ESSAYS ON THE MANAGERIAL IMPLICATIONS OF DIFFERING PRIORS

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL OF BUSINESS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Eric J. Van den Steen

May 2001

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

D. John Roberts
(Principal Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Edward P. Lazear

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Garth Saloner

Approved for the University Committee on Graduate Studies:

Abstract

This dissertation studies managerial implications of the fact that people may openly differ in their beliefs. It consists of three essays.

The first essay is methodological in nature. It considers issues that arise when we allow agents in an economic model to knowingly hold differing beliefs or, more precisely, differing priors. It motivates why this might sometimes be optimal from a methodological point of view, argues that it is fully consistent with the economic paradigm, and counters some potential objections. It then discusses epistemic foundations for the Nash equilibrium, the meaning of efficiency in this context, and alternative ways to set up models with differing priors. With this methodological foundation in place, the next two essays really focus on the managerial implications of differing priors.

The second essay studies the role of organizational beliefs and managerial vision in the behavior and performance of corporations. The paper defines vision operationally as a very strong belief by the manager about the right course of action for the firm. The interaction between employees' beliefs and the manager's vision shapes decisions and determines employees' motivation and satisfaction. Through sorting, the manager's vision also influences organizational beliefs. Under weak conditions, a company's board should select a manager with stronger beliefs than its own. Spurious effects, however, may make vision look better than it really is. The analysis also has implications for theories of corporate culture and strategy.

The third essay shows why rational agents may attribute their own success more to skill and their failures more to bad luck than an outsider, why each agent in a group might think he or she is the best, and why two agents' estimated contributions add up to more than 100 %. Central to the analysis is a simple and robust mechanism that generates endogenous overconfidence in one's own actions. The intuition is that random errors plus systematic choices lead to systematic errors. The paper finally considers organizational implications.

Acknowledgements

The completion of a PhD is a fitting moment to look back and be amazed at the impact other people have had on your life.

These last five years were enriching and, most of the time, fun. Many played a protagonist role in this journey, but one really stands out. John Roberts was my Advisor (with capital A), on work and sometimes on life. I admire enormously how he coached me on research. Most importantly, I will never forget how he helped me through rough times. He made me understand why this relationship is sometimes expressed in family terms. Bedankt, John.

While all econ and strategy faculty in the GSB are impressive, two, next to John, really made a difference to me. Ed Lazear and Garth Saloner are, each in their own way, just amazing and I was fortunate and happy that they shared some of their insights and enthusiasm with me. Thanks so much.

Of course, without a social life with partners in misfortune, PhD studies would be unbearable. The former PhD-LpF crowd, the Belgian-Turkish connection, the engineering geeks, they all helped me (and I hope I helped them) enjoy this journey. Hanno, Philip, Tunay, and Lieven earn a special mention. Cheers.

Finally, the financial support that made this all possible came from Stanford Business School and the Interuniversitair College voor Managementswetenschappen (ICM/CIM) in Belgium. Thanks.

A completed dissertation, however, reflects more than just those last five years. I would never have made it here without the guidance and encouragement of Prof.

Herman Daems, Prof. Paul Baron De Meester, Father Jef Degraeve s.j., Mr. Luc Reynders, and Prof. Lambert Vanthienen. I also wouldn't be who I am (at least on my positive sides) without some good friends, especially Kris, Kris, and Herman. But most important of all was the guidance, support, and love from my family: my parents, my brothers and my sister. If dedicating a dissertation was customary, this would be dedicated to them.

Last but not least, my PhD brought me together with Dei. She suffered through hours of my muttering about research, days of my over-stressed self whenever something important was on the radar screen, and many more of my bad habits. I still can't believe how lucky I've been.

Contents

List of Figures

Chapter 1

Introduction

Since the seminal work of Harsanyi (1967/68) on Bayesian games, information economics has been extremely successful. It has generated insights in important mechanisms such as signaling, screening, and reputation. It has changed our understanding of issues from non-linear pricing, auctions, and incentives to limit pricing and career concerns. It has become a cornerstone of modern economic theory and its implications for business.

Most of information economics assumes that agents hold identical beliefs prior to receiving any private information. This 'common prior assumption' (CPA) allows us to 'zero in on purely informational issues' (Aumann 1987). In doing so, it facilitated the fast development of this field and became almost an axiom or doctrine.

For all its benefits, the CPA also has an important downside. In particular, it excludes the possibility that agents knowingly disagree about the probability of an event. Open disagreement is a major phenomenon, however, especially in organizations. The CPA hinders research on these issues, and thus the reach of economics in the study of organizations.

The first essay of this dissertation, 'Notes on Modeling with Differing Priors', spells this issue out in much more detail. It argues that differing priors allow us to zero in on the implications of open disagreement, to paraphrase Aumann (1987). It also shows how this approach is perfectly consistent with the economic methodology and discusses some further methodological issues, including epistemic foundations for Nash equilibrium and the meaning of efficiency.

After putting in place this methodological foundation, the other two essays consider implications of differing beliefs in two areas of management.

The essay 'Organizational Beliefs and Managerial Vision' shows how vision, defined as a strong belief by the manager, influences employees' actions, motivation, and satisfaction, and determines organizational beliefs through sorting in the labor market. It further shows how it can be optimal for a board to select a manager who is overconfident when judged by the board's beliefs, but also that spurious effects may make vision look better than it really is.

The third essay, 'Skill or Luck', shows that a number of important but seemingly irrational phenomena might in fact be subjectively rational in a world with differing priors. In particular, it suggests an explanation for the fact that people tend to attribute their success more to skill and their failure more to bad luck than an outsider would, for the fact that in some surveys more than 95% of employees rate themselves above the median, and for some closely related 'biases'.

The latter two essays deal in a fully Bayesian-rational model with issues that are usually considered to be outside its realm. This suggests that differing priors might allow us to cover part of the bounded rationality agenda without deviating from the economic paradigm. The advantage of trying this approach before resorting to more ad-hoc modifications of the paradigm is twofold:

1. The economic paradigm is well understood. Staying within its limits ensures a continued uniformity of methodology that allows a more systematic development of knowledge.

2. Much of the developed intuition, methodology and results remains intact.

We hope that this dissertation may contribute to this agenda.

Chapter 2

Notes on Modeling with Differing Priors[*]

2.1 Introduction

Most economic models with incomplete information assume that agents share a common prior. This assumption, usually called the Common Prior Assumption (CPA) or 'Harsanyi doctrine', has been defended on the methodological basis that it 'enables one to zero in on purely informational issues' (Aumann 1987 p14), and on the more philosophical basis that 'differences in probabilities should reflect differences in information *only*' (Aumann 1998; italics in original). A well-known implication of this assumption is that agents cannot 'agree to disagree' (Aumann 1976): it cannot be common knowledge that one agent holds belief x about some event while another agent has belief $y \neq x$. Open disagreement or known differences in beliefs, however, are an important fact of life. Excluding this possibility would seriously limit the reach of economic research.

This paper considers this issue and some methodological issues it raises. The

[*]This chapter benefitted a lot from the discussions I had with Yossi Feinberg and Muhamet Yildiz. The usual disclaimer applies.

main goal is to provide an introduction for applied economists who want to know whether allowing differing priors is methodologically proper and how it relates to the broader economic literature. The next section considers in more detail the argument for allowing agents to hold differing priors. Section 3 extends the epistemological basis for Nash equilibria to the case of differing priors on the payoff space, which covers most situations of interest to us. Section 4 considers the meaning of efficiency in such a subjective world. Section 5, finally, discusses alternative ways of setting up models with differing priors.

While the paper makes some contributions, such as the epistemic results or the perspective on efficiency, some ideas have been forwarded earlier. Morris (1995) and Gul (1998), for example, already argued that there are no fundamental or rationality-based arguments to impose the CPA beyond the methodological argument that it allows to focus on information. The relation between subjective and objective efficiency, as defined in section 2.4, has been studied by Hammond (1983) and Mongin (1995). A different perspective on the efficiency issue was presented by Yildiz (2000). Brandenburger et al. (1992) give a more formal treatment of the relationship between differing priors and informational biases considered in section 2.5.

2.2 The argument for differing priors

Aumann (1976) showed that, under the CPA, it cannot be common knowledge between two agents that they disagree (in the sense that 'one agent has belief x about some event while the other agent has belief $y \neq x$'). An important variation on this result, due to Geanakoplos and Polemarchakis (1982)[1], is that, with generic common priors, it cannot be *mutual* knowledge between two agents that they disagree, in the sense mentioned above. The intuition for this result is simple. If the first agent knows the second agent's belief then he almost always can deduce the other's

[1]In fact, Geanakoplos and Polemarchakis did not literally prove this result, but it is an evident corollary of their proposition 4.

information. Since both agents can do so, they have essentially equivalent information and thus should have identical posteriors. Note how this argument shows how the CPA can effectively negate the impact of private information. The result also shows that we cannot get around this no-disagreement result just by relaxing the common knowledge assumption to mutual knowledge. To model disagreement while preserving the CPA, the information structure must be such that the relevant agents are unaware of the precise form[2] of disagreement at the critical time. This tricks the model to behave 'as if' agents have differing priors. Unless it happens to be the natural set-up, such approach is methodologically not very sound. Furthermore, the resulting conclusions may actually miss the essence of the argument. In chapter 3, for example, the argument is that stronger beliefs can be important *holding information constant*. This idea cannot be expressed in a model with common priors. Finally, it is not always possible to set things up this way. The issue of joint decision making in Feinberg and Scarsini (1999), for example, cannot exist without differing priors. In the absence of differing priors, the self-serving and egocentric biases in chapter 4 cannot persist when people become aware of the beliefs of others. If we want to study the full effects of disagreement, we will therefore have to relax the CPA.

The first and most fundamental argument for dropping the CPA is thus very pragmatic: doing so allows a clean and transparent analysis of the effects of open disagreement or known differences in beliefs. This turns Aumann's argument for the CPA upside down. Where he argued that 'the CPA enables one to zero in on purely informational issues' we would argue that 'differing priors enables one to zero in on open disagreement issues'.

While such pragmatic argument is perfectly valid[3], dropping the CPA also makes

[2]To be precise, the agent might know the *fact* that there is disagreement (as is usually the case in models of incomplete information) but he may not know the precise *form* of disagreement.

[3]Note, in particular, that economists usually defend the rationality assumption on a very pragmatic basis (maximization is analytically a very powerful assumption) as an 'as if' approximation. This rationale for rationality renders discussions on what rational people *should* do rather moot.

sense from a more fundamental perspective. First of all, major implications of the CPA, such as no mutual knowledge of disagreement or no trade/bets based on differing beliefs (Milgrom and Stokey 1982, Sebenius and Geanakoplos 1983) are contradicted by everyday experience. Second, absent any information, rational players have no basis to agree. While the statement that 'after controlling for differences in information, differences in beliefs should be uninformative random variables (from the perspective of the agents)' makes sense, there seems to be no convincing reason that there should be no differences at all. Harsanyi, for example, in his seminal theory of Bayesian games (1967/68) observes 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'. Note that this does not imply that players *must* disagree, only that they *may*. Analogously, Aumann's (1987) argument that players have no rational basis to disagree does not imply that rational agents should always agree.

A popular criticism of relaxing the CPA is that doing so would allow to explain anything. But this criticism should apply with equal force to generally accepted individual differences such as subjective utility, private information, or different actions sets. A search for robustness, critical investigations of what drives the results, and a preference for simple models should prove to be sufficient safeguards. In particular, in this dissertation we relax the CPA only as far as needed to study the implications of differing beliefs.

Another worry is the possibility of very large (or even infinite) bets between people with differing beliefs, as suggested in Harrison and Kreps (1978). On this issue our position is two-sided. On the one hand, implicit bets seem to happen all the time. Consider for example the stock market or the huge stock-option 'incentive' plans in Silicon Valley[4]. On the other hand, it seems perfectly justifiable to exclude

[4]Our model in chapter 3 suggests that the employees of small companies with a vision often

explicit bets in economic models[5] since most people seem to dislike any meaningful explicit bets. This attitude might be explained by contractual difficulties, adverse selection, or the risk that the other agent might affect the outcome of the bet. A model along these lines is considered in Morris (1997).

Working with differing priors has actually a long tradition in economics. The idea of subjective probability can be traced at least to Knight (1921) who distinguished between 'risk', for which there are objective probabilities, and 'uncertainty', for which we have only subjective assessments. In financial economics, starting with Harrison Kreps (1978), differing priors are now used quite often (e.g. Harris and Raviv 1993, Allen and Morris 1998). Behavioral economics and its applications in litigation and bargaining have also relied quite heavily on this approach (e.g. Priest and Klein 1984, Babcock et al. 1995, 1996, 1997, Rabin and Schrag 1999). Note, however, that while behavioral biases such as overconfidence necessitate differing priors, the reverse is not true. In particular, perfectly rational people might equally well entertain differing priors, as we argued earlier. Finally, there is also a recent more theoretical interest in differing priors (e.g. Morris 1995, Feinberg and Scarsini 1999, Feinberg 2000, Yildiz 2000).

When taking this road, however, we have to make sure that our methodology is completely proper. A first issue is the question whether using differing priors is consistent with John C. Harsanyi's seminal work on Bayesian games (1967/68). While most of the analysis in that seminal work uses a common prior, section 15 of that paper points out explicitly that the methodology also works for situations that require differing priors. As long as each player's private information is derived via a lottery from a commonly known, but possibly subjective and individual prior, the methodology goes through. Moreover, Harsanyi explicitly states that the approach

have stronger beliefs in its business model than its owners. Stock options are then essentially a (wealth-constrained) bet between these employees and the owners of the firm.

[5]The possibility of infinite bets is eliminated if people are strictly risk-averse and agents are never 100 % sure about the true state (Harrison and Kreps 1978).

of a Bayesian game with common prior is 'restricted to the case in which player j has no special information suggesting mutual inconsistencies among the probability distributions'. Aumann (1976) and Geanakoplos and Polemarchakis (1982) show that mutual or common knowledge of disagreement amounts to such 'special information'.

A second issue are the epistemic foundations for the use of the (Bayesian) Nash equilibrium as solution concept. The most appealing epistemic foundations (Aumann and Brandenburger 1995) seem to rely explicitly on the common prior assumption. The next section formally shows, however, that these epistemic foundations are preserved under a partial relaxation of the CPA, which covers all 'agree to disagree' cases that are of interest to us. It should be noted that Yildiz (2000) relied implicitly on extensive-form rationalizability which is also preserved under differing priors of the kind considered here.

One could also wonder whether there are useful restrictions, short of the CPA, that may be imposed on differing priors. One such restriction would be to require that all priors have a common support. This means that no agent completely excludes some event that some other agent deems possible. Harrison and Kreps (1978) show that this restriction implies risk-averse agents will not make infinite bets. While such restrictions seem to make sense at face value, they may unnecessarily complicate the modeling and analysis. The information models in section 5, for example, violate this particular restriction.

2.3 Foundations for the use of Nash equilibrium

Aumann and Brandenburger (1995, henceforth AB) consider epistemic foundations for an n-player game and show that, if the players have a common prior, the players' payoff function and their rationality are mutually known, and the players' conjectures on others' actions are commonly known, then these conjectures coincide and form a mixed-strategy Nash equilibrium. They also provide examples that suggest that the CPA can*not* be dispensed with, unless we limit to the case $n \leq 2$ or assume right away that agents *know* (rather than 'have conjectures about') each others' actions.

The rest of this section shows, however, that their epistemic foundations extend to an n-player Bayesian game with differing priors on player types or payoffs. The key is to distinguish the probability space generated by the payoff- or Bayesian types from the space generated by the action- or strategy-types, and then to show that it is only with respect to the second that a common prior is necessary to get the epistemic foundations set forth in AB.

The next subsection introduces the setup and notation used by AB, with some small modifications. It also reiterates their result in formal terms. Subsection 2.3.2 discusses the extension to Bayesian games with partially differing priors.

2.3.1 Set-up and results of Aumann and Brandenburger (1995)

Let, throughout what follows, capitals denote sets, small type their elements, and boldface (sets of) vectors. For some I-dimensional vector $\boldsymbol{x} \in \boldsymbol{X}$, x_i denotes its i'th component, and $\boldsymbol{x}_{-i}$ denotes the $(I-1)$-dimensional vector obtained by removing $\boldsymbol{x}$'s i'th component.

Consider a finite game form $< I, (A_i)_{i=1}^I >$ with A_i denoting the action set of agent $i \in I$, with typical element a_i. Let S_i be the set of i's types with typical element s_i. To each s_i corresponds:

1. a probability distribution μ_{i,s_i} on $\boldsymbol{S}_{-i}$, denoting player-type $(i,\, s_i)$'s belief on the other players' types
2. an action a_{s_i}, denoting the strategy of player-type $(i$,$s_i)$

3. a payoff function $g_{s_i} : \mathbf{A} \to \mathbb{R}$, which is i's payoff function when his type is s_i (where $\mathbf{A}$ denotes the set of possible action profiles).

Note that, for some state $\hat{s}$, $< I, (A_i)_{i=1}^I, (g_{\hat{s}_i})_{i=1}^I >$ defines a normal form game.

Extend μ_{s_i} to a measure $p(\cdot, s_i)$ on $\boldsymbol{S}$ as follows: $\forall \boldsymbol{E} \subset \boldsymbol{S} : p(\boldsymbol{E}, s_i) = \mu_{s_i}(\boldsymbol{E}_{s_i})$, where $\boldsymbol{E}_{s_i}$ is the s_i-section of $\boldsymbol{E}$: $\boldsymbol{E}_{s_i} = \{s_{-i} \in \boldsymbol{S}_{-i} \mid (s_i, s_{-i}) \in \boldsymbol{E}\}$. Implicit in this definition is the assumption that s_i is, with probability one, correct about his own type. A probability distribution P on $\boldsymbol{S}$ is called a common prior if $p(\cdot, s_i) = P[\cdot \mid s_i] \; \forall i, s_i$.

A conjecture ϕ_{s_i} of player i is a probability distribution over $\mathbf{A}_{-i}$ as a function of his type s_i and represents his belief about the other players' actions. Let $[\boldsymbol{a} = \hat{\boldsymbol{a}}]$ denote the event that the action profile is $\hat{\boldsymbol{a}}$, which we will sometimes simplify to $[\hat{\boldsymbol{a}}]$ when the context is clear. We then define $\phi_{s_i}(\hat{\boldsymbol{a}}_{-i}) = p([\hat{\boldsymbol{a}}_{-i}], s_i)$.

Player i is said to be 'rational at some state $\hat{\boldsymbol{s}}$' if $a_{\hat{s}_i}$ maximizes the expectation of $g_{\hat{s}_i}$ with respect to i's conjecture $\phi_{\hat{s}_i}$. Player i is said to 'know an event $\boldsymbol{E}$ at state $\hat{\boldsymbol{s}}$' if $p(\boldsymbol{E}, \hat{s}_i) = 1$, an event which will be denoted $K_i\boldsymbol{E}$. Note that this definition is weaker than the definition of knowledge which requires absolute certainty. $\boldsymbol{E}$ is said to be mutually known if all players know $\boldsymbol{E}$: denoting this event by $K\boldsymbol{E}$, this says that $KE = \cap_{i=1}^I K_i\boldsymbol{E}$. $\boldsymbol{E}$ is said to be commonly known, an event denoted $CK\boldsymbol{E}$, if $\boldsymbol{E}$ is mutually known, and all players know that, and all players know *that*, etc.: $CK\boldsymbol{E} = \cap_{k=1}^\infty K^k\boldsymbol{E}$.

Note that this setup can accommodate Bayesian games, since payoffs are allowed to depend on the player-types.

The key proposition of AB (slightly adapted to fit the current terminology and notation) is then

Proposition 1 (Aumann and Brandenburger (1995) Theorem B) *Let $\hat{\boldsymbol{g}}$ be an n-tuple of payoff functions*[6]*, $\hat{\boldsymbol{\phi}}$ an n-tuple of conjectures. Suppose that the players*

[6]AB referred to this as 'a game', but the terminology usually has a different meaning, so we avoid it.

have a common prior, which assigns positive probability to it being mutually known that $\boldsymbol{g} = \hat{\boldsymbol{g}}$, mutually known that all players are rational, and commonly known that $\boldsymbol{\phi} = \hat{\boldsymbol{\phi}}$. Then, for each j, all the conjectures $\hat{\phi}_i$ of players i other than j induce the same conjecture σ_j about j, and $(\sigma_1, ...\sigma_I)$ is a Nash equilibrium of $\hat{\boldsymbol{g}}$.

2.3.2 Extension to Bayesian games with differing priors on Bayesian types

Extending this result to the situations of interest takes two steps:

1. Show that the result remains valid when the payoffs are subjective.
2. Show that with subjective payoffs the result can be reformulated for Bayesian games with differing priors.

So consider again the earlier set-up with the following modifications:

- With each player-type s_i is now associated an I-profile of payoff functions $\boldsymbol{g}_{s_i} = \times_{j=1}^{I} g_{j,s_i}$ with $g_{j,s_i} : \mathbf{A} \rightarrow \mathbb{R}$ denoting j's expected payoff function *according to i's type s_i* [7].
- Let then $\tilde{\boldsymbol{g}}(\boldsymbol{s}) = \times_{i \in I} g_{i,s_i}$ denote the vector with for each player his own (subjective) expected payoff function, when the type-profile is $\boldsymbol{s}$.
- A player is defined to be 'rational at some state $\hat{\boldsymbol{s}}$' if $a_{\hat{s}_i}$ maximizes the expectation of $g_{i,\hat{s}_i}$ with respect to i's conjecture $\phi_{\hat{s}_i}$.

Then the earlier proposition becomes:

Proposition 2 *Let $\hat{\boldsymbol{g}}$ be an n-tuple of payoff functions, $\hat{\boldsymbol{\phi}}$ an n-tuple of conjectures. Suppose that the players have a common prior on $\boldsymbol{S}$, which assigns positive probability to it being mutually known that $\tilde{\boldsymbol{g}} = \hat{\boldsymbol{g}}$, mutually known that all players are*

[7]This can on itself be considered as a combination of s_i's beliefs on j's types and of s_i's belief on s_j's payoff function g_{s_j,s_i}

rational, and commonly known that $\boldsymbol{\phi} = \hat{\boldsymbol{\phi}}$*. Then, for each* j*, all the conjectures* $\hat{\phi}_i$ *of players* i *other than* j *induce the same conjecture* σ_j *about* j*, and* $(\sigma_1, ...\sigma_I)$ *is a Nash equilibrium of* $< I, (A_i)_{i=1}^I, \hat{\boldsymbol{g}} >$*.*

Proof : Not surprisingly, it turns out that the earlier proof goes through nearly unchanged. In particular, the lemma's 2.6, 4.1, 4.3, 4.4, 4.5, and 4.6 of AB remain valid without any change since they do not relate in any way to $\boldsymbol{g}$.

Lemma 4.2. becomes

Lemma 1 *Let* $\hat{\boldsymbol{g}}$ *be a profile of payoff functions,* $\hat{\boldsymbol{\phi}}$ *an n-tuple of conjectures. Suppose that at some state* $\hat{\boldsymbol{s}}$*, it is mutually known that* $\tilde{\boldsymbol{g}} = \hat{\boldsymbol{g}}$*, mutually known that all players are rational, and mutually known that* $\boldsymbol{\phi} = \hat{\boldsymbol{\phi}}$*. Let* a_j *be an action of player* j *to which the conjecture* $\hat{\phi}_i$ *of some other player* i *assigns positive probability. Then* a_j *maximizes* $\hat{g}_j$ *against* $\hat{\phi}_j$*.*

The proof is identical after substituting everywhere $\hat{g}_j$ for g_j.

Finally, the proof of theorem B of AB goes through after substituting $\tilde{\boldsymbol{g}}$ for $\boldsymbol{g}$, and $\hat{g}_j$ for g_j. ■

The intuition for this modification is straightforward. We consider a player to be rational if he chooses his actions so as to maximize his subjective expected utility. It does not matter what other players expect his utility will be. So we might as well treat his subjective payoff function as 'the' payoff function. That is precisely what the modified propositions and proofs do.

For the second step in the argument, consider a Bayesian game $< N, \mathbf{A}, \boldsymbol{T}, \boldsymbol{\rho}, \boldsymbol{U} >$ with

- N denoting the set of players
- $\mathbf{A} = \times_{n \in N} A_n$ with A_n the action set of player n
- $\boldsymbol{T} = \times_{n \in N} T_n$ with T_n the set of types for player n

- $\boldsymbol{\rho} = \times_{n \in N} \rho_n$ with ρ_n a probability measure on T, denoting player n's prior belief

- $\boldsymbol{U} = \times_{n \in N} \times_{t_n \in T_n} u_{(n,t_n)}$ with $u_{(n,t_n)} : \mathbf{A} \rightarrow \mathbb{R}$ the payoff function of player-type (n, t_n) (i.e. of player n conditional on being of type t_n).

Note that we allow the priors to differ, which requires us to use the posterior-lottery model or Selten model (Harsanyi 1967/68). For this game, a Bayesian Nash equilibrium is simply defined as a Nash equilibrium of the strategic game $< K, \hat{\boldsymbol{A}}, \boldsymbol{V} >$ where

- $K = \times_{n \in N} T_n$ is the set of player-types, considered to be all players in the new game.

- $\hat{\boldsymbol{A}} = \times_{n \in N} \times_{t_n \in T_n} A_n$ where A_n is the action set of player (n, t_n), which is identical over all n's types

- $\boldsymbol{V} = \times_{n \in N} \times_{t_n \in T_n} v_{(n,t_n)}$ with $v_{(n,t_n)} : \hat{\boldsymbol{A}} \rightarrow \mathbb{R}$ being the payoff of player (n, t_n) in function of the overall action profile:

$$v_{(n,t_n)} = \sum_{\hat{a} \in \hat{\boldsymbol{A}}} \sum_{\hat{t} \in T} u_{(n,t_n)}((\hat{a}(n, \hat{t}_n))_{n \in N}) \rho_n(\hat{t} \mid t_n)$$

  where $\hat{a}$ (n, t_n) denotes the action of player-type (n, t_n), for $\hat{a} \in \hat{\boldsymbol{A}}$.

Define now $v_{(n,t_n)}^{(m,t_m)} = \sum_{\hat{a} \in \hat{\boldsymbol{A}}} \sum_{\hat{t} \in T} u_{(n,t_n)}((\hat{a}((n, \hat{t}_n)))_{n \in N}) \rho_m(\hat{t} \mid t_m)$ and let:

- $I = K$

- for $i = (n, t_n)$ and $j = (m, t_m)$, let $g_{i,i} = v_{(n,t_n)}$ and $g_{i,j} = v_{(n,t_n)}^{(m,t_m)}$ denote i's utility respectively according to himself and according to player j, independent of either player's action-types $s_i \in S_i$ or $s_j \in S_j$

We can then directly apply proposition 2, to conclude that the same epistemic conditions give rise to a Bayesian Nash equilibrium for a game with differing priors on the payoff- or Bayesian type-space $\boldsymbol{T}$, but with common prior on the action-type-space $\boldsymbol{S}$.

2.4 Efficiency

The fact that agents have differing priors does not exclude the possibility of an 'objective' or reference prior. With many different realities, however, the concept of ex ante efficiency might seem confusing: who will tell whether a person is better off in expectation? There are two possible answers: either each person decides for himself whether he is better off, or some objective outsider does that for each player in the game. Which criterion should be used depends on the question being asked. In the end, 'man is the measure of all things'.

To make things concrete, consider a situation with N players, a state space Ω, a reference or 'objective' prior ρ, and for each player $n \in N$ a subjective prior ρ_n. Individual agents do not know ρ and honestly believe their individual ρ_n to be the correct prior.

If the issue is whether the agents could improve their situation among themselves, for example by agreeing to change the rules of the game, then we should take the subjective perspective of the individual agents, and the appropriate concept for efficiency is then subjective efficiency.

Definition 1 *An allocation* $(x_n)_{n\in N}$ *is* subjectively efficient *if there does not exist any other allocation* $(\hat{x}_n)_{n\in N}$ *such that*

$$\sum_{\omega\in\Omega} u_n(\hat{x}_n,\omega)\rho_n(\omega) \geq \sum_{\omega\in\Omega} u_n(x_n,\omega)\rho_n(\omega)\ \forall n \in N$$

with strict inequality for some $n \in N$*.*

This can also be considered to be the decentralized version of Pareto efficiency: an outcome is efficient if there does not exist any other outcome that all agents would be willing to sign for.

Note that a subjectively efficient allocation does not always exist. If people disagree on the probability of an event and they are risk-neutral (or one of the agents attaches zero probability to an event that some other agent considers possible and utilities are strictly increasing), then we could increase both agents' utility indefinitely by having them make increasing bets with each other. As mentioned earlier, a restriction on the prior beliefs that their supports must be identical combined with risk-aversion solves this existence problem. Section 2 suggested a more general rationale to exclude explicit bets from economic models.

A very different perspective is that of some social planner who cares about the agents' 'true' utility and who considers making changes to the system. In that case we should use the social planner's belief to evaluate utilities, which can then be considered the reference or 'objective' belief ρ. The right notion is then objective efficiency.

Definition 2 *An allocation* $(x_n)_{n \in N}$ *is* objectively efficient *if there does not exist any other allocation* $(\hat{x}_n)_{n \in N}$ *such that*

$$\sum_{\omega \in \Omega} u_n(\hat{x}_n, \omega)\rho(\omega) \geq \sum_{\omega \in \Omega} u_n(x_n, \omega)\rho(\omega) \ \forall n \in N$$

with strict inequality for some $n \in N$.

This is also the perspective of an 'objective' outsider, such as a researcher.

In some cases we are not interested in efficiency at all. Chapter 3, for example, is only interested in the firm's expected profit. This poses to some degree the same problem. That chapter takes the perspective of an outsider to evaluate profitability. The outsider can be interpreted as the board or the financial markets.

2.5 Two modeling approaches

The most straightforward approach for formulating a model with differing priors, used for example in Yildiz (2000), is to posit directly the individual and reference priors on the variables of interest. To be concrete, consider a game with a set of agents N and a payoff state space $\Omega = (\omega_1, \omega_2)$. The direct approach posits for each agent a belief $\rho_i \in [0, 1]$ that the true state is ω_1, and possibly also a reference prior ρ_0.

A second approach, used in chapter 3, is to construct an information model in which the agents have differing beliefs on the informativeness of their signals. A farmer and a meteorologist, for example, might have a different opinion about the informativeness of observing a low-flying swallow. For the two-state situation above, the simplest version of this approach would posit a common and reference prior $\hat{\rho}$ on those two (payoff) states, and assume that all agents observe an identical signal, which is objectively correct with probability p_0 and which the agents subjectively believe to be correct with probabilities $(p_n)_{n \in N}$. Note that in this case the differing priors refer to the beliefs regarding the informativeness of the signal. This leads to a set of beliefs which can then be used as the subjective priors $\tilde{\rho}_i$ in the main model. Note that, as long as each and every ρ_i is absolutely continuous with respect to $\hat{\rho}$, then there exist a set of p_i such that $\tilde{\rho}_i = \rho_i \, \forall i \in N \cup 0$.

While the second approach is more cumbersome, it has its merits:

- It makes it easier to judge the 'reasonableness' of the priors that are considered in the analysis. In particular, it allows us to link these assumptions to cognitive biases, unknown to the beholder, such as overconfidence or lack of updating (e.g. Nisbett and Ross 1980). Note again, however, that differing priors do not require a presumption of biases in information processing.

- It creates an explicit link with an underlying 'objective reality' which facilitates evaluations in terms of, for example, efficiency.

- Without differing beliefs about the informativeness of signals, the accumulation of information would often lead to a convergence of priors. This construction therefore deals directly with a possible criticism of the differing prior approach. Note that differing beliefs on informativeness can be interpreted as people holding different theories or views of the world.

We consider now two such indirect models in more detail. Both of them deal with instances of biases in information processing.

Common knowledge of overconfidence The first model is motivated by a well-known behavioral bias, overconfidence (see e.g. Nisbett and Ross 1980). The model makes the following assumptions:

- players start with a common prior (on the state space excluding the informativeness of signals), which we assume to be objective
- each player receives a private signal about the true state
- agent i believes his signal is correct with probability $\rho_{i,i}$, while all other agents believe his signal is correct with probability $\rho_{i,-i}$
- the beliefs on the correctness of signals are common knowledge

For a simple parametric example consider a model with two agents and two possible states. Let the agents' (common) prior be completely uninformative, so that they attach equal probability to each of the two states. Assume each gets a signal that is correct with probability p. They are overconfident in that agent i thinks his signal is correct with probability $\beta_i p$, with $\beta_i \geq 1$. After both have received their private signal, their beliefs are simultaneously announced and they update.

When $p = 1/2$, each agent's signal is uninformative (or believed to be so by the other agent). The belief of an agent then depends only on his own signal. In particular, he believes that with probability $\frac{\beta_i}{2}$, the state is indeed the one that his signal suggested.

Update on new facts only, and the role of convincing someone Consider now the following modification of the above model:

- Let p_i be the objective probability that i's signal is correct. Let further $\rho_{i,i} = p_i > \rho_{i,-i}$. This says that each player is appropriately confident about his own signal, but players discount the information they have to infer from other people's beliefs.

- People can explicitly communicate 'objective facts' or signals. When he hears a fact not yet known to him, the receiver uses the objective informativeness of that signal to update his beliefs.

In the extreme case where $\rho_{i,-i} = 1/2$, agents do not update their belief when they are confronted with an opinion that is different from their own (without hearing any supporting facts). They do, however, update their beliefs when they are confronted with new facts that were not yet included in their beliefs. Note that this model now allows for one agent to actively 'convince' another agent by explaining the facts on which his beliefs are based.

Other models Of course other approaches are possible. The primacy and recency effects, for example, suggest that the agents' beliefs will depend on the sequence in which the pieces of information are received. Another interesting variation is Rabin and Schrag (1999) who allow agents to interpret new information so that it tends to confirm their current hypothesis.

2.6 Conclusion

This paper discussed the motivation and methodological implications of allowing agents to hold differing priors. The essence of the argument is that differing priors are often the best approach for studying phenomena in which open disagreement is crucial and that they are completely consistent with the economic paradigm and methodology.

Chapter 3

Organizational Beliefs and Managerial Vision [*]

> *I firmly believe that any organization, in order to survive and achieve success, must have a sound set of beliefs on which it premises all its policies and actions.*
>
> Thomas J. Watson, Jr.
> Former Chairman IBM

3.1 Introduction

Beliefs can shape reality. Organizational beliefs can shape corporate behavior and performance. Donaldson and Lorsch (1983), in their extensive study of top management decisions, stated that 'beliefs and corporate strategy are closely intertwined - at times almost indistinguishably so.' Until 1995, for example, Microsoft and Sun held nearly opposite beliefs regarding the future of computing, which led them to very different strategic choices. Such beliefs are often determined by the vision of the CEO or founder. In fact, practice-oriented studies have concluded that vision

[*]This chapter benefitted a lot from useful comments by Bill Barnett, Robert Burgelman, Kate Dewhurst, Yossi Feinberg, Thomas Hellmann, Jon Levin, Paul Oyer, Andrea Shepard, Brian Viard, Muhamet Yildiz, Ezra Zuckerman, and the participants in the seminars at Harvard, MIT, Northwestern, Stanford, UBC, University of Chicago, and USC. The usual disclaimer applies.

is the key to leadership (Korn Ferry 1989, The Economist Intelligence Unit 1996, Robertson and Walt 1999).

With the notable exception of Rotemberg and Saloner (2000), discussed in detail below, economics has neglected these topics. This gap might be due to the fact that organizational beliefs and vision are thought to be outside the realm of economics. We will argue, however, that these phenomena do fit the economic paradigm and can be studied formally as long as we allow differing beliefs. Moreover, our analysis suggests that the impact of such belief differences is pervasive, so that an economic theory of organizations will have to take them into account.

The model and results. The focus of this paper is on the interaction between the employees' beliefs and those of the manager. 'Vision' is defined operationally[1] as a very strong belief by the manager about the future and about the right course of action for the firm.

The impact of organizational beliefs and managerial vision is studied in the context of a simple model. In this model, employees can spend effort on developing new initiatives. If an employee comes up with a project, his manager has to decide on implementation. If the project gets implemented and is a success, the employee gets part of the revenue through ex-post bargaining. At the time of the project generation and implementation, however, there is uncertainty about what kind of projects (A vs. B) will be successful. The key to the analysis is that the employee and the manager may openly differ in their beliefs about the right course of action. This means that we do not impose the common prior assumption (CPA), an approach that will be justified in more detail.

A stronger belief of the manager will motivate those employees who agree with him to such a degree that they undertake the project that the manager deems optimal. The reason is simply that they get easier approval for the projects they

[1]The relationship of this definition to those in the managerial and psychology literature will be discussed later.

undertake. But, by the same token, it will demotivate those who disagree too much. Analogous effects will increase resp. decrease employee satisfaction. This gives rise to sorting: a firm attracts employees with beliefs that are similar to those of its manager. Such sorting reduces the demotivating effect of vision. This feedback loop suggests that vision might overall be profitable.

To evaluate the profitability of vision, the paper takes the perspective of an outsider, such as the board, with an 'objective' or reference belief. This gives three conclusions. First, in the presence of sorting, vision is profitable under weak conditions. Second, the effect increases in the importance of motivation and initiative, but decreases as market uncertainty goes away. And, third, even when vision is not optimal ex-ante, ex post the best (and worst) firms in the market will be those with a visionary CEO and strong organizational beliefs. This might make vision look better than it really is. The final sections of the paper contain an informal discussion of extensions and related concepts, such as corporate culture and strategy.

The literature. Bennis and Nanus (1985) and Tichy and Devanna (1986), building on the theories of charismatic or transformational leadership (House 1977, Burns 1978), were the first to really focus on managerial vision. Before them, Donaldson and Lorsch (1983) had already documented the importance of managerial and organizational beliefs.

Rotemberg and Saloner (2000) provided the first formal economic model of vision. Extending their work on leadership styles and strategy (Rotemberg and Saloner 1993, 1994), they consider a firm with two employees, or product divisions, working on different projects. Vision in their model is a bias of the manager that makes him favor one project over the other. Such vision improves the incentives of one employee at the cost of reducing the incentives of the other. While the setting is quite different, this effect is similar to the motivation effect in this paper. By allowing all agents to have subjective beliefs, we show that this is only the tip of the iceberg. In particular, vision also influences decisions, satisfaction, hiring, and the organizational beliefs themselves. Goel and Thakor (2000) is somewhat complementary to our analysis. They define overconfidence as underestimating project risks and argue that overconfident people have a higher probability to win in tournaments and thus get elected as leader. They argue further that such overconfidence in managers is good for shareholders since it compensates for their risk aversion[2].

Recently, there have also been some empirical contributions on the effects of vision. Baum, Locke, and Kirkpatrick (1998), for example, find evidence of a positive influence of vision on venture growth and discuss other empirical studies. There is also an extensive related literature, such as that on culture, leadership, or delegation. That literature will be discussed later in the paper.

The next section explains the model setup. It also discusses our notion of vision and compares its definition to that in the literature. Section 3.3 discusses differing priors. Sections 4-6 are the core of this essay. They analyze the impact of organizational beliefs in one-firm and multiple firms contexts, and consider when vision would be profitable. Section 7 discusses the implications for culture and strategy. Section 8 concludes and suggests further topics for research. Appendix A considers some implications for small firms with size restrictions. Appendix B discusses the impact of changes in the assumptions or set-up of the model. All proofs are in Appendix C.

3.2 The model

A sketch of the model Remember the basic model, as sketched in the introduction. Employees try to develop initiatives. The probability that an employee 'comes up with something' is a function $q(e)$ of his effort e. The manager then decides whether to implement it. In making that decision, he considers not only the project's expected revenue but also its organization-wide implementation cost,

[2] For other economic perspectives on leadership see Rotemberg and Saloner (1993) and Hermalin (1998, 1999).

which is a random variable I. If the project gets implemented and is a success, the employee gets part of the revenue through ex post bargaining.

The key element of the model is the presence of uncertainty about which projects will be successful. In particular, at the time they expend effort, employees also have to choose between A- and B-type projects, which are mutually exclusive. The success and revenue of each project type depends on its fit with the (unknown) state of the world $x \in \{a, b\}$. In particular, X-type projects are successful if and only if the state is x. Successful projects will generate a revenue of 1 while failures generate no revenue. Note that the state may include any factor that has a bearing on what the optimal action is, including evolution of the industry, core competences of the firm, or 'the right way of doing things'. All agents in the model have their own subjective belief about the likelihood of each state. These beliefs may differ but are common knowledge. This implies, by Aumann (1976), that the agents must be allowed to start from different priors. It also implies that agents will not update their beliefs merely because they are confronted with a different opinion. This assumption is discussed in more detail in section 3.3.

We will use the notation $\mu_{i,Y}$ for the probability that agent i assigns to the event that state is Y. Employee E, for example, believes that with probability $\mu_{E,A}$ the state is A. The strength of an agent's beliefs will turn out to play an important role in the analysis. We will denote this by $\nu_i = \max(\mu_{i,A}, \mu_{i,B}) \in [1/2, 1]$, i.e. ν_i is the strength of i's belief in the state he considers most likely. We say that an agent has a 'more precise belief' or 'stronger conviction' if ν_i is larger. Finally, p will denote the reference belief or 'objective' probability (that the true state is a) used to evaluate profitability.

We now proceed to a more detailed description of some elements in the model.

Agents, utilities, and beliefs The model has 3 types of agents: firms, managers, and employees. In the analysis of optimal vision, we imagine the firm to

1	2	3	4
Hiring process	Project	Renegotiation	Payoff
1 Employee chooses firm (if there is more than one firm). 2 Firm makes initial wage offer $\tilde{w}$. 3 Employee accepts or rejects. Upon rejection, employee gets outside wage $\underline{w} = 0$.	1 Employee chooses (one and only one) type of project $X \in \{A, B\}$ and invests effort $e \in \mathcal{E}$ (cost $c(e)$ sunk). 2 Employee generates project with probability $q(e)$. 3 Manager observes project type and implementation cost, which is a random variable $I \sim U[0,\ 1]$. 4 Manager decides whether to implement (cost I sunk). 5 Employee and manager observe whether the project will be a success (i.e. they observe the state).	1 Employee can ask for a raise (i.e. decides on wage renegotiation). 2 Manager and employee renegotiate wage (see below). Upon breakdown, employee decides to stay (and get $w = \tilde{w}$) or leave (and get $w = 0$) but the project will be a failure either way.	1 Successful projects generate 1, failures generate 0. 2 Wages paid (according to renegotiation).

Figure 3.1: Timeline of game

be represented by the board (or the owner) who chooses the manager. The board maximizes expected profits using the reference belief p. Each firm also has a manager who hires its employees and decides on implementation. Managers maximize expected firm profit based on their own subjective belief $\mu_{M,Y}$. Employees, finally, choose projects and spend effort on developing them. They maximize their expected revenue net of any cost of effort. In doing so they also use their own beliefs $\mu_{E,Y}$.

Actions and timing The precise timing is indicated in figure 3.1. Stages 1, 2, and 4 are straightforward. The renegotiation in stage 3 is according to Nash bargaining with relative bargaining power $\gamma_E, \gamma_M > 0$, with $\gamma_E + \gamma_M = 1$. This means that a dollar extra will be split (γ_E, γ_M). If bargaining breaks down, the project fails and generates 0 revenue while the employee can choose either to stay with wage $w = \tilde{w}$ or to leave the firm and take his outside wage $\underline{w} = 0$. This renegotiation is just a way to assure that the employee cares about the outcome. We would obtain the same results if, for example, the employee cares about the outcome because it affects his outside options or future promotions. Finally, the results would not change if we also allowed the firm to ask for renegotiation. Appendix 3.B considers further modifications to the setup of the model.

Contractibility Implicit in this timeline are a number of assumptions as to what variables are contractible. In particular, we implicitly assume that the agent's effort e and the project type are economically too complex to contract on. We also assume that employees, without spending any effort, can come up with bad (zero-revenue) projects that are indistinguishable from good ones for an outsider, so that 'coming up with a project' is not contractible. We further let future revenue become (economically) contractible only after the project has been implemented. This can be justified by the difficulty of describing the revenues of a project that does not exist. It then follows that the only possible contract at the start of the game is a fixed-wage contract[3], as described in the timeline. The description of the game also implicitly assumes that the employee's support is needed until the end for the project to become a success, and that he can withdraw that support at will to force a renegotiation. Appendix 3.B discusses how these contractibility and renegotiation assumptions affect the results. That appendix also considers other variations on the basic model.

We also make a number of explicit assumptions:

Assumption 1 *Employees' beliefs are independent draws from a distribution of beliefs F on $[0, 1]$, with continuous density f.*

When indifferent about which firm to apply to, employees randomize between the two firms with equal probability. When indifferent about what action to undertake, employees do as their manager prefers. When indifferent about implementation, managers do as their employee prefers.

Assumption 2

- *The implementation cost I is distributed uniformly on $[0, 1]$.*
- *The probability of success $q(e)$ and the cost of effort $c(e)$ are twice continuously differentiable on $\mathcal{E}$; $0 \in \mathcal{E}$; $\mathcal{E}$ is compact.*

[3]Note that this wage offer $\tilde{w}$ may depend on the employee's belief. Any such dependence, however, gets lost in the later renegotiation.

- $1 \geq q(e) \geq 0$, $q'(e) > 0$, $q''(e) \leq 0$;
- $c''(e) > 0$; $c(0) = c'(0) = 0$; $\lim_{e \to \max \mathcal{E}} c'(e) = \infty$;
- $\gamma_E q(\tilde{e}) \leq c(\tilde{e})$ *where* $\tilde{e} = \inf\{e \in \mathcal{E} \mid q(e) = 1\}$

These are all standard assumptions, except for the last one which assures that the optimal $\hat{e}$ always has room to increase further.

We finally assume that

Assumption 3 *The reference probability* $p \geq 1/2$.

This assumption is without any loss of generality since we can always rename the states and project-types to make sure it holds.

A practical example To fix ideas, think back a few years to the time that the Internet was close to taking off and consider a software product manager who is preparing the next version of his product. His key issue is whether to add and improve traditional features or to focus instead on adding Internet capabilities. The future success of his product may depend crucially on this choice. Complicating matters is the fact that the CEO has the final say on any new release. Consider now the case that the product manager believes the Internet is no more than a fad and developing Internet capabilities a complete waste of resources which might put him fatally behind his competitors. His CEO, however, is a true believer and has made clear to her product managers that they should focus on making their products Internet-ready.

In this case, contracting on direct output is problematic since it is difficult to define what Internet-ready means, what good implementation means, or what the relative importance of different features is. Software development efforts are also difficult to measure objectively. Finally, his product's success is obviously a key factor in the product manager's future wage negotiations (or promotions), but it is difficult to contract on long in advance given the fundamental uncertainties in the industry.

Operational definition of vision As mentioned earlier, we define vision operationally as a strong belief by the manager about the optimal course of action for the firm. A manager who says that 'anything is possible' has no vision, while one who claims that 'in five years handhelds will have completely displaced PC's' conveys a strong sense of vision. In principle, a manager would thus be visionary if he has a stronger belief than the board, i.e. if $\nu_M > \max(p, (1-p))$. Given that we assumed $p \geq 1/2$, however, the interesting case is when the manager has a stronger belief than the reference belief. Our operational definition will thus be that the manager is visionary if $\mu_{M,A} > p$.

This operational definition captures a common element of most descriptions in the literature: vision as a clear idea about (or 'picture' of) the future and the firm's position in that future. Bennis and Nanus (1985), for example, describe it as 'a mental image of a possible and desirable future state of the organization'. Similar definitions are used in most of the literature and dictionaries[4].

The 'vision' studied in this paper is in fact a case of 'overconfidence' (see e.g. Einhorn and Hogarth 1978, Nisbett and Ross 1980). This fits rather well with the 'charismatic leadership theory' in psychology (House 1977, Conger and Kanungo 1988), which showed that 'self-confidence' and 'a strong conviction in the own beliefs' are key characteristics of charismatic and transformational leaders.

While a strong belief about the right course of action is a 'necessary' component

[4]Tichy and Devanna (1986) state that '[Transformational leaders] must provide people with an image of what can be ...'. Kouzes and Posner (1987 p85) define it as 'an ideal and unique image of the future'; Kotter (1990) defined it as 'a description of something (...) in the future, often the distant future, in terms of the essence of what it should be.' The Cambridge Dictionaries Online defines vision as 'the ability to imagine how a country, society, industry, etc. will develop in the future and to plan in a suitable way'. As the term became more popular, however, it sometimes got extended to cover a much broader set of concepts (Quigley 1993) or simply as a synonym for 'admired' (Collins and Porras 1994).

of vision, it does not seem to be 'sufficient'. The management literature argues, for example, that vision also creates 'meaning' or that vision must be attractive (e.g. Bennis and Nanus 1985). By abstracting from these aspects, we do not mean that they are necessarily less important. But we think that doing so is useful on the following grounds. First, the effects seem to be sufficiently independent to allow, at least to the first order, a separate study. Second, such separate analysis is more transparent in terms of cause and effect and allows us to disentangle the implications of specific assumptions. Finally and most importantly, the results we get are very similar to the claims made for managerial vision, which suggests that this very simple definition might well capture the part of the phenomenon that 'does the trick'.

3.3 A note on 'differing beliefs' in economic modeling

The model in section 3.2 differs in one respect from most economic models: the agents knowingly entertain differing beliefs[5] (without having private information). The reason for this assumption is pragmatic: differences in beliefs are at the heart of the issues studied here, and assuming common knowledge of differing beliefs is the most transparent and parsimonious way to study this question[6]. Differing beliefs do not contradict the economic paradigm: while rational agents should use Bayes' rule to update their prior with new information, nothing is said about those priors themselves, which are primitives of the model. In particular, absent any relevant

[5]Whenever we refer to belief differences, or agents entertaining differing beliefs, we mean that 'agents have differing beliefs about a specific event and their beliefs are common knowledge'. In economics, the term 'disagreement' is often used to denote such belief differences. We avoid this term since it suggests conflict. We are definitely not the first to use such differing priors. See for example Harrison and Kreps (1978) or Yildiz (2000).

[6]While formally most of the analysis can be done under standard assumptions, such analysis would miss the essential point: that, *holding information constant*, the strength of beliefs is an important influence; that it can be optimal to have a CEO who has stronger beliefs than the board *even if he does not have more information.*

information agents have no rational basis to agree on a prior. Harsanyi (1967/68), for example, 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'. The best argument for the traditional use of common priors is Aumann's (1987) argument that they allow us to 'zero in on purely informational issues'. Conversely, differing priors allow us to zero in on the implications of open disagreement and differing beliefs.

Chapter 2 considered this issue in more detail[7]. Among other things, it argues against the idea that differing priors might allow us to explain anything, it discounts the theoretical possibility that agents will make very large or infinite bets[8], and 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.

Working with differing priors also raises the issue how to measure expected profits and thus how to determine the optimality of a vision. To that purpose, this paper uses the perspective of an outsider with a 'reference' belief. This outsider can be interpreted as the board or the financial markets. Or this outsider can be interpreted as a truly objective observer who knows the true probability, in which case the analysis studies which firms will really fare better in expectation.

One further remark to facilitate the interpretation of the model is in order. The distribution of beliefs is implicitly assumed to be generated by the following information process. All agents start with a common prior on the state $x \in \{a, b\}$ that puts equal probability on both states. All agents subsequently get a common signal that, for example, the true state is a. The agents, however, have their own

[7]See also Morris (1995) or the discussion between Gul (1998) and Aumann (1998).

[8]Note that such bets are simply not possible in the model under consideration. On the other hand, our model suggests that, often, at least some employees of a 'visionary' company have stronger beliefs in its business model than the owners. Stock options are then essentially a (wealth-constrained) bet between these employees and the owners of the firm.

subjective opinion about the correctness of that signal and these beliefs are common knowledge. In particular, it is commonly known that agent i thinks that the signal is correct with probability $\mu_{i,A}$. Note here that differing beliefs about the correctness of the common signal is just a special case of differing priors[9]. Bayesian updating then leads the agent to believe that the probability of state a is $\mu_{i,A}$. The 'reference' belief p is the belief of the board about the signal. Note that a 'visionary' manager, as defined above, is in fact overconfident relative to the reference belief.

3.4 Decisions, motivation, and satisfaction

The basis of the analysis is an understanding how individual employees react to their manager's beliefs. This is the subject of this section. To that purpose, consider the model of section 3.2 with the firm having only one employee. Let $\mu_{E,Y}$ and $\mu_{M,Y}$ denote the beliefs of the employee and the manager that the correct course of action is Y. Throughout this and the following sections, hats will indicate optimized choice variables and value functions. So $\hat{e}$ is the employee's optimal choice of effort while $\hat{u}$ is his optimized utility. To make the notation more transparent, the dependence of the maximizers on other variables will be suppressed.

We now reason by backwards induction[10]. The renegotiation process will give the firm a gross revenue γ_M if the project is implemented and turns out to be a success, and zero otherwise. Given that the manager can observe the project type, he will thus allow a project Y to be implemented if and only if $\gamma_M \mu_{M,Y} \geq I$. Prior to the revelation of I, the project will thus be implemented with probability $\gamma_M \mu_{M,Y}$, which gives the employee an expected payoff from proposing a project of $\gamma_E \gamma_M \mu_{E,Y} \mu_{M,Y}$.

[9] In particular, agents not only have (prior) beliefs about the state $x \in \{a, b\}$, but also about what game they are playing, how correct their information is, etc. In this particular case, agent i puts probability one on the signal being correct with probability $\mu_{i,A}$, but agent j puts probability one on $\mu_{j,A}$, which might be different.

[10] The proof of proposition 1 shows that all SP equilibria have the same outcomes and are equivalent to one in which the firm offers a wage $\tilde{w} = 0$ and the employee accepts.

In choosing the type of initiative and e, the employee thus solves:

$$\max_{e \in \mathcal{E}, Y \in \{A,B\}} q(e)\gamma_{M}\gamma_{E}\mu_{E,Y}\mu_{M,Y} - c(e)$$

The next proposition now says that whoever has the stronger beliefs or conviction about what should be done, will determine what will be done.

Proposition 1 *If the manager has the stronger conviction then the employee undertakes the action that his manager prefers. Otherwise he follows his own opinion. Formally: if* $\nu_{M} \geq \nu_{E}$ *with* $\nu_{i} = \max(\mu_{i,A}, \mu_{i,B})$*, then* $X = \text{argmax}_{Y \in \{A,B\}} \mu_{M,Y}$*, otherwise* $X = \text{argmax}_{Y \in \{A,B\}} \mu_{E,Y}$*.*

The intuition is simple. If the manager and the employee agree on the optimal action, then E chooses of course that action. If they have different opinions, the employee will have to 'disappoint' one of the two. Since the roles of their beliefs are symmetric in the employee's utility function, it is optimal to 'disappoint' the one who holds the weaker belief (i.e belief closer to 1/2).

Given this symmetry, one might wonder what the difference between the employee and the manager really is: why do we say that managers have a vision while employees 'only' have beliefs? The difference is, first, that the manager influences the decision of the employee but not the other way around and, second, that the manager also influences other employees. On the other hand, it should be noted that not only managers have such influence in actual organizations: the sociological literature on 'gatekeepers' describes precisely how persons with little formal authority who control the access to important resources (such as the assistant to the CEO) can wield a lot of influence (Mechanic 1962). Such cases, however, are not intentional and their impact is most probably less pervasive than that of a manager.

A different way to look at proposition 1 is to say that the manager keeps a strong influence over the project type, even though the decision is formally delegated to the employee[11]. In many non-routine jobs, such indirect authority might be a more effective way to influence the course of action than direct authority, since, among other things, the manager has to get involved only after the project has been successfully developed. For this kind of decision processes, the earlier results then imply that

Corollary 1 ('Visionary managers have more influence.') *The prior probability that the project choice is according to the manager's belief increases in ν_M, the manager's conviction in his view of the world.*

While the manager's opinion has an important influence on the *decisions* of the employee, it is also a key determinant for the employee's *motivation and satisfaction* (or effort and utility). The following proposition essentially says that a stronger belief of the manager will motivate the employee and increase his satisfaction *if* the employee acts according to the manager's beliefs. Such stronger beliefs, however, will *de*motivate an employee who goes against the manager's opinion and will reduce his satisfaction. To state this formally, let N be an open neighborhood of μ_E and μ_M on which the chosen project type X remains identical and let $0 < \mu_{i,A} < 1$ for both agents.

Proposition 2 *Employee effort $\hat{e}$ and satisfaction (or utility) $\hat{u}$ strictly increase in the conviction of the manager $\nu_M = \max(\mu_{M,A}, \mu_{M,B})$ (resp. in the employee's own conviction ν_E) on N if the employee undertakes the action that the manager strictly prefers $X = \text{argmax}_{Y\in\{A,B\}} \mu_{M,Y}$ (resp. that he himself strictly prefers).*

Analogously, employee effort $\hat{e}$ and satisfaction $\hat{u}$ strictly decrease in his manager's conviction (resp. his own conviction) on N if he undertakes the opposite action of what his manager strictly prefers: $X = \text{argmin}_{Y\in\{A,B\}} \mu_{M,Y}$ (resp. of what he himself strictly prefers).

[11] The model might also complement the theory of delegation (Vickers 1985, Prendergast 1993, Aghion and Tirole 1997, Baker Gibbons and Murphy 1999, Zabojnik 2001). The main conjecture would be that, (all else equal) with effort complementary to the probability of success, the project type decision should be taken by the person with the more important non-contractible effort.

The intuition is simple. Suppose that the employee undertakes a project that is the right course of action according to his manager. As the manager is more convinced of that action, the probability that he will implement the project increases. This will increase the expected payoff to the employee from trying to develop the project, which indeed motivates him and gives him higher satisfaction.

This result can be loosely interpreted as follows:

- Employees with no specific opinion on the correct action (μ_E close to 1/2) get more motivated by managers who know precisely what they want, no matter what they want. The same is true for employees whose utility depends only on implementation or approval, and not on the final success (since this case is formally equivalent to setting $\mu_{E,X} = 1$ for the likelihood of whichever action is chosen).

- Employees with a strong opinion about the correct path of action will be very motivated under managers who agree with them (and more so as the manager is more convinced of that opinion). But they will be very *de*motivated under managers with a different opinion.

These statements fit casual empiricism.

3.5 The sorting effects of vision

The motivation and satisfaction effects cause sorting in the labor market[12], which then feeds back into motivation and satisfaction. The basic argument runs as follows.

- Employees get higher satisfaction working for firms that espouse a vision they agree with. Firms get higher profits from employees who agree with their vision, since the latter are more motivated. An efficient labor market should

[12]Note that effects similar to the ones described here can occur in other types of markets. In particular, investors (in financial markets) will be willing to pay more for equity in firms whose managers have beliefs that are similar to their own.

therefore match employees and firms with similar beliefs. Since sorting determines the type of employees a firm attracts, which then influences its profit, this might on itself constitute a sufficient reason for deviating from the 'objective' (or reference) belief.

- Once sorting has taken place the beliefs of the employees and the manager are more aligned. This will decrease or even eliminate the demotivating effect that vision had on some employees, so that vision becomes more effective.

Partial evidence for such sorting comes from sociological studies (e.g. Chatman 1991) that show how employees and firms take into account 'fit' when deciding which firms to join or who to hire. While this evidence relates more to fit in terms of values, it does suggest that such sorting mechanisms operate. We expect the same conclusions to hold for fit in terms of beliefs, especially on the more executive levels of the organization.

To study these effects formally, consider again the model of section 3.2 but let the employee, with belief $\mu_{E,A}$, have the choice between two firms, F_1 and F_2, with managers M_1 and M_2 who have beliefs $\mu_{M_1,A}$ and $\mu_{M_2,A}$, where we assume wlog $\mu_{M_1,A} \geq \mu_{M_2,A}$.

There is again an essentially unique subgame perfect equilibrium, which gives sorting as indicated in figure 3.2.

Proposition 3 *Let* $\breve{\mu} = \frac{1-\mu_{M_2,A}}{\mu_{M_1,A}+1-\mu_{M_2,A}}$. *In any subgame perfect equilibrium, all employees with* $\mu_{E,A} > \breve{\mu}$ *end up being hired by* M_1, *while any employee with* $\mu_{E,A} < \breve{\mu}$ *will be hired by* M_2. $\breve{\mu}$ *decreases in both* $\mu_{M_1,A}$ *and* $\mu_{M_2,A}$.

It can be shown that this allocation of employees is the unique stable matching and the unique element in the core (defined by weak domination) of the corresponding cooperative matching game. To see intuitively what is happening consider first the upper graph of figure 3.2. There are two managers who have approximately

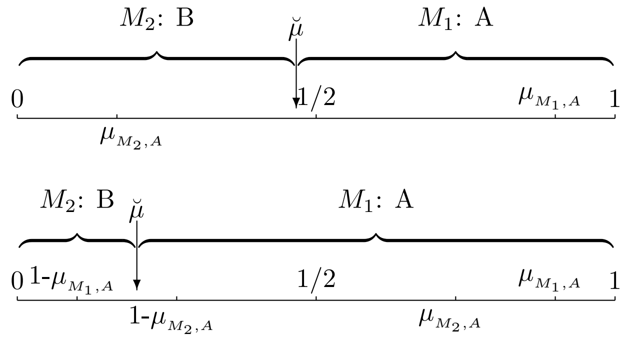

Figure 3.2: Choice of action in function of beliefs

opposite beliefs. Consider the situation of an employee with belief $\mu_E = 1/2$. Personally this employee doesn't see any difference between the two alternatives. All he thus cares about is the probability of implementation. So he will go with the manager with the strongest conviction, which is M_1. Given that his preference is strict we know that the cutoff $\breve{\mu}$ must be strictly to the left of 1/2.

Note two things :

1. The employee with $\mu_E = 1/2$ is closer to M_2 in terms of beliefs, but goes to firm F_1, since M_1 'knows better what he wants'.

2. As M_1 gets more convinced, he becomes more attractive to work for. In particular, an employee that before was indifferent will now go to work for M_1. So $\breve{\mu}$ shift to the left as μ_{M_1} shifts to the right. The same is true for M_2. This gives the lower graph.

The result is also striking in the sense that the firm with the stronger vision attracts precisely these employees who take action according to its manager's beliefs[13].

Corollary 2 *If manager M_1 has the stronger belief, then any employee hired by F_1 will choose the action preferred by its manager, while any employee hired by F_2*

[13]It should be noted that this extreme outcome is partially due to the limited state-space. Nevertheless, even with richer state-spaces, we conjecture that the essence of the result will carry over.

will choose the other action (which then might or might not be preferred by M_2). Formally: if $\nu_{M_1} > \nu_{M_2}$ then $X = \text{argmax}_{Y \in \{A,B\}} \mu_{F_1,Y}$ for F_1.

The intuition is simply that an agent who goes to F_2 and undertakes action A would have been better off going to F_1 while still undertaking A, and vice versa.

The result also says that firm 2 gets 'pushed' into taking the other action, even if its manager thinks it should take the same action as firm 1. It thus follows that firm 2 might be better off hiring a manager with the opposite vision of firm 1, or one whose vision is still stronger. This raises the broader issue how firms will compete with and on vision, a topic of further research[14].

Note, finally, that there is an implicit assumption in this model that firms are not limited in size: they hire any employee that comes their way. In the presence of many candidate-employees, this leads to the rather surprising result that the more visionary firm tends to be larger and have employees with more diverse beliefs. In reality, however, firms are not flexible in terms of their size. Taking into account such limitations would largely eliminate these results. They also tend to disappear as the number of project types increases.

Corollary 2 above combines nicely with the results of section 3.4. There we concluded that an increase in vision could *de*motivate and reduce employees' utility, that is, if they favored the other action so strongly as to go against the manager's opinion. The corollary, however, implies that this negative effect does not apply to the more visionary of the two firms. Since all its employees choose according to the manager's vision, they also get motivated by stronger vision.

Corollary 3 *If M_1 has the stronger belief ($\nu_{M_1} > \nu_{M_2}$) then the effort and utility of F_1's employees increase in $\mu_{M_1,A}$ and thus in ν_{M_1}.*

[14] We conjecture that such competition leads to more extreme visions since firms have an incentive to outbid each other. Some of that might have been going in the early history of e-commerce, with firms competing on business models.

Overall the analysis suggests the following characteristics of a 'visionary organization':

- Employees choose their initiatives without intervention from the top, but nevertheless they choose what management would want them to choose. This strengthens the case for delegation.
- Visionary firms also attract employees that do not really agree with the vision, but who are attracted by its conviction.
- Vision motivates all employees, including those who actually think the other project would be better.

3.6 Profitability of vision

The analysis thus far has uncovered both beneficial and harmful effects of vision. Can we say anything about when a company gains from hiring a CEO with vision? In line with our discussion in section 3.3 on the outsider's perspective, we consider here the question *'Given some reference p, where we assume $1 > p > 1/2$, is the optimal belief of the firm $\mu_{M,A} > p$?'*.

According to the analysis up to this point, the optimal CEO-belief depends on the following forces:

- The motivation/demotivation effect.
- The sorting effect and the influence on the project choice.
- The cost of wrong implementation decisions.

The following subsections consider how these effects combine in specific cases. The conclusions are as follows:

- Absent sorting, no conclusions can be drawn in full generality, though we do obtain clear results for a more restricted but important class of belief distributions.
- When sorting occurs, we show that at least some degree of vision is optimal under very weak conditions.
- Even when vision is not optimal ex-ante, it might seem optimal ex-post. This spurious optimality result suggests some caveats for 'In Search of Excellence'-type of analyses.
- The impact of vision increases in the importance of motivation and initiative, but decreases as the uncertainty goes to zero.

3.6.1 Profitability of vision absent sorting

Consider first the case without sorting. With employees' beliefs drawn from a distribution F, the firm's reference expected profits can be written[15]:

$$\begin{aligned} E[\pi] &= \int_0^{\mu_{M,B}} q(\hat{e})\gamma_M^2 \mu_{M,B} \left((1-p) - \frac{\mu_{M,B}}{2}\right) f(u)du \\ &+ \int_{\mu_{M,B}}^1 q(\hat{e})\gamma_M^2 \mu_{M,A} \left(p - \frac{\mu_{M,A}}{2}\right) f(u)du \end{aligned}$$

Since the balance of forces depends on the distribution of beliefs, we cannot say anything in full generality. There exists, however, an important class of distributions that does allow clear conclusions. Consider in particular the following restriction:

Assumption 4 *All agents think A is the optimal project, i.e.* $\operatorname{supp} F \subset (1/2, 1]$.

[15] Remember that hats indicate optimal choice variables and value functions, and that the dependence of maximizers on parameters has been suppressed. In particular, $\hat{e}$ is function of the type of action taken, γ_E, γ_M, $\mu_{M,A}$ and u.

This assumption will, for example, be satisfied when all employees approximately hold the reference belief. It eliminates all employees who get demotivated or switch actions as the manager gets more convinced. The remaining trade-off is then between the motivation effect and the cost of wrong implementation.

Proposition 4 *Let A4 hold. If* $q(e) \equiv 1$*, then the unique optimal belief is the reference (or 'objective') belief. If* $q(e)$ *is strictly increasing, then vision is strictly optimal.*

The intuition is simple. As long as there is some effect of effort, the motivation effect dominates, since the effect of wrong implementations is second order at $\mu_{M,A} = p$. When the motivation effect is completely absent, then the cost of making wrong decisions will make it optimal to hold the reference belief. Note that this proposition is a partial exception on the general assumption that $q(e)$ is strictly increasing in e.

3.6.2 Profitability of vision with sorting

When sorting occurs, an important cost of vision gets eliminated for the most visionary firm: no employee will get demotivated by the manager's vision. Moreover, at small levels of overconfidence the cost of wrong implementations is second order since it concerns only projects that go marginally the other way. This suggests that 'vision is always good in moderate amounts'. There is still one caveat, however: it is theoretically possible that all potential employees hold beliefs opposite to the reference belief p. A visionary firm ($\mu_M > p$) could then end up with nearly no employees and thus nearly no profits.

To formalize this argument, let the focal firm face one competitor whose manager holds the reference belief p. Consider any of the following two conditions.

Condition 1 *The support of* F *is contained in* $[(1-p), 1]$.

or

Condition 2 *The distribution of beliefs F First Order Stochastically Dominates some symmetric distribution*[16] *and* $1/2 < p < 1$.

This second condition says that the distribution of beliefs weakly favors the side of the reference belief, in the sense that it can be generated from some symmetric distribution by moving some probability mass up. This holds for example when $F(x) \leq 1 - F(1-x)$ or when F is the Beta-distribution $F(x; a, b) = \frac{\int_0^x u^{a-1}(1-u)^{b-1} du}{\int_0^1 u^{a-1}(1-u)^{b-1} du}$ with $0 < b \leq a < \infty$.

The following results confirm that vision is optimal under fairly weak conditions.

Proposition 5 *Under C2 or C1 vision is optimal (against a firm whose manager holds the reference belief).*

Note, however, that this answer is incomplete since we constrained the other firm to hold the reference belief.

3.6.3 Spurious (ex-post) optimality of vision

The fact that many successful firms have visionary CEO's or strong organizational beliefs might be taken as casual evidence for the optimality of vision. Looks may deceive, however. In particular, vision and strong beliefs induce an important selection bias. If you act as if you knew the future and you turn out to be right, then your actions will be ex-post optimal, even if they were ex-ante suboptimal given the objective odds. So we would expect that even when vision is *not* optimal, ex-post the best (but also worst) firms in the market will be those with visionary managers.

To confirm this argument formally, consider an economy with N firms with K employees each. Each employee of firm n faces a choice of action $X_n \in \{A_n, B_n\}$. The state of the world relevant to firm n is an independent draw x_n from $\{a_n, b_n\}$

[16]A distribution F first-order stochastically dominates a distribution G when F is generated from G by adding to every outcome some non-negative random variable. An alternative definition is that $F \leq G$, i.e. some probability mass of G is shifted upwards to obtain F.

with probabilities p and $(1-p)$ respectively, where we assume $1 > p > 1/2$. All employees hold the reference belief $\mu_{E} = p$, which implies A4. Let $q(e) \equiv 1$, so that the reference belief ('no vision') is optimal by proposition 4. The managers' beliefs are independent draws from a distribution of beliefs F with support $[p, 1]$ and with an atom of size $0 < P[p] < 1$ at the endpoint p. Any such draw thus results with probability $P[p]$ in an 'objective' manager. With probability $1 - P[p]$, the draw will be a 'visionary' manager with a belief $\mu_{v} > p$[17]. Assume that the firms face equivalent opportunities: the implementation cost I_k of the k^{th} employee's project is identical for all firms. The following proposition confirms that visionary firms will have extreme results:

Proposition 6 *As the number of firms N and employees per firm K increases, the probability that the best (and worst) firms have visionary managers (as indicated in figure 3.3) and that the profit difference with any firm with an objective manager is strict, converges to one. The probability of being ex-post the best (or worst) performing firm increases in the firm's rank*[18] *in terms of strength of its manager's belief.*

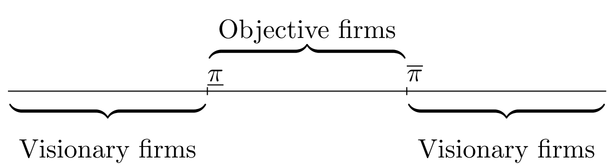

Figure 3.3: The extreme performance of visionary companies

The intuition is exactly the one set forth at the start of this section, and the result confirms essentially the initial conjecture.

[17]Note that, for the sake of getting a simple argument, we implicitly assume that no sorting takes place.

[18]We define rank here as '#firms that have strictly stronger belief + 1'. So the firm with the strongest belief has rank 1, and a firm with rank m has $m-1$ firms that have strictly stronger beliefs. Other definitions are just a matter of changing notation.

Corollary 4 (In Search of Excellence) [19] *For a large enough number of firms and employees per firm, the very best firms in the market have (nearly always) visionary managers.*

This might also explain the observation that many famous 'visionary' managers were actually founders or co-founders of their firm (e.g. Steve Jobs, Sam Walton, Bill Gates, Larry Ellison, Scott McNealy). In particular, the theory here suggests that these people might actually have had too strong beliefs (from an ex ante expected profitability perspective) but turned out to be right. Note also that such extreme believers are willing to spend extreme effort on developing their ideas.

This spurious effect will be stronger as there is more underlying uncertainty (which might explain why 4 out of the 5 names above come from the software sector):

Proposition 7 *The difference in ex-post profitability between the firm of the most visionary manager ($\mu_v = 1$) and that of the closest objective manager increases in the 'objective' uncertainty*[20] $p(1-p)$.

The intuition is simply that objective managers are very cautious in markets with high uncertainty. There is thus much more room for overconfidence to make a difference.

3.6.4 Comparative statics

Which technological and market conditions make vision more or less important? While we can try to answer this question, the results should be treated with great

[19] Although 'In Search of Excellence' (1982) does not refer to it as 'vision', it does conclude that excellent companies are characterized by strong beliefs and values (e.g. 'a belief that most members of the organization should be innovators...') and argues that these values and beliefs are often created by a leader. This is not to say that their results were all spurious. But the effect may have played an important role.

[20] Note that $p(1-p)$ is the variance of a binomial distribution with probability p.

caution. The assumption that only one of the firms can choose its vision seriously affects any comparative statics. A more complete analysis awaits further research.

Consider first how the impact of vision depends on the underlying uncertainty. The most natural measure for 'uncertainty' is $p(1-p)$, the variance of the binomial distribution generated by the reference probability. The basic conjecture is that the impact of vision should decrease as the uncertainty about the true state goes to zero. The argument is simply that there is less room for a manager to be overconfident, and thus for vision to make a difference.

While this intuition is complete in the absence of sorting, things are a bit more complex with sorting. In this case, the overall gain has two components. The first is the gain from inducing sorting with a minimum (limit) deviation from the reference belief, which we call the pure sorting effect. The second is the extra gain from holding a belief that is strictly greater than p. We call this the gain beyond the pure sorting effect. The suggested intuition applies only to the latter.

Proposition 8

- *Absent sorting, the profit gain from vision, if any, converges to zero as $p \to 1$. Formally $\left[\max_{\mu_{M,A} \geq p} E[\pi] - E[\pi \mid \mu_{M,A} = p]\right] \to 0$ as $p \to 1$.*
- *Under C1 or C2 and sorting, the profit gain from vision beyond pure sorting converges to zero as $p \to 1$. Formally $\left[\max_{\mu_{M,A} \geq p} E[\pi] - \lim_{\mu_{M,A} \downarrow p} E[\pi]\right] \to 0$ as $p \to 1$.*

In contrast to proposition 7, the effect here is a real decrease in ex-ante expected profit, instead of a spurious ex-post effect.

The role of 'motivation' in the model also suggests that vision will be more important in sectors where individual non-contractible effort is more important. The problem is to capture the notion of 'effort being more important' without any side effects. We would want to parameterize $q(e)$ and $c(e)$ by η such that $\frac{\partial^2 q(e)}{\partial \eta \partial e} \geq 0$ while at the same time $\hat{e}$ is independent of η and $\left[\frac{\partial q(e)}{\partial \eta}\right]_{e=\hat{e}} = 0$ so as to eliminate indirect effects. In that case, we indeed get

Proposition 9 *Under C1 or C2 and sorting, or under A4 absent sorting, the optimal vision increases as η increases.*

While this confirms the conjecture in principle, it is not clear which practical parameterization would have these properties. On the other hand, the result itself does hold for some common parameterizations with only slightly stronger conditions.

3.7 Implications for theories of organizational culture and business strategy

We now consider some implications for corporate culture and strategy, which are both closely related to organizational beliefs and vision.

Culture After some modification, the model allows an interpretation in terms of differing utility functions instead of differing beliefs. For example, someone who cares about the environment likes to work for a manager with similar preferences since he is more likely to get approval for environmentally friendly projects.As such, it suggests a theory of 'organizational values': why they matter and how they get formed. The sociological and management literature has often defined corporate culture as 'shared beliefs'or as 'shared values' (e.g. Schein 1984, 1985, Kotter and Heskett 1992), which correspond respectively to the original model and the above modification. The 'behavior norms' aspect of corporate culture is then interpreted as a reflection of underlying beliefs or values. A strong belief that 'there is one best way to do things', for example, leads some firms to value uniform practices throughout its worldwide offices, which leads to many implicit and explicit rules about 'how things are done here'. Our theory thus provides a model of corporate culture, which is complementary to existing economic theories of corporate culture (Kreps 1990, Cremer 1993, Lazear 1995, Hermalin 1999). Lazear (1995) considers how culture evolves in an organization, which is complementary to the sorting effect in this model. Cremer's (1993) definition of corporate culture as shared knowledge is closely related to a definition of culture as shared beliefs.

Strategy Rotemberg and Saloner (2000)[21] argue that strategy is a substitute for vision since it can provide similar incentives by restricting which businesses the firm can be in. They further argue that vision is more effective since it allows a finer trade-off[22]. But strategy is also a complement since it can be a means to communicate the vision (Saloner, Shepard and Podolny 2000).

With respect to this interpretation, however, our theory predicts that a CEO would often like to have a stronger belief than he really has. Formulating a strategy that follows the stronger belief would be the answer, but this poses the issue of commitment. Except for hard-wiring the strategy in promotions, incentive schemes and such, the natural approach is to build a reputation for following strategic plans. This, then, would support the categorical nature of strategic plans implicitly assumed by Rotemberg and Saloner: While in principle there is no reason that strategic plans wouldn't be able to mimic the nuance of a vision, it is difficult to build a reputation for following a plan when it is very nuanced.

3.8 Conclusion

This paper argued that managerial vision and organizational beliefs are to a large extent amenable to economic analysis, and that they are an essential part of organization theory. The paper then considered some basic effects of beliefs and vision on decision making, motivation, and satisfaction and showed how vision can shape

[21]See also Zemsky (1994) on the value of intertemporal commitment provided by strategy.

[22]This assumes that strategy is a simple rule that excludes certain types of activities or projects (as it was used operationally by Rotemberg and Saloner 1994). There is in principle no reason, however, why strategy should be so categorical. On the other hand, part of the argument we are about to make is precisely why more nuanced strategies might be very difficult to implement.

organizational beliefs. It finally concluded that vision is profitable under weak conditions, but also identified an important spurious effect that may make vision look better than it really is.

There are many interesting extensions to this work:

- The coordinating role of vision has been left unexplored, although the analysis suggests some possible mechanisms. Since all employees of a visionary firm have similar ideas about the future, they will tend to act in mutually consistent ways. Employees also tend to shade their decisions towards the belief of their manager, making the latter an implicit coordination point. The motivation effect, finally, means that more effort is spent on projects that align with the CEO's vision.

- The current analysis is essentially static. How vision interacts with learning or how a CEO will choose his successor are interesting dynamic issues that are left to explore. The issues of communication, influence, and conviction are also completely absent from this model. Lazear (1995) presented some results in this sense.

- The firms in this model focus all their energy single-mindedly on one course of action[23]. This raises the question under which circumstances the firm would do better to hedge its bets by spending part of its resources on the other option. This is obviously related to the issues of diversity (Athey et al. 2000) and autonomous strategic action by middle management (Burgelman 1983, 1991, 1994, Rotemberg and Saloner 2000).

While we cited already some empirical and 'casual' evidence that supports the theory, real testing remains to be done. The most effective method would probably

[23] This is not necessarily the same as 'exploitation' in the sense of March (1991) or the absence of innovation. In particular, the manager's vision can focus the firm's actions on innovation and exploration, at the cost of exploitation. It is plausible, however, that vision often leads to exploitation at the cost of exploration.

be experimental. This is facilitated by the fact that only the employees' perceptions matter, so that individual experiments suffice. A different approach consists of testing how employee motivation and satisfaction, and firm hiring and firing are affected by the fit in beliefs between the employee and the organization or manager. A direct empirical test of the vision-performance relationship is complicated by its non-monotone form. Testing the second moment prediction (that more visionary firms have more variation in their results) might be more powerful.

From a broader perspective, we think that economic theory has much to gain from studying the consequences of belief differences.

3.A Vision, selective hiring, and size

Sorting can also occur because the firm hires selectively. Consider for example a firm that is alone in a market with N potential employees. Let the employees' beliefs be independent draws from some distribution F. Let the firm have a limited number of $K < N$ positions to fill. Assume in particular that if $n > K$ potential employees accept a wage offer from the firm, then the firm can choose which K employees out of that group it really hires. The following result essentially says that smaller firms tend to have stronger beliefs, higher motivation and satisfaction, and higher expected profits per employee (after correcting for other size related effects).

Proposition 1 *For a given firm-belief* $\mu_M > 1/2$*, fix any* $\epsilon > 0$ *and consider for each number of employees* N *the class of firms with size* $K < \left[1 - F\left(\frac{1-\mu_M}{\mu_M}\right) - \epsilon\right] N$*. Let, for each* N*,* P_N *denote the minimal ex ante probability that for any two firms with sizes* $0 < K_1 < K_2 \leq K$*, the smaller firm has stronger average employee beliefs, higher average effort and satisfaction, and higher expected profits per employee than the larger one. Then* $P_N \to 1$ *as* $N \to \infty$*.*

Moreover, for very small firms, the manager's optimal vision can be weak compared to the beliefs of his employees. In any firm with $K < \left[1 - F(\hat{\mu}_{M,A})\right] N$, all employees will have strictly stronger beliefs than the (optimal) manager. The manager thus plays a bit the 'voice of reason', although he is still overconfident.

It should be noted, though, that these results are very sensitive to the particular assumptions made about the presence of other firms and the ensuing sorting process in the market.

Proof of Proposition 1: We first of all claim that, as $N \to \infty$, there are almost surely at least K employees with belief $\mu_E \geq \frac{1-\mu_{M,A}}{\mu_{M,A}}$. With F_N denoting the empirical distribution of a draw of N employees, we need to show that almost surely $\frac{K}{N} < 1 - F_N\left(\frac{1-\mu_{M,A}}{\mu_{M,A}}\right)$. Since we know that $\frac{K}{N} < 1 - F\left(\frac{1-\mu_{M,A}}{\mu_{M,A}}\right) - \epsilon$ it suffices that, as $N \to \infty$, $\left|F\left(\frac{1-\mu_{M,A}}{\mu_{M,A}}\right) - F_N\left(\frac{1-\mu_{M,A}}{\mu_{M,A}}\right)\right| \leq \frac{\epsilon}{2}$ which follows from the

Glivenko-Cantelli theorem. This allows us to condition the rest of the argument on the event that there are at least K employees with belief $\mu_E \geq \frac{1-\mu_{M,A}}{\mu_{M,A}}$.

We now claim that, conditional on that fact, the firm hires the K employees with the strongest beliefs in A. Let, in abuse of notation, K denote the set of employees hired by the firm and K_L and K_H respectively the subsets of K with beliefs $\mu_E < 1-\mu_{M,A}$ and $\mu_E \geq 1-\mu_{M,A}$, i.e. $K_L = K \cap [0, 1-\mu_{M,A})$ and $K_H = K \cap [1-\mu_{M,A}, 1]$. The firm's profit (from its own subjective perspective) can be written

$$\sum_{K_L} q(\hat{e})\gamma_F^2 \frac{(1-\mu_{M,A})^2}{2} + \sum_{K_H} q(\hat{e})\gamma_F^2 \frac{\mu_{M,A}^2}{2}$$

We now claim that this is maximized when K_L is empty and all employees in K_H have belief $\mu_E \geq F_N^{-1}(N-K)$. In this case, all hired employees undertake A-projects. They also all have beliefs $\mu_E \geq \frac{1-\mu_{M,A}}{\mu_{M,A}}$, so that $\mu_E \mu_{M,A} \geq 1-\mu_{M,A}$ which implies that each hired employee will put in more effort than any non-hired potential employee. Combined with the fact that $\frac{(1-\mu_{M,A})^2}{2} < \frac{\mu_{M,A}^2}{2}$ this implies that we can never be better off by hiring an employee that would undertake B. And for all employees undertaking A, the firm prefers to hire those with the strongest beliefs in A.

Take now any firm with size $K < \left(1 - F(\frac{1-\mu_{M,A}}{\mu_{M,A}}) - \epsilon\right) N - 1$ (and condition on the fact that there are at least $K+1$ employees with belief $\mu_E \geq \frac{1-\mu_{M,A}}{\mu_{M,A}})$). Consider what happens when this firm hires one more employee. That extra employee will have a weakly weaker belief (in A) than any other employee of the firm, so that the median and average belief weakly decrease. The other results follow analogously. ■

3.B Modifications of model and assumptions

This appendix considers how the results are affected by changes in assumptions or in setup.

3.B.1 Contractibility and renegotiation assumptions

Consider first what happens when effort e would become contractible. Assume in particular that, after the employee has accepted the wage offer $\tilde{w}$, the firm can offer an extra effort-based compensation $b(e)$. If the employee rejects, the game just proceeds as before. If the employee accepts, this effort compensation becomes non-renegotiable (while the wage $\tilde{w}$ remains renegotiable)[24]. The following informal argument suggests that all qualitative results are preserved in this case. Let $\hat{e}$ denote the effort that the employee would choose absent any extra compensation scheme. Any compensation scheme $b(e)$ can be replicated by one that induces the same effort, say $\tilde{e}$, and that simply consists of a bonus $\tilde{b} = b(\tilde{e})$ if and only if the employee chooses $e = \tilde{e}$. This bonus must be non-negative (since the employee can always reject $b(e)$ and choose $\tilde{e}$ anyways). It is also straightforward that we must have $\tilde{e} \geq \hat{e}$ (since the firm will never pay anything extra for a lower effort) and $\tilde{b} = [q(\hat{e}) - q(\tilde{e})]\mu_{M,X}\mu_{E,X}\gamma_M\gamma_E - [c(\hat{e}) - c(\tilde{e})]$ (since this is the minimum that the firm has to offer to make the employee willing to choose $\tilde{e}$). It now follows already that the employee's project choice and utility is the same as in the original game. The satisfaction and sorting effects are thus preserved. Moreover, the effort will be strictly larger than before and moves with the manager's ν_i as in the original game.

If instead of effort, we made the project type contractible in the way we just described, then the qualitative effects would again be preserved. The choice of project type will still be influenced by both beliefs although the manager's belief will get more weight. The employee's motivation and satisfaction will still depend on his own and his manager's belief in the action undertaken. So we also get sorting. The case where both project type and effort are contractible in the way described

[24]This corresponds to a situation where $b(e)$ is paid immediately after the effort is spent, while the wage w is paid only at the end of the game.

Note that there are numerous alternatives for modifying the game and that our only goal is to clarify the role of the contractibility assumptions. Therefore, we limit ourselves to some direct modifications and are not too concerned about the realism of the resulting game.

is essentially a combination of both cases, so that we would expect the qualitative results to be again preserved.

A second issue is the non-contractibility of the agent's participation, which leads to the ex-post renegotiation. We noted already that the results extend to the case where the employee gets instead some exogenously determined benefit, such as improved outside options or satisfaction from a successful project. A very different case, however, is that where the firm can make an up-front offer of wage plus bonus, which are then non-renegotiable. The choice of the optimal bonus introduces a second optimization problem in the game, which complicates the analysis. While the original results seem to hold under appropriate restrictions on the third derivatives, a full analysis of this case awaits further research. Alternatively, this game could be simplified by assuming that the size of the bonus is exogenously given, but this brings us back to the above model with exogenously determined benefits.

3.B.2 Other modifications

Consider now some more structural changes to the model. A first important modification is the timing of the renegotiation. We could for example imagine that the firm and the employee renegotiate at the time of implementation (i.e. that the employee's support is critical for implementation). It can be shown that the employee will then undertake the project that the manager considers best, that he spends more effort as the manager has stronger belief, and that he gets the higher expected utility from working for the manager with the stronger beliefs. It thus also follows that vision is optimal. The key change, however, is that the sorting is not based any more on the employee's beliefs (since his wage gets fixed before the project gets realized).

A different set of modifications pertains to the role of employee effort e. In particular, in the model employee utility was strictly increasing and supermodular in e, $\mu_{M,Y}$ and $\mu_{E,Y}$. While this appears to be the more natural case, these properties do not necessarily always hold in modified games[25]. The property that the employees' utility is increasing in the manager's belief in the project he undertakes, tends to hold in most situations. In that case, vision still causes sorting and an increase in satisfaction. The complementarity between e and $\mu_{M,Y}$ however, is more fragile. In some situations, the motivation effect may get lost or even reversed. If so, the optimality of vision depends on the exact strength and interaction of the different effects.

Finally, one might wonder about the impact of the allocation of authority. We consider two cases of interest. First, if the manager were to choose the type of project (while the employee still chooses his own effort level), his criterion would put strictly more weight on his own beliefs. Second, the case in which the employee makes the implementation decision is identical to a situation where the manager happens to have the same belief as the employee. In particular, the analysis implies that the firm would want to hire overconfident employees.

[25]Consider, for example, the following modification. Let the cost of implementation be distributed according to some general distribution function G. Let $q(e)$ denote the probability that the employee's project will be a success conditional on being of the right type (i.e. conditional on fitting the state), instead of the probability that the employee comes up with a proposal. In this case, the employee's overall utility function becomes $\gamma_E \mu_{E,Y} q(e) G(\gamma_M \mu_{M,Y} q(e))$. Complementarity between $\mu_{M,Y}$ and e now depends on the behavior of g'. Another possible modification is that where the effort e is expended after the project is approved (with $q(e)$ then being the probability of success conditional on being of the right type). In this case, there will be no interaction between e and $\mu_{M,Y}$.

3.C Proofs of the propositions

3.C.1 Basic results for one firm

Lemma 1 *All subgame perfect equilibria of this game have the same project types, effort levels and payoffs, and are equivalent to one in which the firm offers* $\tilde{w} = 0$ *and the employee accepts.*

Proof : Let us first determine the full equilibrium by backwards induction. Assume that the firm has made a wage-offer $\tilde{w}$ and the employee has accepted. Let $\breve{w} = \max(\tilde{w}, 0)$. The outside options in the bargaining are $\breve{w}$ for the employee and $-\breve{w}$ for the firm. These are also the final payoffs in case there is no successful project at the start of the renegotiation stage. Furthermore, Nash bargaining when there *is* a successful project will give the employee $w = \gamma_E + \breve{w}$. The employee will thus always ask for renegotiation and end up with this wage while the firm gets $\gamma_M - \breve{w}$. Consider now the firm's decision when it gets a proposal for a project of type X. The firm implements iff $\mu_{M,X}(\gamma_M - \breve{w}) + (1 - \mu_{M,X})(-\breve{w}) - I (= \mu_{M,X}\gamma_M - \breve{w} - I) \geq -\breve{w}$ or iff $\mu_{M,X}\gamma_M \geq I$ i.e. with probability $\mu_{M,X}\gamma_M$.
The employee's payoff upon generating a project is $\mu_{E,X}\gamma_M\mu_{M,X}(\gamma_E + \breve{w}) + (1 - (\mu_{E,X}\gamma_M\mu_{M,X}))\breve{w} = \breve{w} + \gamma_E\gamma_M\mu_{M,X}\mu_{E,X}$ while it is just $\breve{w}$ without any project. So the employee solves $\max_{e \in \mathcal{E}, Y \in \{A,B\}} \breve{w} + q(e)\gamma_E\gamma_M\mu_{M,Y}\mu_{E,Y} - c(e)$. It follows that the employee chooses the project, say X, with the highest $\mu_{E,Y}\mu_{M,Y}$, and then chooses e to solve: $\max_{e \in \mathcal{E}} q(e)\gamma_E\gamma_M\mu_{M,X}\mu_{E,X} - c(e) + \breve{w}$. This is non-negative by A2 and the fact that the employee can set $e = 0$. Since $\underline{w} = 0$, the employee accepts any $\tilde{w}$.
The firm's payoff from offering a wage $\tilde{w}$ is

$$q(\hat{e}) \int_0^{\gamma_M \mu_{M,X}} (\gamma_M \mu_{M,X} - I) dI - \max(\tilde{w}, 0) = q(\hat{e}) \frac{\gamma_M^2 \mu_{M,X}^2}{2} - \max(\tilde{w}, 0)$$

so the firm offers $\tilde{w} \leq 0$, so that $\breve{w} = 0$. Any such wage gives the same payoff. ■

Proof of Proposition 1: Lemma 1 says that the employee will choose the action $Y \in \{A, B\}$ with the highest $\mu_{E,Y}\mu_{M,Y}$. Let $\nu_i > \nu_j$ and let wlog. $A = \text{argmax}_{Y\in\{A,B\}}\, \mu_{i,Y}$. If also $\text{argmax}_{Y\in\{A,B\}}\, \mu_{j,Y} = A$ then $\mu_{E,A}\mu_{M,A} = \nu_i\nu_j > 1/4 > (1-\nu_i)(1-\nu_j) = \mu_{E,B}\mu_{M,B}$, else $\mu_{E,A}\mu_{M,A} = \nu_i(1-\nu_j) > (1-\nu_i)\nu_j = \mu_{E,B}\mu_{M,B}$. In any case, the employee chooses indeed the action preferred by i. If $\nu_M = \nu_E$ and $\mu_M \neq \mu_E$, then, by A1, the employee does as his manager prefers. ■

Proof of Corollary 1: Let, essentially wlog, $\mu_{M,A} > 1/2$, so that $\nu_M = \mu_{M,A}$. The probability that the decision follows the manager's belief is $\int_{1-\mu_{M,A}}^{1} dF$ which increases in $\mu_{M,A}$ and thus in ν_M. ■

Proof of Proposition 2: We first show that, with X denoting the project undertaken by the employee, '$\hat{e}$ and $\hat{u}$ strictly increase in $\mu_{i,X}$ on N'. With $l(e) = \frac{c'(e)}{q'(e)}$, we have that $\hat{e} = l^{-1}(\gamma_E\gamma_M\mu_{E,X}\mu_{M,X})$ so that $\frac{d\hat{e}}{d\mu_{i,X}} = [l^{-1}(\cdot)]'\gamma_E\gamma_M\mu_{-i,X}$ which is strictly positive. This implies the first part of the statement. The second part follows from applying an envelope theorem on the employee's problem $\max_{e\in\mathcal{E}} q(e)\gamma_E\gamma_M\mu_{E,X}\mu_{M,X} - c(e)$.

Assume now that the manager strictly prefers project A, i.e. $\mu_{M,A} > 1/2$, so that $\nu_M = \mu_{M,A}$. If now $X = A$ then $\frac{d\hat{e}}{d\nu_M} = \frac{d\hat{e}}{d\mu_{M,A}} = \frac{d\hat{e}}{d\mu_{M,X}} > 0$. If $X = B$, then $\mu_{M,X} = \mu_{M,B} = 1 - \mu_{M,A} = 1 - \nu_M$, so that $\frac{d\hat{e}}{d\nu_M} = \frac{d\hat{e}}{d\mu_{M,A}} = -\frac{d\hat{e}}{d\mu_{M,X}} < 0$.

The arguments for increases and decreases in utility, and for the analogous relationships with respect to the employee's conviction and project preference, are completely analogous. ■

3.C.2 Sorting

Proof of Proposition 3: Remember that an employee of firm F_i who undertakes Y gets a payoff $q(\hat{e})\gamma_E\gamma_M\mu_{E,Y}\mu_{M_i,Y} - c(\hat{e})$.

We claim first of all that in any SPE, all employees (with $\mu_E \neq \breve{\mu}$) hired by F_1 choose A and all those hired by F_2 choose B. This follows by contradiction: Consider any employee who applies to F_1 but chooses action B. He would be strictly better off applying to F_2 and still undertaking B .

Next, given that F_1 (resp. F_2)-employees choose A (resp. B), an employee strictly prefers F_1 if $\max_{e\in\mathcal{E}} q(e)\gamma_E\gamma_M\mu_{E,A}\mu_{M_1,A} - c(e) > \max_{e\in\mathcal{E}} q(e)\gamma_E\gamma_M\mu_{E,B}\mu_{M_2,B} - c(e)$ or if (by an envelope theorem argument) $\mu_{E,A}\mu_{M_1,A} > \mu_{E,B}\mu_{M_2,B}$ or if $\mu_{E,A} > \breve{\mu}$.

An analogous argument shows that if $\mu_E < \breve{\mu}$ then the employee will definitely choose firm F_2. The fact that $\breve{\mu}$ decreases in $\mu_{M_1,A}$ and $\mu_{M_2,A}$ follows from its definition. ■

Proof of Corollary 2: This follows directly from the proof of proposition 3. ■

Proof of Corollary 3: By the earlier results and assumptions, all employees of F_1 choose A. The corollary then follows from monotone comparative statics and an envelope theorem on the employee's problem.

■

3.C.3 Profitability of vision

Lemma 2 *Absent sorting, the optimal $\mu_{M,A}$ increases in p.*

Proof : It is sufficient to show that $E[\hat{\pi}_O]$ is supermodular in p and $\mu_{M,A}$. The profit equation is:

$$\begin{aligned} E[\hat{\pi}_O] &= \int_0^{\mu_{M,B}} q(\hat{e})\gamma_M^2 \left((1-p)\mu_{M,B} - \frac{\mu_{M,B}^2}{2} \right) f(u)du \\ &+ \int_{\mu_{M,B}}^1 q(\hat{e})\gamma_M^2 \left(p\mu_{M,A} - \frac{\mu_{M,A}^2}{2} \right) f(u)du \end{aligned}$$

where we suppressed notation that indicates that $\hat{e}$ depends on both agents' beliefs and on the action taken. The cross partial of this function in $(p, \mu_{M,A})$ is positive. ■

Restricted belief-support

Proof of Proposition 4: We first want to show that $\hat{\mu}_{M,A} \geq 1/2$. By lemma 2 above, it is sufficient to show this for $p = 1/2$. By contradiction, assume that $\mu_{M,A} < 1/2$ while $p = 1/2$, then firm profits are:

$$E[\hat{\pi}_O] = \int_{1/2}^{\mu_{M,B}} q(\hat{e})\gamma_M^2 \mu_{M,B} \frac{\mu_{M,A}}{2} f(u)du + \int_{\mu_{M,B}}^{1} q(\hat{e})\gamma_M^2 \mu_{M,A} \left(\frac{\mu_{M,B}}{2}\right) f(u)du$$

Consider now what happens if we select instead a manager with belief $\check{\mu}_{M,A} = 1 - \mu_{M,A} > 1/2$.

- Employees who before chose A will still choose A, but their effort strictly increases. This implies that the second term strictly increases.

- Employees who before chose B will now choose A. By the relation between $\mu_{M,A}$ and $\check{\mu}_{M,A}$, the $\mu_{M,X}$ (the manager's belief in the action chosen by the employee) remains the same. $\mu_{E,X}$ on the contrary increases (since by A4 all employees believe more in A than in B), so that again employee effort increases. This implies that the first term increases.

This implies that overall the firm profits increase, so that $\mu_{M,A} < 1/2$ is not optimal. Consider now the case that $q(e) \equiv 1$. The employee sets $\hat{e} = 0$ and undertakes the action that maximizes $\mu_{E,Y}\mu_{M,Y}$. Since $\hat{\mu}_{M,A} \geq 1/2$, profit equals $E[\hat{\pi}_O] = \int_{1/2}^{1} \gamma_M^2 \mu_{M,A} \left(p - \frac{\mu_{M,A}}{2}\right) f(u)du$ which is maximized at $\hat{\mu}_{M,A} = p$. This proves the first part of the proposition. For the second part, the firm profit when $\mu_{M,A} \geq 1/2$ is $E[\hat{\pi}_O] = \max_{\mu_{M,A}} \int_{1/2}^{1} q(\hat{e})\gamma_M^2 \mu_{M,A} \left(p - \frac{\mu_{M,A}}{2}\right) f(u)du$ where the maximum is well defined since the profit function is continuous in $\mu_{M,A}$ on $[1/2, 1]$. The derivative of the integrand (for $\mu_{M,A}$) is strictly positive for $1/2 \leq \mu_{M,A} \leq p$ and continuous in $\mu_{M,A}$. It thus follows that the optimal $\mu_{M,A}$ is strictly larger than p and thus that vision is optimal. ■

3.C.4 Profitability with sorting

Remember that we assume that $1 > p > 1/2$ and that the focal firm F faces one competitor with belief $\mu = p$. We first introduce some notation. Let $\hat{\pi}_H = \max_{\mu_{F,A} \geq p} E[\pi]$ when F attracts all employees with $\mu_{E,A} \geq \breve{\mu}$, and let $\hat{\mu}_{F_H}$ be the corresponding maximizer. Let analogously $\hat{\pi}_L = \max_{\mu_{F,A} \leq p} E[\pi]$ when F attracts all employees with $\mu_{E,A} \leq \breve{\mu}$, and let $\hat{\mu}_{F_L}$ be the maximizer. Note that this implies that $0 \leq \hat{\mu}_{F_L} \leq p \leq \hat{\mu}_{F_H} \leq 1$.

Let $\tilde{\pi}_L$ be the profit of F when $\mu_{F,A} = p$ but F attracts all employees with $\mu_{E,A} < (1-p)$; $\tilde{\pi}_H$ be the profit of F when $\mu_{F,A} = p$ but F attracts all employees with $\mu_{E,A} \geq (1-p)$; $\tilde{\pi}_M$ be the profit of F when $\mu_{F,A} = p$ and employees are allocated randomly between the two firms with equal probability. Note that we always have that $\hat{\pi}_H \geq \tilde{\pi}_H$ and $\hat{\pi}_L \geq \tilde{\pi}_L$.

Finally, let $F^-(x) = \lim_{u \uparrow x} F(u)$ and $F^+(x) = \lim_{u \downarrow x} F(u)$.

Lemma 3 *If $F^-(1-p) < 1$ then $\hat{\mu}_{F_H} > p$. If $F^+(1-p) > 0$ then $\hat{\mu}_{F_L} < p$. Finally, if $F^-(1-p) < 1$ or $F^+(1-p) > 0$ then either $\hat{\pi}_L > \tilde{\pi}_M$ or $\hat{\pi}_H > \tilde{\pi}_M$ or both. If both conditions are satisfied (which is the case when F has full support), then the optimal belief is strictly different from the reference belief.*

Proof : Consider the first part of the lemma, so assume $1 - F^-(1-p) > 0$. Conditional on $\mu_{F,A} \geq p$ and F attracting all the employees with $\mu_{E,A} \geq \breve{\mu}$, its optimal profits are: $\hat{\pi}_H = \max_{\mu_{F,A}} \int_{\breve{\mu}}^1 q(\hat{e})\gamma_F^2 \left(p\mu_{F,A} - \frac{\mu_{F,A}^2}{2}\right) f(u)du$ with $\breve{\mu} = \frac{1-p}{\mu_{F,A}+1-p}$. This profit function is (right)continuously differentiable in $\mu_{F,A}$ on $[p, 1)$. Its right derivative in $\mu_{F,A}$ at $\mu_{F,A} = p$ is:

$$\left[\frac{d\hat{\pi}_H}{d\mu_{F,A}}\right]^+_{\mu_{F,A}=p} = \int_{1-p}^1 q'(\hat{e})\gamma_F^2 \frac{p^2}{2}\frac{d\hat{e}}{d\mu_{F,A}} f(u)du - q(\hat{e})\gamma_F^2 \frac{p^2}{2}\frac{d\breve{\mu}}{d\mu_{F,A}} f(1-p)$$

The second term is non-negative since $\frac{d\breve{\mu}}{d\mu_{F,A}} \leq 0$. The first term is strictly positive since $F(1-p)^- < 1$ and $\frac{d\hat{e}}{d\mu_{F,A}} > 0$. This implies that the optimal $\hat{\mu}_F > p$. Note that this also implies that $\hat{\pi}_H > \tilde{\pi}_H$.

The argument for the second part is analogous and implies $\hat{\pi}_L > \tilde{\pi}_L$.

We now show that if $F^-(1-p) < 1$ or $F^+(1-p) > 0$ then either $\hat{\pi}_L > \tilde{\pi}_M$ or $\hat{\pi}_H > \tilde{\pi}_M$ or both. Just checking definitions of $\tilde{\pi}_L$, $\tilde{\pi}_H$, and $\tilde{\pi}_M$ shows that $\tilde{\pi}_L + \tilde{\pi}_H = 2\tilde{\pi}_M$. But, we always have that $\hat{\pi}_H \geq \tilde{\pi}_H$ and $\hat{\pi}_L \geq \tilde{\pi}_L$ with one of these strict when $F^-(1-p) < 1$ or $F^+(1-p) > 0$. This implies that under that condition $\hat{\pi}_L + \hat{\pi}_H > \tilde{\pi}_L + \tilde{\pi}_H = 2\tilde{\pi}_M$ which implies that $\max(\hat{\pi}_L, \hat{\pi}_H) > \tilde{\pi}_M$.

The very last part follows from the fact that when $F^-(1-p) < 1$ *and* $F^+(1-p) > 0$ then $\hat{\mu}_{F_H} > p$ and $\hat{\mu}_{F_L} < p$. ■

Proof of Proposition 5: For C2, this follows immediately from the lemmas that follow. For C1, it is immediate that the optimal belief must be $\mu \geq p$ since a firm with $\mu < p$ has no employees. Next, there exist some $\mu > p$ that gives the focal firm higher profits than $\mu = p$ (since with $\mu > p$ all the employees prefer the focal firm, while they randomize between the two when $\mu = p$). Finally, the right-derivative (in the manager's belief) of firm profit at $\mu = p$ is strictly positive, so that the optimal belief subject to $\mu \in (p, 1]$ is well-defined. ■

Lemma 4 *Vision is optimal (against a firm with reference beliefs) for any symmetric distribution of beliefs.*

Proof : Fix a symmetric distribution of beliefs F. Note that we always have that $F^-(1-p) < 1$, so that $\hat{\mu}_{F_H} > p$.

Consider first the case that $p = 1-p = 1/2$. By symmetry we have $\hat{\pi}_H = \hat{\pi}_L$ so that vision ($\hat{\mu}_F > p$) is (weakly) optimal.

As p increases, $\hat{\pi}_H$ strictly increases since $\frac{d\hat{\pi}_H}{dp} = \frac{\partial \hat{\pi}_H}{\partial p} = \int_{\breve{\mu}}^1 q(e)\gamma_{F_H}^2 \mu_{F_H,A} f(u)du > 0$, while $\hat{\pi}_L$ (weakly) decreases since $\frac{d\hat{\pi}_L}{dp} = \frac{\partial \hat{\pi}_L}{\partial p} = -\int_0^{\breve{\mu}} q(e)\gamma_{F_L}^2 \mu_{F_L,B} f(u)du \leq 0$. This implies that for all $p > 1/2$, $\hat{\pi}_H > \hat{\pi}_L$. ■

Lemma 5 *Let G and H be distribution functions on $[a,b]$, with H FOSD G. Let $k(\theta,x) = E_{u\sim\theta H+(1-\theta)G}\left[f_3(x,u) \mid f_1(x) \le u \le f_2(x)\right]$ with $\theta \in [0,1]$, $a \le f_1 \le f_2 \le b$ and f_3 u-measurable. Let finally $K(\theta) = \max_{x\in X} k(\theta,x)$ be well-defined for $\theta \in \{0,1\}$.*

If $f_3(x,u)$ increases in u (for fixed x), then $K(1) \ge K(0)$.

Proof : Let $f_3(x,u)$ be increasing in u. Since H FOSD G, the basic theorem on FOSD says that for any fixed $x \in X$, $k(1,x) = E_{u\sim H}\left[f_3(x,u) \mid f_1(x) \le u \le f_2(x)\right] \ge E_{u\sim G}\left[f_3(x,u) \mid f_1(x) \le u \le f_2(x)\right] = k(0,x)$. Let $\hat{x}_H \in \operatorname{argmax}_{x\in X} k(1,x)$ and $\hat{x}_G \in \operatorname{argmax}_{x\in X} k(0,x)$ which exist by assumption. We then have: $K(1) = k(1,\hat{x}_H) \ge k(1,\hat{x}_G) \ge k(0,\hat{x}_G) = K(0)$ which proves the lemma. ■

Lemma 6 *If vision is optimal for some belief-distribution G, then it is optimal for any belief-distribution H that FOSD G.*

Proof : The fact that vision is optimal for some belief-distribution G implies that $\hat{\pi}_{H,G} \ge \hat{\pi}_{L,G}$ where $\hat{\pi}_{H,G} = \max_{\mu_{F_H}\ge p} \int_{\breve{\mu}_H}^{1} q(\hat{e})\gamma_F^2 \left(p\mu_{F_H} - \frac{\mu_{F_H}^2}{2}\right) g(u)du$ with $\breve{\mu}_H = \frac{1-p}{\mu_{F_H}+1-p}$ and $\hat{\pi}_{L,G} = \max_{\mu_{F_L,B}\ge 1-p} \int_0^{\breve{\mu}_L} q(\hat{e})\gamma_F^2 \left((1-p)\mu_{F_L,B} - \frac{\mu_{F_L,B}^2}{2}\right) g(u)du$ with $\breve{\mu}_L = \frac{1-\mu_{F_L}}{p+1-\mu_{F_L}}$.

Define now $\tau_H(\mu_{F_H}, p, \breve{\mu}_H, u) = q(\hat{e})\gamma_F^2 \left(p\mu_{F_H} - \frac{\mu_{F_H}^2}{2}\right)$ if $u \ge \breve{\mu}_H$ and zero otherwise. Define analogously $\tau_L(\mu_{F_L}, p, \breve{\mu}_L, u) = q(\hat{e})\gamma_F^2 \left((1-p)\mu_{F_L,B} - \frac{\mu_{F_L,B}^2}{2}\right)$ if $u \le \breve{\mu}_L$ and zero otherwise. Then we can write $\hat{\pi}_{H,G} = \max_{\mu_{F_H}} \int_0^1 \tau_H(\mu_{F_H}, p, \breve{\mu}_H, u)g(u)du$ and $\hat{\pi}_{L,G} = \max_{\mu_{F_L}} \int_0^1 \tau_L(\mu_{F_L}, p, \breve{\mu}_L, u)g(u)du$. By lemma 5 it suffices to show that τ_H increases and τ_L decreases in u, to conclude that $\hat{\pi}_{H,H} \ge \hat{\pi}_{H,G} \ge \hat{\pi}_{L,G} \ge \hat{\pi}_{L,H}$ which would imply the proposition. The rest of this proof shows that that is indeed the case.

Note, first, that the optimal μ_{F_H} and $\mu_{F_L,B}$ must be such that $\left(p\mu_{F_H} - \frac{\mu_{F_H}^2}{2}\right) > 0$ and $\left((1-p)\mu_{F_L,B} - \frac{\mu_{F_L,B}^2}{2}\right) > 0$ since otherwise profits are non-positive while, in each case, it is always possible to set $\mu_{_F} = p$, which gives strictly positive profits. But then the inequalities follow immediately: For τ_H (using the fact that $\breve{\mu}_H$ is no function of u): the derivative is zero for $u < \breve{\mu}_H$, the function makes a jump upwards at $\breve{\mu}_H$, and the derivative for $u > \breve{\mu}_H$ is $q'(\hat{e})\gamma_{_F}^2 \left(p\mu_{_{F_H}} - \frac{\mu_{F_H}^2}{2} \right) \frac{d\hat{e}}{du}$ which is positive (since $\frac{d\hat{e}}{du}$ is positive for employees who undertake A). An analogous argument for τ_L shows that it is decreasing. ■

The spurious (ex-post) optimality of vision

Proof of Proposition 6:

The probability that the manager of a randomly selected firm has belief $\mu \geq x$ for some x such that $1 > x > p$, is $1 - F(x) > 0$. That fact combined with the fact that $1 > p > 0$ implies that both the event that 'there exists some firm with belief $\mu \geq x$ which turns out to be correct about the true state of the world' and the event that 'there exists some firm with belief $\mu \geq x$ which turns out to be wrong about the true state of the world' are almost surely true in the limit as $N \to \infty$.

The difference in profit between a visionary firm with belief $\mu \geq x > p$ and an 'objective' firm with belief p that turn out to be correct equals $\sum_{k=1}^{K} I_{\{\gamma_{_M} p < I_k \leq \gamma_{_M} \mu\}} (\gamma_{_M} - I_k)$ which is almost surely strictly positive for $K \to \infty$. Analogous arguments show that there are strict differences in profitability between an objective firm that is right and one that is wrong and between an objective firm that is wrong and a (strictly) visionary firm that is wrong. Combined with the earlier conclusion, this proves the first part of the proposition.

For the last part of the proposition, consider a firm that has the m'th rank in terms of strength of belief. The probability that the firm turns out to be the weakly best performing firm equals the probability that this focal firm is correct while the $m-1$ firms with stronger beliefs turn out to be wrong, and is thus $(1-p)^{m-1}p$. The probability that it turns out to be the worst performing is analogously $p^{m-1}(1-p)$.

Both decrease in m, so that they increase as the firm is ranked higher in terms of belief strength. ■

Proof of Corollary 4: This follows from the proof of proposition 6. ■

Proof of Proposition 7: Note that, given that we assumed $p > 1/2$, it is sufficient to prove that that difference increases as p decreases to $\frac{1}{2}$.

Consider first a visionary $\mu_v = 1$ firm that turns out to be right. The objective firm that is closest in terms of profit is just one that is right. The difference in profitability is $\sum_{k=1}^{K} I_{\{\gamma_M p < I_k \leq \gamma_M\}}(\gamma_M - I_k)$ which increases as p decreases since each $\gamma_M - I_k$ term is positive and the number of terms increases as p decreases. The proof for a maximally visionary firm that turns out to be wrong is analogous. ■

3.C.5 Comparative statics

Proof of Proposition 8: For the first part of the proposition, note that with $\hat{\mu}_M = \text{argmax}_{\mu_{M,A} \geq p}(E[\pi])$, we have $p \leq \hat{\mu}_M \leq 1$. Clearly, as $p \uparrow 1$, $\hat{\mu}_M \uparrow 1$. This combined with the continuity of the expected profit $E[\hat{\pi}_O]$, implies the proposition.

For the second part of the proposition, note that vision is optimal so that $p < \hat{\mu}_M \leq 1$. Clearly, as $p \uparrow 1$, $\hat{\mu}_M \to 1$. But this, combined with continuity of the profit function $E[\hat{\pi}_O]$, implies the proposition. ■

Proof of Proposition 9: Consider first the case under C1 or C2 and sorting. By the earlier proposition, vision is optimal. So the profit must be: $E[\hat{\pi}_O] = \int_{\breve{\mu}}^{1} q(\hat{e})\gamma_M^2 \mu_{M,A} \left(p - \frac{\mu_{M,A}}{2}\right) f(u)du$ so that the cross partial

$$\frac{\partial^2 E[\hat{\pi}_O]}{\partial \mu_{M,A} \partial \eta} = \int_{\breve{\mu}}^{1} \frac{\partial^2 q(e)}{\partial e \partial \eta} \gamma_M^2 \mu_{M,A} \left(p - \frac{\mu_{M,A}}{2}\right) \frac{d\hat{e}}{d\mu_{M,A}} f(u)du$$

is positive. The proof for the case under A4 without sorting is analogous. ■

Chapter 4

Skill or Luck*

As always, victory finds a hundred fathers but defeat is an orphan.

Count Galeazzo Ciano

4.1 Introduction

More than 80% of US drivers consider themselves better than the median (Svenson 1981). People tend to attribute success to their own skills and failures to bad luck (Zuckerman 1979). When asked about the percentage of the household work they perform, a husband's and wife's percentages typically sum up to more than 100 % (Ross and Sicoly 1979).

Collectively irrational outcomes sometimes result from perfectly rational individual actions. In particular, this paper considers how rational individuals might come to such inconsistent conclusions. The mechanism is simply that each agent selects actions based on his beliefs about what is important and what is likely to succeed, and evaluates his own and others' actions using those same beliefs. This double use of the same set of criteria favors the focal agent's actions in the evaluation stage. In a probabilistic context when agents choose projects based on their likelihood of

*This chapter benefitted a lot from my discussion with Max Bazerman. The usual disclaimer applies.

success, this translates into an 'endogenous overconfidence effect'. If agents disagree on the probability of success of alternative actions and each agent selects the action that he considers most likely to succeed, then each agent will tend to select precisely the actions about which he is overconfident relative to the rest of the population. Random errors combined with systematic choice thus lead to systematic errors. According to the others, the focal agent will then also overestimate his own expected contribution or his share in the overall result.

Note that in this argument we allowed agents to (knowingly) entertain differing beliefs, i.e. to start with differing priors[1]. While many of the results are therefore stated most naturally in subjective terms, we will sometimes translate the results to objective conclusions by designating one set of beliefs as the 'reference' or 'objective' beliefs.

This topic of self-serving, ego-centric and related biases has a long tradition in the psychology literature. Most of that literature interprets these as motivational biases: people hold unrealistically positive views of themselves since doing so increases their happiness and general well-being. This is essentially a theory of wishful thinking. Some authors have partially challenged that view by forwarding cognitive explanations, in particular for the self-serving bias. Miller and Ross (1975), for example, cite experimental evidence that people are more likely to accept responsibility for expected outcomes and that people generally expect to succeed. They note that these two combine to a self-serving bias.

This paper takes this a step further. It considers how these phenomena might come about in a pure Bayesian-rational framework. For one thing, the model thus provides rational explanations for the observed cognitive behavior cited above[2]. But

[1]For a more extensive discussion of the use of differing priors and the issues this raises, see chapter 2. Note that common knowledge of differing beliefs implies that agents will not update their beliefs merely because they see someone else do something that they themselves consider suboptimal. The end of section 4.4 discusses in how far the results of this paper can be obtained under the CPA.

[2]The sources they cite indicate that Miller and Ross (1975) did not consider these results to be the outcome of Bayesian rational decision making.

it also gives new predictions, such as the endogenous overconfidence effect. Furthermore, we show how the results extend to biases other than the self-serving bias, and how the different biases interact. Finally, the formal modeling also allows for comparative statics. One of these suggests that the typical experiment might actually underestimate the importance of these biases in everyday life.

It is definitely not the intent of this paper to suggest that all apparent biases can be explained in Bayesian rational terms. On the contrary, I personally believe motivational or cognitive biases to be real and important, and some of the experiments cannot be explained by the mechanisms of this paper. But this does not deny the value of studying mechanisms that can have the same effect. In particular, the mechanisms we discuss seem to be very robust. This kind of studies then allows us to predict at least part of the bias and how it is affected by, say, the structure of the task.

While economics has studied the role of self-serving biases in bargaining and litigation (Priest and Klein 1984, Babcock Wang et al. 1997, Babcock Loewenstein et al. 1995), research on the potential causes of such bias and implications in other areas is limited. Zabojnik (2000) presents a model in which people can learn about their abilities at a cost and concludes that in the limit a proportion of the population ends up being overconfident about their abilities. There will, however, also be a proportion that is under-confident[3]. The current paper considers a very different situation in which all agents will almost surely be overconfident. The focus of this paper is also more on other biases, especially the biased attributions of skill and luck, and how they affect inference and learning. There is also a small but growing literature that shows how certain forms of bounded rationality, such as selective recall, can lead to some forms of overconfidence. Benabou and Tirole (1999),

[3]All agents in that model have either high or low ability. No agent will ever be completely sure that he or she has high ability. It follows that all high ability agents are underconfident about their ability. For a more formal argument, note that the beliefs in the model follow a (stopped) martingale.

for example, show how self-serving biases can arise if people have (endogenously chosen) imperfect recall and the recalled information is always bad news. Rabin and Schrag (1999) show how a confirmatory bias may lead to overconfidence. The mechanism underlying much of this paper also bears some similarity to the winner's curse (Capen et al. 1971, Thaler 1988), regression towards the mean (Kahneman and Tversky 1973), decision regret (Harrison and March 1984), and Lazear's (2001) explanation for the Peter Principle (Peter and Hull 1969).

The next section starts the analysis with a very simple model in which each person considers himself the best of the population. Section 3 studies a model in which rational people show a self-serving bias[4], and then considers its implications for learning. The second part of that section considers an extension in which rational people also over-estimate the control they have over the outcome. Section 4 considers how these models translate in cooperative projects with each player taking more than half of the credit and/or taking the praise for himself and putting the blame on the other. Section 5 finally considers managerial implications. All proofs are in appendix 4.B while appendix 4.A considers some variations on the model of section 4.2.

4.2 Everyone is better than average?

Research shows that more than 80% of US drivers consider themselves above the median (Svenson 1981). In another study, less than 5% of employees rated themselves below the median (Meyer 1975).

This section suggests a mechanism through which perfectly rational people might come to these and similar conclusions. To fix ideas, consider the example of driving. Let people care about how well they drive but have different beliefs on what it

[4]We follow most of the psychology literature (e.g. Nisbett and Ross 1980, Fiske and Taylor 1991, Bazerman 1998) in using 'self-serving bias' to refer to the fact that people attribute success to skill and failure to bad luck.

takes to 'drive well'. Then all drivers will act differently, each one following his own beliefs about good driving. But each driver will also evaluate driving styles with the criteria that he used to choose his actions. This double use of criteria will of course favor his own driving style. The rest of this section studies this mechanism with a more formal model.

Let driver quality q be a function of two factors :

1. 'how well you anticipate dangerous situations', denoted by s for 'situations', and

2. 'how well you obey the traffic rules', denoted by r for 'rules'.

Let r and s depend on the amount of attention paid to each, denoted respectively by $a_r \geq 0$ and $a_s \geq 0$. In particular, let

$$q(\alpha, a_r, a_s) = \alpha r(a_r) + (1 - \alpha)s(a_s)$$

where $\alpha \in [0,\ 1]$ denotes the relative importance of following the traffic rules. Let the total amount of attention that a driver can spend be limited to $\bar{a}$, with attention being costless up to $\bar{a}$. Assume further that each person cares about his quality as a driver. In particular, person i's utility function u_i is

$$u_i = q(\alpha_i, a_{r,i}, a_{s,i})$$

where $a_{r,i}$ and $a_{s,i}$ denote i's choice of attention levels, and α_i denotes i's belief about the importance of following traffic rules. We assume all functions to be smooth, r and s to be strictly increasing and strictly concave in attention, with $r'(0) = \infty$ and $s'(0) = \infty$.

Consider now a $[0,\ 1]$ continuum of agents, each with their own individual belief on the relative importance of r and s, i.e. with their own (prior) belief about the value of α. In particular, let person $x \in [0,\ 1]$ believe that $\alpha = \alpha_x = F^{-1}(x)$ with F

a smooth and strictly increasing function[5] with domain a subset of [0, 1] and range [0, 1]. Let that belief be known to all other agents.

In this simple setting we have the following result :

Proposition 1 *Upon observing everyone else's driving behavior, each agent concludes that he/she is (strictly) the best driver of the whole population.*

The suggested interpretation of the model is that driver quality q is an objective but difficult to measure characteristic, for example 'the expected number of accidents caused per 100,000 miles'. The parameter α_i is then agent i's (point-mass) belief about the contribution of each behavior to that outcome. Alternatively, if u_i is the Bernouilli utility function and i is an expected utility maximizer, then α_i can be interpreted as the average value of i's belief on the parameter α. Either way, the model is about how differing priors combined with optimizing behavior lead to *systematically* inconsistent evaluations. An alternative interpretation of the model, which does not fit the overall idea of the paper but could also explain some experimental and empirical results, is that the differing α's are a matter of taste. An 18-year old coming back from a NASCAR race and a 34-year old driving his kid to school evaluate driving styles differently.

There are essentially 3 elements in this model which combine to give the inconsistent assessments :

1. People care about how well they perform.

2. There is disagreement what constitutes good performance. Among other things, this requires that there is no objective and generally agreed-upon measure of performance.

3. The performance is affected by actions and decisions under the agent's control.

[5]If we define [0, 1] to be a probability space with the Lebesgue measure as probability measure, then α_x is a random variable with F as distribution function.

Note in particular that any evaluation inconsistencies disappear in this model if we look at how people rate themselves on a component, say 'anticipating dangerous situations', rather than on the aggregate behavior. This fits with the empirical evidence (Allison et al. 1989, Wade-Benzoni et al. 1997).

In our analysis every agent knows that all other agents also consider themselves better than the median and understands why. While this mutual knowledge and understanding is not completely necessary, it shows explicitly that part of this bias may persist even when people are made aware of the others' self-evaluation.

Appendix 4.A considers two alternative formulations for this model.

4.3 Skill or luck?

People also tend to attribute success to their own skills but to blame failure on bad luck (Zuckerman 1979). This particular asymmetry in attribution is generally called the 'self-serving' bias. As before, we will show how perfectly rational agents may come to such conclusion. The overall mechanism can be decomposed in three effects that we will discuss one by one.

Consider first a situation in which there is some chance that the outcome will unobservably be determined by luck rather than by the agent's action[6]. The inferred ex-post probability that it was luck which caused the success, follows from Bayes law. Let, in particular, q indicate the prior probability that luck does not intervene. Bayes law implies

$$P[\text{action} \mid \text{success}] = \frac{P[\text{success} \mid \text{action}]q}{P[\text{success} \mid \text{action}]q + P[\text{success} \mid \text{luck}](1-q)}$$

This shows that, as the probability of success of the agent's action $P[\text{success} \mid \text{action}]$ increases, it is ex-post more likely that a success was due to a good choice

[6] A more precise formulation of the game follows.

of action. An analogous argument shows that under the same condition, the ex-post probability that failure was caused by (bad) luck also increases. Note that this implies that the asymmetric attribution of success and failure is in itself not irrational *if* the agent believes his action has a higher probability of success than pure 'luck'.

Consider next a situation in which people hold different beliefs about the probability of success of alternative actions. Let each person select the action that he or she thinks has the highest probability of success. Each person will then tend to select precisely these actions about which he or she is overconfident relative to the rest of the population. This effect combines with the earlier one to produce the overall effect that rational people tend to attribute their own success more to skill than outsiders do and their own failures more to bad luck than outsiders do.

Consider, finally, a situation in which there is also uncertainty about the probability that luck will intervene on a particular action. Agents who believe they can do better than luck select actions that give them maximal control. But now an agent will tend to select actions on which he over-estimates his control from the perspective of a randomly selected outsider. This is the same 'random error plus systematic choice gives systematic error' effect as earlier. According to the Bayes expression above an agent will attribute success more to skill as the prior probability q that luck does not intervene, increases. This causes an asymmetry in the self-serving bias: agents display a stronger self-enhancing bias than a self-protective bias. This asymmetry has been documented experimentally (Fiske and Taylor 1991, Miller and Ross 1975).

The rest of this section formalizes these arguments. Consider a set of I agents who each have to decide which action to undertake. Let there be N potential actions $A_n \in \mathbf{A}$ with $1 \leq n \leq N$. Each agent's outcome can be either a success or a failure, denoted S and F respectively. The timing of the game is indicated in figure 4.1. After the agent has chosen which activity to undertake, there is some probability

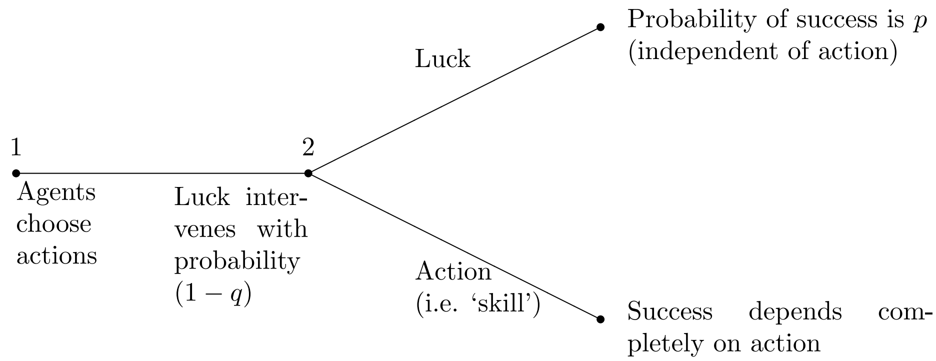

Figure 4.1: Timeline of 'Skill or Luck' game

$(1-q)$ that luck intervenes. If luck intervenes, the probability of success is p, independent of the action. If luck does not intervene, then the outcome depends (probabilistically) on the action chosen by the agent, i.e. it is a matter of 'skill'.

An action's probability of success (conditional on luck not intervening) is unknown. Agents have subjective beliefs about this and these beliefs are commonly known. In particular, let p_n^i denote agent's i's belief regarding the probability of success of action n (conditional on the outcome being determined by skill) and let it be drawn from a distribution F on $[0,\ 1]$, i.e. $p_n^i \sim F[0,\ 1]$. Let $\bar{p} = \sup \operatorname{supp} F$ be the supremum of the support of F, and assume that F is atomless.

Our first result is similar to what we obtained in the last section: agents are relatively overconfident about their likelihood of success. In particular, let Y_i denote the project chosen by agent i and $p_{Y_i}^j$ denote the belief of some agent j that Y_i will be a success (where we allow $j = i$), then

Proposition 2

- *As $N \to \infty$, all agents other than i almost surely consider i to be overconfident in his own project (in the sense of over-estimating the probability of success).*
- *Let $\bar{p} = 1$. As $N \to \infty$, any particular agent i will almost surely believe that his preferred project Y_i will almost surely succeed if luck does not intervene.*

Formally, for any $\epsilon_1, \epsilon_2 > 0$, $\exists \tilde{N}$ *such that* $\forall N \geq \tilde{N}$, $P[p^i_{Y_i} > 1 - \epsilon_1] > 1 - \epsilon_2$.

Note the overall situation :

- Agents have a lot of confidence in their own projects and less so in projects of others.
- Each agent thinks all other agents have too little confidence in his own project and too much confidence in their own projects.

The overconfidence part of the proposition is in relative or subjective terms, in the sense that it says how one agent considers the judgment of the other. It is straightforward to extend this to an objective or absolute statement. To that purpose, introduce a 'reference' (or 'objective') probability of success for action n, denoted $\hat{p}_n$. This can be interpreted, for example, as the true or objective probability or as the belief of the social planner. The following assumption will play a key role.

Assumption 5 *Let* $\hat{p}_n < \bar{p}$*, where* $\bar{p} = \sup \operatorname{supp} F$.

This says that there is some positive probability that the agent will overestimate the probability of success of the action. It is satisfied if, for example, $\hat{p}_n$ itself is drawn from F or if $\hat{p}_n$ is the mean of F. The earlier conclusion now extends. Subjective rationality can lead to objective overconfidence.

Proposition 3 *Let A5 hold. In the limit as* $N \to \infty$*, the agent is almost surely overconfident about his likelihood of success.*

It is very tempting to translate these results to an academic context. In particular, let the actions correspond to research topics. Let each academic be limited to choosing one research topic. Let the contribution to science of any project be measurable on a scale from zero to one. In particular, assume that a successful project contributes one unit to science while a failure contributes zero.

Corollary 1 *In the limit as the number of potential research topics goes to infinity, each academic considers his own expected contribution to be among the highest (i.e. to be 1) while he considers the contribution of others to be distributed according to* F *on* $[0, 1]$*. In particular, he considers the average contribution to be* $\mu_F < 1$ *.*

Any particular academic will also almost surely be considered to overestimate his impact on science by each and every one of his colleagues.

The inflated view of own work that academics sometimes observe among their colleagues (Ellison 2000) might thus be a matter of genuinely differing priors.

Consider now the question of skill versus luck. We will say that 'i attributes success more to skill than j does' if i's estimate of the probability that the outcome was determined by the action (rather than by luck) is higher than j's estimate.

Proposition 4

- *As* $N \to \infty$*, an agent who has a success almost surely attributes that success more to his skill than a randomly selected agent [would attribute the focal agent's success to skill]. Analogously, as* $N \to \infty$*, an agent attributes his failure more to (bad) luck than a randomly selected agent [attributes that failure to bad luck].*

- *Let A5 hold. As* $N \to \infty$*, a successful agent almost surely attributes his success too much to his own skill, while a failing agent almost surely attributes his failure too much to (bad) luck.*

- *Let* $\bar{p} = 1$*. In the limit as* $N \to \infty$ *the agent is almost sure that any failure was due to bad luck, while his belief that success, if he has one, is due to his own skill converges to* $\frac{q}{q+p(1-q)}$

Note in the last expression that the belief that success is due to skill will be closer to 1 as p is smaller or q is larger. This is intuitive : As the probability of success by luck decreases, a success becomes more indicative of skill; As the prior probability that the outcome is due to skill increases, so does the posterior probability.

It is useful to compare this section to the earlier one on driving quality. At first sight, the two models seem very different. Consider, however, the following re-interpretation of section 4.2. Let q refer to the probability of success while a_r is the action (with a_s fixed at $a_s = \bar{a} - a_r$), which is now continuous. It is then clear that both models are essentially the same, except for the fact that we did not allow in the preceding section for the outcome to be observed. This parallelism will become more obvious in section 4.4.

Comparative statics The frequent use of limit results suggests that the number of alternative projects available to the agent might be an important factor. To study this, consider what happens when we let the number of projects increase (while keeping the belief realizations for the already available projects fixed).

Proposition 5 *For any given realization of beliefs, as the number of available projects increases, a successful agent is more likely to attribute his success to skill, while a failing agent is more likely to attribute his failure to bad luck.*

The intuition is simply that, from the focal agent's perspective, more available actions can only improve the subjective probability of success of the action he ends up undertaking. In the other extreme that there is only one possible task, the biases actually disappear.

This comparative static turns out to be quite important. In particular, experiments that test for the self-serving bias typically restrict the actions the subjects can take. In everyday life, people typically have much more freedom. This implies that such structured experiments might under-estimate the practical relevance of these biases. This might also be a reason why, according to some people (e.g. Miller and Ross 1979), the experimental evidence for self-serving biases is weak compared to what 'casual empiricism' suggests. Finally, this comparative static might also be responsible for the observed fact that the biases typically increase as we allow the agents more degrees of freedom (Allison et al. 1989, Wade-Benzoni et al. 1997).

A second comparative static concerns the form of the distribution.

Proposition 6 *In expectation and for $N \to \infty$, an agent's overconfidence (and his other biases) relative to a randomly selected agent increases in a mean-preserving spread of the distribution.*

The intuition is simply that the upper boundary of the support, $\bar{p}$, increases in mean-preserving spreads while the mean remains (by definition) constant. This implies that there will be more over-confidence if there is more underlying uncertainty.

The next subsection explores the implications for learning while subsection 4.3.2 considers an extension in which agents also have some control over the probability that luck intervenes.

4.3.1 The failure to learn from failures

From an average person's perspective, a person with extreme beliefs rejects too easily disconfirming evidence as a random variation. People with different beliefs 'see' different things in the same observations. This fits the experimental evidence (Lord, Ross and Lepper 1979). People with different beliefs therefore also disagree on the amount of updating that should take place. In particular, rational people generally don't learn as much from their own failures as they should from the perspective of a rational outside observer.

To see this formally, assume that projects are either right or wrong. As long as luck does not interfere, right projects are successful, while wrong projects end up in failure. Of course, the possible intervention of luck makes the outcome a less than perfect signal for the quality of the project. In this interpretation, p_n^i is the initial belief of agent i about whether n is a right or wrong project. The key conclusion now is that the agent does not learn enough from his own failures.

Proposition 7 *In the limit as $N \to \infty$ the agent almost surely learns less from his own failure than he should, given the ex-post probability that luck intervened as* *estimated by some randomly selected agent. In particular, if A5 holds then the agent almost surely learns less from his own failures than he should given the 'objective' (or reference) ex-post probability that luck intervened. Analogously, an agent learns 'too much' from his own successes.*

Note that, from a subjective perspective, agents learn precisely what they should learn, since we assumed them to be Bayesian rational. However, from a 'reference' perspective, they discard too much negative evidence as 'random variation'.

4.3.2 An illusion of control

Consider now a situation in which there is not only uncertainty about the probability of success of the actions, p_n^i, but also about the probability that luck will intervene on some particular course of action. In particular, let q_n^i denote agent's i's belief that the outcome of action n is determined by skill and let it be drawn from a distribution G on $[0,\ 1]$, i.e. $q_n^i \sim G[0,\ 1]$. So the agent believes that if action n is undertaken, luck will intervene with probability $1 - q_n^i$. Let G have no atoms. Let $\bar{p} = \sup \operatorname{supp} F > p$ and $\bar{q} = \sup \operatorname{supp} G$.

The expected payoff of action n from the perspective of agent i is

$$q_n^i p_n^i + (1 - q_n^i)p = p + q_n^i(p_n^i - p)$$

where p is still the probability of success when luck determines the outcome. It follows that whenever possible, the agent will choose an action with $p_n^i > p$, i.e. an action that has a higher probability of success than luck. Moreover for every action with $p_n^i > p$, the agent prefers higher q_n^i. This is intuitive: the agent wants maximal control if he thinks he's better than luck. This further suggests that with large enough choice, the agent will end up selecting an action with q_n^i and p_n^i near their respective maxima. That, and its consequences, is essentially the content of the following proposition[7].

Proposition 8 *As $N \to \infty$*

- $P[q^i_{Y_i} > q^j_{Y_i}] \xrightarrow{a.s.} 1$ *i.e. from the perspective of any outsider, i almost surely has a relative illusion of control.*

- *If $\bar{q} = 1$ then $q^i_{Y_i} \xrightarrow{a.s.} 1$ i.e. agent i is almost sure he has complete control over the outcome*

- *If $\bar{q} = 1$ and $\bar{p} = 1$ then $P[Skill \mid Success] \to 1$, while $P[Luck \mid Failure]$ depends on the tails of the distributions of p^i_n and q^i_n, i.e. upon success the agent is almost sure the success is due to his skillful choice of action while his inference from a failure can go either way depending on the specifics of the situation.*

The first part of the proposition says that an agent will over-estimate his control over the situation. The last part of the proposition is one potential explanation for the fact that there is more evidence that people take credit for success than that they deny responsibility (Miller and Ross 1975).

4.4 Praise yourself, blame the others

This section considers the implications of these mechanisms for cooperative ventures. It shows how the agents' self-estimated contributions to cooperative projects may add up to more than 100%, especially when the project turns out to be a success.

We consider two situations that can create such inconsistent estimates. The first model is similar to the model of section 4.3. Only now, there is uncertainty about which of the two agents' action determined the outcome, rather than whether it was the action or luck that did it. The question is of course how responsibility for

[7] Y_i again denotes agent i's preferred action.

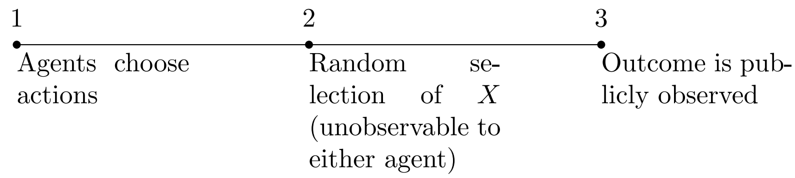

Figure 4.2: Timeline of 'Praise or Blame' game

success and failure is attributed to the two agents. The second model considers the same issue in the context of an extension of section 4.2 to a 2-person situation.

4.4.1 Model I

Let there be two individuals, denoted 1 and 2, who each have to (publicly) choose an action. Let there be N potential actions $A_n \in \mathbf{A}$ with $1 \leq n \leq N$. The overall outcome can be either a success or a failure, denoted S and F respectively. The timing is according to figure 4.2. After each agent has chosen what activity to undertake, it is determined randomly and unobservably whose action will determine the outcome. In particular, the random variable $X \in \{1, 2\}$ indicates whether it is agent 1's or agent 2's action that determines the outcome. Let $P[X = 1] = P[X = 2] = 1/2$. The probability of success of an action is not known, but both agents have beliefs about that probability and these beliefs are common knowledge. In particular, let agent's i's belief on the probability of success of action n be denoted p_n^i and let it be drawn from a distribution F on [0, 1], i.e. $p_n^i \sim F[0,\ 1]$.

We are interested to see how agents make inferences about who is responsible for the outcome, based on whether the outcome is a success or a failure. Let τ_i^S denote agent i's belief about the probability that he was responsible for the outcome conditional on success and τ_i^F analogously for failure. The following proposition says that in the limit the self-estimated contributions add up to more than 100% when the outcome is a success and to less than 100% when it is a failure.

Proposition 9 *As* $N \to \infty$*,*

- $\tau_1^S + \tau_2^S > 100\%$ *a.s.*

- $\tau_1^F + \tau_2^F < 100\%$ *a.s.*

To see how substantial the effect can be, we can calculate the expected sum of (subjectively estimated) contributions. For example , when $p_n^i \sim U[0,\ 1]$, $E[\tau_1^S + \tau_2^S] \xrightarrow{a.s.} 2\ln(2) \approx 1.4$.

4.4.2 Model II

Let there again be two individuals, denoted 1 and 2, who now have to decide simultaneously where to focus their attention. Let there be two tasks, A and B, with the amount of attention that agent i spends on task A denoted $a_{i,A} \geq 0$. Let the total amount of attention of each agent be limited to one unit, i.e. $a_{i,A} + a_{i,B} \leq 1$ and let attention be free. The overall payoff is

$$R = \sum_{i=1,2} \epsilon[\alpha q(a_{i,A}) + (1-\alpha)q(a_{i,B})]$$

with ϵ a mean-one random variable, with a smooth distribution function on $[o, \infty)$, that is not observed by either agent[8]. Assume further that q is smooth, $q' > 0$, $q'(0) = \infty$, and $q'' < 0$. Let α_i denote agent i's subjective belief about the value of α, and let $\alpha_1 \neq \alpha_2$. Finally, let each agent's utility be strictly increasing in R.

After having devoted attention to the tasks and after observing the attention the other agent has devoted to each task, each agent observes R. Define $R_i = \epsilon_i \alpha_i q(a_{i,A})$ to be the share of the payoff that is due to agent i from agent i's perspective (with ϵ_i being i's estimate of ϵ). Then we have that

[8]The presence of ϵ will make sure that the observed outcome R is not a perfect signal for the true value of α, given that each agent observes all other variables in the payoff equation.

Proposition 10 *With probability one, each agent subjectively estimates his contribution to be strictly more than 1/2, so that* $R_1 + R_2 > 1$ *a.s.*

Again, the agent's estimated contributions add up to more than 100%. Moreover the discrepancy increases in the amount of disagreement on α, as can be seen from the following proposition. Let wlog. $\alpha_1 > \alpha_2$.

Proposition 11 *For fixed* α_2 *(or for fixed* α_1*),* $|R_1 + R_2 - 1|$ *increases in* $|\alpha_1 - \alpha_2|$.

The role of differing priors Many results in this paper would not hold if we restricted all agents to hold an identical prior or to know somehow the true prior. In that case, private information would be required to obtain belief differences. But whenever agents can observe others' actions, they can deduce their information sufficiently to correct for a systematic bias. Even when we could assume that agents do not observe each others' actions or otherwise know each others' beliefs, the conclusions would be much weaker since all biases disappear when beliefs become mutual knowledge.

Furthermore, the introduction of private information would unnecessarily complicate the analysis. In the end, the key argument for the use of differing priors is a methodological one: it allows a clean and transparent analysis of differing beliefs or, to paraphrase Aumann (1987), it enables one to zero in on issues caused by differing beliefs.

Chapter 4 gives a more in-depth discussion of this and related issues.

4.5 Managerial implications

This section considers some managerial implications of our results. It should be noted, however, that some parts of what follows are still conjectural and require more formal analysis. For this discussion, it is also useful to note that organizations can sometimes be conceptualized as individuals. Annual reports, for example, also display self-serving biases in their causal attributions (Bettman and Weitz 1983).

Evaluations and incentives Baker, Jensen and Murphy (1988) discuss the impact of the self-enhancing bias (section 4.2) on the effectiveness of incentive systems. In particular, they argue that the bias may explain why superiors are reluctant to give their employees bad performance evaluations: the latter would consider such evaluations unfair and become dissatisfied. They further argue that '[b]iased (...) performance evaluation reduces productivity by reducing the effectiveness of incentives in the organization.'

The current analysis suggests that this conclusion is not always warranted. First of all, the assumption is that superiors tend to give good evaluations to avoid the conflict that arises from employees' dis*satisfaction* with their rating (Baker Jensen and Murphy 1988). Equity theory (Adams 1965), however, suggests that the perception of fairness is also an important determinant of *motivation*. This psychological theory, which has been confirmed experimentally (e.g. Goodman and Friedman 1971, Miles, Hatfield, and Huseman 1989), holds that people try to make sure the balance between efforts (input) and what they get from that effort (outcome) is fair. For example, people who feel they get underpaid will reduce their effort. Compressing evaluations on the better half of the scale might thus actually *improve* the effectiveness of the incentive system: the evaluations seem fair since they match employees' perceptions and such fairness improves motivation. And motivation is what incentives are about. Second, an employer can anticipate his superiors' tendency to avoid negative evaluations and adapt nominal bonuses and incentives to compensate for it. In particular, the nominal bonuses should get steeper and lower. This modification may sometimes completely restore the effectiveness of the incentives.

Not all problems will be eliminated, however. In particular, the effectiveness of the evaluation system in terms of providing feedback and learning will still be reduced. The threat of layoffs as an incentive mechanism is also less effective since an employee with 'above average' ratings will generally not consider himself at risk. For these reasons it is sometimes important to consider remedies. The analysis in section 4.2 showed that the inconsistency is driven by disagreement on the relative importance of different behaviors. One solution is to evaluate the behaviors separately and being explicit on how the results get aggregated to an overall evaluation. A different solution is to give explicit descriptions of good, average, and bad overall behavior. This allows agents to derive for themselves the implicit weight put on different criteria.

Decision processes The endogenous overconfidence effect has important implications for the optimal structuring of decision processes in organizations.

In a famous article, Kaplan (1984) decried the fact that companies often use too high hurdle rates in Discounted Cash-Flow Analysis. Doing so reduces the attractiveness of projects with cash-flows that come far in the future. His argument makes perfect sense if managers can form an unbiased estimate of the future cash flows of suggested projects. The analysis here shows, however, that managers will tend to overestimate the return on their projects, and more so as there is more uncertainty. It is furthermore reasonable to assume that there is more uncertainty about cash-flows as they are further in the future. The use of artificially high discount rates, while imperfect, might be a practical way to compensate for this tendency. For the same reasons, it also makes sense to discount benefits that are difficult to quantify.

We also conjecture that an extension of the analysis would show that, absent informational issues, letting two people jointly choose which project to undertake leads to higher buy-in (in the sense of higher subjective expected probability of success) by the people who are involved in the choice process, but has a relatively smaller impact on the true quality of the decision. The reason is that the two agents will choose a project in which they both believe. But that will also tend to be a project in which both are overconfident. We further conjecture that, instead of involving both in the original selection of the project, having one person first choose the project and subsequently allowing the second to veto it, might sometimes lead to better decisions but at the cost of a lower buy-in by the second person. The reason is that the second person acts as an independent observer who is free from the endogenous overconfidence bias.

Finally, the analysis also has potential implications for the theory of delegation. In particular, it suggests that delegation to an employee might improve efficiency if that employee under-supplies effort and effort is complementary to the probability of success. Along the same vein, centralization of the decision should then be optimal if effort is a substitute to the probability of success. This is similar to Zabojnik (2001).

Organizational learning Section 4.3.1 shows how the self-serving bias might inhibit learning from failure or mistakes (and make decision makers overinterpret the significance of their successes). In cases when learning is important, organizations should look for ways to eliminate these effects.

One way to improve learning is to let someone other than the original decision maker review the failures. Such independent observer does not have the endogenous overconfidence and will therefore tend to be more accurate in his evaluation. On an organizational level, this suggests for example that quality control should not be part of the manufacturing department[9] and suggests a potential function for independent directors. It also provide a rationale beyond mere expertise for the existence of a company like Exponent, Inc. (formerly 'The Failure Group, Inc.') which specializes in general failure analysis. A further example are the FAA reviews of incidents and accidents in air travel and transportation. Note, however, that

[9] In particular, the ISO-9000 certification and the Baldridge Awards require quality control to be directly under the CEO.

this paper also suggests that such outside reviewer will generally be at odds with the original decision maker on the causes of the failure. This is consistent with the fact that the pilot union often rejects the conclusion that an accident was due to pilot error. The resentment that such arguments can generate and its impact on performance should be considered before taking this approach. Moreover, such independent reviews will only work if that outsider has the means to influence the future actions of the decision maker. In the absence of such influence, it might be necessary to replace the original decision maker. This is a very different theory for CEO replacement than the idea that managers get fired because they are not good enough or did not spend enough effort.

Bargaining and litigation The implications of the self-serving bias for bargaining and litigation have been explored in some detail (Priest and Klein 1984, Babcock Wang et al. 1997, Babcock Loewenstein et al. 1995). It has, in particular, been argued that such bias increases the probability of breakdowns in bargaining, of strikes in labor negotiations and of effective litigation. The experimental evidence in this field also largely supports the claims made. These experiments, however, usually do not endogenize the actions that lead to bargaining and litigation. Instead, the experimental subjects get to read a description of the case before or after they have been told which role they will play in the case. This setup excludes essentially the mechanisms that have been discussed here. Our analysis thus suggests that these experimental results may underestimate the magnitude of the real self-serving bias and thus its role in real-life strikes, litigation and bargaining breakdowns.

Team projects and incentives The analysis of section 4.4 has a number of implications for team-work. In particular, the discrepancy in contribution assessments might lead to conflict among the members of the group. The conflict will show up both as disagreement ex-ante on what to do and as disagreement ex-post on who is to praise for success or to blame for failure.

Proposition 11 implies that such conflicts will be stronger as there is more disagreement in the team. This suggests that more homogenous teams will work smoother. This effect must be balanced, however, against the gains that can come from exchange of information and other benefits of heterogenous teams. The same considerations might apply to alliances of firms. These issues require more formal study.

Conflict can also be limited by giving less occasion for members of the team to argue over praise or blame. This might be a reason to use some exogenous distribution rule for team-bonuses (such as equal distribution or distribution according to base wage) rather than trying to divide such bonus according to each member's performance. This compounds the fairness issue and its incentive implications, as discussed above.

We also conjecture that disagreement will be stronger after failure than after a success. This would then predict that teams might break up more easily after a failure.

4.6 Conclusion

Subjective rationality can lead to objective overconfidence. This is probably the most important conclusion of this paper. We expect this effect to be robust and to show up in many other situations because the underlying causes are so simple and ubiquitous. People optimize and, in doing so, select the most extreme alternatives. This 'selection of extremes' then requires a 'regression towards the means'. The true characteristics of the chosen alternative will thus tend to be less extreme than anticipated.

This paper then showed how such overconfidence (and a closely related effect) leads to self-enhancing biases in evaluating one's own and others' behaviors, in explaining success or failure, and in attributing praise and blame in cooperative projects.

The logical next step would be to test the theory. Experimental economics seems to be the most appropriate methodology. While the comparative statics could be a possible starting point, it seems that it would be more effective to test directly the basic mechanism. The main challenge is to generate or accurately measure beliefs.

On a more general level, the paper suggests the value of studying the implications of differing priors. It seems, in particular, that some behaviors that are traditionally considered to be in the realm of bounded rationality might actually be usefully studied by using differing priors.

4.A Alternative formulations for section 4.2

Two characteristics of the model in section 4.2 raise some obvious questions :

1. Attention is costless but limited. What if it were unlimited but costly?

2. The outcome is completely dependent on actions. What if other factors play?

As to the first variation, the results extend, but not unmodified. Consider for example the following specification. Let the driver's quality be $q = \alpha a_r + (1-\alpha)a_s$, and let the agent's utility be $u_i = q(a_{r,i}, a_{s,i}) - c(a_{r,i} + a_{s,i})$ with the cost function quadratic, $c(x) = \frac{x^2}{2}$. The agent's problem now becomes

$$\max_{a_r, a_s} \alpha_i a_r + (1-\alpha_i)a_s - \frac{(a_r + a_s)^2}{2}$$

which solves as follows: If $\alpha > 1/2$ then $a_r = \alpha$ and $a_s = 0$, and if $\alpha < 1/2$ then $a_r = 0$ and $a_s = (1-\alpha)$. From the perspective of any agent with belief $\alpha > 1/2$, the agent with belief $\alpha = 1$ is the best driver. While in this case the results are less extreme, the bias is still present. In particular, with α uniformly distributed on $[0,\ 1]$, it turns out that 80% of the drivers consider themselves better than the median. Overall, we expect this general result to be quite robust.

As for the second variation, we would generally expect the driving quality to vary also with some exogenously given 'driving talents or skills'. This can for example be incorporated by adding a random skill component ϵ to the driving quality that is observed. This will make the result less extreme. The consistency of judgments should increase in the importance of the exogenous component.

4.B Proofs

4.B.1 Everyone is better than average?

Proof of Proposition 1: Consider an individual $y \in [0,1]$. This individual's problem in choosing his attention levels is $\max_{a_r, a_s \geq 0} \alpha_y r(a_r) + (1-\alpha_y)s(a_s) \quad \text{s.t.} \quad a_r + a_s \leq \bar{a}$.

This problem has a unique solution, denoted $(a_{r,y}, \bar{a} - a_{r,y})$ with $a_{r,y}$ strictly increasing in α_y, by the implicit function theorem. It follows that if $\alpha_y \neq \alpha_x$ then $a_{r,y} \neq a_{r,x}$, so that by the definition of $(a_{r,x}, a_{s,x})$ and the uniqueness of the maximum $\alpha_x r(a_{r,x}) + (1-\alpha_x)s(a_{s,x}) > \alpha_x r(a_{r,y}) + (1-\alpha_x)s(a_{s,y})$. ■

4.B.2 Skill or luck

Proof of Proposition 2: Without loss of generality, let $i = 1$.

For the first part of the proposition, the probability that all agents consider agent 1 to be overconfident is $P[p^1_{Y_1} \geq p^2_{Y_1}, ..., p^1_{Y_1} \geq p^I_{Y_1}] = \int P[x \geq p^i_{Y_1}]^{I-1} f_{p^1_{Y_1}}(x)dx = \int F(x)^{I-1} N F(x)^{N-1} dF(x) = \frac{N}{N+(I-1)}$ which converges to 1 as $N \to \infty$.

For the second part of the proposition (again with $i = 1$ wlog.), we need to show that for any $\epsilon_1, \epsilon_2 > 0$, $\exists \tilde{N}$ such that $\forall N \geq \tilde{N}$, $P[p^1_{Y_1} > \bar{p} - \epsilon_1] > 1 - \epsilon_2$.

Note now that $Y_1 = \text{argmax}_{Y \in \mathbf{A}} p^1_Y$ so that $P[p^1_{Y_1} \leq x] \equiv P[p^1_{A_1} \leq x, ..., p^1_{A_N} \leq x] = [F(x)]^N$ so $P[p^1_{Y_1} > \bar{p} - \epsilon_1] = 1 - F_{p^1_{Y_1}}(1 - \epsilon_1) = 1 - [F(\bar{p} - \epsilon_1)]^N$. Since $\epsilon_1 > 0$, $\bar{p} - \epsilon_1 < \bar{p}$ so that $F(\bar{p} - \epsilon_1) < 1$. This implies that $\lim_{N \to \infty} 1 - [F(\bar{p} - \epsilon_1)]^N = 1$ which then proves the second part of the proposition when $\bar{p} = 1$. ■

Proof of Proposition 3: Analogous to the proof of proposition 2. ■

Proof of Proposition 4: Upon success, agent j's estimate that a successful outcome of action Y_i was due to skill is $\frac{p^j_{Y_i} q}{p^j_{Y_i} q + p(1-q)}$ which is strictly increasing in $p^j_{Y_i}$ as long as $p, q > 0$. Now, to show that (as $N \to \infty$) i almost surely attributes success more to skill than j, it suffices that the probability that $p^i_{Y_i} > p^j_{Y_i}$ converges to 1. Note that $P[p^i_{Y_i} > p^j_{Y_i}] = \int P[x > p^j_{Y_i}] f_{p^i_{Y_i}}(x)dx = \int F(x)d[F^N(x)] = \frac{N}{N+1}$ which converges to 1 as $N \to \infty$. The argument for failure is completely analogous. The second part of the proposition is completely analogous.

Consider now the final part of the proposition. An agent i who fails believes that his failure was due to (bad) luck with probability

$$\frac{(1-p)(1-q)}{(1-p)(1-q) + (1-p^i_{Y_i})q} \overset{a.s.}{\to} \frac{(1-p)(1-q)}{(1-p)(1-q) + (1-\bar{p})q}$$

since $p^i_{Y_i} \overset{a.s.}{\to} \bar{p}$ as $N \to \infty$ (by the proof of proposition 2). This equals 1 when $\bar{p} = 1$. Analogously, the belief of a successful agent i that his success was due to skill is

$$\frac{p^i_{Y_i} q}{p^i_{Y_i} q + p(1-q)} \overset{a.s.}{\to} \frac{\bar{p}q}{\bar{p}q + p(1-q)}$$

which equals $\frac{q}{q+p(1-q)}$ when $\bar{p} = 1$. ■

In what follows, we will use ρ_i for the belief of agent i about any event (except for the prior beliefs about the actions that were defined earlier).

Proof of Proposition 5: From the proof of proposition 4, we know that ρ_i[action | success] increases in $p^i_{Y_i}$. Of course, since $Y_i = \text{argmax}_{Y \in \mathbf{A}} p^i_Y$, $p^i_{Y_i}$ weakly increases when we add an extra project to $\mathbf{A}$ (while keeping the beliefs in the other projects unchanged). The first part of the proposition then follows immediately. The second part (regarding role of failure) is completely analogous. ■

Proof of Proposition 6: In expectation, agent i's overconfidence relative to a randomly selected other agent is $p^i_{Y_i} - \mu_F$. As $N \to \infty$, this converges a.s. to $\bar{p} - \mu_F$. With a mean-preserving spread, μ_F remains constant while $\bar{p}$ is non-decreasing. This proves the proposition. ■

The Failure to Learn from Failures

Proof of Proposition 7: We do the proof for the over-learning in case of success and in comparison to the reference belief. The lack of learning in case of failure is completely analogous. The cases with a randomly selected agent are also completely analogous once we observe that, in the limit as $N \to \infty$, a condition analogous to assumption A5 holds almost surely for the belief of the randomly selected agent instead of the reference belief.

Consider agent 1. His belief that his project will succeed is $p^1_{Y_1}$, while the 'reference' belief is $\hat{p}_{Y_1}$. Upon success, agent 1 believes that the outcome was caused by his action (rather than luck) with probability $\rho_1(\text{Ac} \mid S) = \frac{p^1_{Y_1} q}{p^1_{Y_1} q + p(1-q)}$ while the 'reference' belief on that event is $\hat{\rho}(\text{Ac} \mid S) = \frac{\hat{p}_{Y_1} q}{\hat{p}_{Y_1} q + p(1-q)}$. As $N \to \infty$, a proof analogous to that of proposition 2 combined with assumption A5 implies that $p^1_{Y_1} > \hat{p}_{Y_1}$ a.s., so that $\rho_1(A \mid S) > \hat{\rho}(A \mid S)$ a.s.

Consider now how agent 1 updates his belief that his action is right after he observed a success. Denote the event that the agent's action is right by R. We can split the updating in 2 components:

1. determining how likely it is that the successful observation was caused by action or by luck

2. updating his belief conditional on the observation being caused by action or by luck

Let us use ρ to denote the ex-post belief that the observation was caused by the action (rather than by luck). We get

$$\begin{aligned}
\rho_1[R \mid S] &= P[R \mid S\&\text{Ac}]\rho[\text{Ac} \mid S] + P[R \mid S\&L]\rho[L \mid S] \\
&= 1.\rho[\text{Ac} \mid S] + p^1_{Y_1}(1 - \rho[\text{Ac} \mid S]) \\
&= p^1_{Y_1} + (1 - p^1_{Y_1})\rho[\text{Ac} \mid S]
\end{aligned}$$

This increases in $\rho[\text{Ac} \mid S]$. It follows that agent 1 has a stronger belief that his action is correct after a success than what he would have if he used the 'reference' ex-post probability that his success was caused by luck. ■

An illusion of control

We defined $q_n^i \sim G$ to be the probability that luck does not intervene and $p_n^i \sim F$ to be the probability that the action is a success.

Proof of Proposition 8: Define the new random variable $\pi_n^i = p + q_n^i(p_n^i - p) \in [0,\ 1]$, which is the payoff that i expects from choosing action n, and let its distribution function be H, with $\sup \operatorname{supp} H = \overline{\pi} = \bar{p} + \bar{q}(\bar{p} - p)$.

Agent i chooses the action with the highest π_n^i. So $P[\pi_{Y_i}^i \leq x] = P[\pi_1^i \leq x, ..., \pi_N^i \leq x] = H(x)^N$ which converges to zero for each $x < \overline{\pi}$ (since $H(x) < 1$ for $x < \overline{\pi}$). It follows that in the limit as $N \to \infty$, $\pi_{Y_i}^i \overset{a.s.}{\to} \overline{\pi}$. But this implies that $p_{Y_i}^i \overset{a.s.}{\to} \bar{p}$ and $q_{Y_i}^i \overset{a.s.}{\to} \bar{q}$. It follows immediately that $P[q_{Y_i}^i > q_{Y_i}^j] = G(q_{Y_i}^i) \overset{a.s.}{\to} 1$ which proves the first part of the proposition.

The second part follows immediately from the conclusion above that $q_{Y_i}^i \overset{a.s.}{\to} \bar{q}$ and the extra assumption that $\bar{q} = 1$.

As for the third part, we know that

$$\rho_i[\text{action} \mid \text{success}] = \frac{p_{Y_i}^i q_{Y_i}^i}{p_{Y_i}^i q_{Y_i}^i + p(1 - q_{Y_i}^i)} \overset{a.s.}{\to} \frac{1.1}{1.1 + p.0} = 1$$

Furthermore

$$\rho_i[\text{luck} \mid F] = \frac{(1-q)(1-q_{Y_i}^i)}{(1-q)(1-q_{Y_i}^i) + (1-p_{Y_i}^i)q_{Y_i}^i} = \frac{1-q}{1-q+\frac{q_{Y_i}^i}{1-q_{Y_i}^i}(1-p_{Y_i}^i)}$$

the limit of which depends indeed on the relative speed of convergence of $q_{Y_i}^i$ and $p_{Y_i}^i$. ■

4.B.3 Praise yourself, blame the other

Proof of Proposition 9: Let the $\sup \operatorname{supp} F = \bar{p}$. Note first that the proof of proposition 2 implies that $p^i_{Y_i} \overset{a.s.}{\rightarrow} \bar{p}$, so that $P[p^i_{Y_i} > p^j_{Y_i}] = F(p^i_{Y_i}) \overset{a.s.}{\rightarrow} 1$ We thus get

$$\begin{aligned}\rho_i[X = 1 \mid S] &= \frac{\rho_i[S \mid X = 1]\rho(X = 1)}{\rho_i[S \mid X = 1]\rho(X = 1) + \rho_i[S \mid X = 2]\rho(X = 2)} \\ &= \frac{p^i_{Y_i}}{p^i_{Y_i} + p^j_{Y_i}} > 1/2 \quad \text{a.s.}\end{aligned}$$

This implies the first part of the proposition. The second part is analogous. ■

Proof of Proposition 10: Note that agent i solves :

$$\max_{a_{i,A}+a_{i,B}\leq 1; a_{i,A},a_{i,B}\geq 0} \epsilon[\alpha_i q(a_{i,A}) + (1-\alpha_i)q(a_{i,B})]$$

which has a unique maximum, which (by the IFT) is strictly increasing in α_i. It follows that if $\alpha_1 \neq \alpha_2$, then $\alpha_1 q(\hat{a}_{1,A}) + (1-\alpha_1)q(\hat{a}_{1,B}) > \alpha_1 q(\hat{a}_{2,A}) + (1-\alpha_1)q(\hat{a}_{2,B})$ where $\hat{a}_{i,A}$ and $\hat{a}_{i,B}$ are the maximizers for agent i. Now, according to agent 1, his contribution to the outcome is $\frac{\epsilon_1[\alpha_1 q(\hat{a}_{1,A})+(1-\alpha_1)q(\hat{a}_{1,B})]}{\sum_{i=1,2}\epsilon_1[\alpha_1 q(\hat{a}_{i,A})+(1-\alpha_1)q(\hat{a}_{i,B})]}$ which, by the above argument, is a.s. strictly larger than $\frac{1}{2}$.

The argument for agent 2 is completely similar. ■

Proof of Proposition 11: It suffices to show that R_1 and R_2 increase in α_1. But this follows from the definition of R_1 and R_2, the strict concavity of $\alpha_i q(a) + (1-\alpha_i)q(1-a)$ in a, and the fact that $\hat{a}_{i,A}$ is strictly increasing in α_i .

The argument for fixed α_1 is completely analogous. ■

Bibliography

[1] Adams, J., (1965), 'Inequity in Social Exchanges', in L. Berkowitz (ed.), Advances in Experimental Social Psychology, New York, Academic Press

[2] Aghion, P., and Tirole, J., (1997), 'Formal and real authority in organizations, *Journal of Political Economy*, 105(1): 1-29

[3] Allen, F., and Morris, S., (1998), 'Finance Applications of Game Theory', Working Paper 98-23-B, Wharton Financial Institutions Center

[4] Allison, S., Messick, D., and Goethals, G., (1989), 'On Being Better but not Smarter than Others: The Mohammed Ali Effect', *Social Cognition*, 7: 275-296

[5] Athey, S., Avery, C., and Zemsky, P., (2000), 'Mentoring and Diversity', *American Economic Review*, 90(4): 765-786

[6] Aumann, R. J., (1976), 'Agreeing to Disagree', *Annals of Statistics*, 4: 1236-1239

[7] Aumann, R. J., (1987), 'Correlated Equilibrium as an Expression of Bayesian Rationality', *Econometrica*, 55: 1-18

[8] Aumann, R. J., (1998), 'Common Priors : A Reply to Gul', *Econometrica*, 66(4): 929-938

[9] Aumann, R. J., and Brandenburger, A., (1995), 'Epistemic Conditions for Nash Equilibrium', *Econometrica*, 63: 1161-1180

[10] Babcock, L. et al., (1995), 'Biased Judgements of Fairness in Bargaining', *American Economic Review*, 85(5): 1337-1343

[11] Babcock, L., and Loewenstein, G., (1997), 'Explaining Bargaining Impasse: The Role of Self-Serving Biases', *Journal of Economic Perspectives*, 11(1): 109-126

[12] Babcock, L., Wang, X., and Loewenstein, G., (1996), 'Choosing the Wrong Pond: Social Comparisons in Negotiations that Reflect a Self-Serving Bias', *Quarterly Journal of Economics*, 111(1):1-19

[13] Baker, G. P., Gibbons, R. S., and Murphy, K. J., (1999), 'Informal Authority in Organizations', *Journal of Law, Economics, and Organization*, 15(1): 56-73

[14] Baker, G., Jensen, M., and Murphy, K., (1988), 'Compensation and Incentives: Practice versus Theory', *Journal of Finance*, 43(3): 593-616

[15] Baum, R. J., Locke, E. A., and Kirkpatrick, S. A., (1998), 'A longitudinal study of the relation of vision and vision communication to venture growth in entrepreneurial firms', *Journal of Applied Psychology*, 83(1): 43-54

[16] Bazerman, M., (1998), *Judgment in Managerial Decision Making*, 1998, New York: John Wiley & Sons

[17] Benabou, R., and Tirole, J., (1999), 'Self-Confidence: Intrapersonal Strategies', Working paper SSRN

[18] Bennis, W. G., and Nanus, B., (1985), *Leaders : The Strategies for Taking Charge*, 1985, New York: Harper & Row

[19] Bettman, J., and Weitz, B., (1983), 'Attributions in the Board Room: Causal Reasoning in Corporate Annual Reports', *Administrative Science Quarterly*, 28: 165-183

[20] Brandenburger, A., Dekel, E., and Geanakoplos, J., (1992), 'Correlated Equilibrium with Generalized Information Structures', *Games and Economic Behavior*, 4: 182-201

[21] Burgelman, R., (1983), 'A Model of the Interaction of Strategic Behavior, Corporate Context, and the Context of Strategy', *Academy of Management Review*, 8: 61-70

[22] Burgelman, R., (1991), 'Intraorganizational Ecology of Strategy Making and Organizational Adaptation : Theory and Field Research', *Organization Science*, 2: 239-262

[23] Burgelman, R., (1994), 'Fading Memories : A Process Theory of Strategic Business Exit in Dynamic Environments', *Administrative Science Quarterly*, 39: 24-56

[24] Burns, J. M., (1978), *Leadership*, 1978, New York: Harper & Row

[25] Burns, T., and Stalker, G. M., (1961), *The Management of Innovation*, 1961, : Londen : Tavistock

[26] Capen, E., Clapp, R., and Campbell, W., (1971), 'Competitive Bidding in High-Risk Situations', *Journal of Petroleum Technology*, 23: 641-653

[27] Chatman, J., (1991), 'Matching People and Organizations: Selection and Socialization in Public Accounting Firms', *Administrative Science Quarterly*, 36: 459-484

[28] Collins, J. C., and Porras, J. I., (1994), *Built to last : Successful Habits of Visionary Companies*, 1994, New York: Harper & Row

[29] Conger, J. A., and Kanungo, R. N., (1988), *Charismatic Leadership*, 1988, San Francisco, CA: Jossey-Bass Publishers

[30] Cremer, J., (1993), 'Corporate Culture and Shared Knowledge', *Industrial and Corporate Change*, 3(3): 351-386

[31] Curren, M., Folkes, V., and Steckel, J., (1992), 'Explanation for Successful and Unsuccessful Marketing Decisions: The Decision Maker's Perspective', *Journal of Marketing*, 56: 18-31

[32] Donaldson, G., and Lorsch, J. W., (1983), *Decision Making at the Top : The Shaping of Strategic Direction*, 1983, : Basic Books, New York

[33] Einhorn, H. J., and Hogarth, R. M., (1978), 'Confidence in Judgment : Persistence in the Illusion of Validity', *Psychological Review*, 85: 395-416

[34] Ellison, G., (2000), 'Evolving Standards for Academic Publishing: A q-r Theory', Working paper

[35] Feinberg, Y., (2000), 'Characterizing Common Priors in the Form of Posteriors', *Journal of Economic Theory*, 91: 127-179

[36] Feinberg, Y., and Scarsini, M., (1999), 'Rate of Arbitrage and Reconciled Beliefs', Mimeo, Stanford University

[37] Fiske, S., and Taylor, S., (1991), *Social Cognition*, 1991, New York: McGraw-Hill

[38] Geanakoplos, J., and Polemarchakis, H., (1982), 'We Can't Disagree Forever', *Journal of Economic Theory*, 28: 192-200

[39] Goel, A., and Thakor, A., (2000), 'Rationality, Overconfidence, and Leadership', Working Paper, University of Michigan Business School

[40] Goodman, P., and Friedman, A., (1971), 'An Examination of Adams' Theory of Inequity, *Administrative Science Quarterly*, 16: 271-288

[41] Gul, F., (1998), 'A Comment on Aumann's Bayesian View', *Econometrica*, 66(4): 923-927

[42] Hammond, P., (1983), 'Ex-post Optimality as a Dynamically Consistent Objective for Collective Choice under Uncertainty', in *Social Choice and Welfare*, P. Pattanaik and M. Salles (eds.)

[43] Harris, M., and Raviv, A., (1993), 'Differences of Opinion Make a Horse Race', *Review of Financial Studies*, 6: 473-506

[44] Harrison, J., and March, J., (1984), 'Decision Making and Postdecision Surprises', *Administrative Science Quarterly*, 29: 26-42

[45] Harrison, M., and Kreps, D. M., (1978), 'Speculative Investor Behavior in a Stock Market with Heterogenous Expectations', *Quarterly Journal of Economics*, 92(2): 323-336

[46] Harsanyi, J. C., (1967/68), 'Games with Incomplete Information Played by 'Bayesian' Players', Parts I, II and III, *Management Science*, 14: 159-182

[47] Heider, F., (1958), *The Psychology of Interpersonal Relations*, 1958, New York: John Wiley & Sons

[48] Hermalin, B. E., (1998), 'Toward an Economic Theory of Leadership : Leading by Example', *American Economic Review*, 88(5): 1188-1206

[49] Hermalin, B. E., (1999), 'Leading for the Long-Term', Working Paper, Johnson Graduate School of Management, Cornell University

[50] Hermalin, B. E., (1999), 'Economics and Corporate Culture', Working paper, University of California-Berkeley

[51] House, R. J., (1977), 'A 1976 Theory of Charismatic Leadership', in Hunt, J.G. et al. (eds.) *Leadership : The Cutting Edge*

[52] Kahneman, D., and Tversky, A. (1973), 'On the Psychology of Prediction', *Psychological Review*, 80: 237-251

[53] Kaplan, R., (1986), 'Must CIM Be Justified by Faith Alone?', *Harvard Business Review*, (2): 87-93

[54] Knight, F., (1921), *Risk, Uncertainty and Profit*, 1921, Boston and New York:

[55] Korn/Ferry International, , and Columbia Univ. GSB, , (1989), *Reinventing the CEO : 21st Century Report*, 1989, : Korn/Ferry

[56] Kotter, J. P., (1990), *A Force for Change : How Leadership Differs from Management*, 1990, New York: Free Press

[57] Kotter, J. P., and Heskett, J. L., (1992), *Corporate Culture and Performance*, 1992, New York: Free Press

[58] Kouzes, J. M., and Posner, B. Z., (1987), *The Leadership Challenge : How to Get Extraordinary Things Done in Organizations*, 1987, San Francisco, CA: Jossey-Bass Publishers

[59] Kreps, D. M., (1990), 'Corporate Culture and Economic Theory', in Alt J.E., and K.A. Shepsle (eds.) *Perspectives on Positive Political Economy*

[60] Lazear, E. P., (1995), 'Corporate Culture and Diffusion of Values', in Siebert H. (eds) *Trends in Business Organization : Do Participation and Cooperation Increase Competitiveness*

[61] Lazear, E. P., (2001), 'The Peter Principle: Promotions and Declining Productivity', NBER Working Paper W9094.

[62] Locke, E. A. et al., (1991), *The Essence of Leadership*, 1991, : New York : Lexington

[63] Lord, C., Ross, L., and Lepper, M., (1979), 'Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence, *Journal of Personality and Social Psychology*, 37: 2098-2109

[64] March, J. G., (1991), 'Exploration and Exploitation in Organizational Learning', *Organization Science*, 2(1): 71-87

[65] Mechanic, D., (1962), 'The Sources of Power of Lower Participants in Complex Organizations', *Administrative Science Quarterly*, 7(3): 349-364

[66] Meyer, H., (1975), 'The Pay-for-Performance Dilemma, *Organizational Dynamics*, 3:

[67] Miles, E., Hatfield, J., and Huseman, R., (1989), 'The Equity Sensitivity Construct: Potential Implications for Worker Performance', *Journal of Management*, 15(4): 581-588

[68] Milgrom, P., and Roberts, J., (1990), 'The economics of modern manufacturing : Technology, strategy, and organization', *American Economic Review*, 80(3): 511-528

[69] Milgrom, P., and Roberts, J., (1992), *Economics, Organization and Management*, 1992, Englewood Cliffs, New Jersey: Prentice Hall

[70] Milgrom, P., and Stokey, N., (1982), 'Information, Trade and Common Knowledge', *Journal of Economic Theory*, 26: 17-27

[71] Miller, D., and Ross, M., (1975), 'Self-serving Biases in the Attribution of Causality: Fact or Fiction?', *Psychological Bulletin*, 82: 213-225

[72] Mongin, P., (1995), 'Consistent Bayesian Aggregation', *Journal of Economic Theory*, 66(2): 313-351

[73] Morris, S., (1995), 'The Common Prior Assumption in Economic Theory', *Economics and Philosophy*, 11: 227-253

[74] Morris, S., (1997), 'Risk, Uncertainty and Hidden Information', *Theory and Decision*, 42:

[75] Nanus, B., (1992), *Visionary Leadership : Creating a Compelling Sense of Direction for Your Organization*, 1992, San Francisco, CA: Jossey-Bass Publishers

[76] Nisbett, R., and Ross, L., (1980), *Human Inference : Strategies and Shortcomings of Social Judgment*, 1980, Englewood Cliffs, New Jersey: Prentice Hall

[77] Peter, L. J., and Hull, R., (1969),*The Peter Principle*, 1969, New York: W. Morrow

[78] Peters, T. J., and Waterman, R. H., (1982), *In Search of Excellence : Lessons from America's Best-run Companies*, 1982, New York: Harper & Row

[79] Prendergast, C., (1993), 'A Theory of "Yes Men"', *American Economic Review*, 83(4):

[80] Priest, G., and Klein, B., (1984), 'The Selection of Disputes for Litigation', *Journal of Legal Studies*, 13: 1-55

[81] Quigley, J. V., (1993), *Vision : How Leader Develop It, Share It, and Sustain It*, 1993, New York: McGraw-Hill

[82] Rabin, M., and Schrag, J., (1999), 'First Impressions Matter: A Model of Confirmation Bias', *Quarterly Journal of Economics*, 114(1): 37-82

[83] Robertson, A. G., and Walt, C. L., (1999), 'The Leader Within', *Outlook (Andersen Consulting)*, 11(2)

[84] Ross, M., and Sicoly, F., (1979), 'Egocentric Biases in Availability and Attribution', *Journal of Personality and Social Psychology*, 37: 322-336

[85] Rotemberg, J. J., and Saloner, G., (1993), 'Leadership Style and Incentives', *Management Science*, 39: 1299-1318

[86] Rotemberg, J. J., and Saloner, G., (1994), 'Benefits of Narrow Business Strategies', *American Economic Review*, 84(5): 1330-1349

[87] Rotemberg, J. J., and Saloner, G., (2000), 'Visionaries, Managers, and Strategic Direction', *Rand Journal of Economics*, 31(4): 693-716

[88] Saloner, G., Shepard, A., and Podolny, J., (2001), *Strategic Management*, 2001, New York: John Wiley & Sons

[89] Schein, E. E., (1984), 'Coming to a New Awareness of Organizational Culture', *Sloan Management Review*, 3-16

[90] Schein, E. H., (1985), *Organizational Culture and Leadership*, 1985, San Francisco, CA: Jossey-Bass Publishers

[91] Sebenius, J., and Geanakoplos, J., (1983), 'Don't Bet on It: Contingent Agreements with Asymmetric Information', *Journal of the American Statistical Association*, 78(382):

[92] Svenson, O., (1981), 'Are we all less risky and more skillful than our fellow drivers?', *Acta Psychologica*, 47: 143-148

[93] Thaler, R., (1988), 'Anomalies: The Winner's Curse', *Journal of Economic Perspectives*, 2: 191-202

[94] The Economist Intelligence Unit and (Korn/Ferry International) (1996), *Developing Leadership for the 21st Century*, 1996, :

[95] Thompson, S., and Kelley, H., (1981), 'Judgments of Responsibility for Activities in Close Relationships', *Journal of Personality and Social Psychology*, 41(3): 469-477

[96] Tichy, N. M., and Devanna, M. A., (1986), *The Transformational Leader*, 1986, New York: John Wiley & Sons

[97] Vickers, J., (1985), 'Delegation and the Theory of the Firm', *The Economic Journal*, :

[98] Wade-Benzoni, K., Thompson, L., and Bazerman, M., (2000), 'The Malleability of Environmentalism', Working paper, HBS

[99] Yildiz, M., (2000), *Essays on Sequential Bargaining*, PhD Dissertation, Stanford

[100] Zabojnik, J., (2000), 'A Model of Rational Bias in Self-Assessments', Working paper

[101] Zabojnik, J., (2001), 'Centralized and Decentralized Decision-Making in Organizations', *Journal of Labor Economics*, Forthcoming

[102] Zemsky, P., (1994), 'The Management of Organizational Capabilities : A Microeconomic Analysis', PhD Dissertation, Stanford University, Graduate School of Business

[103] Zuckerman, M., (1979), 'Attibution of success and failure revisited, or: The motivational bias is alive and well in attribution theory', *Journal of Personality*, 47: 245-287