EUROPEAN SUMMER SYMPOSIUM IN ECONOMIC THEORY

Generously hosted by
Study Center Gerzensee

Monday 29 June-Friday 10 July 2009

Priors and Desires: A Model of Payoff-Dependent Beliefs

Guy Mayraz (London School of Economics)

We are grateful to the following institutions for their financial and organizational support: Study Center Gerzensee and the Swiss National bank.

Priors and Desires
A Model of Payoff-Dependent Beliefs*

Guy Mayraz†

June 19, 2009

*I am grateful to Erik Eyster Richard Layard for their continual guidance and support, to Wolfgang Pesendorfer and Matthew Rabin for stimulating discussions and valuable suggestions at key points in the writing of the paper, and to seminar participants at LSE, MIT, UBC, UCL, Oxford, Essex, Royal Holloway, Warwick, Hebrew University, Tel-Aviv, Ben-Gurion, Bar-Ilan, Collegio Carlo-Alberto, and Maastricht.

†Centre for Economic Performance, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. http://personal.lse.ac.uk/mayraz. g.mayraz@lse.ac.uk.

Abstract

This paper introduces a decision-theoretic model of beliefs that allows for the possibility that what a person *believes* to be true is affected by what that person *wants* to be true. The substantive assumptions are (i) that distortion requires uncertainty, and that only payoffs over uncertain events matter, and (ii) that belief distortion over an event has to do with what a person has to gain or lose from the event being true. Using these assumptions I derive a simple formula with only one free parameter, which is positive for optimists and negative for pessimists. The key comparative statics are that belief distortion is greater in situation that are important and are where there is a great deal of uncertainty. The representation has the same structure as Bayes Rule, with payoffs playing the same role as normatively relevant information. A key implication is that news that affects the expected payoff consequences of an event may alter beliefs, even when it provides no relevant information about its likelihood.

JEL classification: D01,D03,D80,D81,D83,D84.
Keywords: Optimism, pessimism, subjective beliefs, decision under uncertainty, expectations.

Contents

1 Introduction **3**

2 Model **5**
2.1 Representation theorem 6
2.2 The logit formula 8
2.3 Proof and theorem variations 10

3 The comparative statics of belief distortion **17**
3.1 Relative optimism and pessimism 17
3.2 The comparative statics of uncertainty and stakes 19

4 Non-normative belief updating **20**
4.1 Formal analysis 23
4.2 Motivated cognition and cognitive dissonance 25

5 Conclusion **26**

1 Introduction

Consider the beliefs of parents about their child: would little Johnny get into a good school, marry well, and find a good job? It is common sense—and psychology research supports it—that parents do, in fact, tend to believe that their own little Johnny is more likely to do all these good things than the neighbours' child. The obvious reason, of course, is that parents are not *objective*—their happiness is strongly dependent on their child doing well, and this seems to affect their judgement.

Taking this example seriously suggests that there is a systematic link between a person's beliefs over events (e.g. whether little Johnny would have a successful marriage and obtain a good job) and that person's desire for these events to obtain. Such a link has potentially a great significance for economics. The assumption that beliefs are unbiased is important in many economic models, including not only (i) the optimality of static individual decisions, but also (ii) qualitative predictions in strategic settings—where it is important that the beliefs of different agents are not affected by their different interests, and (iii) dynamic settings—where it is important that beliefs change only as a function of normatively relevant information.

But a link between beliefs and motivation implies first that individuals and groups hold biased beliefs over any event that impacts their payoff (e.g. whether a war would break, whether a financial crisis is likely.). Second, that agents with inherently different interests (e.g. parties to bargaining, an insurance company and a customer) would hold systematically different beliefs from each other, and third, that a change of interests (e.g. as a result of a change in investments) causes a predictable change in beliefs even without any change in normatively relevant information.

Of course, even if such a link exists, it should only really matter to economic modelling of (i) we have empirical evidence that it has a significant effect in important settings, and (ii) we have a good way of modelling its effects. My goal in this paper is to tackle the second of these tasks, i.e. suggest an approach to modelling the effect motivation has on beliefs. The two tasks are nevertheless related, since the model makes clear empirical predictions, and clarifies what the potential implications are in particular settings. It can thus prove useful in facilitating empirical investigation, as well as theoretical modelling.

This modelling is taken in Section 2, and the rest of the paper is then taken with exploring some of the implications of the representation I prove.

The starting point for the model is that an agent has beliefs—modelled as a probability measure, and desires, modelled as a payoff-function, and that the two are related. The goal of the model is to provide a simple and useful description of the mapping that exists between the two. The paradigmatic situation where we expect beliefs to be distorted is where a person is uncertain about some event, and knows he would obtain a significantly higher payoff if the event occurs than if it does not. The assumptions I make allow for belief distortion in this kind of situation, but disallow distortion in certain other situations where we may not expect beliefs to be distorted. I assume (i) that distortion requires uncertainty, and that only payoffs over uncertain events matter, (ii) that large distortions cannot follow from arbitrarily small differences in payoffs, and (iii) that what matters to belief distortion is what a person has to gain or lose from an event being true, interpreted as payoff *differences*[1]. Using this assumption I derive the following representation, in which beliefs are defined using the odds ratio between pairs of constant-payoff events A and B, π_f denotes beliefs with f as the payoff-function, p denotes the beliefs the agent would hold he or she were completely indifferent about which state obtains. $f(B) - f(A)$ is the payoff difference between the two events, and ψ is a *belief-distortion parameter*:

$$\ln \frac{\pi_f(B)}{\pi_f(A)} = \ln \frac{p(B)}{p(A)} + \psi\big(f(B) - f(A)\big) \tag{1}$$

Payoff-dependent belief distortion can thus be described in a simple model, and with only one parameter. If this parameter is positive the agent is optimistic—the higher the payoff the higher the probability. If this parameter is negative the agent is *pessimistic*—the higher the payoff the *lower* the probability. The absolute size of this parameter denotes how optimistic or pessimistic the agent is. Note that the relative belief distortion in the two events depends only on 'local' properties of the two events. In this, as well as in its general form, Equation 1 is analogous to Bayes Rule when expressed in terms of the odds ratio between events. Indeed, it follows from Equation 1 that the beliefs of agents in the model are the same as those of agents, who really believe that payoffs carry relevant information about the events in question. This is very much *not* the right way to interpret real-world optimists and pessimists, but it does provide a convenient intuition to their beliefs. Thus, the beliefs of little Johnny's parents are *as if* they observe

[1] Representations are also provided for distortions that satisfy weaker assumption.

that they really want little Johnny to succeed, but are indifferent about the neighbours' child, and *conclude* (other things being equal) that little Johnny is therefore indeed more likely to do well.

Section 3 considers the comparative statics of the model. To begin with I show that the distribution of payoffs of an agent with a more positive (or less negative) belief distortion parameter stochastically dominates in the likelihood ratio that of an agent with a less positive belief distortion parameter. In particular, relatively more optimistic agents expect payoff to be higher. Later in that section I show the two most important comparative statics: (i) that belief distortion is greater when there is more uncertainty, and (ii) that there is more belief distortion when the stakes are higher. Thus, belief distortion is at its peak in situations that combine importance with uncertainty.

Section 4 turns to one of the most important implications of payoff-dependent belief distortion: that information that affects anticipated payoffs can change beliefs, even if it is normatively irrelevant. Thus, a person learning about a new and promising relationship may become much less positive about the prospects of an existing relationship, simply because the new relationship reduces the desire for the first relationship to succeed. Similarly, parents who learn which school their child is to be allocated to revise their beliefs about that school, simply because the new information makes them care a lot more about that school, and also revise their beliefs in the opposite direction about other schools that are no longer relevant. More generally, I prove that complementarities in the payoff function (e.g. that the two relationships are substitutes) translates into predictable changes in beliefs in one variable given positive news about the other.

2 Model

The premise of this paper is that what a person *believes* to be true may be affected by what the person *wants* to be true. This section develops a representation of this dependence. In order to make the model as general as possible, I do not tie it into any particular model of choice. Instead, I simply assume the existence of beliefs and desires, and seek to model how the former depends on the latter.

I model the domain of beliefs by a set of *states*, each of which represents alternative complete descriptions of the domain in question, and where exactly one state is true (exactly one state *obtains*). The desire for a particular state to be true is modelled by a *payoff-function*, which associates with each state a real number, representing the agent's desire for that state to be the true state. The assumption that beliefs may depend on desires is modelled by a *distortion mapping*, which associates a probability measure to each payoff-function. *Everything else is assumed to be constant.* In particular, the relevant information that the agent has is held constant. The distortion mapping thus represents the *ceteris paribus* effect of desires on beliefs.

The goal is a parsimonious representation of the distortion mapping: one that uses a simple equation with few easily interpretable parameters to characterise the dependence of beliefs on payoffs. There is obviously no such representation for arbitrary distortion mappings. Any hope for a parsimonious representation thus lies with the observation that the belief distortion that occurs in practice is clearly not completely arbitrary. The paradigmatic situation where we expect beliefs to be distorted (assuming they are distorted at all) is where a person is uncertain about some event, and knows he would obtain a significantly higher payoff if the event occurs than if it does not. The assumptions I propose allow for belief distortion in this kind of situation, but disallow distortion in certain other situations where we may not expect beliefs to be distorted. I assume (i) that distortion requires uncertainty, and that only payoffs over uncertain events matter, (ii) that large distortions cannot follow from arbitrarily small differences in payoffs, and (iii) that what matters is what a person has to gain or lose from an event being true, i.e. payoff *differences*.

The representation that is obtained characterises the *ceteris paribus* effect of payoffs. The model does not (and cannot) characterise the relationship between beliefs and objective reality. In applications that otherwise assume that people hold unbiased beliefs it may be natural to assume instead that the agent is unbiased if he or she happens to be indifferent between all states. However, any such assumptions lie outside the scope of the model.

2.1 Representation theorem

Let S denote the state space, and let $F \subset \{f : S \to \mathbb{R}\}$ denote the set of all simple payoff-functions (for a simple payoff-function f, the set of its payoffs $I_f = \{a \in \mathbb{R} : \exists s \in S, f(s) = a\}$ is finite). Let Σ denote the Borel σ-algebra on S, and let Δ denote the set of probability measures on Σ. The probability measure $\mu \in \Delta$ is *non-atomic* if for every event A for which $\mu(A) > 0$ there exists an event $B \subset A$ such that $0 < \mu(B) < \mu(A)$. For $\mu \in \Delta$ and an event A with $\mu(A) > 0$ let $\mu(\cdot|A)$ denote the conditional probability measure on A. I write f, f' for generic payoff-functions, a for a constant payoff-functions that yields utility a in every state.

The primitive of the model is a distortion map $\pi : F \to \Delta$ that associates a non-atomic probability measure with each payoff-function in F. The interpretation is that π_f represents the beliefs the agent would hold if her payoffs were f. The first definition lists some basic properties we may want this distortion to satisfy. The second definition describes the logit formula. The theorem says that the two definitions are equivalent.

Definition 1. $\pi : F \to \Delta$ is a *well-behaved distortion* if the following conditions are satisfied:

A1 (absolute continuity) $\pi_{f'}(A) = 0 \iff \pi_f(A) = 0$.

A2 (independence) If $f = f'$ over a non-null A then $\pi_{f'}(\cdot|A) = \pi_f(\cdot|A)$.

A3 (shift invariance) If $f' = f + a$ then $\pi_{f'} = \pi_f$.

A4 (continuity) If $f_n \to f$ uniformly then $\pi_{f_n}(A) \to \pi_f(A)$.

Both Absolute Continuity and Independence tie belief distortion to subjective uncertainty. Absolute Continuity restricts belief distortion to uncertain events. Independence restricts the set of payoffs that may affect beliefs to states that are consistent with the agent's evidence. The intuition for Independence is that the anticipated gain or loss from an event being true is a function only of payoff in states with positive probability, and hence payoff in states that are inconsistent with the agent's information should not affect belief distortion. Independence has a rough analogue in Luce's Choice Axiom (Luce, 1959)[2]. Shift-invariance states that belief distortion depends only on payoff *differences*, rather than absolute levels. The intuition is that the desire for an event to obtain is a function of what the agent has to gain or lose from the event being true, and that gains and losses correspond to payoff differences. Finally, Continuity is the assumption that significant differences in beliefs can only result from significant differences in payoffs.

[2]Luce's Choice Axiom is defined in the context of a theory of probabilistic choice. It requires that the odds ratio for choosing x over y be independent of what other alternatives are in the choice set.

Definition 2 (Logit Distortion). $\pi : F \to \Delta$ is a *logit distortion* if (i) it satisfies Absolute Continuity, and (ii) there exist a probability measure p (the *undistorted measure*) and a parameter $\psi \in \mathbb{R}$ (the *belief distortion parameter*), such that for all non-null[3] events A and B that have constant payoff under f the following condition holds:

$$\ln \frac{\pi_f(B)}{\pi_f(A)} = \ln \frac{p(B)}{p(A)} + \psi\big(f(B) - f(A)\big) \tag{2}$$

Theorem 1. *$\pi : F \to \Delta$ is a logit distortion if and only if it is a well-behaved distortion.*

2.2 The logit formula

Equation 2 defines π_f via the odds ratio between pairs of constant payoff events. To understand the role of the undistorted probability measure p suppose f is constant, so that the agent is completely indifferent as to which state obtains. In this case $\pi_f = p$, and so p corresponds to the beliefs in conditions of indifference—when the person is *disinterested*. In general, of course, the agent does have a stake in which state obtains, in which case the equation defines distorted beliefs relative to what they would have been if the agent were disinterested. It is important to note that although we do say in common language that a person is *objective* when he or she is disinterested, this sense of being objective is logically unrelated to any notion of objective probability. Indeed, although disinterested beliefs are a natural reference point, no particular relationship between them and objective reality is assumed[4].

Equation 2 is analogous to Bayes Rule when the latter is written in terms of log-odds between pairs of events:

$$\ln \frac{p(B|I)}{p(A|I)} = \ln \frac{p(B)}{p(A)} + \ln \frac{p(I|B)}{p(I|A)}, \tag{4}$$

[3] Absolute Continuity guarantees that the term "non-null" is well-defined without having to specify the payoff-function.

[4] The notion of the undistorted probability measure p is convenient, but is not necessary. An alternative formulation of Equation 2 describes the relationship between two distorted probability measures π_f and π_g with no reference to p:

$$\log \frac{\pi_f(B)}{\pi_f(A)} = \log \frac{\pi_g(B)}{\pi_g(A)} + \psi\Big(\big(f(B) - f(A)\big) - \big(g(B) - g(A)\big)\Big) \tag{3}$$

with p and π_f standing respectively for prior and posterior beliefs, and $\psi f(A)$ and $\psi f(B)$ for the log likelihoods of A and B respectively. The reason Equation 2 takes a particularly simple form when formulated in terms of odds-ratios is because belief-distortion depends only on payoff-*differences*, rather than on absolute payoffs. Similarly, Bayes Rule is simpler when expressed in terms of odds-ratios because belief-update depends only on differences of log-likelihood, rather than absolute levels (which are, indeed, meaningless).

Computing probabilities, rather than odds, can be done—as in Bayes Rule—by partitioning the state-space. An arbitrary constant-payoff event E can be completed into a constant-payoff event partition $\{E_i\}_{i=1}^m$ with $E_1 = E$ and $\cup_i E_i = S$. The distorted probability of E can then be computed using the following formula, analogous to computing posterior probabilities using Bayes Rule:[5]:

$$\pi_f(E) = \frac{p(E)e^{\psi f(E)}}{\sum_i p(E_i)e^{\psi f(E_i)}}. \tag{5}$$

The distorted probability of a generic event F can be computed from the distorted probability of its constituent constant-payoff events.

Equation 2 makes it clear that belief distortion is a function of payoff-differences, which should be interpreted as the *stakes* the agent has in one event rather than another being true. Payoff-differences are weighted by the belief-distortion parameter ψ, which determines both the direction that they affect beliefs, and the importance of their role. If $\psi > 0$ the agent's beliefs are distorted in the direction of events that make the agent better-off. If, instead, $\psi < 0$ the agent's beliefs are distorted in the direction of events that make the agent *worse* off. Finally, if $\psi = 0$ the agent is *objective*—not in the sense of being necessarily correct, but in the sense that beliefs are unbiased by what the agent wants to be true. ψ is defined relative to the representation of payoffs—if payoffs are scaled, ψ has to be scaled in the inverse direction in order for it to represent the same belief distortion mapping.

Equation 2 describes the distorted beliefs of the agent as the linear combination of undistorted beliefs and the weighted stakes in the event being B rather than A. The same distorted beliefs can thus result from different combinations of undistorted beliefs and stakes. For example, an optimist may believe B if undistorted beliefs favour B and the stakes are low, but also

[5]For an independent proof note that by Equation 2, $\pi_f(E_i)/\pi_f(E_1) = (p(E_i)/p(E_1)) \cdot e^{\psi(f(E_i)-f(E_1))}$. Multiplying both sides by $\pi_f(E_1)$ and summing over i yields 1 on the LHS because $\{E_i\}_{i=1}^m$ is a partition. Rearranging yields Equation 5.

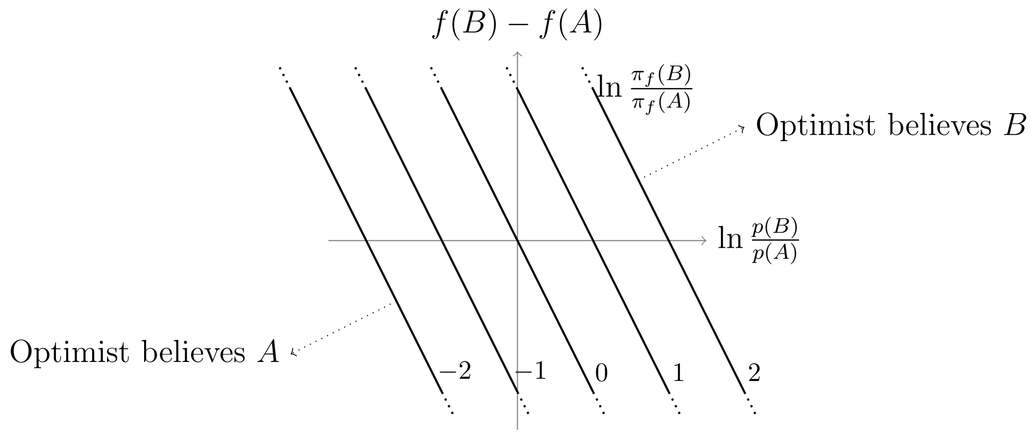


Figure 1: Iso-belief lines as a function of undistorted beliefs and the stakes in the event being B rather than A. Undistorted beliefs are plotted on the x-axis and the stakes are plotted on the y-axis. Iso-belief lines are straight with slope $1/\psi$. This figure is plotted for an optimist. Iso-belief lines for a pessimist slope in the opposite direction, while those of a neutral agent are vertical.

when undistorted beliefs favour A, but are highly uncertain, and the person has large stakes in B being the case (Figure 1).

2.3 Proof and theorem variations

In this section I present the proof of Theorem 1, focusing on the role of different assumptions. I start with the following weaker representation:

Definition 3. $\pi : F \to \Delta$ is a *weak logit distortion* if (i) it satisfies Absolute Continuity, and (ii) there exists a continuous function $v : \mathbb{R} \to \mathbb{R}$ (the *belief distortion function*), such that for all non-null events A and B that have constant payoff under f the following condition holds:

$$\ln \frac{\pi_f(B)}{\pi_f(A)} = \ln \frac{\pi_{f(B)}(B)}{\pi_{f(A)}(A)} + v(f(B)) - v(f(A)) \tag{6}$$

where $\pi_{f(B)}$ and $\pi_{f(A)}$ denote the probability measure associated with the constant payoff-functions yielding $f(B)$ and $f(A)$ respectively.

This representation shares the same structure as that of Definition 2. Equation 2 can be seen as a special case of Equation 6 in which (i) $\pi_a = p$ for all $a \in \mathbb{R}$, and (ii) v is linear (ψ is its slope). As the following theorem shows the representation in Definition 3 depends only on Absolute Continuity and Independence.

Theorem 2. *$\pi : F \to \Delta$ is a weak logit distortion if and only if it satisfies Absolute Continuity and Independence:*

Proof. Absolute Continuity is part of both definitions. In the first direction I need to show that there exists a belief-distortion function v such that for any payoff-functions f and for all non-null constant-payoff events A and B with payoffs a and b respectively,

$$\ln \frac{\pi_f(B)}{\pi_f(A)} = \ln \frac{\pi_b(B)}{\pi_a(A)} + v(b) - v(a). \tag{7}$$

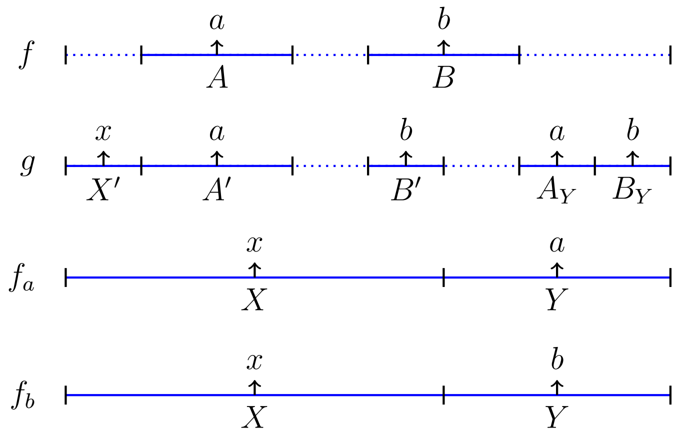
The proof is constructive. Let X be some event such that neither X nor $Y = X^c$ are null, and let $x \in \mathbb{R}$ be some real number. For $y \in \mathbb{R}$ define a payoff-function f_y by $f(X) = x$ and $f(Y) = y$. v is then defined by solving Equation 6 for f_y on the assumption that $v(x) = 0$, i.e:

$$v(y) = \ln \frac{\pi_{f_y}(Y)}{\pi_{f_y}(X)} - \ln \frac{\pi_y(Y)}{\pi_x(X)} \tag{8}$$

Having defined v I need to show that Equation 7 indeed holds, or after rearranging, that

$$\frac{\pi_f(B)}{\pi_b(B)} \Big/ \frac{\pi_f(A)}{\pi_a(A)} = \frac{e^{v(b)}}{e^{v(a)}} \tag{9}$$

The proof involves relating the LHS of Equation 9 to the RHS of Equation 8 for $y = a$ and $y = b$. This is done with the help of non-null auxiliary events $X' \subseteq X, A' \subseteq A, B' \subseteq B$ and $A_y, B_Y \subseteq Y$, and an auxiliary payoff-function g, such that $f(A) = f(A') = g(A') = g(A_Y) = f_a(A_y)$, $f(B) = f(B') = g(B') = g(B_Y) = f_b(B_y)$, and $g(X') = f_a(X') = f_b(X') = f_a(X) = f_b(X)$. See example illustration below:



The reason this construction helps is because of two implications of Independence: (i) that if two payoff-functions agree over two constant-payoff events then the corresponding probability measures agree about the odds ratio between the two events, and (ii) that if two events have the same payoff then their odds ratio is not distorted relative to the corresponding constant payoff-function. This is formalised in the following two lemmas:

Lemma 1. *Let f_1 and f_2 be two payoff-functions and C and D non-null events, and suppose $f_1(C) = f_2(C) = c$ and $f_1(D) = f_2(D) = d$ then $\pi_{f_1}(D)/\pi_{f_1}(C) = \pi_{f_2}(D)/\pi_{f_2}(C)$.*

Proof. The proof is a straightforward application of Independence with $C \cup D$ standing for the event A in the definition of Independence:

$$\frac{\pi_{f_1}(D)}{\pi_{f_1}(C)} = \frac{\pi_{f_1}(D|C \cup D)}{\pi_{f_1}(C|C \cup D)} = \frac{\pi_{f_2}(D|C \cup D)}{\pi_{f_2}(C|C \cup D)} = \frac{\pi_{f_2}(D)}{\pi_{f_2}(C)} \tag{10}$$

□

Lemma 2. *Let h be a payoff-function and C and C' non-null events such that $h(C) = h(C') = c$ then $\pi_h(C)/\pi_c(C) = \pi_h(C')/\pi_c(C')$.*

Proof. The result is straightforward rearranging of the result in Equation 10 with $f_1 = h$ and $f_2 = c$ (the payoff-function yielding a constant payoff of c). □

Given the construction and these two lemmas the main claim follows from the following series of equalities:

$$\begin{aligned}
\frac{\pi_f(B)/\pi_b(B)}{\pi_f(A)/\pi_a(A)} &= \frac{\pi_f(B')/\pi_b(B')}{\pi_f(A')/\pi_a(A')} = \frac{\pi_g(B')/\pi_b(B')}{\pi_g(A')/\pi_a(A')} = \frac{\pi_g(B'_Y)/\pi_b(B'_Y)}{\pi_g(A'_Y)/\pi_a(A'_Y)} \\
&= \frac{\pi_g(B'_Y)/\pi_b(B'_Y)}{\pi_g(X')/\pi_x(X')} \Big/ \frac{\pi_g(A'_Y)/\pi_a(A'_Y)}{\pi_g(X')/\pi_x(X')} \\
&= \frac{\pi_{f_b}(B'_Y)/\pi_b(B'_Y)}{\pi_{f_b}(X')/\pi_x(X')} \Big/ \frac{\pi_{f_a}(A'_Y)/\pi_a(A'_Y)}{\pi_{f_a}(X')/\pi_x(X')} \\
&= \frac{\pi_{f_b}(Y)/\pi_b(Y)}{\pi_{f_b}(X)/\pi_x(X)} \Big/ \frac{\pi_{f_a}(Y)/\pi_a(Y)}{\pi_{f_a}(X)/\pi_x(X)} \\
&= \frac{e^{v(b)-v(x)}}{e^{v(a)-v(x)}} = e^{v(b)-v(a)},
\end{aligned} \tag{11}$$

where the first, fourth and sixth steps apply Lemma 2, the second and fifth use Lemma 1, and the seventh follows from the definition of v.

To complete the proof I need to show that such a construction is possible. This requiring a proof that the five auxiliary events can be defined, such that they are non-null and that the definition of g is consistent. The latter condition requires in addition to previous requirements that $(A' \cup A'_Y) \cap (X' \cup B' \cup B'_Y) = \emptyset$, and $(B' \cup B'_Y) \cap (X' \cup A' \cup A'_Y) = \emptyset$. There are three cases to consider:

1. Both $A \cap Y$ and $B \cap Y$ are not null. Let $X' = X, A' = A'_Y = A \cap Y$ and $B' = B'_Y = B \cap Y$.

2. $A \cap Y$ is not null, but $B \cap Y$ is. Since $Y = X^c$ it follows that $B \cap X$ is not null. By non-atomicity there are disjoint non-null events $A_1, A_2 \subseteq A \cap Y$, and $B_1, B_2 \subseteq B \cap X$. Define $A' = A'_Y = A_1$, $B' = B_1$, $X' = B_2$ and $B'_Y = A_2$.

3. Both $A \cap Y$ and $B \cap Y$ are null, and so both $A \cap X$ and $B \cap X$ are not null. By non-atomicity there are disjoint non-null events $A_1, A_2 \subseteq A \cap X$, and $Y_1, Y_2 \subseteq Y$. Define $A' = A_1$, $X' = A_2$, $B' = B \cap X$, $A'_Y = Y_1$ and $B'_Y = Y_2$.

In all three cases it is easy to see that the definitions are consistent with requirements. The construction is thus valid, and the proof complete.

Finally, for the only-if direction I need to show that Independence follows from Equation 6, i.e. that for all non-null events A, and payoff-functions f and f' that agree on A, $\pi_f(\cdot|A) = \pi_{f'}(\cdot|A)$. The distorted conditional probability of a generic event can be computed from that of its constituent constant-payoff events. It is thus sufficient to show that for non-null constant payoff-events $B \subseteq A$, $\pi_f(B|A) = \pi_{f'}(B|A)$. To see this, complete B into a partition $\{B_i\}_{i=1}^n$ of the state-space, with $B_1 = B$ and such that $A = \cup_{i=1}^m B_i$. Then $f'(B_i) = f(B_i)$ for $i \leq m$. Using Equation 5:

$$\begin{aligned}\pi_f(B|A) &= \frac{\pi_f(B)}{\pi_f(A)} = \frac{p(B)e^{\psi f(B)}}{\sum_{i=1}^n p(B_i)e^{\psi f(B_i)}} / \frac{\sum_{i=1}^m p(B_i)e^{\psi f(B_i)}}{\sum_{i=1}^n p(B_i)e^{\psi f(B_i)}} \\ &= \frac{p(B)e^{\psi f(B)}}{\sum_{i=1}^m p(B_i)e^{\psi f(B_i)}} = \frac{p(B)e^{\psi f'(B)}}{\sum_{i=1}^m p(B_i)e^{\psi f'(B_i)}} = \frac{\pi_{f'}(B)}{\pi_{f'}(A)} = \pi_{f'}(B|A)\end{aligned} \tag{12}$$

□

Before considering the implication of Shift-invariance I introduce a weaker assumption, which captures the idea that only differences in payoff matter to belief distortion, but does not assume that differences form an interval scale:

A3' (impartiality) $\pi_a = \pi_b$ for constant payoff-functions a and b.

The implication of adding Impartiality is obvious:

Proposition 1. *Suppose $\pi : F \to \Delta$ is a weak logit distortion then the following two conditions are equivalent: (i) π satisfies Impartiality, and (ii) there exists a probability measure p such that $\pi_a = p$ for all $a \in \mathbb{R}$, and so*

$$\ln \frac{\pi_f(B)}{\pi_f(A)} = \ln \frac{p(B)}{p(A)} + v(f(B)) - v(f(A)) \tag{13}$$

The additional implication of Shift-invariance is to establish a series of linear constraints on the belief-distortion function v:

Proposition 2. *Suppose $\pi : F \to \Delta$ is a weak logit distortion then the following two conditions are equivalent: (i) π satisfies Shift-invariance, and (ii) there exist $a, b \in \mathbb{R}$, such that $a < b$, and for all $c, d \in \mathbb{R}$ for which $(d - c)/(b - a)$ is a rational number,*

$$v(d) - v(c) = \left(\frac{v(b) - v(a)}{b - a}\right)(d - c) \tag{14}$$

and in addition there exists a probability measure p such that $\pi_a = p$ for all $a \in \mathbb{R}$.

Proposition 2 *almost* establishes that v is linear, but not quite. The catch is that the ratio of $d - c$ to $b - a$ has to be *rational.* Let $=_{\mathbb{Q}}$ denote the equivalence relation defined by $x =_{\mathbb{Q}} x' \iff \exists r \in \mathbb{Q} : x' = x + r$. Then for each real number x, $=_{\mathbb{Q}}$ associates an equivalence class $x_{\mathbb{Q}}$, and when limited to $x_{\mathbb{Q}}$ Proposition 2 implies that v is linear. There is, however, no necessary relationship between v on different equivalence classes. Each $x_{\mathbb{Q}}$ is thus associated with its own and possibly different belief distortion parameter $\psi(x_{\mathbb{Q}})$.

One way of ensuring that v is linear over $\mathbb{R}$, and hence that π is a logit-distortion (Definition 2), is to assume continuity. This works as $\mathbb{Q}$ is dense in $\mathbb{R}$[6]. The first step is the equivalence of the Continuity assumption and the continuity of v:

Proposition 3. *Suppose $\pi : F \to \Delta$ is a weak logit distortion that satisfies Impartiality then π satisfies Continuity if and only if the belief-distortion function v in Equation 6 is continuous.*

The second step proves the claim. Theorem 1 follows as an immediate corollary.

Proposition 4. *Suppose $\pi : F \to \Delta$ is a weak logit distortion that satisfies Shift-invariance then π is a logit-distortion if and only if π satisfies Continuity.*

Continuity is not the only assumption that ensures v is linear. Another alternative assumption is monotonicity:

- A4' (monotonicity) Suppose $a \leq b \leq c$ then for all x either $\pi_{xAa}(A) \leq \pi_{xAb}(A) \leq \pi_{xAc}(A)$, or $\pi_{xAa}(A) \geq \pi_{xAb}(A) \geq \pi_{xAc}(A)$.

Since $\mathbb{Q}$ is dense in $\mathbb{R}$ each real number x can be approached both from below and from above by monotonic sequences of rational numbers, and hence the fact that v is linear over $x_{\mathbb{Q}}$ for all $x \in \mathbb{R}$ implies that v is continuous over $\mathbb{R}$.

Proposition 5. *Suppose $\pi : F \to \Delta$ is a weak logit distortion that satisfies Impartiality then π satisfies Monotonicity if and only if the belief-distortion function v in Equation 6 is monotonic.*

[6] $x_{\mathbb{Q}}$ is dense in $\mathbb{R}$ for all $x \in \mathbb{R}$.

Proposition 6. *Suppose $\pi : F \to \Delta$ is a weak logit distortion that satisfies Shift-invariance then π is a logit-distortion if and only if π satisfies Monotonicity.*

To summarise, Absolute Continuity and Independence together are equivalent to weak logit distortion. Adding Impartiality is equivalent to $\pi_a = p$. Continuity and Monotonicity are equivalent respectively to continuity and monotonicity of v. Finally, Shift-invariance is stronger than Impartiality, and together with *either* continuity or monotonicity implies that π is a logit-distortion[7].

[7] Thus, Continuity and Monotonicity are equivalent in the presence of Absolute Continuity, Independence and Shift-invariance.

3 The comparative statics of belief distortion

Payoff-dependent belief distortion affects beliefs as a function of the anticipated payoff of the events in question. Optimists are biased to believe higher payoffs are likely, while pessimists are biased to expect a lower payoff. In this section I consider the comparative statics of this effect. I start with some necessary definitions, including the distorted and undistorted distribution and expectation of payoffs, and the bias in expected payoff. I then consider the comparative statics of the belief distortion parameter ψ, showing that Equation 2 captures relative optimism and pessimism. I then consider a particular agent with a given belief-distortion parameter, and consider the twin comparative statics of (i) the degree of uncertainty and (ii) the stakes the agent has in what event obtains.

Definition 4 (Distribution of payoffs from a payoff-function)**.** For a payoff-function f let X_f denote the real random variable defined by $X_f(s) = f(s)$. The *undistorted payoff distribution* $F_p(x)$, is the probability distribution of X_f using the undistorted probability measure p. That is, $F(x) = p(X_f \leq x)$. The *distorted payoff distribution* $F_\pi(x) = \pi_f(X_f \leq x)$ is the probability distribution of X_f using the distorted probability measure π_f. The undistorted (distorted) *expected payoff* $\mathcal{E}_p(f)$ and $\mathcal{E}_\pi(f)$ are similarly defined. The *bias in expected payoff* is the difference between the distorted and undistorted expectations: $\mathcal{E}_\pi(f) - \mathcal{E}_p(f)$.

3.1 Relative optimism and pessimism

If the belief distortion parameter ψ is capture the notions of optimism and pessimism, we would like agents with a higher ψ to anticipate a higher payoff. Consider first the case of two events A and B. It follows from Equation 2 that the difference in the log odds ratio between B and A for agents with belief parameters ψ_1 and ψ_2 is

$$\ln \frac{\pi_f^1(B)}{\pi_f^1(A)} = \ln \frac{\pi_f^2(B)}{\pi_f^2(A)} = (\psi_1 - \psi_2) \cdot \big(f(B) - f(A)\big) \tag{15}$$

and so, assuming without loss of generality that B is the better event (i.e. $f(B) \geq f(A)$) then the better event is subjectively more likely for the agent with belief parameter ψ_1 if and only if $\psi_1 \geq \psi_2$.

The binary case is actually as general as we would like, as it implies stochastic dominance in the likelihood ratio, defined below for distributions with the same support (this is always the case because of the continuity of Equation 2):

Definition 5. Suppose $F(x)$ and $G(x)$ are two probability distributions with the same support then $F(x)$ *stochastically dominates* $G(x)$ *in the likelihood ratio*, written $F \succeq_{LR} G$, if $dF(x)/dG(x)$ is non-decreasing for all x in the support of F and G.

Stochastic-dominance in the likelihood ratio is a stronger concept than first-order stochastic dominance. In particular it follows that relative more optimistic agents always expect a higher payoff. These results are summarised in the following proposition:

Proposition 7 (Relative optimism)**.** *Let f be any payoff-function, and let F_π^ψ denote the distorted distribution of the payoff of f for an agent with a distortion parameter ψ. Suppose $\psi_1 \geq \psi_2$ then $F_\pi^{\psi_1} \succeq_{LR} F_\pi^{\psi_2}$.*

As ψ increases (decreases) more and more probability is assigned to better (worse) events. In the limit of $\psi \to \infty$ *all* the probability is put on the best possible outcome. That is, an extremely optimistic agent believes in the best possible world for him. Similarly, an extreme pessimist believes in the *worst* possible world. Note that extreme optimism and pessimism are not themselves logit distortions (they violate absolute continuity, as events resulting in other than the extreme payoff are distorted to zero probability). Yildiz (2007) explores a model in which agents behave like extreme optimists in this sense.

Proposition 8 (Extreme optimism/pessimism)**.** *Let f be a simple payoff-function. Define $a_{min} = \min_{s \in Supp(p)} f(s)$ and $a_{max} = \max_{s \in Supp(p)} f(s)$ to be the minimal and maximal possible prizes respectively, and let $A_{min} = f^{-1}(a_{min})$ and $A_{max} = f^{-1}(a_{max})$ denote the events that these prizes are obtained. Then $\lim_{\psi \to -\infty} \pi_f(A_{min}) = \lim_{\psi \to \infty} \pi_f(A_{max}) = 1$.*

3.2 The comparative statics of uncertainty and stakes

Equation 2 can be written qualitatively as follows:

$$\text{Beliefs} = \text{Undistorted beliefs} + \psi \cdot \text{Stakes} \tag{16}$$

Actual beliefs are thus a tug-of-war between undistorted beliefs on the one hand, and what the agent has at stake on the other. The implied comparative statics for the belief bias are that the bias is bigger (i) the more is at stake, and (ii) the weaker undistorted beliefs are, i.e. the greater is the (undistorted) uncertainty. Belief distortion is greatest, therefore, in a situations that (a) are important, and (b) where there is a great deal of uncertainty.

That this is so can be seen formally most clearly in two simple cases: (i) binary payoffs, and (ii) normally distributed payoffs. In the binary case suppose payoff is $f(E) = v$ if some event E obtains and $f(E^c) = 0$ otherwise. The stakes are thus v and uncertainty is high if $0 \ll p(E) \ll 1$. Using Equation 5 the bias in expected utility terms is given by

$$\begin{aligned} \mathcal{E}_f(f) - \mathcal{E}(f) = (\pi_f(E) - p(E)) \cdot v = \left(\frac{p(E)e^{\psi v}}{1 - p(E) + p(E)e^{\psi v}} - p(E) \right) v \\ = \frac{(e^{\psi v} - 1)p(E)(1 - p(E))}{1 + p(E)(e^{\psi v} - 1)} \cdot v \end{aligned} \tag{17}$$

This expression is, of course, positive (negative) if ψ is positive (negative). As the above qualitative argument suggests, the bias is increases in the stakes v, and decreases as undistorted beliefs approach certainty ($p(E) \to 0$ or $p(E) \to 1$). In the limit of a small belief-distortion parameter ($\psi \approx 0$) the bias is greatest when uncertainty is greatest ($p(E) = 0.5$)[8].

For normally distributed payoffs I first prove the following independently useful proposition:

Proposition 9 (Normally distributed payoffs)**.** *Let f be a payoff-function, such that $F_p(x)$ is normal with mean μ and variance σ^2. Then $F_\pi(x)$ is also normal with the same variance σ^2, but with mean shifted to $\mu + \sigma^2\psi$.*

[8]For $|\psi| \gg 0$ the bias is larger if $p(E)$ is smaller (larger) depending on whether ψ is positive (negative). Intuitively there is more room for a positive (negative) bias when undistorted beliefs are low (high). This is irrelevant when $\psi \approx 0$, and so for such ψ the bias is then greatest when $p(E) \approx 0.5$.

To illustrate the effects of the stakes suppose payoff is proportional to a normally distributed random variable X, i.e. $f(s) = aX(s)$, and let σ^2 denote the undistorted variance of X. Then F_p is normally distributed with variance $a^2\sigma^2$. It follows from Proposition 9 that the bias in distorted beliefs is $\psi a^2\sigma^2$, and is thus increasing both in the stakes (a) and in the level of uncertainty (σ). The comparative statics for the normal case are illustrated in Figure 2.

The evidence for the comparative statics of uncertainty is overwhelming (Kunda, 1990). For the comparative statics of stakes see Weinstein (1980) and Sjöberg (2000). There is also good evidence that *changes* in stakes can cause *changes* in beliefs. The theory and evidence for this are explored in Section 4.

It is interesting to note that models of anticipatory preferences (Akerlof and Dickens, 1982; Brunnermeier and Parker, 2005) produce very different comparative statics. The most obvious comparative statics of such models has to do with the length of time in which the agent enjoys anticipatory utility: the longer this time, the stronger bias. Anticipatory utility models have no obvious comparative statics for the stakes the agent has in a given situation. In fact, if anticipatory utility is increasing less than linearly in the stakes the prediction would be for the *opposite* comparative statics. Finally, the comparative statics for information are unclear, and depend on whether a cost of information distortion is introduced. If there are no such costs then there are cases in which anticipatory utility models would predict *total* bias[9].

4 Non-normative belief updating

Beliefs in the model depend not only on normatively relevant information but also on payoffs. A change in payoffs can therefore lead to a change in beliefs even if there is no change in relevant information. The key is for there to be complementarities in the payoff-function, so that news about one variable affects the anticipated payoff from another. When this is the case, news about the first variable may affect beliefs about the other even if it is normatively irrelevant. I start with a couple of examples:

Example 1. relationship

[9] It makes a big difference for anticipatory utility models whether the decision the agent has to take is continuous or discrete. If the latter, then depending on the optimal trade-off, the agent may either choose just enough bias so as not to make a costly error, or else accept the error and enjoy as much bias as possible.

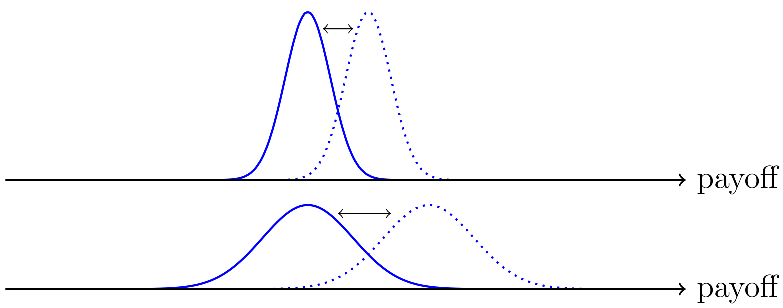


Figure 2: Belief distortion with normally distributed payoffs. The effect of belief distortion is to shift the payoff distribution to the right or left depending on whether the agent is optimistic or pessimistic. The size of the shift is proportional to uncertainty, as measured by the variance of the distribution.

Consider the beliefs of an extremely optimistic person who has is interested in having a successful relationship, and at the moment has two prospective relationships of uncertain quality. The payoff-function is therefore as follows:

	B fails	B works
A fails	0	1
A works	1	1

The key to this example is that the two relationships are substitutes: if one succeeds, the person no longer cares about the prospects for the other. Suppose undistorted beliefs are that each prospective relationship has a probability 1/4 of success, and let $\psi \to \infty$ so as to model extreme optimism. With this extreme level of optimism undistorted and distorted beliefs are as follows:

	B fails	B works	
A fails	9/16	3/16	3/4
A works	3/16	1/16	1/4
	3/4	1/4	

undistorted beliefs

	B fails	B works	
A fails	0	3/7	3/7
A works	3/7	1/7	4/7
	3/7	4/7	

distorted beliefs

Suppose now that the person finds out that the second relationship has succeeded. Then the person becomes indifferent about the first relationship, and there is no longer any belief distortion. Thus, the belief as to the prospects of the first relationship was initially highly distorted, but is no longer distorted after learning the news about the second relationship.

Example 2. school allocation
Consider the case of optimistic parents whose child is to be allocated randomly to either of two schools: Row and Column. For simplicity I assume that the quality of school is binary. Schools can thus be either good (G) or bad (B). With these assumptions there are four possible states of interest: GG, GB, BG and BB. Payoff depends on the quality of the school the child is allocated to, and is 1 if the school is good, and 0 if it is not. Expected payoff as a function of quality of the two schools and the allocation is as follows:

	G	B
G	1	1
B	0	0

Row

	G	B
G	1	0
B	1	0

Column

	G	B
G	1	0.5
B	0.5	0

Unknown

I assume for simplicity that the undistorted probability of the four states is the same: $p(GG) = p(GB) = p(BG) = p(BB) = 0.25$. To make results more readable I assume $\psi = \ln 4$, so that the "likelihood" term is 1, 2, or 4 depending on the payoff[10]. Using Equation 5 distorted beliefs as a function of the allocation are as follows:

	G	B	
G	0.4	0.4	0.8
B	0.1	0.1	0.2
	0.5	0.5	

Row

	G	B	
G	0.4	0.1	0.5
B	0.4	0.1	0.5
	0.8	0.2	

Column

	G	B	
G	4/9	2/9	2/3
B	2/9	1/9	1/3
	2/3	1/3	

Unknown

Beliefs about the Row (Column) school are thus distorted if the child is allocated to the Row (Column) school (and hence the parents care about its quality), but not if the child is allocated to the other school. Before learning the allocation the parents care about both schools, and hence beliefs about both schools are distorted. Learning the outcome of the allocation thus has

[10] $e^{\psi \cdot 0} = e^0 = 0, e^{\psi \cdot 0.5} = e^{\ln 2} = 2, e^{\psi \cdot 1} = e^{\ln 4} = 4.$

the effecting of changing beliefs about the schools, even though the allocation reveals no relevant information about school quality.

To see how this is related to complementarities in the payoff function note that learning that the child is allocated to the row school increases the expected payoff difference between states in which the row school is good and states in which it is bad, and decreases (in fact eliminates) the corresponding payoff difference for the column school. The opposite is true if the child is allocated to the column school. If we define an allocation variable to be 1 if the child is allocated to the Row school and 0 if the child is allocated to the Column school, then this variable is a complement to the quality of the Row school and a substitute to the quality of the Column school.

4.1 Formal analysis

The claim I seek to prove is that beliefs about some variable X are shifted in a predictable direction by good news about another variable Y that is a complement or substitute of X. The general setting I consider assumes that payoff can be written as a separable function of three random variables, so that $f(s) = f(X, Y, Z) = f(X, Y) + f(Z)$, where X and Y are complements or substitutes. For example, the payoff function of the sales agent is given by $f(X, Y) = \max(X, Y)$ for binary $X, Y \in \{0, 1\}$, and X and Y are substitutes. A more general version of this example in which $u = \max(X_1, X_2, \ldots, X_n)$ can be put in the above framework by defining $X = X_1$ and $Y = \max(X_2, \ldots, X_n)$, so that $f(X_1, \ldots, X_n) = \max(X, Y)$. Finally, consider an example such as $u = \sum_i a_i X_i$. Then a_i and X_i are strict complements, so that good news about a_i should raise beliefs about X_i if the agent is optimistic. This example can be put into the framework by defining $X = X_i$, $Y = a_i$ and $Z = \sum_{j \neq i} a_j X_j$.

I denote the undistorted probability distribution of a random variable X by $P_X(x)$ and its distorted probability distribution by $\Pi_X(x)$. I denote the corresponding probability distributions conditional on information I by $P_{X|I}(\cdot)$ and $\Pi_{X|I}(\cdot)$. The proposition is about how news that is about one variable can nonetheless affect beliefs about another unrelated variable. I define news about a variable as information that is independent of other variables, and "good news" as news that changes beliefs about the variable it is about in a predictable direction.

Definition 6 (News about a variable)**.** Suppose X is a random variables then an event I is *news about* X if for all random variables Y that are independent of X relative to p and for all $y \in \mathbb{R}$ the event $Y = y$ is independent of I relative to p. If, furthermore, $P_{X|I} \succeq_{LR} P_X$ then I is *good news about* X. If the inequality is strict then I is *strictly* good news about X.

I define complements and substitutes by adapting a standard definition based on increasing (decreasing) differences (Edgeworth, 1925; Samuelson, 1974; Topkis, 1998) that avoids any assumptions on differentiability:

Definition 7 (Complements and substitutes)**.** Suppose $u = f(X, Y) + f(Z)$ as above. Then X and Y are *complements* (*substitutes*) if for all $t > t'$ in the image of X_2, $f(x, t) - f(x, t')$ is non-decreasing (non-increasing) as a function of x. X and Y are *strict complements (substitutes)* if the differences $f(x, t) - f(x, t')$ are strictly increasing (strictly decreasing) as a function of x.

The main claim now follows. I define the proposition for the case of optimism and complements, but the result generalises for substitutes and pessimism with a sign inversion for both changes. The effect on X of good news about Y is summarised in the following table:

	Optimist	Pessimist
Complements	+	-
Substitutes	-	+

Proposition 10. *Suppose $\psi > 0$, X and Y are complements, and I is good news about Y. Then $\Pi(X|I) \succeq_{LR} \Pi(X)$. Moreover, if (i) X and Y are strict complements, (ii) I is strictly good news about Y, and (iii) the probability distribution of X and Y is not degenerate, then the result is also strict.*

The intuition for this result can be obtained by considering the case of two binary variables. The claim is then that $\pi_f(X = 1|I)/\pi_f(X = 0|I) \geq \pi_f(X = 1)/\pi_f(X = 0)$. Now,

$$\begin{aligned}
\frac{\pi_f(X=1|I)}{\pi_f(X=0|I)} &= \frac{\pi_f(X=1,Y=1|I) + \pi_f(X=1,Y=0|I)}{\pi_f(X=0,Y=1|I) + \pi_f(X=0,Y=0|I)} \\
&= \frac{p(X=1,Y=1|I)e^{\psi f(1,1)} + p(X=1,Y=0|I)e^{\psi f(1,0)}}{p(X=0,Y=1|I)e^{\psi f(0,1)} + p(X=0,Y=0|I)e^{\psi f(0,0)}} \qquad (18) \\
&= \frac{p(X=1)}{p(X=0)} \cdot \frac{p(Y=1|I)e^{\psi f(1,1)} + p(Y=0|I)e^{\psi f(1,0)}}{p(Y=1|I)e^{\psi f(0,1)} + p(Y=0|I)e^{\psi f(0,0)}}
\end{aligned}$$

with a similar result for $\pi_f(X=1)/\pi_f(X=0)$. Taking the difference between the two several terms drop out, and what remains is

$$\Big(p(Y=1|I)p(Y=0)-p(Y=1)p(Y=0|I)\Big)\Big(e^{\psi(f(1,1)+f(0,0))}-e^{\psi(f(0,1)+f(1,0))}\Big) \tag{19}$$

The expression on the left is weakly positive as I is good news about Y, and the expression on the right is weakly positive as X and Y are complements and $\psi > 0$[11].

4.2 Motivated cognition and cognitive dissonance

The prediction of this section is that information that leads to a change in payoffs can cause belief change even when it is not normatively relevant. This corresponds to the essence of the phenomenon known as *cognitive dissonance* (Festinger and Carlsmith, 1959; Cooper and Fazio, 1984), and sometimes known as *motivated cognition* (Kunda, 1990). One way to summarise the key finding of this literature is that the expected value of a variable is increasing in the payoff for high values of the variable in question. This finding fits the predictions of this section on the assumption that most people are optimistic.

For example, Klein and Kunda (1992) found that subjective beliefs about the likelihood playing with or against a given player in a trivia contest influences beliefs about the ability of that player. Letting x denote the ability of the player, and $a \in \{-1, 1\}$ the side on which she will play, the agent's payoff can be modelled by $f(a,x) = ax$. Thus, a and x are *complements*. The prediction would therefore be that an increase in a (telling subjects that the player will be on their team) should lead to a higher assessment of x (an increased subjective valuation of the ability of the player). This prediction agrees with the empirical findings.

Similarly, Berscheid et al. (1976) found that expecting to date a person causes an increased valuation of that person. Here we can let x_i denote the attractiveness of person i, and p_i the probability that person i is to be the date. Then p_i and x_i are complements, and news that $p_i = 1$ is predicted to increase the bias over x_i, again in agreement with empirical findings.

[11] $e^{\psi(f(1,1)+f(0,0))} - e^{\psi(f(0,1)+f(1,0))} \geq 0$ if and only if $e^{\psi(f(1,1)+f(0,0))} \geq e^{\psi(f(0,1)+f(1,0))}$, and since ln is a monotonic function, this is true if and only if $\psi(f(1,1)+f(0,0)-f(1,0)-f(0,1)) \geq 0$.

5 Conclusion

This paper introduces a model of Bayesian decision making where a person's beliefs about the likelihood of different outcomes depend upon the anticipated payoff consequences of those outcomes. This dependence is modelled as a distortion mapping linking the actual beliefs of a person to the beliefs the person would have held if he or she were indifferent what the true state is. The distortion is defined by a real-valued belief-distortion parameter, which is positive for optimists, negative for pessimists, and is zero for people whose beliefs are independent of payoffs.

This simple model can account for a number of phenomena. The predicted relationship between beliefs and payoff is consistent with optimistic and pessimistic bias. Moreover, comparative statics for the level of uncertainty and the stakes in the outcome are consistent with available evidence. When new information results in a change in anticipated payoff the model predicts belief change, even if the new information is normatively irrelevant. These predictions are consistent with evidence on cognitive dissonance (Cooper and Fazio, 1984), as well as evidence from studies of "motivated cognition" Kunda (1990).

The representation derived in the model is mathematically convenient, and can be readily adapted for use in applications. Optimistic bias has been applied to a variety of economic areas, such as financial markets, corporate finance, bargaining, and insurance. One would hope that the model of this paper can lead to better and stronger predictions in some areas. In financial markets in particular there are phenomena that look strongly suggestive of the types of mechanisms explored in this paper. Most obviously, optimistic investors may underestimate risks. In addition, the dynamic predictions of the model may help explain such phenomena as inefficient lack of diversification, and traders avoiding the sale of poorly performing securities.

For reasons of space and focus choice has not been explicitly discussed in this paper. This is an obvious task for a second paper which the author is in the process of writing. Another obvious area for expanding the model is strategic interaction, where there are existing models in the limit of extreme optimism (Yildiz, 2007), but not for more realistic levels of optimism or pessimism. Strategic interaction raises a number of interesting theoretical issues. For example, in isolated strategic interactions optimistic players may both believe the other is making a mistake (they agree to disagree). This, however, cannot be the case if the interaction is repeated to the point that both agents know the true equilibrium payoff. The most interesting case may be intermediate one, in which the payoff in each interaction is a combination of a fixed factor (which agents learn), and a variable factor (which agents are optimistic about). There are also interesting issues to do with interactions between optimistic and unbiased agents and interactions involving pessimists.

Appendix: proofs

Proposition 2

Proof. I start with $(i) \Rightarrow (ii)$. Note first that for all constant payoff-functions a and b, $\pi_b = \pi_a + (b-a)$. Define $p = \pi_b$ and it follows from Shift-invariance that $\pi_a = p$ for all $a \in \mathbb{R}$. Let now X be some non-null event such that $Y = X^c$ is also non-null, and for $y \in \mathbb{R}$ define a payoff-function f_y by $f_y(Y) = y$ and $f_y(X) = a$. Since Shift-invariance implies Impartiality and since π is a weak logit-distortion for all $y \in \mathbb{R}$,

$$\ln \frac{\pi_{f_y}(Y)}{\pi_{f_y}(X)} = \ln \frac{p(Y)}{p(X)} + v(y) - v(a). \tag{20}$$

Thus, Shift-invariance on f_y and $f_y + t = f_{y+t}$ implies that for all $t \in \mathbb{R}$, $v(y+t) - v(a+t) = v(y) - v(a)$. Next, Suppose that $b'' - b' = k(a'' - a')$ for some real numbers a', a'', b' and b'' and for $k \in \mathbb{N}$. Then using the above result,

$$\begin{aligned} v(b'') - v(b') &= \sum_{i=1}^{k} v(b' + k(a'' - a')) - v(b' + (k-1)(a'' - a')) \\ &= \sum_{i=1}^{k} v(a'') - v(a') = k(v(a'') - v(a')) \end{aligned} \tag{21}$$

Finally, by assumption $(d-c)/(b-a)$ is rational. Thus, there exist whole numbers m and n such that $(d-c)/(b-a) = m/n$. Thus, $b = a + 1/n$ and $d = a + m \cdot (1/n)$. By the result in Equation 21 $v(b) - v(a) = n(v(a+1/n) - v(a)$ and $v(d) - v(c) = m(v(c+1/n) - v(c) = m(v(a+1/n) - v(a) =$. Thus,

$$v(d) - v(c) = \frac{m}{n}\big(v(b) - v(a)\big) == \left(\frac{v(b) - v(a)}{b - a}\right)(d - c) \tag{22}$$

In the opposite direction, suppose $f' = f + a$. I need to prove that $\pi_{f'} = \pi_f$. It is sufficient to show that for all constant-payoff events A and B, $\pi_f(B)/\pi_f(A) = \pi_{f'}(B)/\pi_{f'}(A)$. By Equation 6 and the assumption that $\pi_a = p$ for all $a \in \mathbb{R}$ this condition is equivalent to $v(f(B)) - v(f(A)) = v(f(B) + a) - v(f(A) + a)$. Let $a = f(A), b = f(B), c = f(A) + a$ and $d = f(B) + a$. Then $d - c = 1 \cdot (b - a)$ and hence Shift-invariance follows from Equation 14. $\square$

Proposition 3 The proof is trivial, and hence omitted.

Proposition 4

Proof. In the first direction, by Proposition 2 v is linear over $\mathbb{Q}$. Without loss of generality suppose $v(r) = v(r_0) + \psi(r - r_0)$ for all $r, r_0 \in \mathbb{Q}$. I need to show that equality holds also for $x \in \mathbb{R} \setminus \mathbb{Q}$. Since $\mathbb{Q}$ is dense in $\mathbb{Q}$ there exists a sequence $\{r_n\}_{n=1}^{\infty}$ of rational numbers that converges to x. By Continuity v is continuous, and hence

$$v(x) = \lim_{n\to\infty} v(r_n) = \lim_{n\to\infty} v(r_0) + \psi(r_n - r_0) = v(r_0) + \psi(x - r_0) \qquad (23)$$

The only if direction is immediate since by Proposition 3 it is sufficient to show that v is continuous, and $v(x) = v(r_0) + \psi(x - r_0)$ is a continuous function. $\square$

Proposition 5

Proof. Suppose the first set of inequalities hold. By Equation 13 $\pi_{sax}(A) \leq \pi_{bAx}(A) \leq \pi_{cAx}(A)$. Since probabilities are positive numbers and $\pi_f(A^c) = 1 - \pi_f(A)$ for all f, it follows that $\frac{\pi_{aAx}(A)}{\pi_{aAx}(A^c)} \leq \frac{\pi_{bAx}(A)}{\pi_{bAx}(A^c)} \leq \frac{\pi_{cAx}(A)}{\pi_{cAx}(A^c)}$. Plugging these inequalities into Equation 13 yields $v(a) - v(x) \leq v(b) - v(x) \leq v(c) - v(x)$, and hence $v(a) \leq v(b) \leq v(c)$ all for $a, b, c \in \mathbb{R}$ such that $a < b < c$. In other words, v is monotonically increasing. The other case is analogous. $\square$

Proposition 6

Proof. Explained in text. $\square$

Proposition 9

Proof. Let $f_\pi(x) = dF_\pi(x)/dx$ denote the probability density function of the payoff of f under the distorted probability measure π_f. Then,

$$\begin{aligned} f_\pi(x) &\propto \left(\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)e^{\psi x} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2-2\psi\sigma^2 x}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\{\mu+\psi\sigma^2\})^2-2\psi\sigma^2\mu-\psi^2\sigma^4}{2\sigma^2}} \propto \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\{\mu+\psi\sigma^2\})^2}{2\sigma^2}} \end{aligned} \tag{24}$$

$F_\pi(x)$ is thus a normal distribution with mean $\mu + \sigma^2\psi$ and variance σ^2. ☐

Proposition 7

Proof. By Equation 5 $dF_\pi^\psi(x)/dF_\pi^{\psi'}(x) = (p(x)e^{\psi x})/(p(x)e^{\psi' x}) = e^{(\psi-\psi')x}$ which is a non-decreasing function of x if $\psi \geq \psi'$. The claim thus follows from the definition of stochastic dominance in the likelihood ratio. ☐

Proposition 8

Proof. The proof is immediate from Equation 2 by considering the events A_{min} and A_{min}^c and taking the limit $\psi \to -\infty$, and similarly for the events A_{max} and A_{max}^c with the limit $\psi \to \infty$. ☐

Proposition 10

Proof. I need to prove that for all pairs of possible values $x > x'$ for X, $\pi_f(X = x|I)/\pi_f(X = x'|I) \geq \pi_f(X = x)/\pi_f(X = x')$. Now,

$$\begin{aligned} \frac{\pi_f(X=x|I)}{\pi_f(X=x'|I)} &= \frac{\sum_y \sum_z \pi_f(X=x, Y=y, Z=z|I)}{\sum_y \sum_z \pi_f(X=x', Y=y, Z=z|I)} \\ &= \frac{\sum_y \sum_z p(X=x, Y=y, Z=z|I)e^{\psi u(x,y,z)}}{\sum_y \sum_z p(X=x', Y=y, Z=z|I)e^{\psi u(x',y,z)}} \\ &= \frac{p(X=x)}{p(X=x')} \cdot \frac{\sum_y \sum_z p(Y=y)p(Z=z)e^{\psi(u(x,y)+u(z))}}{\sum_y \sum_z p(Y=y)p(Z=z)e^{\psi(u(x',y)+u(z))}} \\ &= \frac{p(X=x)}{p(X=x')} \cdot \frac{\sum_y p(Y=y|I)e^{\psi u(x,y)}}{\sum_y p(Y=y|I)e^{\psi u(x',y)}} \end{aligned} \tag{25}$$

where the second step uses Equation 5, the third step the assumption that X, Y and Z are independent, that I is news about Y, and that $u(x, y, z) = u(x, y) + u(z)$. Similarly,

$$\frac{\pi_f(X = x)}{\pi_f(X = x')} = \frac{p(X = x)}{p(X = x')} \cdot \frac{\sum_{y'} p(Y = y')e^{\psi u(x,y')}}{\sum_{y'} p(Y = y')e^{\psi u(x',y')}} \tag{26}$$

Thus the claim is true if and only if

$$\frac{\sum_y p(Y = y|I)e^{\psi u(x,y)}}{\sum_y p(Y = y|I)e^{\psi u(x',y)}} - \frac{\sum_{y'} p(Y = y')e^{\psi u(x,y')}}{\sum_{y'} p(Y = y')e^{\psi u(x',y')}} \geq 0 \tag{27}$$

or

$$\sum_y \sum_{y'} p(Y = y|I)p(Y = y') \cdot \left(e^{\psi(u(x,y)+u(x',y'))} - e^{\psi(u(x,y')+u(x',y))}\right) \geq 0 \tag{28}$$

This expression is antisymmetric in y and y'. The terms with $y = y'$ therefore drop out, and terms with $y' > y$ can be combined with the corresponding term for which $y' < y$. Thus, the following condition is equivalent:

$$\begin{aligned} &\sum_y \sum_{y'<y} \big(p(Y = y|I)p(Y = y') - p(Y = y'|I)p(Y = y)\big) \\ &\qquad \cdot \left(e^{\psi(u(x,y)+u(x',y'))} - e^{\psi(u(x,y')+u(x',y))}\right) \geq 0 \end{aligned} \tag{29}$$

Now, $p(Y = y|I)p(Y = y') - p(Y = y'|I)p(Y = y)$ is non-negative by the fact that $y' < y$ and the assumption that I is good news about y, and $e^{\psi(u(x,y')+u(x',y))}$ is non-negative by the assumption that $\psi \geq 0$, X and Y are complements, and the fact that e^x is a monotonically increasing function. It therefore follows that the entire expression is non-negative, thereby concluding the main proof.

Finally, note that if $\psi > 0$, I is strictly good news about Y and X and Y are strict complements, then all the expressions are strictly positive. The condition that X and Y are non-degenerate ensures that the claim is not empty, and that the sums contain at least one term. When all these conditions hold the inequality is therefore strict. □

References

Akerlof, G. and Dickens, W. (1982). The Economic Consequences of Cognitive Dissonance, *The American Economic Review* **72**(3): 307–319.

Berscheid, E., Graziano, W., Monson, T. and Dermer, M. (1976). Outcome dependency: Attention, attribution, and attraction, *Journal of Personality and Social Psychology* **34**(5): 978–988.

Brunnermeier, M. and Parker, J. (2005). Optimal Expectations, *The American Economic Review* **95**(4): 1092–1118.

Cooper, J. and Fazio, R. (1984). A new look at dissonance theory, *Advances in experimental social psychology* **17**: 229–266.

Edgeworth, F. (1925). The Pure Theory of Monopoly, *Papers Relating to Political Economy* **1**: 111–142.

Festinger, L. and Carlsmith, J. (1959). Cognitive consequences of forced compliance., *J Abnorm Psychol* **58**(2): 203–10.

Klein, W. and Kunda, Z. (1992). Motivated person perception: constructing justifications for desired beliefs, *Journal of experimental social psychology* **28**(2): 145–168.

Kunda, Z. (1990). The case for motivated reasoning, *Psychological Bulletin* **108**(3): 480–498.

Luce, D. (1959). *Individual Choice Behavior*, Wesley, New York.

Samuelson, P. (1974). Complementarity, *Journal of Economic Literature* **12**: 1255–1289.

Sjöberg, L. (2000). Factors in risk perception, *Risk Analysis* **20**(1): 1–11.

Topkis, D. (1998). *Supermodularity and Complementarity*, Princeton University Press.

Weinstein, N. (1980). Unrealistic optimism about future life events, *Journal of Personality and Social Psychology* **39**(5): 806–820.

Yildiz, M. (2007). Wishful Thinking in Strategic Environments, *Review of Economic Studies* **74**(1): 319–344.