Overconfidence and diversification

Heller, Yuval

Tel-Aviv University

26 September 2010

Overconfidence and Diversification

Yuval Heller (Job Market Paper)

Abstract

Experimental evidence suggests that people tend to be overconfident in the sense that they overestimate the accuracy of their own predictions. In this paper we present a principal-agent model in which principal's interest in diversification motivates him to hire overconfident agents. We show that the induced overconfidence satisfies experimental stylized facts. In addition, we show that overconfidence is a unique evolutionarily stable strategy, and that it can Pareto-improve social welfare. Finally, we demonstrate applicability by demonstrating why CEOs hire overconfident intermediate managers, and why investors prefer overconfident entrepreneurs.

[1] This work is in partial fulfillment of the requirements for the Ph.D. in Mathematics at Tel-Aviv University, and it was supported by the Israel Science Foundation (grant number 212/09). I would like to thank Eilon Solan for his careful supervision, and for the continuous help he has offered. I would also like to express my deep gratitude to Eddie Dekel, Tzachi Gilboa, Zvika Neeman, Ady Pauzner, Ariel Rubinstein, Ran Spiegler, and Roee Teper for many useful comments, discussions and ideas. An earlier version of this paper was called "Overconfidence and risk dispersion".

December 6, 2010

1 Introduction

Many experimental studies ask participants to answer a two-alternative (usually general knowledge) question, and to report their confidence (subjective probability) that they answered the question correctly. The typical result in such experiments is that people are overconfident: their confidence systematically exceeds their true accuracy. In this paper, we present a theoretical model that studies the relation between overconfidence and diversification, and shows why principals prefer to hire overconfident agents in a variety of economic interactions. Applying the model to a biological framework, gives a new evolutionary foundation for overconfidence.

The introduction is structured as follows. The following subsection presents the main experimental findings about overconfidence. Subsection 1.2 presents motivating examples for our model, which is described in subsection 1.3. Subsection 1.4 discusses some extensions of our model, and Subsection 1.5 shows how our model is used in a few examples.

1.1 Experimental Findings on Overconfidence

The observed overconfidence in experiments usually satisfies a few recurrent properties (or effects). In the following paragraphs we describe the main observed properties. A more thorough discussion of the related experimental, empirical and theoretical literature is given in Section 2.

One of the main findings in the experimental literature is that the degree of overconfidence depends on the difficulty of the task - the *hard-easy effect*. The more difficult the task, the greater the observed overconfidence (Lichtenstein, Fischhoff, and Phillips, 1982; Moore and Healy, 2008). A few papers suggest that the hard-easy effect and apparent overconfidence in general may be the result of choosing unrepresentative hard questions in experiments (Gigerenzer, Hoffrage, and Kleinbolting, 1991), or regression toward the mean and boundary effects in the presence of unbiased judgmental random errors (Erev, Wallsten and Budescu, 1994; Soll, 1996). Later experiments demonstrate that people still present overconfidence (and the hard-easy effect), though to a less extent, when representative questions are used (which are randomly sampled from a natural set) and when unbiased judgmental random errors are taken into account in the analysis (see, e.g., Budescu, Wallsten, and Au, 1997; Klayman et al., 1999; Glaser, Langer and Weber, 2010).

Another finding is the *false certainty effect*: people are often wrong when they are certain in their private information. In the experiment of Fischhoff, Slovic, and Lichtenstein (1977) participants severely underestimated the probability they erred in seemingly easy questions. Specifically, the error probability of 10% of the questions was estimated by the subjects to be extremely low (less than 1:1,000), while the true error probability in these questions was approximately 10%. The participants had sufficient faith in their confidence judgments to be willing to stake money on their validity.

Griffin and Tversky (1992) suggest that many observed patterns of overconfidence (and underconfidence) can be explained by the *strength-weight effect*: "people focus on the strength or extremeness of the available evidence with insufficient regard for its weight or credence. This mode of judgment yields overconfidence when strength is high and weight is low, and underconfidence when strength is low and weight is high." (Griffin and Tversky, 1992, p. 411). They also show that people tend to *underuse base rate:* to overweight the strength of an evidence, and to underestimate the prior probability (similar findings appear in Brenner et al., 1996; Novemsky and Kronzon, 1999; Koehler, Brenner and Griffin, 2002; Brenner, Griffin, and Koehler, 2005).

Another interesting property is that people evaluate *external uncertainty* and *internal uncertainty* differently (Howell and Burnett, 1978; Kahneman and Tversky, 1982). External uncertainty is generated by sources external to the agent, such as: tossing a coin, the outcome of a future football game and the behavior of a volcano mountain. Internal uncertainty is generated by the state of knowledge of the agent. For example: whether or not Mont Blanc is the tallest mountain in Europe. The experimental literature suggests that people exhibit overconfidence mostly for evaluations of internal uncertainty (see, e.g., Budescu, and Du, 2007,). This overconfidence for internal uncertainty can be interpreted as overestimation of the accuracy of private information.

Some experiments (e.g., Gigerenzer, Hoffrage and Kleinbolting, 1991; Griffin and Tversky, 1992) compare people's confidence in giving correct answers by two methods: 1) each answer is evaluated separately (case by case evaluation), and 2) after answering several questions, participants are asked to evaluate the frequency of correct answers (set-based evaluation). These papers show that people exhibit less overconfidence (or even underconfidence) when evaluating set-based frequencies.

Finally, Sieber (1974) suggests that when more is at stake, people present more overconfidence. In her experiment, two groups of students were compared. One group was told that they were taking an their mid-term examination (important test). The other group was told that the test is not mid-term, but would be used to coach them to mid-term (unimportant test). The two groups had similar number of correct answers, but the group with the important test presented more overconfidence.

1.2 Illustrating Examples

The following three examples are presented to illustrate the abstract model that is described in the following subsection. Examples 1 and 2 are closely-related and deal with similar economic interactions, while Example 3 is qualitatively different and deals with a biological setup.

Example 1 *Consider a firm that operates in several regions. The success of the firm in each region depends on the marketing strategy of the local manager. A marketing strategy has higher success probability if it fits either the national trend in consumer's preferences, or* *the local trend in the region. Industry's experts evaluate the expected national trend, and each manager privately estimates the expected local trend in his region. Each manager is interested in maximizing the success probability in his region. The risk-averse CEO of the firm wishes to maximize the number of regional successes.*

Example 2 *Consider an angel investor who invests money in several founders of startup software companies (entrepreneurs). The success of each software depends on the platform in which it is developed to (e.g., Smartphone, social network, personal computer, etc.). Analysts evaluate which platforms are more likely to be popular. Each entrepreneur estimates which platform is more adapted to his software. Each entrepreneur wishes to maximize his success probability. The investor has diminishing marginal payoff from the total number of successful startup companies.*

Example 3 *Each individual has to choose his occupation. Popular opinion predicts which occupation is more likely to yield high fitness (e.g., fish population seems large, and being a fisherman seems good). Each individual privately estimates which action is best adapted to his skills (e.g., whether or not he has good hunting skills). Each type (gene) in the population induces a strategy for its members how to choose their occupation. When inducing such a strategy the type has a tradeoff between two objectives: 1) maximizing the fitness of each member, and 2) diversifying risk among its members.*

1.3 Model and Results

Our basic model (Section 3) includes a principal and many agents. Each agent is characterized by bias function g that determines how he evaluates the accuracy of private information: if the true accuracy of a noisy private signal (probability of being correct) is p, the agent believes it to be $g(p)$. The strategic interaction includes two stages. At stage 1 the principal observes the bias functions of the agents, and chooses which agents to hire. At stage 2, all agents receive a common public noisy signal with a known accuracy, and each agent i receives an independent private noisy signal with accuracy p_i, which he perceives as $g(p_i)$.[2] Then each agent chooses whether to follow the public signal or the private signal.[3] An agent receives high payoff if the signal he followed was correct (success) and low payoff otherwise (failure). The payoff of the principal is a concave increasing function of the average payoff of the agents.

Following the public signal bears a common risk - either all agents that follow it are successful or all of them are unsuccessful. This creates a conflict of interests between calibrated agents ($g(p) = p$) who maximize their probability of success, and the principal who also wishes to diversify risk among the agents. Our first result (Theorem 4) shows if the number of agents is sufficiently large, then this conflict is optimally resolved by hiring overconfident agents.

[2] The accuracies of the public signal and of each private signals are independently drawn from known distributions.

[3] The assumptions that the public signal is evaluated without a bias and that each agent has only two actions are without loss of generality as discussed in the following subsection.

That is, there is a continuous and increasing bias function g^*, which always overestimates the perceived accuracy of private information ($g^*(p) > p$ for every $0 < p < 1$, see Figure 1 in Section 3), such that if all agents have this bias function, it approximately induces the first-best outcome for the principal (the outcome he would achieve if he could receive all the private signals and directly control agent's actions). We further show (Theorem 6) that g^* is unique in the following sense: all other bias profiles, including heterogeneous profiles in which agents have different bias functions, induce strictly worse outcomes. [4]

Our third result (Theorem 7) shows that more risk-aversion induces more overconfidence. That is, if principal I is more risk-averse than principal II, then he hires more overconfident agents. Our next result (Theorem 8) presents further comparative statics under the mild assumption that the principal's utility satisfies decreasing absolute risk-aversion. It shows that the principal hires more overconfident agents if: 1) the payoff for success is higher; 2) the payoff for failure is lower; and 3) the task is harder in the sense that there is higher probability to receive less accurate private signals (first-order stochastic domination, see Figure 4). These results are in accordance with the stylized facts described earlier: the first two results fit the findings of Sieber (1974), and the third result fits the hard-easy effect.

Our last result (Theorem 9) assumes that the principal has a constant relative risk aversion (CRRA) utility . It shows that the optimal overconfidence bias only depends on the (relative) potential gain - $\frac{H-L}{L}$, where H is the payoff for success and L is the payoff for failure, and that a larger potential gain induces more overconfidence (see Figure 5). In addition we show that ratio between the perceived error probability and the true error probability of the private information is decreasing in the difficulty level. Moreover, if the payoff for failure is close to 0, then this ratio is very low for seemingly easy tasks in accordance with the false certainty effect.

1.4 Variants and Extensions

In Section 4 we present a few variants and extensions of the basic model. We first show that our assumption that the number of agents is exogenously given and large, can be relaxed. We allow the principal to choose the number of agents, and show that he prefers to hire many agents. In particular, we show that the principal always prefer to hire $k \cdot n$ agents instead of only n agents. [5]

Later, we show that our results hold when agents are informed experts who recommend the principal which action to choose. In addition, we show that our model can also describe situations in which agents have more than two actions. Specifically, each agent chooses an action out of a fixed set, and each signal is interpreted as describing which of these actions is the best one. We show that in this setup, having the optimal bias function g^* can also be interpreted as underusing base rates (described in Subsection 1.1).

[4] As demonstrated in Example 5, when the number of agents is small, our results do not hold.

[5] Example 11 shows that that the principal may prefer hiring 2 agents instead of 3.

Next, we adapt our results to a setup that is more appropriate to Examples 1-3, and is also valid when the agents can share their information before taking their actions. Specifically, each agent is successful by either choosing the best action (on average), or by choosing an action that is most-adapted for his characteristics. The public signal describes which action is more likely to be the best, and the private signal points on the action that is more likely to be most-adapted.

We then show that our assumption that agents only have bias for their private signals (and not for the public signal) is without loss of generality. When agents can have both biases, the unique optimal bias function g^* is the relative overestimation of private information with respect to the public information. This is interpreted as overconfidence, and it is closely related to the stylized facts discussed earlier: 1) the public signal may be induced by the aggregation of many "weak" signals of different agents; such aggregation induces less overconfidence due to the strength-weight effect; and 2) it is plausible that history of past games played by different agents encourage agents to treat the uncertainty of the public signal as external, and evaluate it by set-based frequencies, which decreases the overconfidence with respect to the public signal. A similar interpretation of overconfidence can be found in other models in the literature (see, e.g., Bernardo and Welch, 2001; Grubb, 2009).

Later, we extend our results to a setup where private information is costly, and each agent privately invests effort in improving the accuracy of his private signal. Finally, we show that our results hold in setups where the agents are more risk-averse than the principal.

1.5 Applying the Model in Examples

In Section 5 we apply our model in an evolutionary framework. As described earlier (Example 3), in each generation, each agent chooses an action, and this choice influences his fitness: he receives high payoff if he chooses the best action or his most-adapted action and low payoff otherwise. Each type (gene) induces a (possibly random) bias function for its members. In the long run the unique surviving type would be the one that maximizes the expectation of the logarithm of the average fitness of its members (see, e.g., Mcnamara, 1995, and the discussion in Section 5). Thus adapting our results to this framework, explains why overconfidence is evolutionarily stable. We further discuss predictions for different levels of overconfidence in different societies (see, e.g., the experimental findings in Yates et al., 2002), and the plausibility that the induced overconfidence satisfies the patterns that are observed in experiments (such as, the false certainty effect).

It is interesting to compare our evolutionary model for overconfidence with existing related evolutionary models. In Bernardo and Welch (2001)'s model a small proportion of individuals are overconfident, while the rest of the population are calibrated. Being overconfident reduces the personal fitness of an individual, but it substantially improves the fitness of his group, by inducing positive information externality in *cascade* interaction. Under the assumption that the evolutionary dynamics combines both group and individual selection, the evolutionarily stable profile includes a minority of overconfident individuals. Contrary to that, being overconfident in our model directly improves the type's fitness (without information externality), and it induces an evolutionarily stable profile in which all individuals are overconfident. Waldman (1994) shows that "second-best" adaptations can be evolutionarily stable with sexual inheritance, and demonstrates that the combination of overconfidence (overestimating self-ability) with excess disutility from effort is a "second-best" adaptation. Contrary to that, in our model overconfidence induces the "first-best" outcome, and does not compensate for another error.

In Section 6 we discuss Examples 1-2. Recall that Example 1 demonstrates why a risk-averse CEO prefers to hire overconfident intermediate mangers, who induce better diversification for the CEO. Observe that choosing overconfident intermediate managers is bad from the perspective of the risk-neutral shareholders, as it reduces the expected profit of the firm. This conflict of interests with the shareholders prevents the CEO from achieving diversification by formally allocating monetary incentives to calibrated intermediate mangers. It is interesting to compare this result with the model of Gervais, Heaton and Odean (2010), in which, given that the CEO is risk-averse, it is optimal for the risk-neutral shareholders, if the CEO is overconfident, and overestimates his ability to reduce risks.[6]

Recall that Example 2 shows why investors would invest their money in overconfident entrepreneurs, who induce better diversification for the investor.[7] This gives a new explanation to the high level of overconfidence that entrepreneurs present in experimental studies (see, Cooper , Woo, and Dunkelberg ,1988; Busenitz and Barney, 1997).

Section 6 also includes an example that shows how overconfidence can induce Pareto-optimal social outcomes, which Pareto-improve the outcome that is induced by calibrated agents.

1.6 Structure of the Paper

The structure of the remaining of this paper is as follows. Section 2 discusses the related literature. Section 3 presents the basic model and the results. A few extensions and variants are discussed in Section 4. Section 5 applies our model in an evolutionary framework. Section 6 includes examples for the applicability of the model in economic interactions.

2 Related Literature

The term "overconfidence" has been widely used in psychology since the 1960s, and in the economics and finance literature since the 1990s. Google Scholar reports on 876 papers that

[6] Goel and Thakor (2008) also study how a risk-averse CEO's overconfidence enhances firm's value.
[7] Observe, that expensive monetary incentives are needed to encourage a calibrated entrepreneur (who holds a large share of his startup companies) to reduce the investor's total risk and follow a nosier private signal.

include this term in their titles and about 40,000 papers that include it anywhere in the text (September 2010). In this section we briefly discuss a small portion of this literature.

The interested reader is referred to the following surveys on overconfidence: the classical survey of Lichtenstein, Fischhoff and Phillips (1982), which summarizes overconfidence literature in the 1960s and 1970s; the survey of Griffin and Brenner (2004) that summarizes the theoretical controversies about overconfidence, and the recent survey of Skala (2008).

2.1 Definitions of Overconfidence and Experimental Literature

The term overconfidence has been defined by three main ways in the literature. The most popular definition describes overconfidence as a systematic calibration bias, for which the assigned probability that the answers given are correct exceeds the true accuracy of the answers (see e.g., Oskamp, 1965; Lichtenstein, Fischhoff and Phillips, 1982; Brenner et al., 1996; Dawes and Mulford, 1996). This systematic bias is interpreted as overestimation of the accuracy of private information. As mentioned earlier, this is the definition we use in this paper.

A related definition of overconfidence is excessive certainty regarding the accuracy of one's beliefs about an uncertain continuous quantity. Researchers examining this effect typically ask their participants questions with numerical answers (e.g., "How long is the Nile River?"), and then have participants estimate (usually 90%) confidence intervals. Overconfidence is measured by the rate of surprises, i.e., the percentage of true values falling outside the confidence intervals. The typical finding (see Lichtenstein, Fischhoff and Phillips, 1982; Russo and Schoemaker, 1992) is that people tend to present substantial overconfidence: 90% confidence intervals contain on average only 50% of the true values. [8]

The third definition of overconfidence describes the phenomenon in which people believe themselves to be *better than average*. A review of this literature can be found in Alicke and Govorun (2005). A typical finding in this literature is the oft-quoted finding of Svenson (1981) that 77% of Swedish subjects felt they were safer drivers than the median. This bias is closely related to overly positive self-evaluations and to over-optimism about the future. Taylor and Brown (1988) report such phenomena to be positively correlated with different criteria of mental health. Recently, Moore (2007) and Benoit and Dubra (2008) suggest that most of the experimental findings of the better than average phenomenon can also be explained by a fully-rational Bayesian model.

Training improves overconfidence but usually only to a limited extent. Russo and Schoemaker (1992) show that asking people job relevant questions reduced overconfidence from 50% to 30% (for 90% confidence interval). Weather forecasters, who typically have several years of experience in assessing probabilities and receiving an immediate feedback, are quite well

[8] People also present overconfidence for 50% confidence intervals and for free-choice intervals, but this overconfidence is substantially smaller (Soll and Klayman, 2004; Teigen and Jorgensen, 2005).

calibrated (Lichtenstein, Fischhoff and Phillips, 1982; and also expert Bridge players - see Keren, 1987). Other experts such as physicians and professional traders, tend to present substantial confidence biases (see, e.g., Koehler, Brenner and Griffin, 2002; Glaser, Langer, and Weber, 2010).

Empirical data suggests that people present overconfidence not only in the lab but also in real-life situations. Russo and Schoemaker (1992) report the following example: "newly hired geologists were wrong much more than their levels of confidence implied. For instance, they would estimate a 40% chance of finding oil, but when ten such wells were actually drilled, only one or two would produce." Henrion and Fischhoff (2002) show that scientists systematically underestimate uncertainty in measurements of physical constants. Chuang and Lee (2006) empirically evaluate data on prices of firms in NYSE and AMEX during 1963-2001 and find evidence in support of the overconfidence hypothesis: investors overestimate accuracy of private information. Finally, Grubb (2009) analyzes consumer tariff choices and usage decisions of cellular services, and show that the consumers are overconfident in their ability to estimate their future demand for cellular services.

Examples predict overconfidence among entrepreneurs and managers. Cooper, Woo and Dunkelberg (1988) demonstrated that entrepreneurs perceive their prospects much more favorable than the true odds, and Busenitz and Barney (1997) showed that entrepreneurs substantially overestimate the probability that their answers for general-knowledge questions are correct. Recently, Ben-David, Graham and Harvey (2010) demonstrated that financial managers overestimate their ability to predict stock market returns.

2.2 Financial and Economic Models

In this subsection we briefly survey some related financial and economic models that deal with overconfidence.

A few papers study motivational reasons for overconfidence. Bénabou and Tirole (2002) present a multiple-self model, in which a rational agent tries to deceive his future self to be overconfident (overestimate his ability), in order to motivate him to undertake more ambitious goals and persist in the face of adversity. Compte and Postlewaite (2004) present a model in which positive emotions can improve performance, and individuals use biases in information processing that enhance their welfare. Köszegi (2006) and Weinberg (2009) model a decision maker who in addition to having preferences over material outcomes, also derives "ego" utility from positive self-image. In such a setup, moderate overconfidence raises the expected wealth.

Other papers study the evolutionary process that is generated by wealth that flows between investors in an asset market, and investigate the conditions in which overconfidence can survive or even dominate the market. Blume and Easley (1992) and Wang (2001) present models in which investors have high level of risk-aversion (or high discount factor), and overconfident investors can dominate the market due to trading more aggressively in the right way. Gervais and Odean (2001) show how a tendency of a trader to take too much credit for successes leads relatively-inexperienced successful traders to become overconfident. With more experience, investors better recognize their abilities. In markets where inexperienced traders continuously enter and old traders die, there will always be overconfident traders, and these traders will tend to control more wealth than their less confident peers.

Van den Steen (2004) models "rational overconfidence". Agents have an unbiased random error when evaluating their success probability for each possible action. When such agents face a choice among a few actions, they are more likely to select actions for which they overestimate the probability of success. Thus they will tend to be overconfident about the likelihood of success of the actions they undertake.

A few papers study the influence of overconfident agents on different markets. Odean (1998) shows that overconfidence among investors in financial markets increases expected trading volume, increases market depth, and decreases the expected utility of overconfident trader. Sandroni and Squintani (2007) show the the presence of some overconfident agents qualitatively change the equilibrium and the policy implications in insurance markets with asymmetric information.

3 Model and Results

3.1 Model

Let $I = \{1, ..., n\}$ be a set of agents. A typical agent is denoted by $i \in I$. The unknown state of nature determines the value of the tuple of random variables $\left(\mathbf{q}, (\mathbf{p}_i)_{i\in I}, \mathbf{m}_q, (\mathbf{m}_i)_{i\in I}\right) \in \left([0,1] \times [0,1]^I \times \{0,1\} \times \{0,1\}^I\right)$ as follows:

- $\mathbf{q} \sim f_{\mathbf{q}}$, where $f_{\mathbf{q}}$ is a continuous pdf (probability density function) with a full support: $f_{\mathbf{q}}(q) > 0$ for every $q \in [0,1]$.[9] $\mathbf{q}$ is interpreted as the accuracy of the public signal.
- For each $i \in I$, $\mathbf{p}_i \sim f_{\mathbf{p}}$, where $f_{\mathbf{p}}$ is a continuous pdf with full support: $f_{\mathbf{p}}(p) > 0$ for every $p \in [0,1]$. Let $F_{\mathbf{p}}$ be its cumulative distribution function (cdf). $\mathbf{p}_i$ is interpreted as the accuracy of the private signal of agent i. The variables $\left(\mathbf{q}, (\mathbf{p}_i)_{i\in N}\right)$ are independent.
- $\mathbf{m}_q$ is equal to 1 with probability $\mathbf{q}$ (and 0 otherwise). $\mathbf{m}_q = 1$ ($\mathbf{m}_q = 0$) is interpreted as the event where the public signal is correct (incorrect), and following it yields all agents high (low) payoff.
- For each $i \in I$, $\mathbf{m}_i$ is equal to 1 with probability $\mathbf{p}_i$ (and 0 otherwise). $\mathbf{m}_i = 1$ ($\mathbf{m}_i = 0$) is interpreted as the event where the private signal of agent i is correct (incorrect), and following it would yield agent i high (low) payoff. The variables $\left(\mathbf{m}_q, (\mathbf{m}_i)_{i\in N}\right)$ are independent.

[9] The full support assumption is given to simplify the presentation of the results. The results are qualitatively unaffected by relaxing this assumption.

The strategic interaction between the principal and the agents includes two stages. At stage 1 the principal (who has no information on the state of nature) chooses a profile of bias functions $(g_i)_{i\in I}$. Each function $g_i : [0,1] \to [0,1]$ determines the bias of agent i when estimating accuracy levels of private signals. That is, if the private signal of agent i has accuracy $\mathbf{p}_i$, he mistakenly believes it to have accuracy $g_i(\mathbf{p}_i)$.[10] The choice of the bias profile $(g_i)_{i\in I}$ is interpreted as follows: there is an infinite pool of potential agents with all possible bias functions. The principal can observe these biases, and choose $|I|$ agents with any given profile of bias functions.[11] After stage 1, all agents are publicly informed about the value of $\mathbf{q}$ (the accuracy of the public signal), and each agent i with bias function g_i, is (mis-)informed that the value of $\mathbf{p}_i$ is $g_i(\mathbf{p}_i)$.

At stage 2 each agent i chooses an action $a_i \in \{a_{pub}, a_{pri}\}$,[12] where a_{pub} (a_{pri}) is interpreted as following the public (private) signal. The payoff of agent i is as follows:

$$\mathbf{u}_i(a_{pub}) = u_i(a_{pub}) = \begin{cases} H & \text{if } \mathbf{m}_q = 1 \\ L & \text{if } \mathbf{m}_q = 0 \end{cases}, \text{ and } \mathbf{u}_i(a_{pri}) = \begin{cases} H & \text{if } \mathbf{m}_i = 1 \\ L & \text{if } \mathbf{m}_i = 0 \end{cases},$$

where $H > L > 0$. That is, the agent receives high payoff (H) if the signal he followed was correct (success), and low payoff (L) otherwise (failure). Let $D = \frac{H-L}{L}$ be the (normalized) potential gain: the ratio between the potential gain from following a good signal ($H - L$) and the minimal guaranteed payoff (L).

Our assumption that f_p and f_q are continuous guarantee that the inequality $\mathbf{q} \neq g(\mathbf{p}_i)$ holds with probability 1. Thus, each bias profile $(g_i)_{i\in I}$ induces a strictly-dominating strategy profile for each agent i: following the public signal a_{pub} if $\mathbf{q} > g_i(\mathbf{p}_i)$, and following the private signal if $\mathbf{q} < g_i(\mathbf{p}_i)$.[13] Let $\mathbf{u}_i(g_i) = u_i(g_i, \mathbf{p}_i, \mathbf{q}, \mathbf{m}_q, \mathbf{m}_i)$ be the random variable that describes the payoff of agent i while using this strictly-dominating strategy.

The payoff of the principal, $u\left((g_i)_{i\in I}\right)$, is a vN-M (John von-Neumann and Oscar Morgenstern, 1944) strictly concave increasing function of the average payoff of the agents:

$$u\left((g_i)_{i\in I}\right) = \mathbf{E}_{(\mathbf{p}_i)_{i\in I}, \mathbf{q}, \mathbf{m}_q, (\mathbf{m}_i)_{i\in I}}\left(h\left(\frac{1}{n}\sum_{i\in I}\mathbf{u}_i(g_i)\right)\right)$$

where $h' > 0$ and $h'' < 0$ in $[0,1]$.

Bias profile $(g_i^*)_{i\in I}$ is ϵ-optimal (for $\epsilon > 0$) if it yields the best payoff up to ϵ: $u\left((g_i^*)_{i\in I}\right) > u\left((g_i)_{i\in I}\right) - \epsilon$ for every profile $(g_i)_{i\in I}$. Let the first-best payoff the game, be the payoff that

[10] The assumption that each agent only has bias with respect to his private signal (but not with respect to the public signal) is without loss of generality, as discussed in Subsection 4.6.

[11] The number of agents the principal hires is exogenously given in the basic model. In Subsection 4.1 we extend the model to allow the principal to choose the number of hired agents.

[12] The assumption that each agent has only two actions is essentially without loss of generality, as discussed in Subsections 4.3 and 4.5.

[13] and playing arbitrary if $\mathbf{q} = g(\mathbf{p}_i)$ (a 0-probability event).

can be achieved by the principal if he would obtain all the private signals and have full control over the agents' actions. A bias profile ϵ-induces the first-best payoff, if its payoff is as good as the first-best payoff up to ϵ.

Bias profile $(g_i)_{i\in I}$ is homogeneous (or symmetric) if all agents have the same bias function: $\forall i, j \in I,\ g_i = g_j$. With some abuse of notations, we identify a function $g : [0,1] \rightarrow [0,1]$ with the homogeneous profile $(g)_{i\in I}$. We say that g is an *optimal bias function* (for large number of agents) if for every $\epsilon > 0$, there is large enough n_0 such that for any game with at least n_0 agents, g is an ϵ-optimal profile. Similarly, we say that g *induces the first-best payoff* (for large number of agents) if for every $\epsilon > 0$, there is large enough n_0 such that for any game with at least n_0 agents, g ϵ-induces the first-best payoff.

3.2 Overconfidence as a unique Optimal Bias Profile

The following theorem characterizes the optimal bias function (all proofs are given in the appendix). It shows that there exists a unique optimal bias function g^* that reveals overconfidence: $g^*(p) > p$ for every $0 < p < 1$. Moreover, this overconfidence bias induces the principal's first-best payoff.

Theorem 4 *There exists a unique optimal bias function g^*, which induces the first-best payoff, with the following properties:*

(1) g^* is continuous, $g^*(0) = 0$, and $g^*(1) = 1$.
(2) g^* is increasing: $\frac{dg^*(p)}{dp} > 0$ for every $0 < p < 1$.
(3) Overconfidence: $g^*(p) > p$ for every $0 < p < 1$.

The proof of all the results are given in the appendix. The intuition for Theorem 4 is as follows. There is a conflict of interest between a calibrated agent ($g(p) = p$) who maximizes his probability of success, and the principal who wishes some agents with $\mathbf{p}_i < \mathbf{q}$ to follow the private signal in order to diversify and to reduce the variance of the number of successes. Choosing which signal agent i would follow in the principal's first-best action profile, generally depends on the entire realized profile of private accuracies: $(\mathbf{p}_1, .., \mathbf{p}_n)$. However, when there are many agents, the realized empirical distribution of accuracies is very close to its prior distribution $f_{\mathbf{p}}$. Thus, approximately, the optimal choice of agent i only depends on the realizations of $\mathbf{p}_i$ and $\mathbf{q}$. Specifically, for every q, there is some threshold level $g^{-1}(q) < q$ such that it is approximately optimal for the principal if an agent would follow the private signal if and only if $\mathbf{p}_i > g^{-1}(q)$. These threshold construct the optimal bias function $g(p)$.

The following example shows that Theorem 4 is not valid when the number of agents is small. In this example, there is an asymmetric bias profile that induces higher payoff than the best bias function, and in addition the first-best outcome is substantially better than what can be induced by bias profiles.

Example 5 *There are two agents. Let the low payoff be zero ($L = 0$), and the high payoff* *one* $(H = 1)$*. Let the principal's utility be:*[14]

$$h(x) = \begin{cases} 2x & \text{if } x < 0.5, \\ 1 & \text{if } x \geq 0.5. \end{cases}$$

That is, the principal wishes that at least one agent succeeds (but does not care if both agents succeed or only one of them). Let the distribution of the accuracy of the private signals be uniform on $(0, 0.5)$*:* $f_{\mathbf{p}} \sim Uniform(0, 0.5)$*. Consider the case in which the accuracy of the public signal is* 0.7*. One can see that the best bias function (i.e., symmetric bias profile) is one such that (approximately)* $g(0.34) = 0.7$*,*[15] *and that it induces payoff 0.75. The principal can achieve higher payoff of 0.775 by using the following optimal heterogeneous bias profile: one agent always follow the public signal while the other agent always follow the private signal. The principal's first best payoff is even higher - 0.8, and it is achieved by observing both private accuracies, and having the agent with the higher (lower) accuracy follow the private (public) signal.*

Figure 1 demonstrates how an optimal bias function g^* looks like (for potential gain $D = \frac{H-L}{L} = 3$, principal's utility $h(x) = ln(x)$, and a uniform distribution for the accuracy of the private signal).

Theorem 4 shows uniqueness in the set of homogeneous bias profiles. That is, it shows that any other homogeneous bias profile induces a worse outcome than g^*, given that the number of agents is sufficiently large. The following theorem extends the uniqueness also to the set of heterogeneous profiles. It shows that every heterogeneous profile can be replaced with an homogeneous profile that induces a strictly better outcome, given that the number of agents is sufficiently large.

Two auxiliary definitions are needed for stating Theorem 6. Bias profile $(g_i)_{i \in I}$ is *heterogeneous* if there a set $Q \subseteq [0, 1]$ with a positive Lebesgue measure such that for each $q \in Q$, $\min_i (g_i)^{-1}(q) < \max_i (g_i)^{-1}(q)$. With some abuse of notations, we identify the bias profile $(g_i)_{i \in I}$ with the following bias profile in a game with $k \cdot |I|$ agents: agents $\{1, ..., k\}$ have bias function g_1, agents $\{k+1, ..., 2k\}$ have bias function g_2, ..., agents $\{k \cdot (|I| - 1) + 1, ..., k|I|\}$ have bias function $g_{|I|}$.

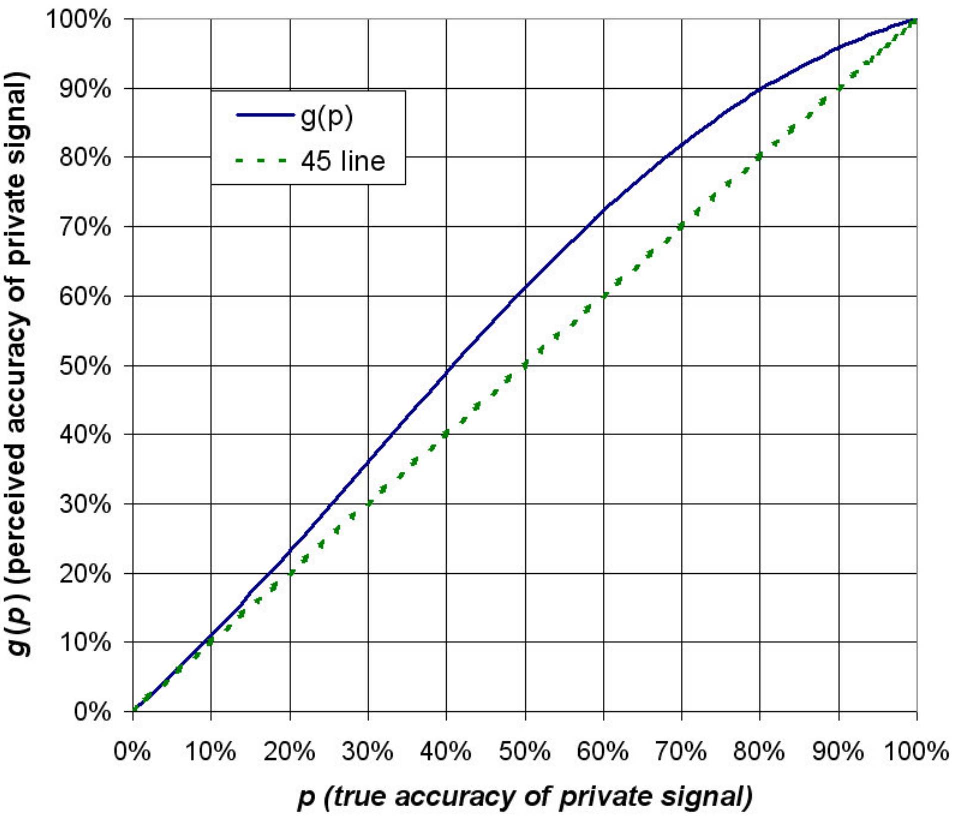
Theorem 6 *Let* $(g_i)_{i \in I}$ *be an heterogeneous profile. Then there is* k_0 *such that there is an homogeneous profile that induces a strictly better payoff than* $(g_i)_{i \in I}$ *in the game with* $k \cdot |I|$ *agents for every* $k \geq k_0$*.*

The intuition for Theorem 6 is as follows. Let g be a bias function (an homogeneous profile) that induces the same expected number of agents that follow the public signal for every

[14] To simplify the example we use a weakly concave and increasing function h, and a distribution $f_{\mathbf{p}}$ without full support. The example can be adapted such that h would be strictly concave and increasing and $f_{\mathbf{p}}$ would have full support.

[15] $p_0^* = 0.34$ maximizes the expression: $F^2(p_0) \cdot 0.7 + 2 \cdot (1 - F(p_0)) \cdot F(p_0) \cdot (0.7 + 0.3 \cdot \mathbf{E}(p|p > p_0)) + (1 - F(p_0))^2 \left(1 - (1 - \mathbf{E}(p|p > p_0))^2\right)$.

Figure 1. An Example for an Optimal Confidence-Bias Function
($D = 3$, $h(x) = ln(x)$, uniform distribution)



$0 < q < 1$. One can show that the expected average accuracy level of the private signals that are followed is strictly higher given g than given the heterogeneous profile $(g_i)_{i \in I}$. If the number of agents is sufficiently large, then the law of large numbers imply that$g$ induces a strictly better payoff.

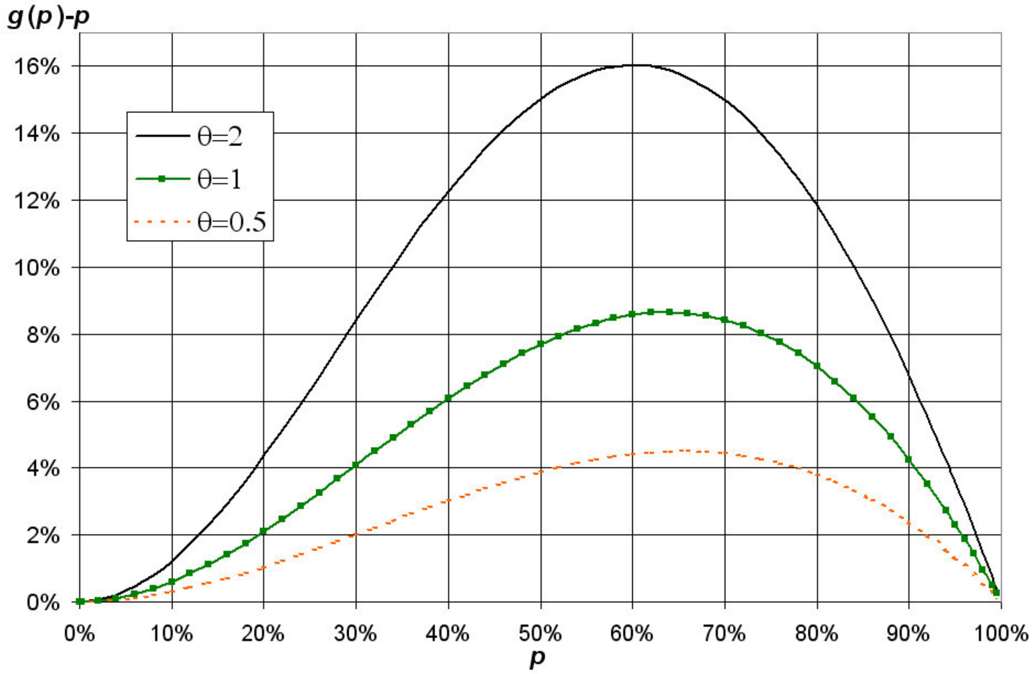
3.3 Characterization and Comparative Statics

The following proposition shows that more risk-aversion induces more overconfidence. That is, if principal I is more risk-averse than principal II, then he hires agents with higher level of overconfidence. The intuition is similar to that of Theorem 4.

Proposition 7 *Assume $h_1 = \psi \circ h_2$, where ψ is concave and increasing. Let g_1^* (g_2^*) be the unique optimal bias function given that the principal's utility is h_1 (h_2). Then, $g_1^*(p) > g_2^*(p)$ for every $0 < p < 1$.*

We demonstrate the above result in Figure 2. It assumes that principal's utility has constant relative risk aversion (CRRA, see below), and it shows the optimal overconfidence bias ($g^*(p) - p$) for different levels of relative risk aversion: $\theta = 2$, $\theta = 1$ (i.e., $h(x) = ln(x)$), and $\theta = 0.5$. The figure assumes that the potential gain is: $D = \frac{H-L}{L} = 2$, and that the accuracy

Figure 2. Optimal overconfidence ($g^*(p) - p$) for different risk-aversion levels (D=2, Uniform pdf)



of the private signal is uniformly distributed in $[0, 1]$).

Our next result presents further comparative statics under the mild assumption that the principal's utility satisfies decreasing absolute risk-aversion. Specifically, it shows that the optimal level of overconfidence is higher if:

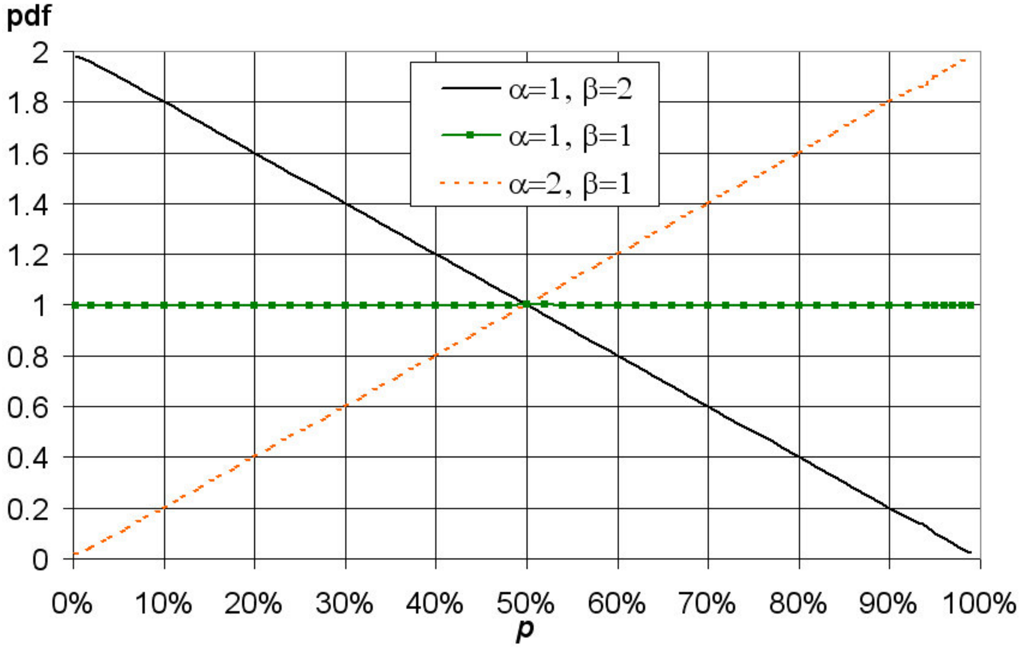
(1) The high payoff for success (H) is higher.
(2) The low payoff for failure (L) is lower.
(3) The task is harder, in the sense that there is higher probability to receive less accurate private signals (first-order stochastic domination). That is, it shows that our model predicts the hard-easy effect (presented in the introduction).

Theorem 8 *Assume that the principal's utility h satisfies decreasing absolute risk-aversion. That is, Arrow-Pratt coefficient of absolute risk-aversion $r_A(x) = -\frac{h''(x)}{h'(x)}$ is a decreasing function of x. Then higher level of optimal overconfidence ($g_2^*(p) < g_1^*(p)$ for every $0 < p < 1$) is induced by:*

(1) Higher payoff for success: $H_2 > H_1$.
(2) Lower payoff for failure: $L_2 < L_1$.
(3) Harder tasks (less accurate private signals): If $\mathbf{p}_2$ has first order stochastically dominance over $\mathbf{p}_1$.

The intuition of the first two results is as follows. The higher the difference between H and L, the higher the risk of having many agents follow the public signal (as the variance of the payoff becomes larger). This induces the principal to hire more overconfident agents, to compensate of the excess risk.

Figure 3. Different Private Accuracy Distributions



The intuition of the last result (harder tasks induces more overconfidence) is that the principal wishes that the agents with the highest levels of private accuracy would follow their private signals. When there is higher probability to receive less accurate private signals, each accuracy level p_i is more likely to be one of the highest levels.

The last result is demonstrated in Figure 4, which shows the induced optimal overconfidence $g^*(p) - p$ for different prior Beta distributions for the accuracy of the private signal (with $\theta = 1$ and $D = 2$), which are presented in Figure 3: 1) decreasing distribution ($\alpha = 1, \beta = 2$, expectation - 33%), 2) uniform distribution ($\alpha = 1, \beta = 1$, expectation - 50%), and 3) increasing distribution ($\alpha = 2, \beta = 1$, expectation - 67%). Observe that the last distribution stochastically dominates the second distribution, which stochastically dominates the first distribution. The figure assumes that the principal's utility is: $h(x) = ln(x)$, and that the potential gain is: $D = \frac{H-L}{L} = 2$.

Finally, our last result deals with the case in which the principal has constant relative risk aversion (CRRA). That is:

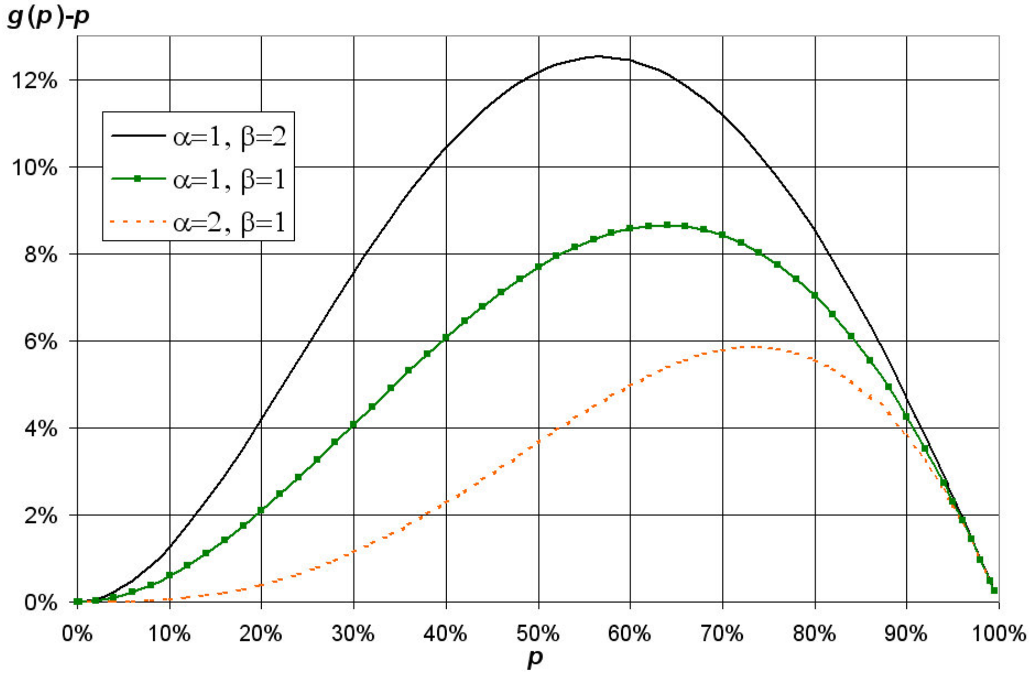
$$h(x) = \begin{cases} \frac{x^{1-\phi}}{1-\phi} & \text{if } \phi > 0, \phi \neq 1, \\ ln(x) & if\ \phi = 1, \end{cases}$$

where $\phi > 0$ is the relative risk aversion of the principal.

The following theorem shows that with the CRRA assumption the induced optimal overconfidence satisfies the following properties:

(1) The optimal level of overconfidence only depends on the potential gain $D = \frac{H-L}{L}$ (and not on L and H directly). Specifically, larger potential gain induces more overconfidence

Figure 4. Overconfidence for Different Private Accuracy Distributions ($\theta = 1$, $D = 2$)



(see Figure 5). This fits the experimental findings of Sieber (1974), which were discussed in Subsection 1.1.

(2) The ratio between the perceived error probability and the true error probability of the private information is decreasing in the difficulty level. Moreover, if the potential gain D is large, then agents are often wrong when they are certain in their private information (false certainty effect, see Figure 8 in Section 5).

Theorem 9 *Assume that the principal has a CRRA utility function with parameter θ. Then the unique optimal bias function is:*

$$g^*(p) = \frac{Bp}{1 - p + Bp}, \quad \text{where } B = \left(1 + \frac{D \cdot F_{\mathbf{p}}(p)}{1 + D \cdot \int_p^1 x \cdot f_{\mathbf{p}}(x)\, dx}\right)^{\phi},$$

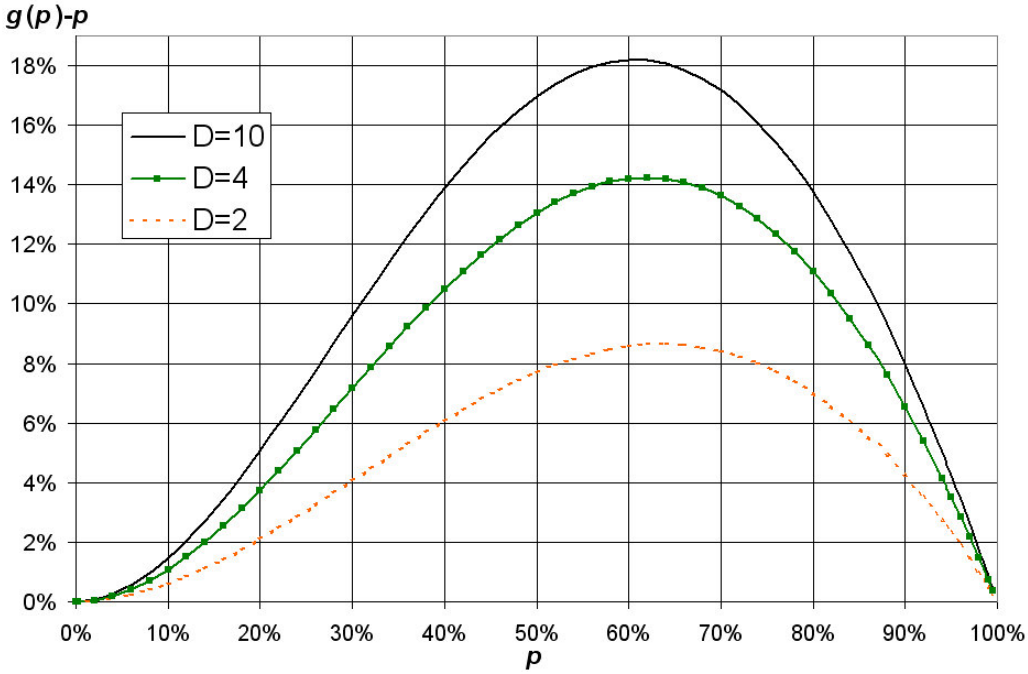
and it satisfies the following properties:

(1) Overconfidence ($g^*(p) - p$) is increasing in the potential gain $D = \frac{H-L}{L}$.

(2) The ratio between the perceived and the true probability that the private signal is incorrect ($\frac{1-g(p)}{1-p}$) is decreasing in p, and it converges to $(D+1)^{\phi} = \left(\frac{H}{L}\right)^{\phi}$ as p converges to 1.

Figure 5 demonstrates the first result of Theorem 9. It shows the induced optimal overconfidence $g^*(p) - p$ for different levels of potential gains: $D = 10$, $D = 4$ and $D = 2$ (with $\theta = 1$ and a uniform distribution).

Figure 5. Overconfidence for different potential gain levels ($\theta = 1$,Uniform pdf)



4 Variants and Extensions

This section includes a few variants and extensions of the basic model, which relax some of our assumptions, and allow to apply it in a broader variety of applications (as discussed in the following sections). In each subsection we only present the differences between the described variant and the basic model.

4.1 Choosing the Number of Agents

In the basic model we assumed that the number of agents is large. In this subsection we relax this assumption. Specifically, we allow the principal to choose the number of agents he employs, and then to choose the bias profile of the hired agents. We show that it is optimal for the principal to hire large number of agents.

The following proposition shows that for every n the principal strictly prefers to hire $k \cdot n$ agents than n agents.

Proposition 10 *For each $n \geq 1$ and $k \geq 2$ the principal can induce a strictly better outcome when the number of agents is $k \cdot n$ than when it is n.*

The intuition of Proposition 10 is that having more agents allow the principal to achieve better diversification. Each bias profile $(g_i)_{i \in I}$ with n agents can be replaced with a similar profile with $k \cdot n$ agents, in which each bias function g_i is induced by k agents. It can be shown that the random number of success in the game with $k \cdot n$ agents second order stochastically dominates the number of success in the game with n agents, and thus it is more preferred by the principal.

The following example shows that the principal may prefer employing 2 agents than employing 3 agents.

Example 11 *(Example 5 revisited) Let $L = 0$, $H = 1$, $f_{\mathbf{p}} \sim$ uniform$(0, 0.5)$, $\mathbf{q} = 0.7$ and let the principal's utility be:*

$$h(x) = \begin{cases} 2x & \text{if } x < 0.5, \\ 1 & \text{if } x \geq 0.5. \end{cases}$$

Recall (Example 5) that when there are two agents the principal can achieve payoff 0.775 by using an asymmetric bias profile: one agent always follow the public signal while the other agent always follow the private signal. When there are three agents, the principal's best payoff is only 0.75, and it is achieved by having two agents always follow the public signal, and one agent always follow his private signal. The intuition of the preference for having 2 rather then 3 agents is as follows. The definition of the payoff function h implies that the principal mainly cares that at least 0.5 of his agents would succeed. It is easier to achieve it when there are only 2 agents (1 of them should be successful) rather then when there are 3 agents (and 2 of them should be successful).

4.2 Agents as Experts

Consider a variant of the basic model in which at stage 2 each agent recommends an action (which signal to follow), and the principal chooses the profile of actions $(a_i)_{i \in N}$ based on these recommendations. That is, each agent i is an informed expert, who advises the principal what to do in his local environment i, based on his private information. Each expert's payoff remains the same as in the basic model: high payoff if the recommended signal was correct, and low payoff otherwise.

If all agents are calibrated ($g(p) = p$), then too many agents would recommend the principal to follow the public signal (all experts i with $\mathbf{p}_i < \mathbf{q}$). The principal can gain higher payoff than in the basic model, by not following some of the recommendations. However, his inability to separate agents with inaccurate private signals ($\mathbf{p}_i$ is substantially smaller than $\mathbf{q}$) from agents with more accurate private signals limits his payoff.

One can see that this variant yields the same optimal bias function g^*. This is because agents that follow g^* induce the principal's first-best payoff. Specifically, they act as if they have the same utility as the principal including his interest in diversification. Thus, the principal will always choose to follow the recommendations of such g^*-biased experts.

4.3 Choice between k alternatives

The basic model assumes that agents have only two actions: follow the public signal or follow the private signal. In this subsection we show that this assumption is especially without loss of generality. Specifically, we present a variant of the model in which each agent has to choose an action out of a fixed set of k alternatives. Let A be the set of actions of each agent: $1 < |A| = k < \infty$. The unknown state of nature determines a single best action $a_{best} \in A$. In order to keep the model simple and tractable, we assume that each action has the same prior probability $(\frac{1}{k})$ to be the best action.

In this variant, the public signal includes two parts: 1) an action $\mathbf{a}_{pub} \in A$, and 2) an accuracy level $\mathbf{q}$ - the probability that $\mathbf{a}_{pub} = \mathbf{a}_{best}$. If the public signal is incorrect ($\mathbf{a}_{pub} \neq \mathbf{a}_{best}$), then $\mathbf{a}_{pub}$ is uniformly distributed among all other actions: $A \backslash \{a_{best}\}$. Similarly, the private signal of each agent i includes two parts: 1) an action in $\mathbf{a}_{pri} \in A$, and 2) an accuracy level $\mathbf{p}_i$- the probability that $\mathbf{a}_{pri} = \mathbf{a}_{best}$, and conditional on $\mathbf{a}_{pri} \neq \mathbf{a}_{best}$, $\mathbf{a}_{pri}$is uniformly distributed in $A \backslash \{\mathbf{a}_{best}\}$. Each signal is independent of each other signal (given the value of $\mathbf{a}_{best}$).

We assume that $f_{\mathbf{q}}$ and $f_{\mathbf{p}}$ are continuous pdfs with full support on $x \in \left[\frac{1}{K}, 1\right]$. That is, given that action a is being signaled by either the public or the private signal, it is more likely that a is the best action. At stage 2 each agent i chooses an action $a_i \in A$. The payoff of agent i is:

$$u_i(a_i) = \begin{cases} H & \text{if } a_i = a_{best}, \\ L & \text{if } otherwise. \end{cases}$$

Observe that it is a dominating strategy for each agent to choose to either follow the public or the private signal (and not to choose an action that is not recommended by any of the signals). All of our results (Theorems 4-9) hold in this setup as well.

4.4 Underusing Base Rates

In this subsection we present a different interpretation to the variant described in the previous subsection, and show that the optimal bias profile can be interpreted as insensitivity to prior probability or underusing base rates.

For simplicity of presentation we assume that there are only two actions: $A = \{l, r\}$. Let C be the event that the public signal is correct. We interpret q as its prior probability ($q = P(C)$). Let D be the event that the private signal is different than the public signal. To maximize his payoff, an agent should follow the public signal if the posterior probability $P(C|D)$ is at least 0.5, and follow the private signal if $P(C|D) < 0.5$.[16] A calibrated agent ($g(p) = p$)

[16] If both signals recommend the same action, then the agent should choose this action.

calculates the posterior probability correctly:

$$P(C|D) = \frac{P(D|C)\,P(C)}{P(D|C)\,P(C) + P\left(D|\bar{C}\right)P\left(\bar{C}\right)} = \frac{(1-p)\cdot q}{(1-p)\cdot q + p\cdot(1-q)}$$

An agent with bias g plays as if his posterior probability is:

$$P_g(C|D) = \frac{P(D|C)\,P_g(C)}{P(D|C)\,P_g(C) + P\left(D|\bar{C}\right)P_g\left(\bar{C}\right)} = \frac{(1-p)\cdot g^{-1}(q)}{(1-p)\cdot g^{-1}(q) + p\cdot\left(1-g^{-1}(q)\right)}$$

Thus, agents with the optimal bias function g^* play as if they are insensitive to the prior distribution in the following sense: when the true prior is $P(C) = q > 0.5$, the agent plays as if the prior is a mixture between a uniform prior and the true prior: $0.5 \leq P_{g^*}(C) = (g^*)^{-1}(P(C)) < P(C)$. Specifically, if the prior is changed from $(0.5, 0.5)$ to $(q, 1-q)$, the agent is only partially sensitive to this change, and he plays as if the prior probability is $\left((g^*)^{-1}(q), 1-(g^*)^{-1}(q)\right)$ where $0.5 \leq (g^*)^{-1}(q) < q$. Similar effects are experimentally demonstrated in Griffin and Tversky (1992), Brenner et al. (1996), Novemsky and Kronzon (1999), Koehler, Brenner and Griffin (2002), and Brenner, Griffin, and Koehler (2005).

4.5 Choice between k alternatives with i-adapted actions

The variant of choice between k alternatives presented in Subsection 4.3 has a potential drawback in setups in which agents can communicate among themselves and share their private signals. In such setups, agents could obtain the true value of the best action (a_{best}) with high probability by combining a large number of independent private signals. This is solved in the following variant in which private signals are related to i-adapted actions (and not to the best action).

In this variant each agent has to choose an action $a \in A$ ($1 < |A| = k < \infty$). The unknown state of nature determines a single best action $\mathbf{a}_{best} \in A$, and for each agent $i \in N$, it determines a single i-adapted action $\mathbf{a}_{i-adpated} \in A$ (which has a prior uniform distribution). The values of $\left(\mathbf{a}_{best}, (\mathbf{a}_{i-adapted})_{i \in N}\right)$ are independent. Action $\mathbf{a}_{i-adapted}$ is interpreted as an action that is well-adapted to the specific characteristics of agent i or for the specific properties of his local environment.

The public signal includes two parts: 1) an action $\mathbf{a}_{pub} \in A$, and 2) an accuracy level $\mathbf{q}$ - the probability that $\mathbf{a}_{pub} = \mathbf{a}_{best}$. If the public signal is incorrect ($\mathbf{a}_{pub} \neq \mathbf{a}_{best}$), then $\mathbf{a}_{pub}$ is uniformly distributed among all other actions: $A \backslash \{a_{best}\}$. Similarly, the private signal of each agent i includes two parts: 1) an action in $\mathbf{a}_{pri} \in A$, and 2) an accuracy level $\mathbf{p}_i$- the probability that $\mathbf{a}_{pri} = \mathbf{a}_{i-adapted}$, and conditional on $\mathbf{a}_{pri} \neq \mathbf{a}_{i-adpatedt}$, $\mathbf{a}_{pri}$is uniformly distributed in $A \backslash \{\mathbf{a}_{i-adapted}\}$. Each signal is independent of each other signal (given the value of $\mathbf{a}_{best}$).

At stage 2 each agent i chooses an action $a_i \in A$. The payoff of agent i is:

$$u(a_i) = \begin{cases} H & \text{if } a_i \in \{a_{best}, a_{i-adpated}\}, \\ L & \text{otherwise.} \end{cases}.$$

Observe that the expected payoff when following an incorrect signal is equal to $\tilde{L} = L + \frac{H-L}{k}$, because there is probability $\frac{1}{k}$ that the recommended public (private) signal is equal to $a_{i-adapted}$ (a_{best}). All of our results (Theorems 4-9) hold in this setup as well (with respect to the revised low payoff $\tilde{L}$), and in this setup, they are not sensitive to the assumption that agents cannot share their private signals.

4.6 Bias With Respect to the Public Signal

In the basic model we assume that agents can only have confidence bias with respect to their private signals, but not with respect to the public signal. In this subsection, we observe that this assumption is without loss of generality.

Consider a more general model, where the bias of each agent i be described by two functions $(g_{i,1}, g_{i,2})$ from $[0,1]$ to $[0,1]$, where $g_{i,1}$ is his bias with respect to his private signal (accuracy p_i is perceived as $g_{i,1}(p_i)$) and $g_{i,2}$ is his bias with respect to the public signal (accuracy q is perceived by agent i as $g_{i,2}(q)$). Observe that the choice of agent i between the two signals only depends on the composite function $(g_{i,2})^{-1} \circ g_{i,1}$. This is because agent i chooses to follow the public signal if $g_{i,1}(\mathbf{p}_i) < g_{i,2}(\mathbf{q}) \Leftrightarrow (g_{i,2})^{-1} \circ g_{i,1}(\mathbf{p}_i) < \mathbf{q}$. This implies that our results remain the same in this extension. Specifically, the optimal profile is such that each agent i has bias functions $(g_{i,1}, g_{i,2})$ that satisfy $(g_{i,2})^{-1} \circ g_{i,1} = g^*$, where g^* has the properties that were characterized in Theorems 4 and 9.

Thus one can interpret g^* in the basic model as the bias in estimating accuracy of private information relative to public information. A natural question that emerges is why we interpret this bias in favor of private information as overconfidence. Similar interpretations can be found in other models in the literature (see, e.g., Bernardo and Welch, 2001; Grubb, 2009). We motivate this interpretation by the following arguments:

(1) Assume that a long history of principal-agents games has been played in the past by different agents. Information about past public signals would encourage agents to evaluate the accuracy of the public signal by set-based evaluations, and the historical data. The experimental literature suggests that such evaluations tend to reduce overconfidence (see e.g., Gigerenzer, Hoffrage and Kleinbolting, 1991; Griffin and Tversky, 1992). In addition, if the environment is relatively stationary, simple learning rules would allow the agents to be well calibrated with respect to the public signal. On the other hand,

it is plausible that past detailed information about the success or failure of each agent and his private signal is unavailable. If agents are relatively new in this game, then they would not have empirical data to make set-based frequency evaluations for the private signals.

(2) The public signal may be induced by the aggregation of many "weak" pieces of information that are shared by different agents. This especially fits the variant with i-adapted actions presented in Subsection 4.5. Each agent has a "weak" local signal about the global state (a_{best}), and by aggregating their information, agents create the public signal. Thus the public signal has high "weight" (many different sources) but low "strength" (each source has low accuracy). Griffin and Tversky (1992) show that people tend to underestimate accuracy of high-weight/low-strength signals (such as the public signal), and overestimate the accuracy of low-weight/high-strength signals, such as the private signals (each private signal has a single source with relative high accuracy).

(3) Internal v.s. external uncertainty - The public signal is related to uncertainty about the outside world (external uncertainty). The private signal (in the variant with i-adapted actions) is related to uncertainty about the characteristics and knowledge of the agent (internal uncertainty). The psychological literature suggests that people evaluate these two kinds of uncertainty differently, and that internal uncertainty induces more overconfidence (Budescu, and Du, 2007, P. 1741).

Alternatively, the above arguments can be used to show that the induced confidence bias in our model is in accordance with the observed stylized facts in the experimental literature: strength-weight effect, hard-easy effect, and internal-external uncertainty.

4.7 Costly Private Signals

The basic model assumes that private signals are costless. In this subsection we relax this assumption and extend our results to a more general framework that allows private signals to be costly. In the extended model, an independent random variable $0 \leq \mathbf{t}_i \sim f_{\mathbf{t}} \leq 1$ is assigned to each agent $i \in N$. Variable $\mathbf{t}_i$ is interpreted as the effectiveness of agent i in acquiring private information about his environment.

After agents are publicly informed about the value of $\mathbf{q}$ (the accuracy of the public signal), each agent is privately informed of $\mathbf{t}_i$. Then, each agent privately chooses an effort level $0 \leq e_i \leq 1$, and receives a private signal with accuracy level $p_i = p(e_i, \mathbf{t}_i)$, where p is an increasing function (in both parameters), and it is concave in the effort level e_i. The payoff of each agent is $H - (H - L) \cdot e_i$ if the signal he followed was correct, and $L - (H - L) \cdot e_i$ otherwise. The rest of the model is the same as the basic model.

Let $p_{\mathbf{t}_i} \in [0, 1]$ be the accuracy level that maximizes $p(e_i, \mathbf{t}_i) - e_i$. The distribution of effectiveness levels $f_{\mathbf{t}}$ induces a distribution of maximizing accuracy levels $f_{p_{\mathbf{t}}}$. The following proposition asserts that our results also hold in this extended model, where $f_{p_{\mathbf{t}}}$ replaces $f_{\mathbf{p}}$.

Proposition 12 *The extended model with costly signals admits an optimal bias function g^*,* *which is the same as the optimal bias function g^* of the basic model with $f_{\mathbf{p}} = f_{p_{\mathbf{t}}}$.*

4.8 Risk-Averse Agents

In the basic model the utility of each agent is equal to

$$\mathbf{u}_i\left(a_{pub}\right) = \begin{cases} H & \text{if } \mathbf{m}_q = 1, \\ L & \text{if } \mathbf{m}_q = 0, \end{cases} \text{ and } u_i\left(a_{pri}\right) = \begin{cases} H & \text{if } \mathbf{m}_i = 1, \\ L & \text{if } \mathbf{m}_i = 0, \end{cases}$$

and the utility of the principal is a concave function of the average utility of the agents. Thus, in the basic model the principal is more risk-averse than the agents (for example, when there is a single agent, the principal's utility is a concave function of the agent's utility), which may seem implausible in some applications.

However, this assumption can be relaxed without changing the results as follows. We reinterpret $\mathbf{u}_i$ as a monetary payoff, and we allow the the utility of agent i to be any monotone function of this monetary payoff: $h_i\left(\mathbf{u}_i\right)$. Specifically, our results (Theorem 4-9) also holds if each agent has utility function $h_i\left(x\right)$ that is more concave then the principal's utility $h\left(x\right)$.

5 Overconfidence and Evolutionary Stability

In this section we apply our model in an evolutionary setup (extending Example 3, which were briefly discussed in the introduction), explain why overconfidence is evolutionarily stable, and discuss the implications of the model in this setup.

5.1 Evolutionary Model

Consider a large population of agents. In each generation, each agent has to choose an action, and this choice influences his fitness (number of offspring in the next generation). For example, this may describe choice of occupation, living area, how to provide food for the family, or how to raise and educate the offspring.

The unknown state of nature determines an action that is best on average (e.g., being a fisherman is good due to large fish population), and for each individual it determines which action is most-adapted to his characteristics (e.g., Alice has good hunting skills). All agents receive a public noisy signal on the best action, and each agent privately receives a noisy signal on his most-adapted action. Each agent obtains high fitness for choosing the best (on-average) action or for choosing his most-adapted action, and low fitness otherwise (as formulated in Subsection 4.5).

The population includes a set of genetic types, and each agent is a member of one of these types. Each type induces a (possibly random) confidence bias function g for its members. It is well known (Lewontin and Cohen, 1969; Mcnamara, 1995; Robson, 1996) that in the long run the type that maximizes the expectation of the geometric mean of the fitness will prevail the population. This is equivalent to maximizing the expected logarithm of the average fitness of its members.

This fits our model as follows. Each agent is an individual with utility (fitness):

$$\mathbf{u}_i = \begin{cases} H & \text{if } a_i \in \{a_{best}, a_{i-adapted}\}, \\ L & \text{otherwise,} \end{cases}$$

and the principal (type) has utility $u = \mathbf{E}\left(ln\left(\frac{1}{N}\sum_i \mathbf{u}_i\right)\right)$, where the sum is taken over all its members. In this setup, our main results show that the unique evolutionarily stable type induces its members to be overconfident (in order to resolve the conflict of interest between an individual who maximizes his fitness, and a gene is interested in diversification among its members).

5.2 Levels of Overconfidence in Different Cultures

Yates et al. (2002) report substantial differences in the levels of observed overconfidence in different cultures. They summarize results from several studies, and show that Asians tend to present more overconfidence then Westerners (in the sense of overestimating the accuracy of private information).

The above evolutionary model describes a society where the state of nature uniformly influences all agents (that is the best action is the same for all agents). Alternatively, one can think of societies in which the best action may differ among different subgroups (such as tribes, or geographical area). This can be modeled, for example, by making the best actions of different agents correlated, but not necessarily the same. One can show, that lowering the correlation induces less concave gene's utility. By Theorem 9, this causes individuals to be less overconfident.

Such differences can be explained by our model, as the result of different evolutionary histories. Specifically, our model predicts that in different societies people would present different levels of overconfidence, based on:

- The typical accuracy of private information (higher probability to receive noisier private signals induces more overconfidence).
- The typical potential gain (higher potential gain $D = \frac{H-L}{L}$ induces more overconfidence).

Remark 13 *Our evolutionary model describes a society where the state of nature uniformly influences all agents; that is the best action is the same for all agents. Alternatively, one can think of societies in which the best actions of different agents are correlated, but not* *necessarily the same. One can show, that lowering the correlation between the best actions of different agents, induces a less concave gene's utility. By Theorem 9, this causes individuals to be less overconfident.*

5.3 Decreasing Overconfidence for $p > 0.5$

In most of the experimental literature participants are asked to choose one of two possible answers for a question, and estimate the probability (from 50% to 100%) that they had answered the question correctly (see, e.g., Lichtenstein, Fischhoff and Phillips, 1982). These papers suggest that overconfidence is decreasing with the difficulty for every $p > 0.5$.

The following figures demonstrate that a similar behavior is induced in our model for plausible private signal accuracy distributions $f_{\mathbf{p}}$ (and any value of the potential gain D). Figure 6 presents three Beta distributions for the accuracy of the private signal in the evolutionary history: 1) uniform distribution ($\alpha = 1, \beta = 1$, expectation - 50%), 2) single-peaked distribution around 20% ($\alpha = 2, \beta = 5$, expectation - 29%), and 3) decreasing distribution ($\alpha = 1, \beta = 3$, expectation - 25%). Figure 7 shows the induced overconfidence from these three distributions (for potential gain $D = 2$, qualitative results are insensitive to the value of D). The uniform distribution induces overconfidence which slightly increases between 50%-60%, and then monotonically decreasing in the accuracy level p (that is, decreasing in the difficulty level). The other two distributions are approximately decreasing for every $p > 50\%$.

5.4 False Certainty Effect

Fischhoff, Slovic and Lichtenstein (1982) experimentally demonstrate the false certainty effect. Participants in their experiments were asked to choose the most likely answer for a general-knowledge question, and then to indicate their degree of certainty that the answer they had selected was correct. Across several different question and response formats, participants underestimated the error probability of seemingly easy questions: 1) the error probability of 10% of the questions was estimated to be extremely low (less than 1:1,000), while the true error probability was approximately 10%, and 2) the error probability of other 10% of the questions was estimated as 1% while the true error probability for these questions was approximately 20%, Participants had sufficient faith in their confidence judgments to be willing to stake money on their validity.

As shown earlier (Theorem 9) such an effect can emerge in our model for relatively high potential gains. For example, Figure 8 shows the perceived error probability and the true error probability for potential gain $D = 30$, and for three prior beta distributions for the accuracy of the private signals: 1) uniform distribution ($\alpha = 1$, $\beta = 1$, expectation - 50%), 2) single-peaked distribution around 20% ($\alpha = 2$, $\beta = 5$, expectation - 29%, see figure 6), and 3) single-peaked symmetric distribution ($\alpha = 3$, $\beta = 3$, expectation - 50%). The figure demonstrates the false certainty effect in our model, especially for the two single-peaked

Figure 6. Examples for Plausible Private Signal Accuracy Beta Distributions

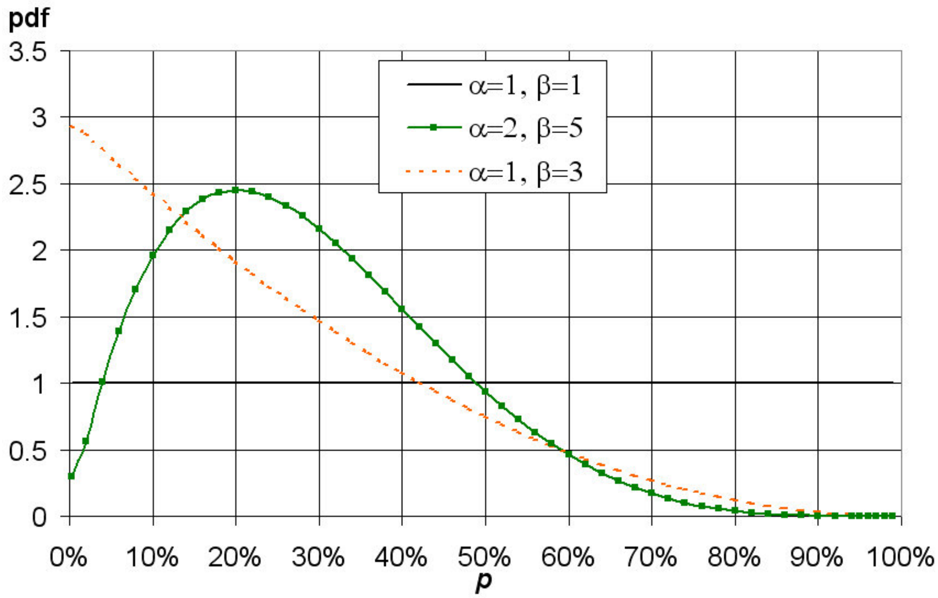
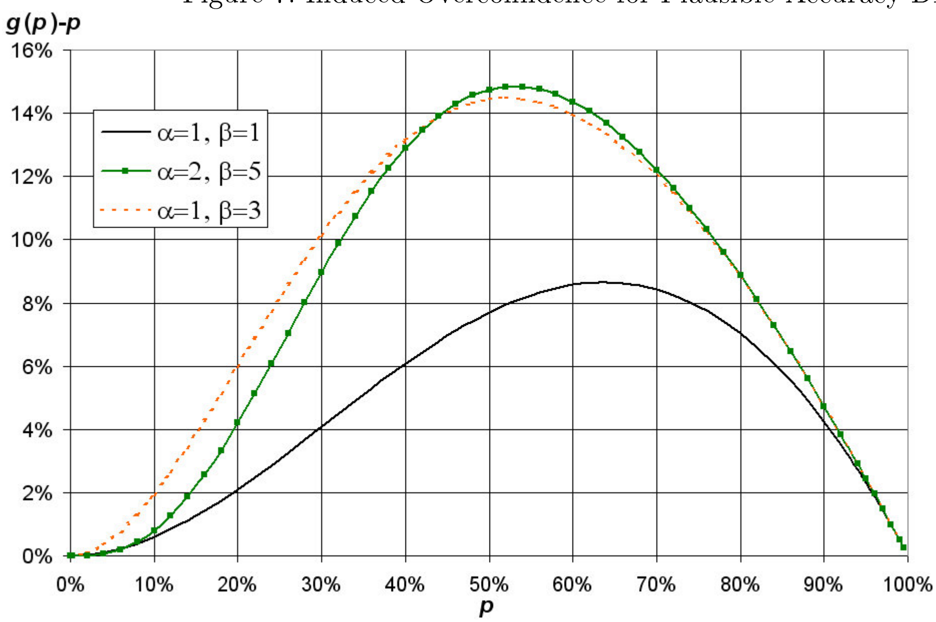


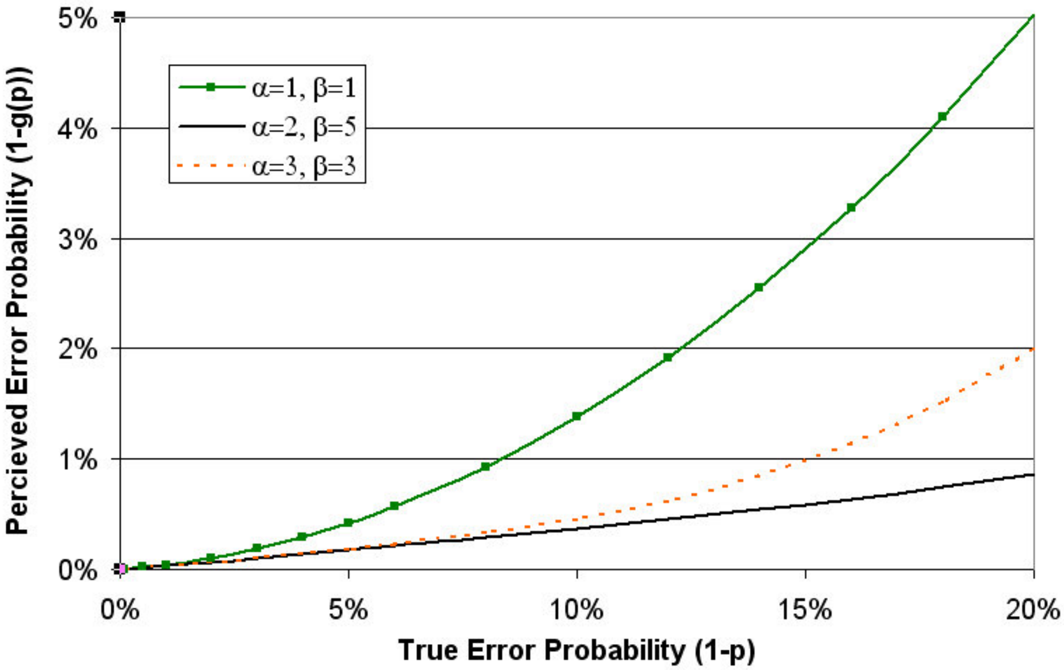
Figure 7. Induced Overconfidence for Plausible Accuracy Distributions



distributions: when the true error probability is 20% the perceived error probability is 1-2% (5% for the uniform distribution), and when the true error probability is 10% the perceived probability is less than 0.5% (1% for the uniform distribution).

Our assumption that the potential gain is high (30 - i.e., a good action yields 30 times more fitness than a bad one) may seem too extreme. However, one can extend our results into a setup where the the potential gain $\mathbf{D}$ is a random variable, and that its distribution has some positive small weight on high values. The optimal $\mathbf{D}$-dependent bias function $g^*(p|\mathbf{D})$ depends on the realization of $\mathbf{D}$. In many situations it seems plausible to assume that each type induces a single bias function $g^*(p)$ for all values of $\mathbf{D}$ because either: 1) it is too

Figure 8. Perceived vs. True Error Probability ($D = 30$)



complicated to induce numerous bias functions $g^*(p|\mathbf{D})$, or 2) individuals do not know the realization of the potential gain when they choose their actions. Observe that for relatively low levels of potential gain $\mathbf{D}$ and low error probabilities (high p-s) the difference in the long-run type's utility from either choosing the private or the public signal is small (both yield high payoff). However, when the potential gain $\mathbf{D}$ is high, the utility of the type is substantially influenced by the individual's choice, even when the error probability is low. Thus, for low error probabilities, the single optimal confidence bias function $g^*(p)$ would be close to the value of $g^*(p|\mathbf{D} = D)$ of a high realization of $\mathbf{D}$.

6 Examples

In this section we present a few examples to demonstrate the applicability of our model in economic interactions. In each example we identify the risk-averse principal and the agents, and sketch the interaction between them. The first two examples were briefly discussed in the introduction.

Our model shows that overconfidence can resolve conflict of interests between principal and agents. This conflict can also be resolved by giving the agents appropriate monetary incentives (e.g., a higher payoff for following the private signal). Thus, in order to demonstrate the plausibility of our model, we also shortly discuss the difficulties in implementing such monetary incentives in these examples.

6.1 CEO and Intermediate Managers (Example 1)

Consider a firm that operates in several markets, and its operation in each market is managed by a different agent (intermediate manager). The success of the firm in each market depends on the actions of the agent. For example, the firm operates in several regions, each agent is the manager in charge of the operation in one of these regions, and success depends on his marketing strategy. A marketing strategy has higher success probability if it either fits the national trend in consumer's preferences, or by being well adapted to the local trend in the region. Alternative examples are: 1) each agent is a product manager of a firm that manufactures several products; 2) each agent is an editor (or a producer) of a publishing company (or a film studio), and 3) each agent is a researcher in a research and development department of a firm or a non-for-profit organization.

The payoff of each agent is an increasing (possibly concave) function of the firm's profit is in his market. The payoff of the CEO is an increasing concave function of the total profit of the firm in all the markets (that is, the CEO is risk-averse with respect to the total profit). Applying our model to this setup, shows that the CEO would prefer to hire overconfident intermediate managers.

The CEO could also solve the conflict of interests with agents by appropriate monetary incentives (bonus policy). However, implementation of such a policy would require the agreement of the firm's shareholders. If the shareholders are risk-natural (for example due to having a diversified portfolio), they would not approve such a policy. On the other hand, the choice of overconfident agents can be done by the CEO without formally informing the shareholders. In addition, there are situations in which the agents are not employees of the firm. For example, they might be independent local distributors, and competition with other manufacturers may restrict the plausible contracts between the manufacturer and the distributors, such that the distributor's payoff must be highly correlated with the local profit.

6.2 Investor and Entrepreneurs (Example 2)

Consider an investor (the principal, an angel investor or a manager of a venture capital fund) who invests money in several entrepreneurs - founders of startup companies. Each such entrepreneur (agent) receives high payoff if his startup succeeds and low payoff otherwise. The investor has a concave increasing utility that depends on the number of successful startup companies (diminishing marginal payoff from successes). When the investor interviews his entrepreneurs (before choosing them), he obtains a signal on their confidence-bias.

The probability of success of each startup company depends on the product's design. For concreteness, consider the case where the product is a software, and success depends on the platform in which the software is developed to (e.g., member-restricted web site, Smartphone application, social network, tablet PC application, etc.). During the software development phase (after the investor already chose his entrepreneurs) everyone receives a noisy public signal which platform is more likely to be "hot" (best), and each entrepreneur receives a private signal which platform is most-adapted to the special characteristics of his software.

In this situation there is a conflict of interest between the investor who wishes to diversify risk among the different entrepreneurs, and the entrepreneur who only wishes to maximize his probability of success. This conflict can be resolved if the investor chooses overconfident entrepreneurs. Solving this conflict with monetary incentives to calibrated entrepreneurs would be expensive: if each entrepreneur holds a large share of his startup company, then only very large monetary incentives would encourage him to follow a nosier private signal.

Our model presents a new explanation why entrepreneurs tend to have high levels of overconfidence (see, e.g., Arnold C. Cooper, Woo, and Dunkelberg, 1988). In addition, it implies that entrepreneurs in different areas would present different levels of overconfidence. Specifically, entrepreneurs in areas in which typical investors are individuals and small area-specific funds would tend to be more overconfident, than entrepreneurs in areas in which the typical investors are large multi-area funds or the government.

Our result does not depend on the assumption that there is a single investor. It can be extended to a setup where there are many (risk-averse) investors and many entrepreneurs. Due to risk aversion, each investor would divide his money among several entrepreneurs, and the qualitative result would remain the same: all investors prefer to invest in overconfident entrepreneurs.

6.3 Overconfidence and Social Welfare

Consider a society, where each agent may act by either following a public signal or a private signal. This action influences agent i's productivity $\mathbf{x}_i$: high output if he followed a correct signal and low output otherwise. The payoff of each agent is a function $u_i = h\left(\mathbf{x}_i, \sum_j \mathbf{x}_j\right)$ of his output $\mathbf{x}_i$ and the total output $\sum_j \mathbf{x}_j$. The function h is assumed to be (strictly) increasing and concave in both parameters. For example, this is the case if a fixed amount of each agent's output is taxed and is being used for producing a public good. Alternatively, it might be that the output of each agent has a direct positive externality on the other agents.

Calibrated agents (without confidence-bias) would follow the public signal too often, and obtain a Pareto-inefficient outcome, in which the variance of the total productivity $\sum_j \mathbf{x}_j$ is too high. Applying our results to this setup shows that if all agents are moderately overconfident, then they can achieve a Pareto-efficient outcome, which is Pareto-superior with respect to the outcome induced by calibrated agents. This may explain development of social norms in favor of moderate overconfidence (e.g., "self trust is the first secret of success", Ralph Waldo Emerson, 1803-1882).

A Proofs

A.1 Proof of Theorem 4

Theorem 4 There exists a unique optimal bias function g^*, which induces the first-best payoff, with the following properties:

(1) g^* is continuous, $g^*(0) = 0$, and $g^*(1) = 1$.
(2) g^* is increasing: $\frac{dg^*(p)}{dp} > 0$ for every $0 < p < 1$.
(3) $g^*(p) > p$ for every $0 < p < 1$ (overconfidence).

PROOF. The proof includes two parts. The first part shows that the first-best outcome of the principal can be approximately induced by a bias function. The second part characterizes this optimal bias function g^*, and shows its uniqueness.

Approximating the first-best payoff by a bias function

We begin by dealing with the "first-best" case in which the principal receives all the signals $(\mathbf{p}_i)_{i\in I}$ and chooses the actions of all the agents. Without loss of generality the first-best strategy is a function ϕ that chooses a threshold $p = \phi(\mathbf{q}; \mathbf{p}_1, ..., \mathbf{p}_n)$, such that each agent i with higher (lower) accuracy level $\mathbf{p}_i \geq p$ ($\mathbf{p}_i < p$) follow the private (public) signal. The first-best expected payoff is equal to:

$$h\left(L + (H - L)\cdot\left(\frac{\#\{i|\mathbf{p}_i < p\}}{n} + \frac{\sum_{\mathbf{p}_i \geq p}\mathbf{p}_i}{n}\right)\right)$$

if the public signal is correct, and equal to

$$h\left(L + (H - L)\cdot\left(\frac{\sum_{\mathbf{p}_i \geq p}\mathbf{p}_i}{n}\right)\right)$$

otherwise. To simplify notation let $f = f_p$ and $F = F_p$. By the law of the large numbers for sufficiently large number of players (n) with high probability:

$$\frac{\#\{i|\mathbf{p}_i < p\}}{n} \approx F(p),$$

and

$$\sum_{\mathbf{p}_i \geq p}\mathbf{p}_i \approx \int_p^1 x\cdot f(x)\,dx.$$

Thus, the expected payoff of the first-best action profile is well approximated by:

$$u = q \cdot h\left(L + (H-L)\cdot\left(F(p) + \int_p^1 x\cdot f(x)\,dx\right)\right) \tag{A.1}$$
$$+ (1-q)\cdot h\left(L + (H-L)\cdot\int_p^1 x\cdot f(x)\,dx\right) + o(\epsilon).$$

Consider the bias function $g(p) = g^*(p)$ that is defined as follows: $p = g^{-1}(q)$ is the threshold that maximizes Eq. A.1 (neglecting the error term $o(\epsilon)$). By the above arguments, such bias function ϵ-induces the first-best payoff.

Characterizing the unique optimal bias function $g^*(p)$

We now calculate the value of $p = (g^*)^{-1}(q)$ that maximizes Eq. A.1 (neglecting the error term $o(\epsilon)$). One can verify that $\alpha(0) = 0$ and $\alpha(1) = 1$. For every $0 < q < 1$ we find $p = (g^*)^{-1}(q)$ by derivation:

$$\frac{du}{dp} = q\cdot h'\left(L + (H-L)\cdot F(p) + \int_p^1 x\cdot f(x)\,dx\right)(f(p) - p\cdot f(p))(H-L)$$
$$+ (1-q)\cdot h'\left(L + (H-L)\cdot\int_p^1 x\cdot f(x)\,dx\right)(-p\cdot f(p))(H-L).$$

Assuming an internal solution ($\frac{du}{dp} = 0$) yields:

$$\frac{h'\left(L + (H-L)\cdot\int_p^1 x\cdot f(x)\,dx\right)}{h'\left(L + (H-L)\cdot\left(F(p) + \int_p^1 x\cdot f(x)\,dx\right)\right)} = \frac{q - q\cdot p}{p - q\cdot p}. \tag{A.2}$$

The fact that $h' > 0$ (increasing) and $h'' < 0$ (concavity) implies that the left hand-side (l.h.s.) of Eq. A.2 is a strictly increasing function of p. One can verify that the right-hand side (r.h.s.) is a strictly decreasing function of p, and that for small enough p the r.h.s. is larger than the l.h.s., while for $p = 1$ the l.h.s. is larger than the r.h.s. . Thus for each $0 < q < 1$ there is a unique solution to Eq. A.2 $0 < p = (g^*)^{-1}(q) < 1$, which is a continuous function of q.

The fact that the l.h.s. is always larger than 1 (due to the concavity of h) implies that $g^{-1}(q) < q$ for every $0 < q < 1$ (the overconfidence property). Increasing q by $\delta > 0$ while holding p constant, would not change the l.h.s. of Eq. A.2, while the r.h.s. would increase. This implies that $(g^*)^{-1}(q+\delta) > (g^*)^{-1}(q)$ (because the l.h.s. is decreasing in p), and thus $g^{-1}(q)$ is a strictly increasing function of q.

One can verify that $\frac{du}{dp} > 0$ for every $p < (g^*)^{-1}(q)$, and $\frac{du}{dp} < 0$ for every $p > (g^*)^{-1}(q)$. Thus, any other bias threshold $p \neq (g^*)^{-1}(q)$ would yield a strictly lower expected payoff.

The above arguments show that the profile in which all agents have bias g^* induces (up to ϵ) the first-best outcome for the principal (and thus it is ϵ-optimal), that g^* has all the required properties (continuous, increasing and overconfidence), and that g^* is unique in the following sense: any other bias function $\tilde{g}$ such that $\tilde{g} \neq g^*$ on a set with a positive Lebesgue measure yields a strictly lower payoff, assuming the number of agents is sufficiently large.

A.2 Proof of Theorem 6

Theorem 6 Let $(g_i)_{i\in I}$ be an heterogeneous profile. Then there is k_0 such that for every $k \geq k_0$, there is an homogeneous profile that induces a strictly better outcome than $(g_i)_{i\in I}$ in the game $k \cdot |I|$ with agents.

PROOF. For simplicity of notation let $p_i(q) = (g_i)^{-1}(q)$. Let $\tilde{g}$ be the following bias function (homogeneous bias profile): for each $q \in [0,1]$ let $\tilde{p}(q) = (\tilde{g})(q)^{-1}$ be the unique solution to the following equation:

$$F(\tilde{p}(q)) = \sum_{i\in I} \frac{1}{n} (F(p_i(q)))$$

That is, $\tilde{g}$ is a bias function that averages the heterogeneous profile $(g_i)_{i\in I}$.

For each q, the payoff of the heterogeneous profile $(g_i)_{i\in N}$, given that the accuracy of the public signal $\mathbf{q}$ is equal to$q$ is:

$$h\left(L + (H-L) \cdot \left(\frac{\#\{i|\mathbf{p}_i < p_i(q)\}}{n} + \frac{\sum_{\mathbf{p}_i \geq p_i(q)} \mathbf{p}_i}{n}\right)\right)$$

if the public signal is correct, and it is

$$h\left(L + (H-L) \cdot \left(\frac{\sum_{\mathbf{p}_i \geq p_i(q)} \mathbf{p}_i}{n}\right)\right)$$

otherwise. The expected payoff of $(g_i)_{i\in N}$ conditional on $\mathbf{q} = q$ is equal to

$$h\left(L + (H-L) \cdot \left(\frac{\sum_{i\in I} F(p_i(q))}{n} + \frac{\sum_{i\in I} \int_{p_{i(q)}}^1 x f(x)\, dx}{n}\right)\right)$$

if the public signal is correct, and it is equal to

$$h\left(L+(H-L)\cdot\left(\frac{\sum_{i\in I}\int_{p_{i(q)}}^{1}xf(x)\,dx}{n}\right)\right)$$

otherwise. The expected payoff of the homogeneous bias profile $\tilde{g}$ is

$$h\left(L+(H-L)\cdot\left(F(p(q))+\int_{\tilde{p}(q)}^{1}xf(x)\,dx\right)\right)$$

if the public signal is correct, and equal to

$$h\left(L+(H-L)\cdot\left(\int_{\tilde{p}(q)}^{1}xf(x)\,dx\right)\right)$$

otherwise. As $F(p(q))=\sum_{i\in I}\frac{1}{n}(F(p_i(q)))$, the homogeneous profile has an higher expected payoff if

$$\frac{1}{n}\sum_{i\in I}\int_{p_{i(q)}}^{1}xf(x)\,dx<\int_{\tilde{p}(q)}^{1}xf(x)\,dx. \tag{A.3}$$

For simplicity of notation let $\tilde{p}=\tilde{p}(q)$ and $p_i=p_i(q)$. Eq. A.3 yields:

$$\frac{1}{n}\sum_{i\in I}\left(\int_{p_i}^{1}xf(x)\,dx-\int_{\tilde{p}}^{1}xf(x)\,dx\right)<0.$$

This is equivalent to:[17]

$$\frac{1}{n}\sum_{i\in I}\int_{p_i}^{\tilde{p}}xf(x)\,dx<0,$$

which is equivalent to:

$$\frac{1}{n}\sum_{i\in I}(F(\tilde{p})-F(p_i))\cdot\mathbf{E}\left(\mathbf{p}|\min(p_i,\tilde{p})\le\mathbf{p}\le\max(p_i,\tilde{p})\right)<0. \tag{A.4}$$

Observe that:

$$\frac{1}{n}\sum_{i\in I}(F(\tilde{p})-F(p_i))=0,$$

[17] Using the notation that $\int_a^b f(x)\,dx=-\int_b^a f(x)\,dx$ when $b<a$.

and that $E(\mathbf{p}|\min(p_i,\tilde{p}) \le \mathbf{p} \le \max(p_i,\tilde{p}))$ is strictly increasing in p_i and thus strictly decreasing in $(F(\tilde{p}) - F(p_i))$. This implies that Inequality A.4 holds.

The above arguments show that for each q such that $\min_i (g_i)^{-1}(q) < \max_i (g_i)^{-1}(q)$, $\tilde{g}$ has higher expected value than $(g_i)_{i\in I}$, conditional on $\mathbf{q} = q$. The fact that $(g_i)_{i\in I}$ is an heterogeneous bias profile (i.e., that $\min_i (g_i)^{-1}(q) < \max_i (g_i)^{-1}(q)$ in a set with positive Lebesgue measure), implies that $\tilde{g}$ has higher expected value than $(g_i)_{i\in I}$, without conditioning on the value of $\mathbf{q}$. By the law of large numbers, if the number of agents is sufficiently large then it implies that with high probability $\tilde{g}$ would have induce strictly larger payoff then $(g_i)_{i\in I}$, and thus would be more preferred by the principal.

A.3 Proof of Proposition 7

Proposition 7 Assume $h_1 = \psi \circ h_2$, where ψ is concave and increasing. Let g_1^* (g_2^*) be the unique optimal bias function given that the principal's utility is h_1 (h_2). Then, $g_1^*(p) > g_2^*(p)$ for every $0 < p < 1$.

PROOF. The expected payoff of a principal with utility $\psi \circ h$ is:

$$u = q \cdot \psi\left(h\left(L + (H-L)\cdot\left(F(p) + \int_p^1 x\cdot f(x)\,dx\right)\right)\right) \\ + (1-q)\cdot\psi\left(h\left(L + (H-L)\cdot\int_p^1 x\cdot f(x)\,dx\right)\right) + o(\epsilon).$$

Looking for an internal solution ($\frac{du}{dp} = 0$) yields:

$$\frac{\psi'(h(\Theta))\cdot h'(\Theta)}{\psi'(h(\Theta + (H-L)\cdot F(p_0)))\cdot h'((\Theta + (H-L)\cdot F(p_0)))} = \frac{q - q\cdot p}{p - q\cdot p}. \tag{A.5}$$

Where $\Theta = L + (H-L)\cdot\int_p^1 x\cdot f(x)\,dx$. Observe that for any $0 < p < 1$ the l.h.s. of Eq. A.5 is larger than the l.h.s. of Eq. A.2 due to the concavity of ψ. This implies that $q = g^*(p)$ has to be larger (as the r.h.s. is an increasing function of q). That is, a more risk-averse principal induces higher level of overconfidence.

A.4 Proof of Theorem 8

Theorem 8 Assume that the principal's utility h satisfies strictly decreasing absolute risk-aversion. That is, Arrow-Pratt coefficient of absolute risk-aversion $r_A(x) = -\frac{h''(x)}{h'(x)}$ is a decreasing function of x. Then higher level of optimal overconfidence ($g_2^*(p) > g_1^*(p)$ for every $0 < p < 1$) is induced by:

(1) Higher payoff for success: $H_2 > H_1$.
(2) Lower payoff for failure: $L_2 < L_1$.
(3) Harder tasks (less accurate private signals): If $\mathbf{p}_2$ has first order stochastically dominance over $\mathbf{p}_1$.

We begin by showing the following simple lemma:

Lemma 14 $f_a(y) = \frac{h'(y)}{h'(y+a)}$ is a strictly decreasing function of y for each $a > 0$.

PROOF. (Lemma 14) Observe that the assumption that h satisfies decreasing absolute risk-aversion implies that for every $y, a > 0$:

$$r_A(y) > r_A(y+a) \Leftrightarrow -\frac{h''(y)}{h'(y)} > -\frac{h''(y+a)}{h'(y+a)}$$

$$\Leftrightarrow \frac{h''(y)}{h'(y)} < \frac{h''(y+a)}{h'(y+a)} \Leftrightarrow h''(y) \cdot h'(y+a) - h''(y+a) \cdot h'(y) < 0.$$

Observe that:

$$f_a'(y) = \frac{h''(y) \cdot h'(y+a) - h''(y+a) \cdot h'(y)}{(h'(y+a))^2}.$$

The above arguments prove that $f_a(y) = \frac{h'(y)}{h'(y+a)}$ is strictly decreasing.

PROOF. (Theorem 8) We now apply Lemma 14 to prove each of the claims in Theorem 8:

(1) Comparing the l.h.s. of Eq. A.2 in the case where $H_2 > H_1$ yields:

$$\frac{h'\left(L + (H_2 - L) \cdot \int_p^1 x \cdot f(x)\, dx\right)}{h'\left(L + (H_2 - L) \cdot \left(F(p) + \int_p^1 x \cdot f(x)\, dx\right)\right)} >$$

$$\frac{h'\left(L + (H_2 - L) \cdot \int_p^1 x \cdot f(x)\, dx\right)}{h'\left(L + (H_1 - L) \cdot F(p) + (H_2 - L) \cdot \left(\int_p^1 x \cdot f(x)\, dx\right)\right)} >$$
$$\frac{h'\left(L + (H_1 - L) \cdot \int_p^1 x \cdot f(x)\, dx\right)}{h'\left(L + (H_1 - L) \cdot \left(F(p) + \int_p^1 x \cdot f(x)\, dx\right)\right)}.$$

The first inequality holds due to the concavity of h, and the second inequality holds due to Lemma 14, with $y_1 = L + (H_1 - L) \cdot \int_p^1 x \cdot f(x)\, dx$, $y_2 = L + (H_2 - L) \cdot \int_p^1 x \cdot f(x)\, dx > y_1$, and $a = (H_1 - L) \cdot F(p)$. Thus, for each $0 < p < 1$ the l.h.s. of Eq. A.2 is larger for H_2, and this implies that the r.h.s. has to be larger, and thus $q_2 = g_2^*(p) > g_1^*(p) = q_1$.

(2) Comparing the l.h.s. of Eq. A.2 in the case where $L_2 < L_1$ yields:

$$\frac{h'\left(L_2 + (H - L_2) \cdot \int_p^1 x \cdot f(x)\, dx\right)}{h'\left(L_2 + (H - L_2) \cdot \left(F(p) + \int_p^1 x \cdot f(x)\, dx\right)\right)} >$$

$$\frac{h'\left(L_2 + (H - L_2) \cdot \int_p^1 x \cdot f(x)\, dx\right)}{h'\left(L_2 + (H - L_1) \cdot F(p) + (H - L_2) \cdot \left(\int_p^1 x \cdot f(x)\, dx\right)\right)} >$$

$$\frac{h'\left(L_1 + (H - L_1) \cdot \int_p^1 x \cdot f(x)\, dx\right)}{h'\left(L_1 + (H - L_1) \cdot \left(F(p) + \int_p^1 x \cdot f(x)\, dx\right)\right)}.$$

The first inequality holds due to the concavity of h, and the second inequality holds due to Lemma 14, with $y_1 = L_1 + (H - L_1) \cdot \int_p^1 x \cdot f(x)\, dx = H \cdot \int_p^1 x \cdot f(x)\, dx - L_1 \left(1 - \int_p^1 x \cdot f(x)\, dx\right)$, $y_2 = L_2 + (H - L_2) \cdot \int_p^1 x \cdot f(x)\, dx - H \cdot \int_p^1 x \cdot f(x)\, dx - L_2 \left(1 - \int_p^1 x \cdot f(x)\, dx\right) > y_1$, and $a = (H - L_1) \cdot F(p)$. Thus, for each $0 < p < 1$ the l.h.s. of Eq. A.2 is larger for L_2, and this implies that the r.h.s. has to be larger, and thus $q_2 = g_2^*(p) > g_1^*(p) = q_1$.

(3) Comparing the l.h.s. of Eq. A.2 in the case where $\mathbf{p}_2$ has first order stochastically dominance over $\mathbf{p}_1$ (i.e., $F_1(p) \leq F_2(p)$ for every p yields):

$$\frac{h'\left(L + (H - L) \cdot \int_p^1 x \cdot f_2(x)\, dx\right)}{h'\left(L + (H - L) \cdot \left(F_2(p) + \int_p^1 x \cdot f_2(x)\, dx\right)\right)} =$$

$$\frac{h'\left(L + (H - L) \cdot \int_p^1 (1 - F_2(x))\, dx\right)}{h'\left(L + (H - L) \cdot \left(F_2(p) + \int_p^1 (1 - F_2(x))\, dx\right)\right)} >$$

$$\frac{h'\left(L + (H - L) \cdot \int_p^1 (1 - F_2(x))\, dx\right)}{h'\left(L + (H - L) \cdot \left(F_1(p) + \int_p^1 (1 - F_2(x))\, dx\right)\right)} >$$

$$\frac{h'\left(L + (H - L) \cdot \int_p^1 (1 - F_1(x))\, dx\right)}{h'\left(L + (H - L) \cdot \left(F_1(p) + \int_p^1 (1 - F_1(x))\, dx\right)\right)}.$$

The first inequality holds due to the concavity of h, and the second inequality holds due to Lemma 14, with $y_1 = L + (H - L) \cdot \int_p^1 (1 - F_1(x))\, dx$, $y_2 = L + (H - L) \cdot \int_p^1 (1 - F_2(x))\, dx > y_1$, and $a = (H - L) \cdot F_1(p)$. Thus, for each $0 < p < 1$ the l.h.s. of Eq. A.2 is larger for F_2, and this implies that the r.h.s. has to be larger, and thus $q_2 = g_2^*(p) > g_1^*(p) = q_1$.

A.5 Proof of Theorem 9

Theorem 9 Assume that the principal has a CRRA utility function with parameter θ. Then:

$$g^*(p) = \frac{Bp}{1-p+Bp}, \text{ where } B = \left(1 + \frac{D \cdot F_{\mathbf{p}}(p)}{1 + D \cdot \int_p^1 x \cdot f_{\mathbf{p}}(x)\, dx}\right)^{\phi},$$

and it satisfies the following properties:

(1) overconfidence $(g^*(p) - p)$ is strictly increasing in the potential gain $D = \frac{H-L}{L}$.
(2) The ratio between the perceived and the true probability that the private signal is incorrect $(\frac{1-g(p)}{1-p})$ is strictly decreasing in p, and it converges to $\left(\frac{H}{L}\right)^{\phi}$ as p converges to 1.

PROOF. We begin by proving the above formula for $g^*(p)$.

Placing $h'(x) = x^{-\phi}$ in Eq. (A.2) yields:

$$\frac{\left(L + (H-L) \cdot \int_p^1 x \cdot f(x)\, dx\right)^{-\phi}}{\left(L + (H-L) \cdot \left(F(p_0) + \int_p^1 x \cdot f(x)\, dx\right)\right)^{-\phi}} = \frac{q - q \cdot p}{p - q \cdot p} \Rightarrow$$

$$\left(\frac{\frac{L}{L} + \frac{(H-L)}{L} \cdot \left(F(p) + \int_p^1 x \cdot f(x)\, dx\right)}{\frac{L}{L} + \frac{(H-L)}{L} \cdot \int_p^1 x \cdot f(x)\, dx}\right)^{\phi} = \frac{q - q \cdot p}{p - q \cdot p} \Rightarrow$$

$$\left(1 + \frac{\left(\frac{H-L}{L}\right) \cdot F(p)}{1 + \left(\frac{H-L}{L}\right) \cdot \int_p^1 x \cdot f(x)\, dx}\right)^{\phi} = \frac{q - q \cdot p}{p - q \cdot p}.$$

Substituting $D = \frac{H-L}{L}$ gives:

$$\left(1 + \frac{D \cdot F(p)}{1 + D \cdot \int_p^1 x \cdot f(x)\, dx}\right)^{\phi} = \frac{q - q \cdot p}{p - q \cdot p}. \tag{A.6}$$

Let B be defined as follows:

$$B = B(p, f, D) = \left(1 + \frac{D \cdot F(p)}{1 + D \cdot \int_p^1 x \cdot f(x)\, dx}\right)^{\phi}.$$

Observe that: 1) $B > 1$, 2) B is increasing in p, and 3) when $p \to 1$ B converges to $(1+D)^{\phi}$ (and $\frac{dB}{dp}$ converges to 0). Placing B in Eq. (A.6) yields:

$$B = \frac{q - q \cdot p}{p_0 - q \cdot p} \Rightarrow Bp - qBp = q - qp.$$

Isolating q gives:

$$q = \frac{Bp}{1 - p + Bp}.$$

Substituting q by $g^*(p)$ gives:

$$g^*(p) = \frac{Bp}{1 - p + Bp}.$$

We continue by proving the two properties:

(1) We have to show that $g^*(p)$ is increasing in the potential gain D. Observe that the l.h.s. of Eq. (A.6) is increasing in D because:

$$\frac{d}{dD}\left(\frac{D \cdot F(p)}{1 + D \cdot \int_p^1 x \cdot f(x)\, dx}\right) =$$

$$\frac{F(p)\left(1 + D \cdot \int_p^1 x \cdot f(x)\, dx\right) - D \cdot F(p) \int_p^1 x \cdot f(x)\, dx}{\left(1 + D \cdot \int_p^1 x \cdot f(x)\, dx\right)^2} =$$

$$\frac{F(p)}{\left(1 + D \cdot \int_p^1 x \cdot f(x)\, dx\right)^2} > 0.$$

This implies that $p = \alpha(q)$ is strictly decreasing in D, and thus $g^*(p) = \alpha^{-1}(p)$ is strictly increasing in D.

(2) Observe that $1 - g(p)$ is equal to:

$$1 - g(p) = \frac{1 - p}{1 + p(B - 1)}.$$

Thus $\frac{1-g(p)}{1-p}$ is equal to:

$$\frac{1 - g(p)}{1 - p} = \frac{1}{1 + p(B - 1)},$$

which is decreasing in p, and when $p \longrightarrow 1$ it converges to:

$$\frac{1}{B} = \frac{1}{(1 + D)^{\phi}} = \left(\frac{H}{L}\right)^{\phi}.$$

A.6 Proof of Proposition 10

Proposition 10 For each $n \geq 1$ and $k \geq 2$ the principal can induce a strictly better outcome when the number of agents is $k \cdot n$ than when it is n.

PROOF. Let $(g_i)_{i \in I}$ be a bias profile in the game with $n = |I|$ agents. Recall that for each agent $i \in I$, $\mathbf{u}_i$ is the random payoff of agent i with bias function g_i, and that the principal's payoff is $h\left(\frac{1}{n}\sum_{i\in I}\mathbf{u}_i\right)$. Consider $(g_i)_{i\in I}$ as a profile in the game with $k\cdot n$ agents (where each k agents have one of the bias functions g_i). This profile induces the following payoff:

$$h\left(\frac{1}{n}\sum{}_{i\in I}\frac{1}{k}\sum_{j=1}^{k}\mathbf{u}_{(i-1)\cdot k+j}\right),$$

where for each i, the variables $\left(\mathbf{u}_{(i-1)\cdot k+j}\right)_{j=1,\ldots,k}$ are identically distributed. Observe that $\frac{1}{n}\sum{}_{i\in I}\frac{1}{k}\sum_{j=1}^{k}\mathbf{u}_{(i-1)\cdot k+j}$ second-order stochastically strictly dominates $\frac{1}{n}\sum_{i\in I}\mathbf{u}_i$. By the concavity of h, it implies that the principal strictly prefers the outcome in the game with $k\cdot n$ agents. Thus, any outcome in the game with n agents is strictly dominated by an outcome in the game with $k\cdot n$ agents.

A.7 Proof of Proposition 12

Proposition 12 The extended model with costly signals admits a unique optimal bias function g^*, which is the same as the optimal bias function g^* of the basic model with $f_{\mathbf{p}} = f_{p_{\mathbf{t}}}$.

PROOF. We begin by calculating the first-best profile in a game with many agents $n >> 1$. Without loss of generality for each public accuracy $q\in[0,1]$, there is some effectiveness value $t_0=\alpha(q)$ such that the optimal payoff can be induced by all agents using the same threshold strategy: 1) agents with low effectiveness ($\mathbf{t}_i < t_0$) do not invest any effort and follow the public signal, and 2) agents with high effectiveness ($\mathbf{t}_i \geq t_0$) invest some effort and follow the private signal.

Consider an agent with high effectiveness: $\mathbf{t}\geq t_0$. His expected payoff from investing effort e is $L+(H-L)\cdot(p(e,t)-e)$. This is maximized in e_t^* that satisfies $\frac{d(p(e,t))}{de}=1$ (a unique maximizer exists due to the concavity of $p(e,t)$). Let $p_t^*=p(e^*,t)$. For large enough n, if all agents with high effectiveness invest effort e_v^*, it ϵ-maximizes the principal's payoff (by the law of large numbers).

Let $p_0=p_{t_0}^*$ be the accuracy level of an agent with threshold effectiveness value t_0. The choice of an optimal threshold t_0 is equivalent to the problem solved in Subsection 4.3 - finding the optimal accuracy threshold p_0. Thus the optimal bias function g^* of the basic model (Subsection 4.3) is also optimal in the extended model (with $f_{\mathbf{p}}=f_{p_{\mathbf{t}}}$).

References

[1] **Alicke Mark D., Olesya Govorun.** 2005. "The better-than-average effect". In: *The self in social judgment*, ed. Mark D. Alicke, David A. Dunning, and Justin I. Krueger, New York:

Psychology Press.

[2] **Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey.** 2010. "Managerial Miscalibration" NBER Working Paper #w16215, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1648015.

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

[4] **Benoit, Jean P. and Juan Dubra**. 2009. "Overconfidence?" Unpublished manuscript. http://ssrn.com/abstract=1088746

[5] **Blume, Lawrence, and David Easley**. 1992. "Evolution and Market Behavior" *Journal of Economic Theory*, 58, 9-40.

[6] **Bernardo, Antonio E., and Ivo Welch.** 2001. "On the evolution of overconfidence and entrepreneurs" *Journal of Economics & Management Strategy*, 10(3), 301-330.

[7] **Brenner, Lyle, Dale Griffin, and Derek J. Koehler.** 2005. "Modeling patterns of probability calibration with random support theory: diagnosing case-based judgment" *Organizational Behavior and Human Decision Processes*, 97: 64-81.

[8] **Brenner, Lyle A., Derek J. Kohler, Varda Liberman, and Amos Tversky.** 1996. "Overconfidence in probability and frequency Judgments: a critical examination" *Organizational Behavior and Human Decision Processes,* 65(3): 212-219.

[9] **Budescu, David V., Ning Du**. 2007. "Coherence and consistency of investors' probability judgments" *Management Science*, 53(11): 1731-1744.

[10] **Budescu, David V., Thomas Wallsten, and Wing Tung Au**. 1997. "On the importance of random error in the study of probability judgment. Part II: applying the stochastic judgment model to detect systematic trends" *Journal of Behavioral Decision Making,* 10: 173-188.

[11] **Busenitz, Lowel W., Jay B. Barney**. 1997. "Differences between entrepreneurs and managers in large organizations: biases and heuristics in strategic decision-making" *Journal of Business Venturing*, 12: 9-30.

[12] **Chuang, Wen-I, and Bong-Soo Lee**. 2006. "An empirical evaluation of the overconfidence hypothesis" *Journal of Banking and Finance* 30, 2489-2515.

[13] **Compte, Oliver, and Andrew Postlewaite**. 2004. "Confidence-Enhanced Performance" *The American Economic Review*, 94(5), 1536-1557.

[14] **Cooper, Arnold C., Carolyn Y. Woo, and William C. Dunkelberg**. 1988. "Entrepreneurs' perceived chances for success" *Journal of Business Venturing* 3, 97-108.

[15] **Erev, Ido, Thomas S. Wallsten, David V. Budescu**. 1994. "Simultaneous over and underconfidence: the role of error in judgment processes" *Psychological Review*, 101(3): 519-527.

[16] Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein. 1977. "Knowing with certainty: the appropriateness of extreme confidence" *Journal of Experimental Psychology: Human Perception and Performance,* 3(4): 552-564.

[17] **Gervais, Simon, J.B. Heaton, Terrance Odean**. 2010. "Overconfidence, Compensation Contracts, and Capital Budgeting" *Journal of Finance*, forthcoming.

[18] **Gervais, Simon, Terrance Odean**. 2001. "Learning to be overconfident" *The Review of Financial Studies*, 14(1): 1-27.

[19] **Gigerenzer, Gerd, Ulrich Hoffrage, and Hienz Kleinbolting.** 1991. "Probabilistic mental models: a Brunswikian theory of confidence" *Psychological Review*, 98(4): 506-528.

[20] **Glaser, Markus, Thomas Langer, and Martin Weber**. 2010. "Overconfidence of professionals and lay people: individual differences within and between tasks?" unpublished manuscript. http://ssrn.com/abstract=712583

[21] **Goel, Anand M, Anjan V. Thakor**. 2008. "Overconfidence, CEO selection and Corporate Governance" *The Journal of Finance*, LXIII(6): 2737-2783.

[22] **Griffin, Dale, and Lyle Brenner.** 2004. "Perspectives on probability judgment calibration" In: *Blackwell Handbook of Judgment and Decision Making*, ed. Derek J. Koehler and Nigel Harvey, 177–99, Boston: Blackwell Publishing.

[23] **Griffin, Dale, and Amos Tversky.** 1992. "The weighing of evidence and the determinants of confidence" *Cognitive Psychology,* 24: 411-435.

[24] **Grubb, Michael D.** 2009. "Selling to overconfident consumers" *American Economic Review*, 99(5): 1770-1807.

[25] **Henrion, Max, and Baruch Fischhoff.** 2002. "Assessing uncertainty in physical constants." In: *Heuristics and Biases: The Psychology of Intuitive Judgment, ed.* T. Gilovich, Dale Griffin and Daniel Kahneman, 666-677, Cambridge: Cambridge University Press.

[26] **Howell, William C., and Sarah A. Burnett.** 1978. "Uncertainty measurement: a cognitive taxonomy" *Organizational Behavior and Human Performance* 22, 45-68.

[27] **Kahneman, Daniel and Amos Tversky.** 1982. "Variants of Uncertainty" *Cognition*, 11: 143-157.

[28] **Keren, Gideon**. 1987. "Facing uncertainty in the game of Bridge: a calibration study" *Organizational Behavior and Human Decision Processes*, 39: 98-114.

[29] **Klayman, Joshua, Jack B. Soll, Claudia González-Vallejo, and Sema Barlas**. 1999. "Overconfidence: it depends on how, want and whom you ask" *Organizational Behavior and Human Decision Processes*, 79(3): 216-247.

[30] **Koehler, Derek J., Lyle Brenner, and Dale Griffin**. 2002. "The calibration of expert judgment: Heuristics and biases beyond the laboratory. In: *Heuristics and Biases: The Psychology of Intuitive Judgment, ed.* T. Gilovich, Dale Griffin and Daniel Kahneman, 686-715, Cambridge: Cambridge University Press.

[31] **Köszegi, Botond**. 2006. "Ego utility, overconfidence and task choice." *Journal of European Economic Association*, 4(4): 673-707.

[32] **Lewontin Richard C., and D. Cohen**. 1969. "On population growth in a randomly varying environment." *Proceedings of the National Academy of Sciences of USA*, 62: 1056-60.

[33] **Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D. Phillips**, 1982. "Calibration of probabilities: the state of the art to 1980." In *Judgment Under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slovic, and Amos Tversky, Cambridge: Cambridge University Press.

[34] **Mcnamara, John M**. 1995. "Implicit frequency dependence and kin selection in fluctuating environments" *Evolutionary Ecology,* 9: 185-203.

[35] **Moore. Don A.** 2007. "Not so above average after all: when people believe they are worse than average and its implications for theories of bias in social comparison" *Organizational Behavior and Human Decision Processes*, 102: 42-58.

[36] **Moore, Don A., and Paul J. Healy.** 2008. "The trouble with overconfidence" *Psychological Review* 2008, 115 (2): 502-517.

[37] **von Neumann, John, and Oscar Morgenstern.** 1944 (third edition, 1953). *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

[38] **Novemsky, Nathan, and Shirit Kronzon**. 1999. "How are base-rates used, when they are used: a comparison of additive and Bayesian models of base-rate use" *Journal of Behavioral Decision Making*, 12: 55-69.

[39] **Odean, Terrance**. 1998. "Volume, volatility, price, and profit when all traders are above average." *The Journal of Finance*, LIII(7): 1887-1934.

[40] **Oskamp, Stuart.** 1965. "Overconfidence in case-study judgments" *Journal of Cognitive Psychology*, 29(3), 261-265.

[41] **Robson Arthur J**. 1996. "A biological basis for expected and non-expected utility" *Journal of economic Theory*, 68: 397-424.

[42] **Russo, Edward J., Paul J.H. Schoemaker**. 1992. "Managing overconfidence" *Sloan Management Rev.* 33: 7-17.

[43] **Sandroni, Alvaro, and Francesco Squintani**. 2007. "Overconfidence, insurance and paternalism" *The American Economic Review*, 97(5), 1994-2004.

[44] **Sieber, Joan E**. 1974. "Effects of decision importance on ability to generate warranted subjective uncertainty" *Journal of Personality and Social Psychology*, 30(5): 688-694.

[45] **Skala, Dorota.** 2008. "Overconfidence in psychology and finance - an interdisciplinary literature review" *Bank i Kredy*, 4: 33-50.

[46] **Soll, Jack B.** 1996. "Determinants of Overconfidence and Miscalibration: The Roles of Random Error and Ecological Structure" *Organizational Behavior and Human Decision Processes*, 65(2): 117-137.

[47] **Soll, Jack B., Joshua Klayman**. 2004. "Overconfidence in interval estimates" *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(2), 299-314.

[48] **Svenson, Ola**. 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychological* 47, 143-148.

[49] **Taylor, Shelley E., Jonathan D. Brown**. 1988. "Illusion and well-being: a social psychological perspective on mental health" *Psychological Bulletin,* 103(2), 193-210.

[50] **Teigen, Karl H., and Magne Jorgensen**. 2005. "When 90% confidence intervals are 50% certain: on the credibility of credible intervals" *Applied cognitive Psychology*, 19: 455-475.

[51] **Van Den Steen, Eric**. 2004. "Rational overoptimism (and other biases)" *American Economic Review*, 94(4): 1141-1151.

[52] **Waldman, Michael.** 1994. "Systematic errors and the theory of natural selection" *American Economic Review*, 84(3): 482-497.

[53] **Wang, F. Albert**. 2001. "Overconfidence, investor sentiment, and evolution" *Journal of Financial Intermediation* 10, 138-170.

[54] **Weinberg, Bruce A.** 2009. "A model of overconfidence" *Pacific Economic Review*, 14(4): 502-515.

[55] **Yates, J. Frank, Ju-Whei Lee, Winston R. Sieck, Incheol Choi, and Paul C. Price.** 2002. "Probability judgment across cultures." In: *Heuristics and Biases: The Psychology of Intuitive Judgment, ed.* T. Gilovich, Dale Griffin and Daniel Kahneman, 271-291, Cambridge: Cambridge University Press.