* \mid \overline{x})g(b^* \mid \overline{x})(\overline{x} - p) = [1 - F(a^* \mid \underline{x})]G(b^* \mid \underline{x})g(b^* \mid \underline{x})(p - \underline{x}) \tag{9}$$

$$[1 - F(a^* \mid \overline{x})]G(b^* \mid \overline{x})f(a^* \mid \overline{x})(h(\overline{x}) - p) = [1 - F(a^* \mid \underline{x})]G(b^* \mid \underline{x})f(a^* \mid \underline{x})(p - h(\underline{x})) \tag{10}$$

Remark. In the zero-sum case, i.e. whenever $h(x) = x$, positive trade can be sustained. To see this most simply, allow to vary p in $(\underline{x}, \overline{x})$ without touching on the other parameters.

(9) and (10) can simultaneously hold for interior a^*, $b^* \in (0,1)$ and for some well chosen p as soon as

$$\frac{g(b^* \mid \overline{x})}{g(b^* \mid \underline{x})} = \frac{f(a^* \mid \overline{x})}{f(a^* \mid \underline{x})}$$

which can obviously be met for a range of interior a^*, b^*.[28]

In the next Proposition, I consider the fully rational benchmark.[29]

Proposition 7 *Under MLRP, when both sellers and buyers are rational, an interior equilibrium if it exists takes the following form. Buyers accept to transact whenever their observed attribute realization a exceeds a^R, sellers accept to transact whenever their observed attribute realization b is below b^R where the thresholds a^R and b^R are defined by:*

$$[1 - F(a^R \mid \overline{x})]g(b^R \mid \overline{x})(\overline{x} - p) = [1 - F(a^R \mid \underline{x})]g(b^R \mid \underline{x})(p - \underline{x}) \tag{11}$$

$$G(b^R \mid \overline{x})f(a^R \mid \overline{x})(h(\overline{x}) - p) = G(b^R \mid \underline{x})f(a^R \mid \underline{x})(p - h(\underline{x})) \tag{12}$$

In the next Proposition, I consider the scenario in which sellers are rational while buyers follow the ABEE heuristic. Such a scenario may be appropriate to describe situations in which the owners of the goods (the sellers) are more familiar with the goods they own than outside agents such as new buyers. Equilibrium is then characterized as follows.

Proposition 8 *Under MLRP, assuming that sellers are rational and buyers follow the ABEE heuristic, any interior equilibrium if it exists takes the following form. Buyers accept to transact whenever their observed attribute realization a exceeds a^{**}, sellers accept to transact whenever their observed attribute realization b is below b^{**} where the thresholds a^{**} and b^{**} should satisfy:*

$$[1 - F(a^{**} \mid \overline{x})]g(b^{**} \mid \overline{x})(\overline{x} - p) = [1 - F(a^{**} \mid \underline{x})]g(b^{**} \mid \underline{x})(p - \underline{x}) \tag{13}$$

$$[1-F(a^{**} \mid \overline{x})]G(b^{**} \mid \overline{x})f(a^{**} \mid \overline{x})(h(\overline{x})-p) = [1-F(a^{**} \mid \underline{x})]G(b^{**} \mid \underline{x})f(a^{**} \mid \underline{x})(p-h(\underline{x})) \tag{14}$$

[28] If $f = g$, then any $a^* = b^* \in (0,1)$ can be sustained for well chosen p.
[29] Observe that there may be multiple equilibria with positive trade in this case.

The next Proposition establishes that there is most trade when both the sellers and the buyers follow the ABEE heuristic and least trade when both the sellers and the buyers are rational. More precisely,

Proposition 9 *Under MLRP, most trade is achieved when both the sellers and the buyers follow the ABEE heuristic. Moreover, for any interior equilibrium in the rational case, there is an equilibrium with at least as much trade in the cognitive scenario in which sellers are rational and buyers follow the ABEE heuristic.*

What Proposition 9 establishes is that either there is no trade in all the equilibria of the cognitive scenarios considered in Propositions 6, 7, and 8 or else there is a unique interior equilibrium when both the sellers and the buyers follow the ABEE heuristic, and the corresponding volume of trade is the highest among all possible equilibria covered by Propositions 6, 7, and 8. Moreover, there is less trade in the rational case than in the scenario in which only the sellers are rational.

Intuitively, the reason for these results is as follows. A more tolerant trading behavior of one side of the market induces a more tolerant trading attitude on the other side. When one side follows the ABEE heuristic as opposed to the rational heuristic, it becomes more tolerant (as shown in Section 3). As a result, when both the sellers and the buyers follow the ABEE heuristic there is most trade, and when both the sellers and the buyers are rational there is least trade.

I now specialize the above results. In the zero-sum case, $h(x) = x$ for $x = \underline{x}$ and $\overline{x}$, the no-trade theorem applies in the rational paradigm (see Sebenius and Geanakoplos (1983)), but trade can take place in the ABEE world, as was noted after Proposition 6. In some instances, there may also be a positive trade equilibrium in the world in which only buyers follow the ABEE heuristic, but when it exists, it necessarily exhibits less trade than the interior equilibrium (which must then exist) of the ABEE world. In some sense, the rationality of sellers reduces the selection bias causing the overconfidence of ABEE buyers.

In the case in which $h(x) > x$, the above results reveal that the ABEE world may induce higher welfare given that under such specifications, the more trade the better for welfare. It should be noted though that in the case in which sellers would be perfectly informed of the quality x as in Akerlof (1970), there would be no trade in the ABEE world given that sellers would only agree to sell low quality goods and sampling from those transactions that were approved by sellers could only lead buyers to conclude that it is not worth buying. For trade to occur here, it is essential that sellers are not perfectly informed of the quality.[30]

6 Discussion

In this Section, I discuss various extensions of the basic model which can serve as the basis for possible avenues for future research.

6.1 A rational approach to selection bias

In the basic model, investors extrapolate from the sample of implemented projects as if these were representative of all projects. Instead, one may consider the behaviors of sophisticated yet inexperienced investors who would have to make the best inference from the (biased) sample they have access to.

There are various possible scenarios one may consider. Assume first that all investors are sophisticated and rightly know this as well as the fact that the chance that two investors would consider the same attribute is negligible. By looking at the empirical distribution of the characteristics θ of all implemented projects, such investors can identify that conditional on the return x all attributes are distributed independently of each other according to the density function $f(a \mid x)$. I claim that assuming all investors use the same threhsold strategy, such investors can recover the probabilities $l(x)$ that an arbitrary project would have return x, and as a result they can behave optimally, i.e. invest whenever $a > a^B$ and not invest otherwise where a^B is the rational threshold defined in subsection 3.2.

To see this, consider a generic investor who would postulate that all other investors use

[30] The finding that there is more trade in the ABEE world is to be contrasted with the finding in Esponda (2008) that naivete may exacerbate adverse selection in Akerlof-type situations. A key difference is that buyers cannot vary the price in my setting and selection bias comes from the attitude of other buyers, and not from a misunderstanding about how a change in the price affects the quality of the goods put for sale.

the threshold strategy z, consisting in investing if $a > z$ and not investing otherwise. For any postulated threshold z, considering the relative proportion of implemented projects with return x would allow such an investor to estimate the corresponding proportion $l_z(x)$ that a project would have return x given that the proportion of implemented projects with return x should then be up to an x-independent scalar equal to $l_z(x)[1-F(z \mid x)]$. Based on $l_z(\cdot)$ and $f(\cdot \mid \cdot)$, the investor could then determine the Bayesian threshold $a^B(z)$ as characterized in subsection 3.2 but now for probabilities of x being defined by $l_z(x)$ (instead of $l(x)$). Given that $z \to a^B(z)$ can be shown to be monotonically decreasing under MLRP, there is at most one fixed point $a^B(z) = z$, which must coincide with the Bayesian benchmark considered in subsection 3.2.

Another scenario one may consider is one in which there is a mix of sophisticated and naive investors where the naive investors behave as in the main model and the sophisticated investors try to make the best inference from what they observe. If the proportion of sophisticated investors is known to sophisticated investors, then an argument similar to the one just seen would allow the sophisticated investors to behave optimally as in the rational benchmark considered in subsection 3.2. If however only bounds on the proportion of sophisticated investors are known then sophisticated investors may not be able to infer how to play optimally. If they are sufficiently ambiguity averse, such investors may invest less than what would be optimal. One may also consider a third scenario in which sophisticated investors would not be sure regarding the degree of correlation of the attributes considered by two different investors in the vein of subsection 4.3. Variations of this sort should be the subject of future research.

The take-away of this is as follows. Modelling decision makers in the face of biased samples as econometricians would do (I referred to such decision makers as sophisticated) either leads to view such decision makers as being fully rational in the sense defined in subsection 3.2 or as facing an ambiguous environment, which would typically not give rise to the overconfidence documented in the managerial literature. Understanding better how a mix of sophisticated and naive investors would interact should be the subject of future research, especially when sophisticated investors cannot be assumed to infer how to behave in a Bayesian optimal way.

6.2 Multiple types of attributes

In the above investment problem, all attributes were distributed according to the same density. Consider the variant in which there are two types of attributes: the a-attributes distributed according to $f(a \mid x)$ and the b-attributes distributed according to $g(b \mid x)$ where both f and g are assumed to satisfy MLRP, and conditional on x any two attributes (whatever the type) are distributed independently of each other as in the main model.

Which type of attribute is preferable?

A question of interest is whether it is preferable (both from an individual and welfare viewpoints - they coincide here) that decision makers rely on a-attributes or b-attributes. In the classic Bayesian settting, if b is more informative than a about x (i.e. if a is a garbling of b), then Blackwell's analysis tells us that it is preferable that decision makers rely on a-attributes. In the sampling economy considered here, the comparison is not so clear given that decisions are made suboptimally on the basis of the observation.

To illustrate that it need not be best to consider the most informative attribute, consider the example developed in subsection 4.1 and assume that either attributes of type a are available or else uninformative attributes of type b (i.e. such that $g(b \mid \underline{x}) = g(b \mid \overline{x})$ for all b) are available. On the one hand, if attributes a are used, when $4 > \frac{c-\underline{x}}{\overline{x}-c} > \frac{3}{2}$, decision makers make aggregate losses, as noted in subsection 4.1. On the other hand, if attributes b are used, decision makers either invest always or they never invest, and if they invest always, they must be making gains in aggregate (given that the assessment of the expected return of any b-project would then coincide with the true aggregate expected return - there is no selection bias in this case). It follows that when all decision makers use the uninformative attributes b, they cannot make losses in aggregate and thus the uninformative attributes are preferable to the informative ones when $4 > \frac{c-\underline{x}}{\overline{x}-c} > \frac{3}{2}$.[31]

Equilibrium approach to the choice of attribute

Instead of comparing what agents obtain when they all use a-attributes or when they all use b-attributes, one may consider instead an equilibrium approach in the vein of

[31] From a more abstract perspective, one can show that there can be no ranking of any two distributions of attribute that applies uniformly for all shapes of the selection bias. That is, in our setting with selection bias, there is no extension of the partial ordering developed by Blackwell in the standard rationality case.

theoretical biology. That is, attributes of type a or b are now compared taking as given the distribution of attributes considered by other investors and an attribute a or b is retained only if it delivers the highest (true) expected payoff in the corresponding sampling environment.

Such an approach could now give rise to situations in which populations of decision makers would be split between considering various types of attributes despite being exposed to the same underlying investment decision problem.

To illustrate this, consider the same environment as just considered with a-attributes defined as in subsection 4.1 and b-attributes being uninformative. Suppose that when all decision-makers use a-attributes, they make aggregate losses. This ensures that it cannot be that all decision makers rely on a-attributes, since in such a configuration b-attributes -which would lead not to invest because there are aggregate losses- would deliver a higher expected payoff than a-attributes, thereby leading to a contradiction. On the other hand, if only b attributes are used, a decision maker proceeding as in the ABEE heuristic and considering an a attribute would not be subject to selection bias and would thus be making a higher expected profit than a decision maker relying on a b-attribute given that a is more informative than b.[32] This simple argument shows that equilibrium must be such that a-attributes and b-attributes give the same expected payoff and that a positive share of the population relies on each type of attribute.

6.3 When the investment rate affects the informativeness of available attributes

So far, the sampling heuristic was a source of inefficiency in the context of investment decisions. There may be various reasons why overconfidence can in some instances be beneficial (see, for example, Kyle and Wang (1997) or Compte and Postlewaite (2004)). To illustrate this possibility within the above ABEE investment framework, consider the following scenario. Suppose that the type of attribute that can be observed by decision makers is affected by the overall investment rate in the economy. More precisely, suppose that, as there is more entrepreneurial activity, the attribute observed by investors becomes

[32] Some trembling behavior is required to have positive investment in the b-regime.

more informative about the profitability of projects. A rationale for this may be that as there is more entrepreneurial activity in the economy, collecting information about projects becomes easier due to economies of scale.

It is not difficult to see that such kinds of externalities may, in some cases, give rise to multiple equilibria with some equilibria having more entrepreneurial activity, more informative attributes on which entrepreneurs base their investment decisions and more selection bias. In some specifications, the high entrepreneurial activity equilibrium may dominate the low entrepreneurial activity equilibria and also the best equilibrium that would arise in a setting with fully rational decision makers (the latter may be true even if there is a unique equilibrium). Such cases illustrate the potential benefit of the sampling heuristic as an instrument to induce more investment so as to better take advantage of the positive externality induced by a higher entrepreneurial activity.

To illustrate more concretely this, consider again the example of subsection 4.1 and suppose now that when attribute a is considered, investors make gains in aggregate (i.e., assume that $\frac{c-\underline{x}}{\overline{x}-c} < \frac{3}{2}$) while the expected return over all projects falls short of the cost c (i.e. $\frac{c-\underline{x}}{\overline{x}-c} > 1$). Assume also that either a-attributes are accessible or else only uninformative b-attributes are available. Finally, suppose that for a-attributes to be accessible, it should be that the share of projects in which there is investment is in between $\frac{1}{2}(1-a^B)+\frac{1}{2}(1-(a^B)^2)$ and $\frac{1}{2}(1-a^*)+\frac{1}{2}(1-(a^*)^2)$ where a^* and a^B were defined in subsection 4.1.

In such a scenario, in the rational world, there is a unique equilibrium in which a-attributes are not accessible and there is no investment.[33] In the ABEE world, there is an equilibrium in which a-attributes are available and decision makers invest according to the a^* threshold.[34] This equilibrium dominates the rational equilibrium from a welfare viewpoint given that $\frac{c-\underline{x}}{\overline{x}-c} < \frac{3}{2}$. There is also, in the ABEE world, the same equilibrium as in the rational case, thereby illustrating that the range in which multiple equilibria may arise in the two cognitive scenarios need not coincide.

[33] The absence of investment is the consequence of the assumption that when in the complete dark (i.e. faced with b only) it is not profitable to invest given the overall expected return of projects. The observation that a-attributes cannot be accessible is that if they were the share of investment would only be $\frac{1}{2}(1-a^B)+\frac{1}{2}(1-(a^B)^2)$, thereby not allowing a-attributes to be accessible.

[34] Assuming a-attributes are accessible, the share of investment would be $\frac{1}{2}(1-a^*)+\frac{1}{2}(1-(a^*)^2)$, thereby ensuring that a-attributes are accessible.

The above considerations suggest a positive link between the level of entrepreneurial activity and the level of managerial overconfidence. It would be of interest to explore more systematically if such a link prevails empirically.[35]

[35] Koellinger et al. (2007) is suggestive of such a correlation but more work is required to establish whether the causality suggested here is at work.

Appendix

Proof of Proposition 4.

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

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

Proof. H increasing in a follows directly from MLRP.

H increasing in z follows from the observation that $\frac{f(z|x)}{(1-\lambda)(1-F(z|x))+\lambda(1-F(a^B|x))}$ is decreasing in x, which is proven in the same way as the decreasing hazard rate property.[36]

H decreasing in λ for $z\leq a^B$ follows because $\frac{F(a^B|x)-F(z|x)}{\lambda(F(a^B|x)-F(z|x))+1-F(z|x)}$ is increasing in x for $z\leq a^B$, which follows because $\frac{1-F(a^B|x)}{1-F(z|x)}$ 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 $a^*(\lambda)$ is smaller than a^B follows easily by noting that $H(a^B,a^B,\lambda)\geq H(a^B,0,0)$.

Proving that $a^*(\lambda)$ is decreasing follows easily by noting that for an interior solution

$$H(a^*(\lambda),a^*(\lambda),\lambda)=c$$

and thus if $\lambda'>\lambda$, $H((a^*(\lambda),a^*(\lambda),\lambda')\leq c$ (by the monotonicity of H in λ), which implies that $a^*(\lambda')\leq a^*(\lambda)$ (by the monotonicity of H in a and z). **Q.E.D.**

Proof of Proposition 6

A buyer observing attribute a assesses the expected profitability of the asset according to

$$\frac{[1-F(a^*\mid\overline{x})]G(b^*\mid\overline{x})f(a\mid\overline{x})h(\overline{x})+[1-F(a^*\mid\underline{x})]G(b^*\mid\underline{x})f(a\mid\underline{x})h(\underline{x})}{[1-F(a^*\mid\overline{x})]G(b^*\mid\overline{x})f(a\mid\overline{x})+[1-F(a^*\mid\underline{x})]G(b^*\mid\underline{x})f(a\mid\underline{x})}$$

[36]Specifically, integrate $f(a_1\mid x_1)f(a_0\mid x_0)\geq f((a_0\mid x_1)f(a_1\mid x_1)$ (which holds for all $a_1\geq a_0$, $x_1\geq x_0$) in a_1 from a_0 to 1 and multiply by $1-\lambda$ and integrate in a_1 from a^B to 1 and multiply by λ to obtain that

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

as required.

where the term $[1 - F(a^* \mid x)]G(b^* \mid x)$ stands for the chance that there is a transaction in a sampled asset with profitability x. Equating this with p at $a = a^*$ (as required for an interior equilibrium) yields (10). The same logic applied to sellers yields (9). **Q. E. D.**

Proof of Propositions 7 and 8:

The difference between Proposition 7 and Proposition 6 lies in the way a seller assesses the expected profitability of an asset with attribute b that a buyer would agree to buy. This assessment is given by

$$\frac{[1 - F(a^R \mid \overline{x})]g(b \mid \overline{x})\overline{x} + [1 - F(a^R \mid \underline{x})]g(b \mid \underline{x})\underline{x}}{[1 - F(a^R \mid \overline{x})]g(b \mid \overline{x}) + [1 - F(a^R \mid \underline{x})]g(b \mid \underline{x})}$$

where the term $[1 - F(a^R \mid \overline{x})]$ stands for the conditioning on the fact that the buyer agrees to buy the asset (this is the traditional winner's curse correction). Equating this with p at $b = b^R$ yields (13). The other expressions are derived similarly. **Q. E. D.**

Proof of Proposition 9:

Start from an interior equilibrium (a^{**}, b^{**}) in the one-sided world in which only buyers follow the ABEE heuristic. Solving in b equation

$$G(b \mid \overline{x})[1 - F(a^* \mid \overline{x})]g(b \mid \overline{x})(\overline{x} - p) = G(b \mid \underline{x})[1 - F(a^* \mid \underline{x})]g(b \mid \underline{x})(p - \underline{x})$$

yields $b > b^{**}$ (by MLRP which implies that $G(b \mid \overline{x})/G(b \mid \underline{x})$ increases in b).

When b increases, the solution in a of the following equation

$$[1 - F(a \mid \overline{x})]G(b \mid \overline{x})f(a \mid \overline{x})(h(\overline{x}) - p) = [1 - F(a \mid \underline{x})]G(b \mid \underline{x})f(a \mid \underline{x})(p - h(\underline{x}))$$

decreases by MLRP.

Iterating from (a^{**}, b^{**}) these two operations leads to converge to an equilibrium in the ABEE world, hence (a^*, b^*) with $a^* < a^{**}$ and $b > b^{**}$.

The comparison between rational equilibria and equilibria in the one-sided world are done similarly. Start from a rational equilibrium (a^R, b^R). Solving in a

$$[1 - F(a \mid \overline{x})]G(b \mid \overline{x})f(a \mid \overline{x})(h(\overline{x}) - p) = [1 - F(a \mid \underline{x})]G(b \mid \underline{x})f(a \mid \underline{x})(p - h(\underline{x}))$$

yields an $a < a^R$.

When a is decreased, the solution in b of the following equation

$$[1 - F(a \mid \overline{x})]g(b \mid \overline{x})(\overline{x} - p) = [1 - F(a \mid \underline{x})]g(b \mid \underline{x})(p - \underline{x})$$

increases by MLRP. We conclude again by iterating these two operations. **Q. E. D.**

References

[1] Akerlof, G. (1970): 'The market for lemons: Quality uncertainty and the market mechanism,' *Quarterly Journal of Economics* 84, 488-500.

[2] Angrist J. and G. Imbens (1994): 'Identification and estimation of local average treatment effects,' *Econometrica* 62, 467-475.

[3] Bénabou R. and J. Tirole (2002): 'Self-confidence and personal motivation,' *Quarterly Journal of Economics* 117, 871-915.

[4] Camerer C. and D. Lovallo (1999): 'Overconfidence and excess entry: An experimental approach,' *American Economic Review* 89, 306-318.

[5] Compte O. and A. Postlewaite (2004): 'Confidence-enhanced performance,' *American Economic Review*, 94, 1536-1557.

[6] Cooper A., C. Woo, and W. Dunkelberg (1988): 'Entrepreneurs' perceived chances of success,' *Journal of Business Venturing* 3, 97-108.

[7] Daniel, K. and D. Hirshleifer (2015): 'Overconfident investors, predictable returns, and excessive trading,' *Journal of Economic Perspectives* 29, 61-88.

[8] Denrell, J. (2003): 'Vicarious learning, undersampling of failure, and the myths of management,' *Organization Science* 14, 227-243.

[9] Esponda, I. (2008): 'Behavioral equilibrium in economies with adverse selection', *American Economic Review*, 98, 1269-1291.

[10] Esponda, I. and D. Pouzo (2014): 'Conditional retrospective voting in large elections,' mimeo.

[11] Esponda, I. and E. Vespa (2015): 'Endogenous sample selection and partial naiveté: A laboratory study,' mimeo.

[12] Ettinger, D. and P. Jehiel (2010): 'A theory of deception,' *American Economic Journal: Microeconomics*, 2, 1-20.

[13] Eyster, E. and M. Piccione (2013): 'An approach to asset pricing under incomplete and diverse perceptions,' *Econometrica*, 73, 1623-1672.

[14] Eyster E. and M. Rabin (2005): 'Cursed equilibrium,' *Econometrica*, 73, 1623-1672.

[15] Han, B. and D. Hirshleifer (2015): 'Social transmission bias and active investing,' mimeo.

[16] Heaton J. B. (2002): 'Managerial optimism and corporate finance,' *Financial Management* 31, 33-45.

[17] Heckman J. (1979): 'Sample selection bias as a specification error,' *Econometrica* 47, 153-161.

[18] Hogarth R. and N. Karelaia (2012): 'Entrepreneurial success and failure: Confidence and fallible judgment,' *Organization Science* 23, 1733-1747.

[19] Jehiel P (2005): 'Analogy-based expectation equilibrium,' *Journal of Economic Theory* 123, 81-104.

[20] Jehiel P and F. Koessler (2008): 'Revisiting games of incomplete information with analogy-based expectations,' *Games and Economic Behavior*, 62, 533-557.

[21] Jehiel P and D. Samet (2007): 'Valuation equilibrium,' *Theoretical Economics* 2, 163-185.

[22] Jehiel P. and L. Samuelson (2012): 'Reputation with analogical reasoning,' *Quarterly Journal of Economics* 127, 1927-1969

[23] Kyle A. and A. Wang (1997): 'Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?' *Journal of Finance* 52, 2073-2090.

[24] Koellinger P., M. Minniti and C. Schade (2007): '"I think I can, I think I can": Overconfidence and entrepreneurial behavior,' *Journal of Economic Psychology* 28, 502-527.

[25] Köszegi, B. (2006): 'Ego utility, overconfidence, and task choice,' *Journal of the European Economic Association* 4, 673-707.

[26] Malmendier U. and G. Tate (2005): 'CEO overconfidence and corporate investment,' *Journal of Finance*, 60, 2661-2700.

[27] Manski C. (1993): 'Adolescent econometricians: How do youth infer the returns to schooling?,' in *Studies of Supply and Demand in Higher Education*, ed. by C. Clotfelter and M. Rothschild Chicago: University of Chicago Press, 43-57.

[28] Manski C. (2004a): 'Social learning from private experiences: The dynamics of the selection problem,' *Review of Economic Studies* 71, 443-458.

[29] Manski C. (2004b): 'Measuring expectations,' *Econometrica* 72, 1329-1376.

[30] Milgrom, P. and N. Stokey (1982): 'Information, trade, and common knowledge,' *Journal of Economic Theory* 26, 17-27.

[31] Mullainathan, S. J. Schwartzstein and A. Shleifer (2008): 'Coarse thinking and persuasion,' *Quarterly Journal of Economics*, 123, 577-619.

[32] Odean, T. (1998): 'Volume, volatility, price, and profit when all traders are above average,' *Journal of Finance* 53, 1887-1934.

[33] Osborne M. and A. Rubinstein (1998): 'Games with procedurally rational players,' *American Economic Review* 88, 834-849.

[34] Piccione, M. and A. Rubinstein (2003): 'Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns,' *Journal of the European Economic Association*, 1, 212-233.

[35] Rabin, M. and J. L. Schrag (1999): 'First impressions matter: A model of confirmatory bias,' *Quarterly Journal of Economics* 114, 37-82.

[36] Sebenius, J. K. and J. Geanakoplos (1983): 'Don't bet on it: Contingent agreements with asymmetric information,' *Journal of the American Statistical Association* 78, 424-426.

[37] Spiegler, R. (2015): 'Bayesian networks and boundedly rational expectations,' mimeo.

[38] Tversky A. and D. Kahneman (1974): 'Judgment under uncertainty: Heuristics and biases,' *Science* 185, 1124-1131.

[39] Uscbasaran, D., P. Westhead, and M. Wright (2011): 'When serial entrepreneurs don't learn from failure,' *Harvard Business Review* 89, 26-26.

[40] Van den Steen E. (2004): 'Rational overoptimism (and other biases),' *American Economic Review* 94, 1141-1151.