HARVARD | BUSINESS | SCHOOL

Overconfidence by Bayesian Rational Agents

Eric Van den Steen

Working Paper

11-049

Overconfidence by Bayesian Rational Agents

Eric Van den Steen[*]

November 2010

Abstract

This paper derives two mechanisms through which Bayesian-rational individuals with differing priors will tend to be relatively overconfident about their estimates and predictions, in the sense of overestimating the precision of these estimates. The intuition behind one mechanism is slightly ironic: in trying to update optimally, Bayesian agents overweight information of which they over-estimate the precision and underweight in the opposite case. This causes overall an over-estimation of the precision of the final estimate, which tends to increase as agents get more data.

1   Introduction

People, even experts, generally overestimate the precision of their estimates and predictions (Oskamp 1965, Fischhoff, Slovic, and Lichtenstein 1977, Lichtenstein, Fischhoff, and Phillips 1982, Wallsten, Budescu, and Zwick 1993, Barberis and Thaler 2003). For example, in an experiment that is often repeated in a class context (in slightly different form) to give students first-hand experience with this phenomenon, Alpert and Raiffa (1969) asked a large number of students to estimate an inter-fractile range for unknown quantities. For example, for the (presumably unknown) quantity "total egg production in the US in 1965" students were asked to estimate the range such that the true value would fall in that range with a 98% probability. If students were appropriately confident, then on average only two percent of them should miss the true value with their range. Instead, the percentage is typically much higher, often even exceeding 50%. The subjects thus clearly over-estimate the precision of their estimates. The seminal, and still very relevant, work in this area was Oskamp (1965), who had (experienced) clinical psychologists and psychology students make predictions about one particular person's actions and behavior based on factual information about that individual. (The person about who the questions were asked had been the subject of a detailed psychological case study, so that information about his actual actions and behavior was available.) As the subjects got more information about that person, they could revise earlier answers and were also asked to report their confidence level in their assessments. The evidence showed that the subjects were considerable overconfident about their assessments and became *more* overconfident as they got more information.

Apart from its important role in personal and business decision making – as illustrated, for example, by Russo and Schoemaker (1992) and Grubb (2009) – this bias is particularly relevant in financial settings. Daniel, Hirshleifer, and Subrahmanyam (1998, 2001), for example, showed that overconfidence can lead to excess volatility and to predictability of stock returns. Scheinkman and Xiong (2003) showed how it can lead to financial bubbles, a frequent cause of financial crises (Allen and Carletti 2010), which has recently taken on extra significance.

While the psychology literature and the economics literature have forwarded explanations for the phenomenon of overconfidence, as I discuss below, this paper proposes two new mechanisms that are structurally different from the existing explanations. Both mechanisms in this paper assume that people are perfectly Bayesian rational, but they do not require that people necessarily share a common prior, a critical assumption that I discuss in more detail later. Agents can thus openly disagree in their beliefs, in the sense of having differing priors.[1] To keep the analysis simple and transparent, I will consider separately the cases where people disagree on the mean versus where they disagree on the precision (or informativeness) of their estimates and of new data. The paper then shows that overconfidence arises naturally in each of these two settings, through two different mechanisms. The first mechanism is the fairly direct result that when agents' prior beliefs differ (only) in their mean but have otherwise – conditional on that mean – identical distributions, each agent will be overconfident from the other's perspective. The reason is that the two agents will have different confidence sets (in the sense of the minimal subsets of the parameter space such that the agent is confident that the true value will fall into that subset with probability α or more). But then one agent's confidence in the other's α-confidence set will be lower than α, at least weakly and often strictly so. While this result already establishes a fairly direct link between differing priors and overconfidence, the paper derives a second and more striking result for the case of disagreement about the variance: Bayesian updating can *endogenously generate* overconfidence when the agents have differing priors. The intuition behind the latter result is as follows. In trying to find the best estimate, Bayesian rational individuals put more weight on information that is more precise. As a consequence, they put too much weight on the information of which they over-estimate the precision and not enough on that of which they under-estimate the precision. This generates an overall overconfidence, *even when* their original precision estimates are on average unbiased, i.e., even when their original variance estimates are on average correct.

Note that, by themselves, neither mechanism generates the hard-easy effect, i.e., the effect that people tend to be relatively more overconfident on harder assessments tasks than on easier assessment tasks (Lichtenstein and Fischhoff 1977). However, both models turn out to be consistent with this important effect, in the sense that they both generate a hard-easy effect under some additional assumptions. One such assumption is that subjects are more likely to disagree (about means or variances) for hard assessment tasks than for easy assessment tasks. While this assumption seems reasonable, it is obviously an empirical question whether it holds in the data.

Literature   The psychology literature has suggested a number of explanations for overconfidence, which all differ (at least somewhat) in their predictions and comparative statics. Most of these explanations derive overconfidence from other forms of bounded rationality or from other biases. One such explanation is self-enhancement: people entertain inflated views of their own abilities since doing so improves their mental health (Taylor and Brown 1988). If we take 'making a precise

---

[1]To make sure that I don't introduce a bias directly through the assumptions, I will assume that in each case all agents' prior beliefs are independently drawn from the same distribution, so that no agent is originally biased relative to any other agent.

estimate' as the relevant ability, then this self-enhancement bias leads indeed to an overestimation of the precision of one's estimates. In terms of comparative statics, however, there seems to be no immediate reason why a bias from this source would change, for example, with the amount of information available or with the degree to which people disagree. A second explanation, forwarded by Tversky and Kahneman (1974), is that the 'anchoring and adjustment' bias could lead to overconfidence, at least for the interfractile estimates. The idea here is that people start by formulating their best estimate – i.e, they try to estimate the mean or median – and then adjust that 'best estimate' to find the higher and lower fractiles. 'Anchoring and adjustment' would then indeed lead to insufficient adjustments from the mean/median to the fractiles and thus to an overly narrow confidence interval. But it seems that overconfidence through this mechanism should be very sensitive to the procedure for eliciting the confidence intervals. Moreover, and relatedly, it is also less clear how this mechanism would explain the results of Oskamp (1965) where the confidence level is not an 'adjustment' to their base estimate but a separate estimate. Finally, as with the self-enhancement explanation, this bias seems again independent of the amount of information or the degree to which people disagree. Another explanation was suggested by Einhorn and Hogarth (1978) who argue that overconfidence might be caused by the documented tendency of people to look for confirming evidence, rather than for disconfirming evidence, for their particular hypothesis. If so, the resulting balance between a lot of confirming evidence and a lack of disconfirming evidence might seem to lead the person to hold overly strong beliefs in her original hypothesis. However, as already pointed out by Klayman and Ha (1987), such bias for confirmation does not generate overconfidence without an extra assumption: if a person is well-calibrated about the informativeness of a confirmation when only looking for confirmation, such bias does not cause overconfidence but simply leads to a suboptimal collection of evidence.[2] To get overconfidence out of this bias requires the extra assumption that the person places too much trust in confirming evidence, as suggested, for example, by Lord, Ross, and Lepper (1979). Moreover, it is unclear how this mechanism works for interval estimates and for situations where no new evidence is collected, as in Alpert and Raiffa (1969). Erev, Wallsten, and Budescu (1994), finally, showed how some overconfidence results may actually be a statistical artifact caused by regression towards the mean. The idea is as follows. Assume that people always make some error in estimating their confidence. For nearly any distribution of objective confidence levels, a person with a subjective 99% confidence level is more likely to have started from an objective confidence level between 0 and 99% and have made an upwards error than to have started from an objective confidence level between 99% and 100% and have made a downwards error. This is a form of regression towards the mean leading to more extreme subjective confidence levels. While this mechanism is also driven by errors in the variance estimates, it is not generated by trying to combine data in the optimal way, as the mechanism in section 5. It therefore has very different properties and comparative statics. First, and most importantly, the Erev, Wallsten, and Budescu (1994) effect only works for confidence levels above the mean confidence level. It goes in the *opposite* direction below the mean level: a person asked to give her 10% confidence interval will have a hitrate higher than 10% according to the Erev, Wallsten, and Budescu (1994) mechanism. And there is some confidence level, typically close to 50%, for which subjects will be exactly well-calibrated. That is obviously very different from the global overoptimism derived in this paper. Second, as with most other mechanisms, there is again

---

[2]In particular, a rational person who, for some exogenous reason, only samples from potentially confirming evidence would correct his updating for this bias and thus overall be unbiased. More precisely, the person would discount when he finds confirming evidence (since he knows he had only been looking for confirming evidence) and play up any lack of such evidence.

no reason why more data points would increase overconfidence.

There is also a relatively small economics literature that has largely followed an approach very similar to that of the psychology literature by trying to derive overconfidence from other biases. Rabin and Schrag (1999), for example, show how a confirmatory bias – the tendency of people to *interpret* ambiguous evidence as confirming their current hypothesis – may lead to overconfidence. Brunnermeier and Parker (2005) show that if agents can choose their beliefs optimally – making a trade-off between the immediate utility from having favorable beliefs against the costs of making mistakes due to mistaken beliefs – that overconfidence will sometimes result. All these explanations thus either also assume that agents can purposely bias their beliefs (for example, by simply choosing advantageous beliefs or by remembering favorable things better than other) or rely on some other form of cognitive bias (for example, a bias in updating or a bias in looking for evidence). This paper, on the contrary, assumes that people are perfectly Bayesian rational with *unbiased* differing priors, in the sense that the prior beliefs of the subjects and observers are drawn from the same distribution so that neither is ex-ante biased relative to the other. There is also an interesting connection between this paper and Kogan (2009) who shows that people may underweight others' opinions, relative to their own, as a rational response to the others' overconfidence. Due to that strategic compensation, perfectly Bayesian rational agents may act 'as if' they are overconfident. An important difference with this paper is that such agents would not actually be overconfident if they are well-calibrated with respect to others' overconfidence. His analysis also makes it clear that the connection between disagreement and overconfidence is a two-way street: while this paper explores how disagreement may lead to overconfidence, overconfidence may also cause disagreement both directly and through a rational response to others' overconfidence.

The second mechanism in this paper is related to the mechanisms in Van den Steen (2004). That paper showed that rational agents with differing priors will be (subjectively) overoptimistic about their chances of success: when trying to choose the action that is most likely to succeed, such agent is more likely to select an action on which he overestimated, rather than underestimated, the likelihood of success. While the process of forming a confidence judgment is very different from that of choosing an action, the two papers share the idea of relying more on things for which one overestimates the benefits and less on things from which one underestimates the benefits, which then leads to some kind of bias. But the meaning, implications, and insights are obviously quite different in this context of information processing.

Contribution The key contribution of this paper is to propose two new and simple mechanisms through which Bayesian rational people with unbiased differing prior beliefs will come to systematically overestimate the precision of their estimates and to study when these mechanisms are prevalent. An important result is that, for at least one of the two mechanisms, the overconfidence bias tends to increase with the addition of new information (even if the person cannot herself select the additional evidence).

Sections 2 and 3 discuss respectively the differing priors assumption and various ways to operationalize and measure the concept 'overconfidence'. Sections 4 and 5 develop the two channels through which disagreement may cause overconfidence. Section 6 discusses these results while section 7 concludes.

4

2 Differing Priors

The models in this paper assume that agents can openly disagree, i.e., they can agree to disagree. This implies that agents must have differing prior beliefs (Aumann 1976), which captures the fact that people may have different mental models, intuition, or world views. This assumption, while not so common in economics, has a respectable tradition, including Arrow (1964), Harrison and Kreps (1978), Varian (1989), and Morris (1994), and has recently seen increased interest (Acemoglu, Chernozhukov, and Yildiz 2006, Hong and Stein 2007, Boot, Gopalan, and Thakor 2008, Geanakoplos 2009). More detailed practical and theoretical discussions can be found in Morris (1995), the discussion between Gul (1998) and Aumann (1998), Van den Steen (2001), and Van den Steen (2010a). Among other things, Morris (1995) observes that introducing differing priors does not allow to 'explain anything' any more than introducing differing utility functions, information sets, action sets etc. ; Harrison and Kreps (1978), Morris (1994), and Van den Steen (2010b) discount the potential issue of large bets on the state of the world; Van den Steen (2001) shows that the epistemic foundations for Nash equilibria in the sense of Aumann and Brandenburger (1995) extend to this context with differing priors on the payoff-space, and Van den Steen (2010a) explores the differences between differing priors and private benefits.

To simplify the exposition and analysis, I assume in this paper that all disagreement originates in differing priors and that no agent has private information. The absence of private information implies that agents will not change their beliefs when confronted with others who hold different opinions: agents simply believe that such other people are mistaken. Analogously, knowledge of the distribution of beliefs in the population does not lead agents to change their beliefs: agents interpret that distribution of beliefs as the distribution of how others are mistaken. The assumption that others have no private information is obviously extreme and made for analytical convenience.[3]

An obvious question in this context is where such differing priors come from. There are essentially two answers. Building on the fact that the prior of this period is the posterior of last period, the first answer is that any (unconscious) error in updating will lead to differing priors, even when starting originally from a common prior. If, for example, agents sometimes forget what information is already included in their prior and are not fully aware of such forgetfulness, then agents will have differing priors for later periods. This answer (implicitly) interprets the Bayesian model as an 'as if' model or as the best local approximation of human inference (while allowing for the possibility of tiny deviations that don't matter for immediate decisions but that may accumulate over time). But even when the Bayesian model is interpreted strictly as a perfect positive model of human inference, it is not necessarily in conflict with differing priors. In particular, the Bayesian model specifies how we use new information but not what beliefs we happen to start from. Absent any relevant information, agents have no rational basis to agree on a prior. In particular, if there were a rational basis, then we should have a theory on what the uninformed prior should look like. We actually have not one but multiple theories, and they disagree with each other. For example, when applied to the parameter $p \in [0,1]$ of a Binomial distribution, Bayes (1763) and Laplace (1774) suggested the uniform distribution as the uninformed prior belief, Haldane (1932) (on the basis of invariance arguments) suggested the improper prior $p^{-1}(1-p)^{-1}$, Jeffreys (1939) suggested the widely used 'Jeffreys ignorance prior' $p^{-1/2}(1-p)^{-1/2}$, and Zellner (1971) (on the basis of information theoretical principles) suggested $1.6186p^p(1-p)^{(1-p)}$, among others. Such disagreement

---

[3] Although analytically considerably more complex, the results would also obtain in a setting with both differing priors and private information. Morris (1994) considers a trade setting with both differing priors and private information and shows how it leads to bid-ask spreads.

suggests that a 'common prior' is a non-trivial assumption. Harsanyi (1968) furthermore observed that 'by the very nature of subjective probabilities, even if two individuals have exactly the same information and are at exactly the same high level of intelligence, they may very well assign different subjective probabilities to the very same events'. An agent's prior will reflect his personal world view, which may be rooted in things such as that individual's genetics, early (and by now unconscious) life experiences, and things we accept on 'authority'. For example, most of us have rather strong beliefs in the findings of science, such as the fact that the Earth revolves around the Sun and evolution theory. We know of these findings 'on authority': few of us have any first hand observations whether the Earth circles the sun (or the other way around) or whether matter indeed consists of atoms and few have studied real fossils. In past times, people had equally strong beliefs about contradicting theories, also 'on authority', and some people still do.[4]

To be maximally consistent, the results in this paper are formulated completely in subjective terms: 'how does agent i think about the inferences and decisions of agent j'. While this is definitely the most consistent approach, it would also be possible to introduce a 'reference belief,' such as the belief of the researcher or social planner, to derive 'objective' comparisons. This does raise the philosophical question, however, how it is possible that we as the observers know the true distribution while the agents themselves don't. Moreover, it can lead to some confusion about the distinction between the reference prior and a common prior. For these reasons, I decided to formulate everything in subjective terms. For objective results, one only needs to pick one agent and promote this person's beliefs to the reference beliefs.

I will return to the differing assumption in section 6, to discuss its role in the results.

3   Measuring Overconfidence

Before proceeding to the formal analysis it is useful to discuss in more detail the meaning of 'overconfidence' and ways to formally measure it.

As a starting point, consider the experimental setup in the style of Alpert and Raiffa (1969) that I mentioned in the Introduction. The subjects get asked, for example, 10 questions and have to indicate their 90% confidence interval on each. If the subjects were well-calibrated, the average hit rate – i.e., the percentage of times that the correct value falls in the indicated range – should be about 9 out of 10. The fact that the *realized* average is typically much lower is then taken as an indication of overconfidence. This experiment thus measures overconfidence in terms of the subjectively expected and realized hit rates. In other words, the measure of overconfidence is the relative confidence that one agent – the observer who sees the realized hit rate – has in the other agent's confidence interval.

This is clearly not the only possible measure. In particular, the often used informal 'definition' of overconfidence as the 'tendency of people to overestimate the precision of their estimates' (Daniel, Hirshleifer, and Subrahmanyam 1998, Scheinkman and Xiong 2003, Grubb 2009) suggests the ratio of *precisions* as a measure. An intermediate measure with good analytical properties is, of course, the ratio of the estimated variances of the estimator. These variance and precision-based measures, however, are somewhat problematic when the focal agent and the observer have different expected values. Consider, for example, the two distributions in figure 1. Conditional on their respective means, both have the same variance and precision. So it is not very clear whether the precision- or variance-based definition of overconfidence would imply that one belief is overconfident relative
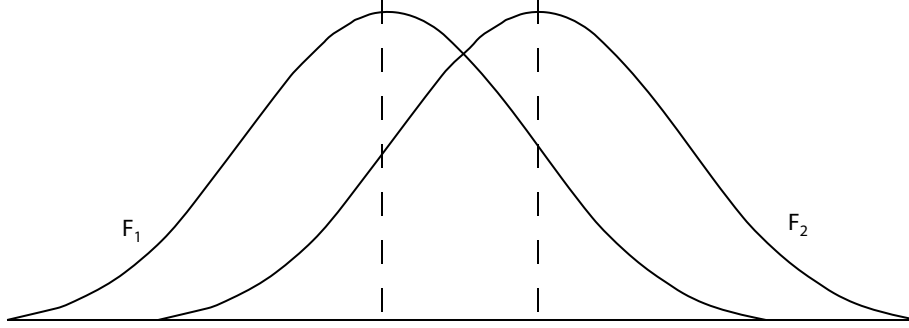
---

[4]See www.fixedearth.com.

Figure 1: Variance versus Hit Rate

to the other. However, and that is the basic result of the next section and will be explained in more detail there, *each* of these beliefs is overconfident from the perspective of the *other* when looking at hit rates. The difference in conclusion is driven by the fact that the variance (and by extension the precision) loses its usual meaning of variation *around* a central tendency when the two distributions have different means. For these and other reasons, the difference between effective and expected confidence (or hit rates) seems to be the most general and the most meaningful measure of overconfidence.

When, however, the focal agent and the observer have the same expected value and parametrically similar belief distributions, then these alternative measures are strictly increasing non-linear transformations of each other. In that case, all measures will always result in the same rankings. They are also consistent under almost sure convergence (since that works fine with non-linear transformations).[5] For the purposes of this paper, they are therefore equivalent measures under these conditions.

Based on these observations, I take in this paper a very practical approach. Section 4, where the focal agent's and observer's means differ, uses the 'hit-rate' of the confidence interval to measure overconfidence. Section 5, where the focal agent and observer have identical parametric distributions and identical means – since they look at the same estimator – measures overconfidence instead by the relative variance of the estimator, since this measure has much better analytical properties. As mentioned earlier, these measures are equivalent for the purposes of this paper in their respective context.

---

[5]With smaller data sets, where we need to work with expectations, the situation is more complex, but the potential differences do not play a role for the results in this paper.
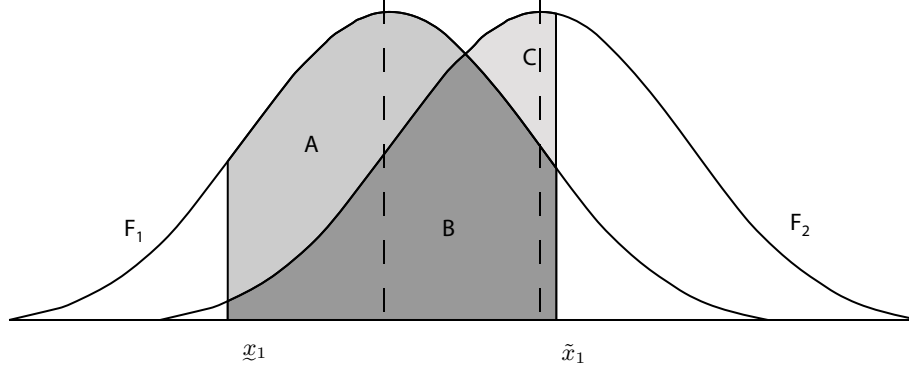
7

Figure 2: Overconfidence with differing means but identical variance

4   Overconfidence through Disagreement on the Expected Value

People may disagree both about the expected value and about the precision of their estimates and of new data. To keep the analysis simple and transparent, I will treat these two sources of disagreement separately and eliminate, each time, any other form of disagreement from the formal analysis. I will thus first consider, in this section, the case that there is disagreement about the mean and about the mean only. I will then consider, in the next section, the polar opposite case where there is disagreement about the precision and about the precision only.

This section then shows how disagreement about the expected value leads almost automatically to each agent thinking the other is overly confident in her assessments. To do so, I study two agents whose belief distributions are identical except for their respective means and look at how one agent assesses the other's confidence intervals. (Since the agents and beliefs are completely symmetric, their assessments will be too.)

Consider thus a setting with two agents and a true state of the world, $x \in \mathbb{R}$, which is unknown to the agents. Each agent has a subjective belief about x. In particular, agent i believes that $x \sim F_i$ with continuous density function f_i and expected value $\overline{x}_i$. Assume now that the agents' beliefs F_i differ only by their mean, so that $F_i(x) = F_j(x + \overline{x}_j - \overline{x}_i)$. Let now $F(x) = F_i(\overline{x}_i + x) = F_j(\overline{x}_j + x)$ be a mean-zero normalized version of these distributions. Say that f is strictly single-peaked when $\exists \check{x}$ s.t. $f(x)$ is strictly increasing for $x < \check{x}$ and strictly decreasing for $x > \check{x}$.

Let $C_i^\alpha \subset \mathbb{R}$ be agent i's α-level confidence interval – the smallest set such that, according to i, it contains the true value with probability α (or more) – and defined formally as follows

$$C_i^\alpha = \operatorname*{argmin}_{\{C \subset \mathbb{R} \mid \int_C dF_i \geq \alpha\}} \int_C du$$

Let further $\tilde{\alpha}_{j,i} = \int_{C_i^\alpha} dF_j$ be agent j's confidence in i's α-confidence interval, i.e., j's belief about the probability that the true value will fall in that interval and thus j's belief how much confidence anyone, including agent i, should have in i's α-confidence interval.

The following proposition then says that agent i is overconfident from the perspective of agent j (who I will denote as the 'observer' to make the setting and statement more clear). Moreover, the level of overconfidence increases in the amount of disagreement, under the (strong sufficient) condition that f is strictly single-peaked.

Proposition 1 *Agent i is always weakly overconfident according to the observer j: $\alpha \geq \tilde{\alpha}_{j,i}$. Moreover, if f is strictly single-peaked and the agents' expected values differ ($\overline{x}_i \neq \overline{x}_j$) then i is strictly overconfident according to j: $\alpha > \tilde{\alpha}_{j,i}$. Finally, if f is strictly single-peaked and symmetric then the overconfidence $\alpha - \tilde{\alpha}_{j,i}$ increases monotonically and strictly in the disagreement in expected value, $|\overline{x}_i - \overline{x}_j|$.*

Figure 2 shows the result graphically. According to agent 1, the confidence level of the interval $[\underline{x}_1, \tilde{x}_1]$ equals the area $A + B$. According to agent 2, however, the confidence level is only $B + C$. The proposition essentially implies that $C \leq A$.

The conditions for getting strict overconfidence and increasing overconfidence are in fact strong sufficient conditions. While these results seem to hold quite generally, it has proved tricky to find conditions that nicely characterize these properties. But even though strict single-peakedness is much stronger than what is needed, it is still satisfied by many common distributions.

The intuition for this result is – like the result in the next section – based on (implicit) optimization by the agent. While it may be somewhat less transparent and less elegant than the intuition in the next section, it is nevertheless worth stating. The logic goes as follows. When an agent determines a confidence interval, she must find the smallest possible interval to reach a certain level of confidence. One way to think about this is that this agent implicitly ranks infinitesimal subintervals of identical length according to their (subjective) probability and then picks the ones that she believes have the highest probability. But two agents who differ in their expected value will also differ in their assessment of the probability of almost any such subinterval. When picking all the intervals in which she has most confidence, the agent is more likely to pick intervals of which she over-estimates the likelihood relative to the other agent's assessment, than subintervals of which she under-estimates the likelihood relative to the other agent's assessment. Overall, then, the focal agent will have more confidence in her confidence interval than the other agent, i.e., than the observer.

To get a feel whether this mechanism may actually contribute to our understanding of overconfidence in a substantive way, I used some simulations to see what it takes to generate the level of overconfidence that is observed in typical data. To that purpose I considered a setting where both the agents' beliefs and the distribution of their expected value follow a normal distribution.[6] In that setting, there is only one parameter on which the expected level of relative overconfidence depends. That parameter is the ratio of the variance of the means-distribution to the variance of an individual agent's belief, i.e., how much disagreement is there about the mean relative to the confidence an individual agent has in her own estimate.

Table 1 shows the hitrate (for a 90% confidence interval) and the relative hitrate ($\tilde{\alpha}_{j,i}/\alpha$) in function of the ratio of variances, where the variance of the means distribution is denoted $\sigma_{\overline{x}}^2$ and that of the beliefs is denoted σ_i^2. The levels of overconfidence generated definitely seem to be in the

---

[6]This seems to be a very reasonable assumption for settings where the estimates are either unrestricted or when they are far away from their restrictions.

Variance Ratio ($\sigma_i^2/\sigma_{\overline{x}}^2$)	Hitrate for 90% Confidence interval	Relative Hitrate ($\tilde{\alpha}_{j,i}/\alpha$)
.5	.28	.32
1	.46	.51
2	.58	.64

Table 1: Hitrate for 90% Confidence Interval in function of the Variance Ratio

same order of magnitude as those observed in real data. This suggests that this mechanism may be playing a substantive role in the observed overconfidence.

For an empirical (or experimental) test of the theory, probably the best starting point is the comparative static that more disagreement about the means leads to more overconfidence. This could be tested directly in existing data. A drawback of this approach is that the measure of disagreement will likely confound private information and differing priors. A cleaner but more involved experimental approach is to eliminate first disagreement that comes from differences in private information.[7] The remaining disagreement is then driven by differing priors, for which the comparative static holds.

While this overconfidence mechanism by itself does not explain the hard-easy effect, it does generate this effect when combined with an extra assumption. In particular, if – for equivalent levels of confidence – hard tasks are associated with more disagreement about the mean value than easy tasks, then agents will be relatively more overconfident about hard tasks than about easy tasks. That is shown formally in the following immediate corollary of Proposition 1.

Corollary 1 *If, for a fixed distribution of beliefs F, the agents' disagreement about the mean values ($|\overline{x}_i - \overline{x}_j|$) is larger for harder tasks than for easier tasks, then an agent will be relatively more overconfident on a harder task than on an easier task, i.e., $\tilde{\alpha}_{j,i}/\alpha$ will be smaller for a harder task than for an easier task.*

While this seems like a fairly reasonable assumption, it requires empirical validation in the data. This mechanism can thus contribute to our understanding of the hard-easy effect.

A natural question to ask is what happens to this type of overconfidence when the agents get new information. The answer depends critically on the degree to which the agents agree or disagree on the meaning of that new information. To analyze that issue, consider a setting where the two agents have a prior belief that follows a standard normal distribution but with different means. Agent 1 believes that x is distributed according to $N(\overline{x}_1, 1)$, while agent 2 believes that $x \sim N(\overline{x}_2, 1)$. Both agents observe the *same* set of signals – so neither agent has any private information – but they may disagree on the meaning of that information for the value of x, i.e., they have differing prior beliefs about the meaning of the signal. In particular, the agents observe a series of K public signals $\{s_k\}_{k=1}^{k=K}$. Each agent i believes that s_k is a biased signal for x in the sense that $s_k \sim N(x + \epsilon_{i,k}, 1)$. But the agents now openly disagree on how $E[s_k]$ relates to x, in the sense that it may be common knowledge that $\epsilon_{1,k} \neq \epsilon_{2,k}$ even though no agent has private information. For analytical purposes, I will assume that the $\epsilon_{i,k}$ have an empirical distribution

---

[7]One way to do so is to group the subjects in small groups and let them publicly display their confidence range to each other. The subjects may then – publicly and iteratively – update their displayed confidence range, based on their observations of others' confidence ranges. After a – typically very short – time, subjects will stop updating and each thus settles on a final range. At this point, any differences are differences about which the subjects agree to disagree. These thus correspond to differing priors (for the rest of the experiment). This procedure has interesting parallels and relations to the experimental procedure of Kogan (2009).

$\epsilon_{i,k} \sim N(0, V/2)$ so that $\Delta\epsilon_k = \epsilon_{1,k} - \epsilon_{2,k} \sim N(0, V).$[8] Note that this an *empirical* distribution: each agent i considers her own $\epsilon_{i,k}$ as the correct value with probability 1 and thus interprets the distribution of $\Delta\epsilon_k$ as the distribution of *how wrong other agents are.* Note also that the V can be interpreted as a measure for the degree of disagreement between the agents.

The following proposition then shows that whether overconfidence increases or decreases with new information depends on the level of disagreement about that information.

Proposition 2 *If $V > 0$ then the agent is always relatively overconfident, even in the limit as $K \to \infty$. Moreover, there exists a $\hat{V}$ so that the agent is strictly more overconfident in the limit as $K \to \infty$ than without any extra signals ($K = 0$) if and only if $V > \hat{V}$.*

If $V = 0$ – so that there is completely no disagreement on the signals – then the relative overconfidence disappears (i.e., $\tilde{\alpha}_{j,i}/\alpha \uparrow 1$) as the number of signals $K \to \infty$.

This result thus shows that this mechanism is (potentially) consistent with the experimental result that people sometimes get more overconfident when they get more data. In fact, it suggests a very particular comparative static for this result. But the result also shows that disagreement about the prior means alone is not sufficient to generate persistent overconfidence: agents also need to disagree on the meaning of additional information.

While the overconfidence result of this section establishes a fairly direct link between differing priors and overconfidence, the next section presents a much more striking relationship.

5   Endogenously Generated Overconfidence

In this section, I now turn to the case of disagreement about the *precision* of new data and will show that (subjective) Bayesian rationality might actually *cause* overconfidence: in trying to find the best estimate, agents convert unbiased errors into systematic biases. To derive that result, I will study a setting in which all agents observe the same signals but they openly disagree on each signal's informativeness or relevance. For transparency and analytical tractability, I now simplify the world to the polar opposite of the last section: agents start with a common prior about the state variable and agree on the values of the signals but have differing priors about the precision of each signal.

To show formally that this leads to overconfidence, consider, as before, a setting with 2 agents and a state of the world, $x \in \mathbb{R}$, that is unknown to the agents. These two agents now start with a common prior about the location of x, with both agents' prior belief being that x is normally distributed $N(x_0, \sigma_0^2)$ with x_0 some arbitrary point and σ_0^2 some exogenously given variance.[9]

All agents observe the same identical signals $x_k = x + \epsilon_k$, $k \in \{1, \ldots, K\}$, where ϵ_k is an observation error. The observation errors ϵ_k are (commonly known to be) independently and normally distributed with common mean 0. The variances of these ϵ_k distributions, on the other hand, are unknown. As a consequence, the precision and thus the relevance of each particular signal is unknown. Instead, each agent has a subjective belief about the variance of each signal

---

[8]Note that the disagreement is still about the *means* of the signals. The next section studies the effect of disagreement about the *variance* of the signals.

[9]Although σ_0 is here assumed to be exogenously given, it could also, for example, be drawn from G like the other variances. Alternatively I could assume that agents start from a diffuse prior. Yet another alternative is that each agent i's prior is normally distributed $N(x_0, \sigma_{0,i}^2)$ with x_0 some arbitrary point and $\sigma_{0,i}^2$ drawn from G, as specified later. In this case, the agents thus also disagree on the variance of the prior. The same results obtain in all these settings.

x_k and agents may disagree about these variances: agent i believes that $\epsilon_k \sim N(0, \sigma_{k,i}^2)$ and thus $x_k \sim N(x, \sigma_{k,i}^2)$ where it may be that $\sigma_{k,i} \neq \sigma_{k,j}$ for agent j. These beliefs, and thus also the fact that agents have differing priors, are for simplicity common knowledge. To simplify the notation, I will use $\sigma_{0,i} = \sigma_{0,j} = \sigma_0$ for the variance of the common prior. The $\sigma_{k,i}^2$, finally, are independent draws from some non-degenerate distribution G on $(0, \infty)$. Note that this distribution G is *not* a prior belief, but just the empirical distribution of the agents' beliefs. In other words, knowledge of this distribution does not affect any agent's beliefs about $\sigma_{k,i}$: each agent considers his own belief as the right one and interprets G as information about how other agents are mistaken in their beliefs. Note also that while the agents' prior beliefs and beliefs about ϵ_k are normally distributed, this empirical distribution of the beliefs about the variances G may be different from the normal distribution.

Given her information, agent i can form a belief about the true value of x. Since each agent i is assumed to be Bayesian rational, her final belief will be normally distributed according to $N(\hat{x}_i, \hat{\sigma}_i^2)$ with

$$\hat{x}_i = \sum_{k=0}^{K} \lambda_{k,i} x_k$$

$$\hat{\sigma}_i^2 = \sum_{k=0}^{K} \lambda_{k,i}^2 \sigma_{k,i}^2$$

where $\lambda_{k,i} = \frac{\tau_{k,i}^2}{\sum_{k=0}^{K} \tau_{k,i}^2}$ with $\tau_{k,i} = \frac{1}{\sigma_{k,i}}$.

Since the agents' beliefs are common knowledge, agent j knows i's estimator, i.e. he knows what weights $\lambda_{k,i}$ agent i will put on the different bits of information. According to agent j, the distribution of that estimator $\hat{x}_i = \sum_{k=0}^{K} \lambda_{k,i} x_k$ is $N(x, \hat{\sigma}_{j,i}^2)$ with

$$\hat{\sigma}_{j,i}^2 = \sum_{k=0}^{K} \lambda_{k,i}^2 \sigma_{k,j}^2$$

the variance on i's estimator according to agent j.[10] From agent j's perspective, $\hat{x}_i$ is thus an ex ante unbiased but inefficient estimator. Moreover, and that is key to the results of this paper, *agent j disagrees with i on the variance of that estimator*. In particular, comparing $\hat{\sigma}_{j,i}$ with $\hat{\sigma}_i$ measures i's overconfidence according to j. The following proposition then says that i is indeed overconfident in his own judgment from the perspective of agent j.

Proposition 3 *For any amount of information $K \geq 1$, agent i is in expectation deemed to be overconfident by agent j, i.e., $E\left[\frac{\hat{\sigma}_{j,i}^2}{\hat{\sigma}_i^2}\right] > 1$.*

The intuition for this result is as described in the introduction: to find her best estimate, agent i weights the different pieces of information in inverse proportion to the (subjective) variance of each. But since agents disagree on the variance, agent i will tend to overweight information of which she under-estimated the variance (relative to an observer such as agent j) and underweight information of which she over-estimated the variance (relative to agent j). Overall, she will underestimate the variance of her estimator from agent j's perspective. As in the case of the confirmation bias,

---

[10]Note that I look here at j's belief conditional on the variances and $\hat{x}_i$, but not conditional on the realizations of the x_k. Alternatively, this is i's expected overconfidence (from j's perspective) over all realizations of x_k.

the overconfidence is caused by differential weights that are put on new information. But the confirmation bias favors information that fits the original beliefs – i.e., it puts more weight on signals that are closer to the prior mean – whereas this mechanism favors information that is deemed more precise - i.e., it puts more weight on signals for which the (subjective) variance is smaller.
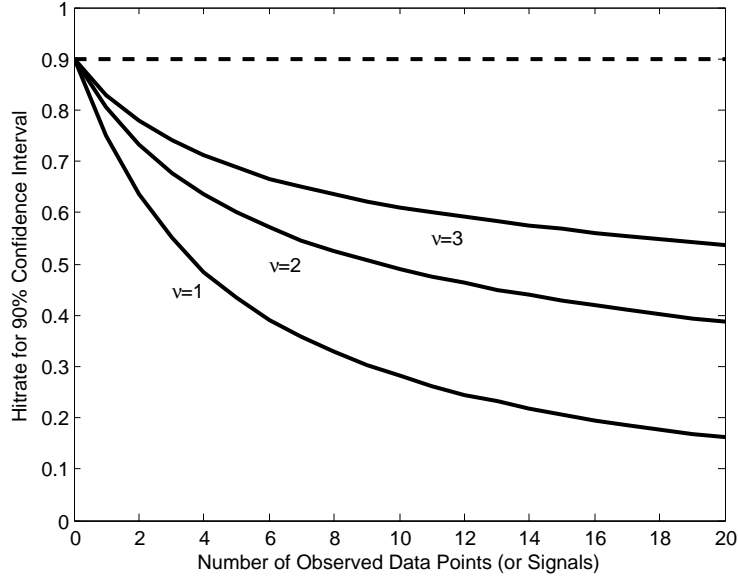

Figure 3: The hitrate for the 90% confidence interval in function of the amount of data observed when beliefs-variances are distributed according to Chi-squared distribution with mean ν.)

An important question is, again, whether the size of the effect is sufficient for it to play a substantive role in the overconfidence that is observed in data. To that purpose, I conducted some simulations to see what it takes for this mechanism to generate the degree of overconfidence that is observed in data. For the distribution of the beliefs about the variances of new signals, I used the one-parameter chi-squared distribution (with mean ν).[11] Figure 3 shows the results of the simulation. On the horizontal axis is the number of extra signals or data that the agents observe. On the vertical axis is the simulated hitrate of the 90% confidence interval. The horizontal dashed line would be the hitrate for a well-calibrated person with no overconfidence. The convex downsloping lines are the results for simulations with different means ν. The simulation shows that the overconfidence effects of this mechanism can be significant and that the mechanism can thus potentially contribute substantively to our understanding of overconfidence.

This and other simulations also suggest two results that one might have conjectured based on the intuition for the mechanism. The first is that what really matters for the amount of overconfidence

---

[11]I chose the chi-squared distribution since it is the most common distribution with support $[0, \infty)$ and with only one parameter (so that the effect of the parameter choice is uni-dimensional).

generated through this mechanism is the ratio of the standard deviation to the mean of the variance-belief distribution. In particular, a few data points for which the agent considerably underestimates the variance will lead to strong overconfidence because they will get a disproportionate weight in the estimation.

The second, and very important, result is that the degree of overconfidence in the simulation actually gets *worse* as the agents get more data. This very surprising result has also been found in experimental data, such as Oskamp (1965). The intuition for why this happens here is that more data means more occasions to make errors and thus to bias the overall confidence level. This is somewhat similar to the result in Van den Steen (2004) that overoptimism increases with the number of alternatives from which the agent can choose. Another way to see this is that the source of the overconfidence is the updating mechanism itself: if the overconfidence is caused by the updating mechanism, then at least some updating is necessary to get overconfidence. Whereas both the intuition and the simulations suggest that this result holds in full generality, a general formal proof has been elusive, though Appendix A.1 gives some formal results that go in this direction.

The fact that the overconfidence may increase with the amount of data has some important potential implications. For one thing, this mechanism may be a source of resistance to change that becomes stronger over time. In particular, as a person has more experience with a specific method, he may (through this mechanism) become more overconfident in his assessment of the method's effectiveness. When circumstances change, such person will be slower to update his beliefs and thus to accept change. Another potential implication is that it may actually be more useful to aggregate a number of experts' snap judgments than to let one expert spend a lot of time formulating a carefully weighted conclusion. Finally, one should also be careful about management's confidence in decisions that they have been considering for a long time: they are more likely to be overconfident about decisions that they have been considering for a long time than about those that they were forced to take in relatively short order.

Since this overconfidence mechanism is driven by Bayesian updating, a natural question is whether violations of Bayesian rationality would eliminate the result. The good thing here is that this mechanism requires only a minimum level of Bayesian sophistication by the agents: as long as agents tend to attach more weight to signals that they believe to be more precise (and make errors in estimating these precisions), the mechanism will go through. That is a fairly low standard. In the end, however, the question whether people are sufficiently Bayesian for the mechanism to play a substantive role is an empirical one.

The most effective test of the theory seems to be a direct test of the mechanism itself. One potential approach would be an experiment along the lines of Oskamp (1965), where subjects get additional information over time and are asked to state their confidence levels, complemented with questions about the perceived informativeness of the additional data. Any changes in (over)confidence can then be correlated to these precision estimates. Confidence should increase more with data that are deemed more precise or more relevant. And overall overconfidence should increase in the disagreement among the agents about the informativeness of the data. An interesting variation on this setup would be to actively control the subjects' perception of the signals' precision, along the lines of, for example, Kogan (2009). One procedure would be to feed the subjects both signals and their estimated precision and to tell the subjects that the precision estimates were made by experts who may make errors in estimating the precision but who are on average correct.

A final consideration is again the hard-easy effect. While the model does not necessarily or inherently lead to a hard-easy effect, it is consistent with it under an extra assumption. The first assumption under which the mechanism leads to a hard-easy effect is that there is more disagreement

about the precision of signals for 'hard' tasks than for 'easy' tasks. The second assumption that works is if hard tasks require an agent to combine more data to get to an α-confidence interval than easy tasks, since more data lead to more overconfidence. Whereas both assumptions seem very reasonable, it is again an empirical question whether they hold in the data. The point here is not to argue that this model explains the hard-easy effect, but rather that it is consistent with it under seemingly reasonable assumptions.

6   Discussion

The result of section 4 suggested a fairly direct link between differing priors and overconfidence. A natural question, given that agents also have differing priors in section 5, is whether the result of section 5 is indirectly caused by, or a special case of, the result of section 4. The answer is negative. First, the parallel of the result of section 4 in the model of section 5 is overconfidence about ϵ_k rather than overconfidence about the agent's belief about x. Second, the overconfidence of section 4 did not depend on any updating, while the result of section 5 *does* depend on the updating mechanism. To see this more explicitly, note that a non-rational agent who derives his estimates using random weights or fixed weights would be appropriately confident, despite the differing priors about ϵ_k. It is even possible to create estimators about which the agents would be underconfident[12]. The overconfidence conclusions are thus really driven by the Bayesian updating mechanism.

A second question that arises quite often with this analysis is whether the same mechanism and results could be obtained in a common priors setting. For example, wouldn't private information that agents cannot communicate not give the same results? This is actually not the case, and differing priors really seem to be necessary to get these results (via the same mechanisms). With common priors, agents could always derive each others' information from a difference in confidence, and would then update their own beliefs using that information. Overall, they would on average be appropriately confident. To see this another way, consider the result of proposition 3 for the case when f is strictly single peaked and the expected values differ. In that case, the proposition implies that it is common knowledge that the two agents (agree to) disagree on the appropriate confidence in agent i's confidence interval. But by Aumann (1976), this implies that the agents *must* have differing priors. More generally, the martingale property (applied to beliefs) implies that there cannot be (in expectation) any systematic bias in beliefs if agents start from a common prior.

The paper also has implications for estimating and correcting overconfidence. The most important implication is that overconfidence is more likely in settings where people openly disagree and especially when they disagree on the significance of particular information. In such setting, a larger correction will be necessary. A further implication relates to the way in which experts' estimates are aggregated. The most common way to aggregate such estimates is to simply average them. The results of this paper suggest, instead, to try to average on a lower, more detailed level: average the experts' estimates on the rough data and then combine these average data-estimates into the final estimate.

---

[12]Just picking the data point with the largest (subjective) variance gives an ex ante unbiased estimator. By exactly the opposite effect of the one in section 5, however, agents will tend to overestimate the variance in this estimator, and thus be underconfident.

7   Conclusion

This paper shows that overconfidence is quite a natural bias for Bayesian rational agents when they may entertain differing priors. The two mechanisms that I derived are simple and robust, as long as people attach more importance to information that they consider to be more informative. One of the mechanisms also provided a potential explanation for a very intriguing experimental result on overconfidence, i.e. that people sometimes get more overconfident as they get more data.

The mechanisms in this paper are obviously not the only possible source of overconfidence, but they suggest that very little is necessary to make even perfect Bayesian-rational agents be subject to overconfidence.

A   Appendix

A.1   Endogenous Overconfidence as Function of Data

The paper mentioned that a general formal proof for the result that the endogenously generated overconfidence of Section 5 increases with the number of signals has been elusive. Nevertheless, both the simulations and the intuition suggest that this result holds in general. This appendix provides some formal results that point in that direction.

The first result shows that the polar opposite is definitely not true: the overconfidence bias does not disappear with more data. In particular, the following proposition says that the overconfidence persists even with an infinite amount of new information.

Proposition 1 *There exists $\beta > 1$ such that as the amount of information $K \to \infty$, $\frac{\hat{\sigma}^2_{j,i}}{\hat{\sigma}^2_i} \overset{a.s.}{\to} \beta > 1$.*

Proof :   For the proof of the proposition, note that

$$
\frac{\hat{\sigma}^2_{j,i}}{\hat{\sigma}^2_i} = \frac{\frac{\sum_{k=0}^K \tau^4_{k,i}\sigma^2_{k,j}}{\left(\sum_{k=0}^K \tau^2_{k,i}\right)^2}}{\frac{\sum_{k=0}^K \tau^4_{k,i}\sigma^2_{k,i}}{\left(\sum_{k=0}^K \tau^2_{k,i}\right)^2}} = \frac{\sum_{k=0}^K \tau^4_{k,i}\sigma^2_{k,j}}{\sum_{k=0}^K \tau^4_{k,i}\sigma^2_{k,i}} = \frac{\frac{\sum_{k=0}^K \tau^4_{k,i}\sigma^2_{k,j}}{K}}{\frac{\sum_{k=0}^K \tau^4_{k,i}\sigma^2_{k,i}}{K}}
$$

$$
\overset{a.s.}{\to} \frac{E[\tau^4_{k,i}\sigma^2_{k,j}]}{E[\tau^4_{k,i}\sigma^2_{k,i}]} = \frac{E[\tau^4_{k,i}]E[\sigma^2_{k,j}]}{E[\tau^2_{k,i}]} = \frac{E[\tau^4_{k,i}]E[\sigma^2_{k,i}]}{E[\tau^2_{k,i}]} > 1
$$

where the last equality follows from the fact that $E[\sigma^2_{k,i}] = E[\sigma^2_{k,j}]$, and the very last inequality follows from applying a minor extension of Jensen's inequality (to get a strict inequality) twice. In particular, $E[\tau^4_{k,i}] > E[\tau^2_{k,i}]^2$ and

$$
E[\sigma^2_{k,i}] = E\left[\frac{1}{\tau^2_{k,i}}\right] > \frac{1}{E[\tau^2_{k,i}]}
$$

so that $E[\tau^4_{k,i}]E[\sigma^2_{k,i}] > E[\tau^2_{k,i}]$.   ∎

Since these are limit results, they imply analogous results for the other ways of measuring overconfidence.

The following proposition, finally, shows that at least over some range, the overconfidence will increase with more data.

Proposition 2 *For any distribution G, the expected overconfidence of the agent increases monotonically over at least some range of data: there exist $\underline{k}, \overline{k} \in \mathbb{N}$ with $0 \le \underline{k} < \overline{k}$ such that $E\left[\frac{\hat{\sigma}^2_{j,i}}{\hat{\sigma}^2_i}\right]$ increases monotonically as the number of signals available to the agent increases from $\underline{k}$ to $\overline{k}$.*

Proof :   Note that with $\underline{k} = 0$, $E\left[\frac{\hat{\sigma}^2_{j,i}}{\hat{\sigma}^2_i}\right] = E\left[\frac{\sigma^2_0}{\sigma^2_0}\right] = 1$. Further, Proposition 3 says that for any $K \ge 1$, $E\left[\frac{\hat{\sigma}^2_{j,i}}{\hat{\sigma}^2_i}\right] > 1$. This establishes the proposition.   ∎

A.2 Proofs

Proof of Proposition 1: Note that $\int_{C_i^\alpha} du = \int_{C_j^\alpha} du$, i.e., the α-confidence intervals of agents i and j have exactly the same length, by definition of F_i, F_j, C_i^α, C_j^α and an argument by contradiction. Assume now (by contradiction) that $\int_{C_i^\alpha} dF_j > \alpha$. Then, since F_i and F_j are atomless, there exists a $C \subset C_i^\alpha$ with $\int_C du < \int_{C_i^\alpha} du = \int_{C_j^\alpha} du$ and $\int_C dF_j > \alpha$. But this contradicts the definition of C_j^α.

For the second part of the proposition, note that – for a strictly single-peaked density – C_i^α must be of the form $[\underline{x}_i, \tilde{x}_i]$ with $\underline{x}_i < \check{x}_i < \tilde{x}_i$. Moreover, $f_i(\underline{x}_i) = f_i(\tilde{x}_i)$ because if this condition were not satisfied, then a smaller interval can be found by shortening the interval on one side and lengthening it slightly less on the other side. This further implies that C_i^α is uniquely determined. Finally, $\tilde{x}_i - \underline{x}_i = \tilde{x}_j - \underline{x}_j$ by the fact that F_i and F_j differ only in their mean and by the fact that F_i and F_j are *strictly* single-peaked. It then follows that

$$\frac{d}{du} \int_{\underline{x}_i+u}^{\tilde{x}_i+u} dF_i(x) = f_i(\tilde{x}_i + u) - f_i(\underline{x}_i + u) \begin{cases} < 0 & \text{when } u > 0 \\ > 0 & \text{when } u < 0 \end{cases}$$

To see why this is true, consider the case that $u > 0$. If $\underline{x}_i + u \le \check{x}_i$ then f is strictly increasing between $\underline{x}_i$ and $\underline{x}_i + u$ (and it is also strictly decreasing between $\tilde{x}_i$ and $\tilde{x}_i + u$ for $u > 0$) so that $f(\underline{x}_i+u) > f(\underline{x}_i) = f(\tilde{x}_i) > f(\tilde{x}_i+u)$. If $\check{x}_i < \underline{x}_i+u$ then $f(\check{x}_i) > f(\underline{x}_i+u) > f(\tilde{x}_i+u)$ since both $\underline{x}_i+u$ and $\tilde{x}_i+u$ are in the decreasing part of f and $\underline{x}_i+u < \tilde{x}_i+u$. This also further implies the second and third parts of the proposition. ∎

Proof of Corollary 1: This follows immediately from Proposition 1. ∎

Proof of Proposition 2: Fix some level of confidence $\alpha \in (0,1)$ and some number of signals $K > 0$. Let $\gamma > 0$ be defined $\int_{-\gamma}^{\gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \alpha$, so that $\int_{-\gamma\sigma}^{\gamma\sigma} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz = \alpha$. The posterior averages are $\overline{x}_{1,K} = \frac{\overline{x}_1 + \sum_{k=1}^K (s_k - \epsilon_{1,k})}{K+1}$ and $\overline{x}_{2,K} = \frac{\overline{x}_2 + \sum_{k=1}^K (s_k - \epsilon_{2,k})}{K+1}$, while the variance (from each agent's own perspective) is $\frac{1}{K+1}$ for both since each agent believes that her signal results from combining K signals and a prior that all have variance 1. Let $\Delta x = \overline{x}_2 - \overline{x}_1$ then $\overline{x}_{2,K} - \overline{x}_{1,K} = \frac{\Delta x}{K+1} + \frac{\sum_{k=1}^K \Delta \epsilon_k}{K+1}$. Now let $\frac{\sum_{k=1}^K \Delta \epsilon_k}{\sqrt{K+1}} = v_k \sim N(0, \frac{K}{K+1}V)$.

The α-confidence interval of agent 2 is then $C_2^\alpha = [\overline{x}_{2,K} - \frac{\gamma}{\sqrt{K+1}}, \overline{x}_{2,K} + \frac{\gamma}{\sqrt{K+1}}] = [\overline{x}_{1,K} + \frac{\Delta x}{K+1} + \frac{v_k}{\sqrt{K+1}} - \frac{\gamma}{\sqrt{K+1}}, \overline{x}_{1,K} + \frac{\Delta x}{K+1} + \frac{v_k}{\sqrt{K+1}} + \frac{\gamma}{\sqrt{K+1}}]$ so that

$$\int_{C_2^\alpha} dF_1 = \int_{\overline{x}_{1,K} + \frac{\Delta x}{K+1} + \frac{v_k}{\sqrt{K+1}} - \frac{\gamma}{\sqrt{K+1}}}^{\overline{x}_{1,K} + \frac{\Delta x}{K+1} + \frac{v_k}{\sqrt{K+1}} + \frac{\gamma}{\sqrt{K+1}}} \frac{1}{\sqrt{2\pi}} \frac{1}{1/\sqrt{K+1}} e^{-\frac{1}{2}\left(\frac{u-\overline{x}_{1,K}}{1/\sqrt{K+1}}\right)^2} du$$

or, with a change of variable $y = \frac{u - \overline{x}_{1,K}}{1/\sqrt{K+1}}$,

$$\int_{C_2^\alpha} dF_1 = \int_{\frac{\Delta x}{\sqrt{K+1}} + v_k - \gamma}^{\frac{\Delta x}{\sqrt{K+1}} + v_k + \gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

Its expected value converges monotonically to

$$E[\int_{C_2^\alpha} dF_1] = E_v \left[\int_{v-\gamma}^{v+\gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \right]$$

where $v \sim N(0, V)$. When $V = 0$ so that $v = 0$, this equals α, so that the overconfidence decreases and vanishes in the absence of disagreement over the signals. Consider henceforth $V > 0$. This gives, after a change of variable $u = v/\sigma$ with $\sigma = \sqrt{V}$,

$$E[\int_{C_2^\alpha} dF_1] = \int_u \int_{\sigma u - \gamma}^{\sigma u + \gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \, dy \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du$$

Let now $Z(u) = \int_{\sigma u - \gamma}^{\sigma u + \gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \, dy$. Calculus shows that Z reaches its (strict) maximum at $u = 0$ where $Z(0) = \alpha$, so that $E[\int_{C_2^\alpha} dF_1] < \alpha$ and overconfidence persists in the limit. Moreover, except at $u = 0$ where Z is constant and equal to α, Z strictly and monotonically decreases everywhere in σ and goes to zero as $\sigma \to \infty$. It follows that $E[\int_{C_2^\alpha} dF_1] \downarrow 0$ as $\sigma \to \infty$. It further follows that there is a critical $\hat{\sigma}$ such that in the limit overconfidence is higher than the overconfidence of the prior if and only if $\sigma > \hat{\sigma}$. ∎

Proof of Proposition 3: To prove that $E\left[\frac{\hat{\sigma}_{j,i}^2}{\hat{\sigma}_i^2}\right] > 1$ for any amount of information $K \geq 1$, it suffices to prove that, with $E[\sigma_{k,j}^2] = \bar{\sigma}^2$,

$$E[\frac{\hat{\sigma}_{j,i}^2 - \hat{\sigma}_i^2}{\hat{\sigma}_i^2}] = E[\frac{\sum_{k=0}^{K} \lambda_{k,i}^2 (\sigma_{k,j}^2 - \sigma_{k,i}^2)}{\sum_{k=0}^{K} \lambda_{k,i}^2 \sigma_{k,i}^2}]$$

$$= E\left[\frac{\frac{\sum_{k=0}^{K} \tau_{k,i}^4 (\sigma_{k,j}^2 - \sigma_{k,i}^2)}{(\sum_{k=0}^{K} \tau_{k,i}^2)^2}}{\frac{\sum_{k=0}^{K} \tau_{k,i}^4 \sigma_{k,i}^2}{(\sum_{k=0}^{K} \tau_{k,i}^2)^2}}\right] = E\left[\frac{\sum_{k=0}^{K} \tau_{k,i}^4 (\bar{\sigma}^2 - \sigma_{k,i}^2)}{\sum_{k=0}^{K} \tau_{k,i}^4 \sigma_{k,i}^2}\right] = E\left[\frac{\sum_{k=0}^{K} \tau_{k,i}^4 (\bar{\sigma}^2 - \sigma_{k,i}^2)}{\sum_{k=0}^{K} \tau_{k,i}^2}\right] > 0$$

To that purpose, it is sufficient to prove that, for any indicator m,

$$\int \frac{\tau_{m,i}^4 (\bar{\sigma}^2 - \sigma_{m,i}^2)}{\sum_{k=0}^{K} \tau_{k,i}^2} g(\sigma_{m,i}^2) d\sigma_{m,i}^2 > 0$$

Note now that, with z a random variable with density function g and mean $\bar{z}$, and $h(z)$ a strictly increasing function, $\int h(z)(z - \bar{z}) g(z) \, dz > 0$ since the covariance of z and $h(z)$ is positive. It is thus sufficient to prove that

$$\frac{\tau_{m,i}^4}{\sum_{k=0}^{K} \tau_{k,i}^2}$$

is strictly decreasing in $\sigma_{m,i}^2$, or strictly increasing in $\tau_{m,i}^2$. But this follows from the fact that its derivative for $\tau_{m,i}^2$ is

$$\frac{2\tau_{m,i}^2 \sum_{k=0}^{K} \tau_{k,i}^2 - \tau_{m,i}^4}{\left(\sum_{k=0}^{K} \tau_{k,i}^2\right)^2} = \tau_{m,i}^2 \frac{2\sum_{k=0}^{K} \tau_{k,i}^2 - \tau_{m,i}^2}{\left(\sum_{k=0}^{K} \tau_{k,i}^2\right)^2} > 0$$

This concludes the proof. ∎

19

References

ACEMOGLU, D., V. CHERNOZHUKOV, AND M. YILDIZ (2006): "Learning and Disagreement in an Uncertain World," MIT Department of Economics Working Paper No. 06-28.

ALLEN, F., AND E. CARLETTI (2010): "New Theories to Underpin Financial Reform," Working Paper, University of Pennsylvania.

ALPERT, M., AND H. RAIFFA (1969): "A Progress Report on the Training of Probability Assessors," Unpublished Report.

ARROW, K. J. (1964): "The Role of Securities in the Optimal Allocation of Risk-bearing," *Review of Economic Studies*, 31(2), 91– 96.

AUMANN, R. J. (1976): "Agreeing to Disagree," *Annals of Statistics*, 4, 1236– 1239.

——— (1998): "Common Priors : A Reply to Gul," *Econometrica*, 66(4), 929– 938.

AUMANN, R. J., AND A. BRANDENBURGER (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161– 1180.

BARBERIS, N. C., AND R. H. THALER (2003): "A Survey of Behavioral Finance," in *Handbook of the Economics of Finance 1B*, ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Elsevier North Holland.

BAYES, T. (1763): "An Essay towards Solving a Problem in the Doctrine of Chances," *Philosophical Transactions of the Royal Society of London*, 53, 370–418.

BOOT, A. W. A., R. GOPALAN, AND A. V. THAKOR (2008): "Market Liquidity, Investor Participation, and Managerial Autonomy: Why Do Firms Go Private?," *Journal of Finance*, 63(4), 2013–2059.

BRUNNERMEIER, M. K., AND J. A. PARKER (2005): "Optimal Expectations," *American Economic Review*, 95(4), 1092– 1118.

DANIEL, K., D. HIRSHLEIFER, AND A. SUBRAHMANYAM (1998): "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53(6), 1839– 1885.

DANIEL, K., D. HIRSHLEIFER, AND A. SUBRAHMANYAM (2001): "Overconfidence, Arbitrage, and Equilibrium Asset Pricing," *The Journal of Finance*, 56(3), 921–965.

EINHORN, H. J., AND R. M. HOGARTH (1978): "Confidence in Judgment : Persistence in the Illusion of Validity," *Psychological Review*, 85, 395– 416.

EREV, I., T. WALLSTEN, AND D. BUDESCU (1994): "Simultaneous over- and underconfidence: The role of error in judgment processes," *Psychological Review*, 101(3), 519–527.

FISCHHOFF, B., P. SLOVIC, AND S. LICHTENSTEIN (1977): "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology: Human Perception and Performance*, 3, 552– 564.

GEANAKOPLOS, J. (2009): "The Leverage Cycle," Cowles Foundation Discussion Paper No. 1715.

GRUBB, M. D. (2009): "Selling to Overconfident Consumers," *American Economic Review*, 99(5), 1770–1807.

GUL, F. (1998): "A Comment on Aumann's Bayesian View," *Econometrica*, 66(4), 923–927.

HALDANE, J. (1932): "A Note on Inverse Probability," *Proceedings of the Cambridge Philosophical Society*, 28, 55–61.

HARRISON, M., AND D. M. KREPS (1978): "Speculative Investor Behavior in a Stock Market with Heterogenous Expectations," *Quarterly Journal of Economics*, 92(2), 323–336.

HARSANYI, J. C. (1968): "Games with Incomplete Information Played by 'Bayesian' Players, I-III, Part III. The Basic Probability Distribution of the Game," *Management Science*, 14(7), 486–502.

HONG, H., AND J. C. STEIN (2007): "Disagreement and the Stock Market," *Journal of Economic Perspectives*, 21(2), 109–128.

JEFFREYS, H. (1939): *Theory of Probability*. Oxford University Press, Oxford.

KLAYMAN, J., AND Y.-W. HA (1987): "Confirmation, Disconfirmation, and Information in Hypothesis Testing," *Psychological Review*, 94(2), 211–228.

KOGAN, S. (2009): "Distinguishing the effect of overconfidence from rational best-response on information aggregation," *Review of Financial Studies*, 22(5), 1889–1914.

LAPLACE, P. S. (1774): "Mémoire sur la Probabilité des Causes par les Évènemens," *M'emoires de Math'ematique et de Physique, Present'es à lAcadémie Royale des Sciences*, 6, 621656.

LICHTENSTEIN, S., AND B. FISCHHOFF (1977): "Do those who know more also know more about how much they know?," *Organizational B*, 20, 159–183.

LICHTENSTEIN, S., B. FISCHHOFF, AND L. D. PHILLIPS (1982): "Calibration of Probabilities: The State of the Art to 1980," Cambridge University Press.

LORD, C., L. ROSS, AND M. LEPPER (1979): "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology*, 37, 2098–2109.

MORRIS, S. (1994): "Trade with Heterogeneous Prior Beliefs and Asymmetric Information," *Econometrica*, 62(6), 1327–1347.

——— (1995): "The Common Prior Assumption in Economic Theory," *Economics and Philosophy*, 11, 227–253.

OSKAMP, S. (1965): "Overconfidence in Case-Study Judgments," *The Journal of Consulting Psychology*, 29, 261–265.

RABIN, M., AND J. SCHRAG (1999): "First Impressions Matter: A Model of Confirmation Bias," *Quarterly Journal of Economics*, 114(1), 37–82.

RUSSO, J. E., AND P. J. H. SCHOEMAKER (1992): "Managing Overconfidence," *Sloan Management Review*, 33(2), 7–17.

SCHEINKMAN, J. A., AND W. XIONG (2003): "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 111(6), 1183– 1219.

TAYLOR, S. E., AND J. D. BROWN (1988): "Illusion and Well-being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin*, 103(2), 193– 210.

TVERSKY, A., AND D. KAHNEMAN (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124–1131.

VAN DEN STEEN, E. (2001): "Essays on the Managerial Implications of Differing Priors," PhD Dissertation, Stanford University.

——— (2004): "Rational Overoptimism (and Other Biases)," *American Economic Review*, 94(4), 1141– 1151.

——— (2010a): "Culture Clash: The Costs and Benefits of Homogeneity," *Management Science*, Forthcoming.

——— (2010b): "Interpersonal Authority in a Theory of the Firm," *American Economic Review*, 100(1), 466–490.

VARIAN, H. R. (1989): "Differences of Opinion in Financial Markets," in *Financial Risk: Theory, Evidence, and Implications*, ed. by C. S. Stone, pp. 3– 37. Springer.

WALLSTEN, T. S., D. V. BUDESCU, AND R. ZWICK (1993): "Comparing the Calibration and Coherence of Numerical and Verbal Probability Judments," *Management Science*, 39(2), 176– 190.

ZELLNER, A. (1971): *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York.