OPTIMISM AND PESSIMISM WITH EXPECTED UTILITY

By

David Dillenberger, Andrew Postlewaite, and Kareen Rozen

October 2011

COWLES FOUNDATION DISCUSSION PAPER NO. 1829

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Optimism and Pessimism with Expected Utility*

David Dillenberger
University of Pennsylvania

Andrew Postlewaite†
University of Pennsylvania

Kareen Rozen‡
Yale University

October 2011

Abstract

Savage (1954) provided a set of axioms on preferences over acts that were equivalent to the existence of an expected utility representation. We show that in addition to this representation, there is a continuum of other "expected utility" representations in which for any act, the probability distribution over states depends on the corresponding outcomes. We suggest that optimism and pessimism can be captured by the stake-dependent probabilities in these alternative representations; e.g., for a pessimist, the probability of every outcome except the worst is distorted down from the Savage probability. Extending the DM's preferences to be defined on both subjective acts and objective lotteries, we show how one may distinguish optimists from pessimists and separate attitude towards uncertainty from curvature of the utility function over monetary prizes.

*First version August 2011. We thank Eddie Dekel, Itzhak Gilboa, Edi Karni, Mark Machina, Larry Samuelson, Tomasz Strzalecki, and Peter Wakker for helpful discussions and suggestions.

†Postlewaite thanks the National Science Foundation for support.

‡Rozen thanks the National Science Foundation for support.

If one has really technically penetrated a subject, things that previously seemed in complete contrast, might be purely mathematical transformations of each other.
—John von Neumann (1955, p. 496).

1. Introduction

Consider a decision maker (DM) who is faced with gambles on whether a coin flip will come up heads. He is told that if the outcome is heads (H) he will get \$100 and if tails (T) he will also get \$100. When asked what he thinks the probability of H is, he responds .5. He is then told about another gamble in which the outcome for T is unchanged but the outcome for H is increased to \$1000, and is asked what he thinks the probability of H is now. He responds that he thinks the probability of H is now .4. When asked how he can think the probability of H can differ across the two gambles when it is the same coin, DM simply says that random outcomes tend to come out badly for him. After being offered a third gamble that gives \$100 for T and \$10,000 for H, he says that faced with that gamble, he thinks the probability of H is .2.

When faced with a choice between any two gambles, each of which specifies the amount received conditional on the realized state, DM says that he maximizes expected utility. He has a utility function over money, and for any two gambles (x_1, x_2) and (y_1, y_2), he will have two probability distributions over the states, $p(x_1, x_2)$ and $p(y_1, y_2)$. DM's probability assessments reflect his belief that luck is not on his side. For each gamble he computes its expected utility under the associated probability, and then chooses the gamble with the higher expected utility.

Confronted by such a DM, one might well judge him irrational. But would that judgment change if one discovered that the DM's revealed preferences satisfy Savage's axioms? We show below that for any preferences over acts that satisfy Savage's axioms, there will be representations of those preferences as described in the paragraph above: there will be a utility function over outcomes and, for any act, a probability distribution over states that depends on the payoffs the act generates, with preferences given by expected utility. Furthermore, the probability distribution depends on the payoffs as in the example above: the probability of the state with the good outcome is smaller than the Savage probability, and it decreases when the good outcome is replaced by an even better outcome.

We suggest that a DM who describes his decision-making process as above can be thought of as pessimistic. When good outcomes get better, he thinks it is less likely he will win them. In addition to the multitude of pessimistic representations of preferences that satisfy Savage's axioms, there is a continuum of "optimistic" representations. Here, the probability distribution on states associated with a given act will, as in the pessimistic case, depend on the outcomes associated with the act, but the probability of the state with the best outcome is higher than under the Savage probability distribution, and will increase if that outcome is replaced with an even better one.

We may still want to characterize the DM above as being irrational, but notice that we cannot make that determination on the basis of his choices: his preferences over acts are the same as those of a person who uses an analogous decision process using the Savage representation utility function and associated "standard" probability distribution. Any distinction between the rationality of the Savage representation and the alternative representation must be done on the basis of the underlying process by which the DM makes decisions and not only on the decisions themselves.

We lay out the model in Section 2 and demonstrate how pessimistic and optimistic representations can be constructed for the case in which there are two states, where our basic idea is most transparent. In Section 3 we show how the two-state case can be extended to an arbitrary finite number of states. Section 4 examines how one may be able to distinguish optimism from pessimism, and how we may use our model to separate attitude to uncertainty from utility over prizes. We conclude by discussing further implications of using different representations for the same underlying preferences.

2. Optimism and pessimism with two states

There are two states of nature, s_1 and s_2. Let $X \subset \mathbb{R}$ be a set of monetary prizes. Consider a DM whose preferences over the set of (Savage) acts satisfy Savage's axioms, and who prefers more money to less.[1] Formally, an act is a function $l : \{s_1, s_2\} \to X$. For notational convenience, in the text we simply denote an act by an ordered pair of state contingent payoffs, $x = (x_1, x_2)$, where x_i is the payoff received in state i. Let $v(x_1, x_2) = p_1 u(x_1) + p_2 u(x_2)$ represent the DM's preferences over the act (x_1, x_2). Here $p_i \in (0, 1)$ is the subjective, stake-independent, probability that he assigns to state i, and $p = (p_1, p_2)$ is the probability distribution.

We now consider a different representation of the same preferences, in which the probabilities are *stake-dependent*, that is, the probability assigned to state i is $P_i(x_1, x_2; p)$. We look for an alternative representation $\hat{v}$ of the form

$$\hat{v}(x_1, x_2) = P_1(x_1, x_2; p)\, \hat{u}(x_1) + (1 - P_1(x_1, x_2; p))\, \hat{u}(x_2)\,. \tag{1}$$

[1] Although Savage's original work applies only to the case where the state space is not finite, it has been shown how to derive a Savage-type representation when there are only a finite number of states (see, e.g., Wakker (1984) or Gul (1992)).

Recall that $\hat{v}$ and v represent the same preferences if and only if each is a monotonic transformation of the other. Consider a strictly increasing function (and for simplicity, differentiable) $f : \mathbb{R} \to \mathbb{R}$, and define $\hat{v}(x_1, x_2) := f(v(x_1, x_2))$. Then, we seek probabilities $P_i(x_1, x_2; p)$ and a utility over prizes $\hat{u}$ such that (1) is satisfied. By considering the case that the outcomes in each state are the same (that is, the case of constant acts), note that (1) implies that $\hat{v}(z, z) = \hat{u}(z) = f(v(z, z)) = f(u(z))$ for all z. Then the desired representation (1) simplifies to

$$\hat{v}(x_1, x_2) = f(v(x_1, x_2)) = P_1(x_1, x_2; p) f(u(x_1)) + (1 - P_1(x_1, x_2; p)) f(u(x_2)).$$

Solving for $P_1(x_1, x_2; p)$, we get

$$P_1(x_1, x_2; p) = \frac{f(v(x_1, x_2)) - f(u(x_2))}{f(u(x_1)) - f(u(x_2))}$$

for $x_1 \neq x_2$. Note that $P_1(x_1, x_2; p)$ is always between zero and one because, by properties of expected utility, $v(x_1, x_2)$ is always between $u(x_1)$ and $u(x_2)$. As $x_1 \to x_2$, $P_1(x_1, x_2; p)$ converges to p_1. Naturally, $P_2(x_1, x_2; p) := 1 - P_1(x_1, x_2; p)$.

When $x_1 > x_2$, the denominator of $P_1(x_1, x_2; p)$ is positive. Thus, when f is convex, Jensen's inequality implies that

$$P_1(x_1, x_2; p) \leq \frac{p_1 f(u(x_1)) + (1 - p_1) f(u(x_2)) - f(u(x_2))}{f(u(x_1)) - f(u(x_2))} = p_1;$$

that is, the probability of the larger prize is distorted down. Similarly, when f is concave, the probability of the larger prize is distorted up. (Analogous characterizations hold when $x_2 > x_1$: the probability of the smaller prize is distorted up when f is convex, and distorted down when f is concave). Stated differently, the pessimist holds beliefs that are first-order stochastically dominated by the standard Savage distribution, while the optimist holds beliefs that first-order stochastically dominate it.

For specific classes of convex and concave functions, we can say more.

Proposition 1. *Suppose $X \subseteq \mathbb{R}^2_+$, $x_1 \neq x_2$, and $r \in \{2, 3, \ldots\}$. Then $\frac{\partial P_i(x_1, x_2; p_1)}{\partial x_i} < 0$ for $f(z) = z^r$ and $\frac{\partial P_i(x_1, x_2; p_1)}{\partial x_i} > 0$ for $f(z) = z^{\frac{1}{r}}$.*

The proof appears in the appendix. The case $r = 1$ would correspond to the standard Savage formulation in which there is no stake-dependent probability distortion. When $f(z) = z^r$, the DM's probability assessments reflect a stronger notion of pessimism. The better the consequence in any state, the less likely he thinks that this state will be realized. In particular, improving the best outcome reduces his assessment of its probability (as in the example in the introduction), and worsening the worst outcome increases his assessment of its probability. When $f(z) = z^{\frac{1}{r}}$ the comparative statics are flipped. For the optimist, the better is the best outcome, the more likely the DM thinks it is; and the worse is the worst outcome, the less likely he thinks it is. By construction, however, choice behavior in either case is indistinguishable from that of a DM with a Savage-type representation.

Remark 1. *Based on Aumann's 1971 exchange of letters with Savage (reprinted in Drèze (1987)), the following argument has often been used to point out that the Savage representation could have multiple state-dependent expected utility representations, leaving the (single) probability distribution ill-defined. Consider a Savage representation of the form* $p_1 u(x_1) + p_2 u(x_2)$. *Then for any* $\widehat{p}$ *with the same support as* p, *this expression is equal to* $\widehat{p}_1 \widehat{u}(x_1, s_1) + \widehat{p}_2 \widehat{u}(x_2, s_2)$, *where* $\widehat{u}(x_s, s) := \frac{p_s u(x_s)}{\widehat{p}_s}$ *is a state-dependent utility function. Now imagine using the same idea to generate stake-dependent probabilities. Fix any strictly positive utility function over prizes,* $\overline{u}$. *Then* $p_1 u(x_1) + p_2 u(x_2)$ *is the same as* $\overline{p}(s_1, x_1)\overline{u}(x_1) + \overline{p}(s_2, x_2)\overline{u}(x_2)$, *where* $\overline{p}(s, x_s) := \frac{p_s u(x_s)}{\overline{u}(x_s)}$. *But note that* $\overline{p}(s, x_s)$ *is not a probability distribution unless we normalize it by* $\overline{p}(s_1, x_1) + \overline{p}(s_2, x_2)$, *a scaling factor which generically depends on the stakes in all states (that is, unless* u *is a scalar multiple of* $\overline{u}$*). The resulting utility representation, therefore, will not longer represent the same preferences as the original Savage representation.*

3. The general case

The two-state case showed how one could find a continuum of (nonlinearly transformed) utility functions and an associated distortion of probabilities for each, so that under any of these representations, the certainty equivalent of every act is the same as that under the original Savage representation. We show next that this can be done for an arbitrary finite number of states.[2]

Let $S = \{s_1, ..., s_n\}$ be the set of states and let $x = (x_1, ..., x_n) \in R^n$ be an act, where x_i corresponds to the outcome in state s_i. For ease of exposition, we will assume that $x_1 > \cdots > x_n$; it is straightforward to handle the general case at the cost of more complicated notation. Consider a Savage expected utility representation, with p the probability vector and u the utility function over prizes. We look for a representation of the form

$$\hat{v}(x) = \sum\nolimits_{i=1}^{n} P_i(x; p)\hat{u}(x_i), \tag{2}$$

[2] Alternatively, it can be done for simple (finite support) acts on a continuum state space.

where $P_i(x;p)$, $i = 1, \ldots, n$, constitute a stake-dependent probability distribution, and $\hat{v}$ represents the same preferences as the Savage expected utility function. As before, this implies that $\hat{u}(z) = f(u(z))$ for some increasing function f, and that

$$\hat{v}(x) = f(v(x)) = \sum\nolimits_{i=1}^{n} P_i(x;p) f(u(x_i)). \tag{3}$$

Including the above equation and the obvious restriction that $\sum_{i=1}^{n} P_i(x;p) = 1$, we have two equations with n unknowns. While this sufficed for a unique solution in the case $n = 2$, when $n \geq 3$ there will generally be many ways to construct a probability distortion, corresponding to different ways a DM might allocate weight to events.

We will demonstrate one simple construction of an optimistic distortion; the case of pessimism is analogous. For an act x, let $c(x;\tilde{u},\tilde{p})$ be the certainty equivalent of x given a utility function $\tilde{u}$ and a probability vector $\tilde{p}$, that is, $u(c(x;\tilde{u},\tilde{p})) = \sum_{i=1}^{n} \tilde{p}_i \tilde{u}(x_i)$. Let $\hat{u} = f(u)$ for an increasing and concave transformation function f. Since $\hat{u}$ is more risk averse than u, $c(x;\hat{u},p) < c(x;u,p)$. Define $P_\lambda(x) = \lambda \cdot (1,0,\ldots,0) + (1-\lambda) \cdot p$ to be the convex combination of the Savage probability p and the distribution $(1,0,\ldots,0)$, which puts probability one on the best state. Note that $c(x;\hat{u},P_0(x)) = c(x;\hat{u},p) < c(x;u,p)$ and that $c(x;\hat{u},P_1(x)) = x_1 > c(x;u,p)$. It is clear that $c(x;\hat{u},P_\lambda(x))$ is continuous and strictly increasing in λ, so for any x there is a unique $\lambda^* \in (0,1)$ such that $c(x;\hat{u},P_{\lambda^*}(x)) = c(x;u,p)$.[3] This construction ensures that for each act x, the certainty equivalent given $\hat{u}$ and $P(x;p)$ is the same as the certainty equivalent given the Savage representation. Thus $\hat{u}$ and $P(x;p)$ form an alternate representation of the preferences represented by the original Savage representation, u and p. The constructed representation can be thought of as optimistic; it is straightforward to see that for every act, the distorted probability distribution over states first-order stochastically dominates the Savage distribution p. For any $r \in (0,1)$, if $f(z) = z^r$ then the probability distribution can be shown to satisfy one side of the comparative statics exhibited in the case $n = 2$: as the worst outcome becomes worse, the DM is more convinced that the best outcome will occur (see the supplement for the proof).

There are, of course, many ways to modify this construction. Any function $P_\lambda(x)$, $\lambda \in [0,1]$, that satisfies (i) $P_0(x) = p$; (ii) $P_1(x) = (1,0,\ldots,0)$; and (iii) $P_\lambda(x)$ first-order stochastically dominates $\tilde{P}_{\lambda'}(x)$ if $\lambda > \lambda'$ could have been used to construct $P(x;p)$. Thus for each of the continuum of concave transformations $\hat{u}$ of the Savage utility function u, there is a continuum of probabilities $P(x;p)$ that could be coupled with $\hat{u}$ to generate an alternative

[3] More generally, if the best payoff occurs in state j, replace $(1,0,\ldots,0)$ above with the distribution that puts probability one on state j. If for a given x the best payoff occurs under several states, any distribution for which the support is contained in the set of those best states would do. If x_j is the best outcome, $\lambda^* = \frac{f\left(\sum_{i=1}^{n} p_i u(x_i)\right) - \sum_{i=1}^{n} p_i f(u(x_i))}{f(u(x_j)) - \sum_{i=1}^{n} p_i f(u(x_i))}$. Since f is increasing and concave, $\lambda^* \in (0,1)$.

optimistic representation of the preferences over acts. For a pessimistic distortion, we would use instead an arbitrary increasing and convex transformation f to generate $\hat{u} = f(u)$. Since u is more concave than $\hat{u}$, $c(x; \hat{u}, p) > c(x; u, p)$. Then, for example, the stake-dependent distribution $P(x; p)$ could be generated by a convex combination of the Savage distribution p and $(0, \ldots, 0, 1)$, the distribution which puts probability one on the worst outcome. In this case, p stochastically dominates the constructed $P(x; p)$. Alternatively, one can construct a probability distortion using a different method altogether (an example of such a construction can be found in the supplemental appendix).

4. Implications on a larger domain

We next examine how one may be able to distinguish optimism from pessimism, and how we may use our model to separate attitude to uncertainty from utility over prizes. Suppose the DM's preferences $\succeq$ are defined over the domain of choice $\Psi = \mathcal{L}^1 \cup \mathcal{F}$, where $\mathcal{L}^1$ is the set of (purely objective) simple lotteries over the set of prizes X, and $\mathcal{F}$ is the set of (purely subjective) Savage acts over X.[4] Assume that the DM satisfies the axioms of Savage over subjective acts, leading to a subjective expected utility representation with Bernoulli function v. Assume also that the DM satisfies the axioms of vNM over objective lotteries, leading to an expected utility representation with Bernoulli function u.

Suppose we observe that the DM is ambiguity averse, in the sense that he is more risk averse in uncertain settings than in objective settings; that is, $v = f(u)$ for some increasing and concave function f.[5] Note that it is impossible to keep u and v the same without assuming identical risk attitude in both settings, as risk preferences are entirely characterized by the utility for prizes under the conventions of the expected utility representation. A natural focal point, however, is for the DM's utility over prizes (which captures his tastes for the ultimate outcomes) to be consistent across the objective and subjective domains; that is, $u = v$. Simply put, the prizes are the same in both domains; it is only the probabilities that differ in the two situations. The behavioral implication is that the DM is indifferent

[4] Note that this domain is essentially a strict subset of the domain of Anscombe and Aumann (1963), in which the outcome of an act in every state is an objective lottery. This domain is similar to the one used in Chew and Sagi (2008). Using their language, the sets $\mathcal{L}^1$ and $\mathcal{F}$ can be thought of as two different sources of uncertainty, on which the DM's preferences may differ. This domain allows us to talk about ambiguity while avoiding the multistage complications of Anscombe and Aumann (1963)'s model.

[5] This corresponds to Ghirardato and Marinacci's (2002) definition of ambiguity aversion, used in the context of Anscombe-Aumann acts in Grant, Polak, and Strzalecki (2009). According to that definition, the DM is more risk averse in uncertain settings than in objective settings if there exists a probability distribution p over S, such that for all $\pi \in \mathcal{L}^1$ and $l \in \mathcal{F}$, $l \succeq \pi$ implies that $\mu_{l,p} \succeq \pi$, where $\mu_{l,p}$ is the objective lottery under which the prize $l(s)$ is received with probability $p(s)$. The intuition behind this axiom is that if the DM prefers an act to a given lottery, it would also be better to simply receive that "act" with the objective probabilities that would ultimately be specified by the Savage distribution.

between a degenerate lottery that yields the prize z with certainty and the degenerate act that yields z in each state of nature.

Our model allows comparing the risk attitude of the DM on the subjective and objective domains while attributing to him the same utility for prizes in the two domains, simply by viewing him as a pessimist. In particular, if $v = f(u)$ and f is concave, then there exists a unique *convex* transformation $g = f^{-1}$ such that $g(v) = u$.[6] Given the results above, the Savage representation with utility function v is equivalent to a pessimistic representation using the more convex u. That is, we may attribute the DM's greater risk aversion in the subjective domain to pessimistic probability assessments, rather than a change in his utility for prizes. Under this convention, his utility over prizes is always u, which is elicited in the presence of well-defined risks. Analogously, if a DM is discovered to be less risk averse in the subjective domain than in the objective one, he may be viewed as an optimist using the same utility function over prizes in both domains. Much in the same way that probabilities are identified under Savage's convention of *state*-independent utility for prizes (as discussed in Remark 1), optimism and pessimism are identified under a convention of *source*-independent utility for prizes.

5. Discussion and related literature

Should we care? We have demonstrated the existence of alternative representations of preferences over acts that satisfy Savage's axioms and that involve stake-dependent probabilities and a stake-independent utility function. Should we care about such alternate representations?

It is useful to distinguish between a *utility representation* (or *model*), which is a construct for imagining how a DM makes decisions, and *choice behavior*, which is the observable data. The standard point of view is that the representation is nothing more than an analytically convenient device to model a DM's choices. In this approach, termed *paramorphic* by Wakker (2010), the representation does *not* suggest that a DM uses the utility function and a probability distribution to make choices, or indeed, that a DM even has a utility function or probability distribution. An alternative approach is that the models we employ should

[6] More formally, the assumptions above imply that for any $\xi \in \Psi$, there are increasing transformations h and $\widehat{h}$ such that the DM's preference is represented by

$$U(\xi) = \begin{cases} h\left(\sum_x \pi(x) u(x)\right) & \xi = \pi \in \mathcal{L}^1 \\ \widehat{h}\left(\sum_s p_s v(l(s))\right) & \xi = l \in \mathcal{F} \end{cases}$$

Indifference between a degenerate lottery and act would imply $u(x) = h^{-1}\left(\widehat{h}(v(x))\right) := g(v(x))$. If the DM is more risk averse in uncertain settings than in objective settings, then the transformation g is convex.

not only capture the choices agents make, but should match the underlying processes in making decisions. Wakker (2010) lays out an argument for this approach, which he terms *homeomorphic*. In his words, "we want the theoretical parameters in the model to have plausible psychological interpretations." This stance is also common in the behavioral economics literature, where mental processes and psychological plausibility are of particular interest.[7]

As Dekel and Lipman (2010) note, a utility representation is, at minimum, useful for organizing our thoughts around the elements of that representation (e.g., in terms of probabilities, utilities, and expectations).[8] Indeed, the expected utility representation has strongly influenced the view of "rational" behavior. Gilboa, Postlewaite, and Schmeidler (2009) argue that given the axiomatic foundations provided by Savage, "one may conclude that we should formulate our beliefs in terms of a Bayesian prior and make decisions so as to maximize the expectation of a utility function relative to this prior." This normative statement links Savage's axioms to a particular process of decision making. In a recent survey, Gilboa and Marinacci (2011) express a similar idea, according to which Savage's axiomatic foundations essentially say,"Even if you don't know what the probabilities are, you should better adopt some probabilities and make decisions in accordance with them, as this is the only way to satisfy the axioms."

In this paper, we show that a DM may want to satisfy Savage's axioms but still view the world with (our notion of) optimism or pessimism, and in particular, without having a stake-independent prior. We should emphasize that this is a positive issue of how people actually think and behave. Looking at implied choice behavior, a model that assumes expected utility maximization accommodates people who might think in an "irrationally" pessimistic or optimistic way. Only by characterizing the full set of representations can we know the full range of possible decision processes that are consistent with the analysis in our standard models. This means that one need not always modify the standard model to include psychologically plausible decision processes. Hey (1984), for example, introduces a notion of pessimism and optimism very similar to our own: an optimist (pessimist) revises up (down) the probabilities of favorable events and revises down (up) the probabilities of unfavorable events. Hey incorporates consequence-dependent probabilities in a Savage-like

[7] A similar discussion appears in Karni (2011). Karni distinguishes between the definitional meaning of subjective probabilities, according to which subjective probabilities define the DM's degree of belief regarding the likelihood of events, and the measurement meaning, according to which subjective probabilities measure, rather than define, the DM's beliefs. That is, the DM's beliefs are cognitive phenomena that directly affect the decision-making process.

[8] They further argue that the "story" of a model is relevant and may provide a reason for preferring one model to the other, even if the two models predict the same choices. Dekel and Lipman emphasize that while the story's plausibility (or lack thereof) may affect our confidence in the predictions of the model, it cannot refute or confirm those predictions; and that even if the story suggested by the representation is known to be false, it may still be valuable to our reasoning process.

representation, which can generate behavioral patterns that are inconsistent with expected utility because additional restrictions are not placed on the distorted probabilities. The notion that optimism and pessimism are inconsistent with Savage's axioms is implicit in his analysis, whereas our paper suggests that this is not necessarily the case.[9]

There are good arguments for the approach that takes the elements of the representation as actual entities in themselves. Consider a situation in which a DM may have little or no information about the relative likelihoods of outcomes associated with different choices she confronts. An expert who is informed about those likelihoods could determine which of the choices is best if he knew the DM's utility function. Through a sequence of questions about choices in a framework that the DM understands, the expert can, in principle, elicit the utility function, which can then be combined with the expert's knowledge about the probabilities associated with the choices in the problem at hand in order to make recommendations. Wakker (2008, 2010) and Karni (2009) treat problems of this type in the context of medical decision making. Under this point of view, it may be important to understand which representation is being elicited. If a DM had stake-dependent pessimistic beliefs but was assumed to have a "standard" Savage representation, the elicited utility function would exhibit greater risk aversion than the true utility function. Analogously, if the DM was optimistic, the elicited utility function would exhibit less risk aversion than her true utility function.

Related literature. The observation that the Savage-type representation and the optimist (or pessimist) case support the same underlying preferences, and hence cannot be distinguished by simple choice data, is related to general comments about model identification. In a series of papers, Karni (2011 and references therein) points out that the identification of probabilities in Savage's model rests on the (implicit) assumption of state-independent utility, and proceeds to propose a new analytical framework within which state independence of the utility function has choice-theoretic implications.[10] In the context of preference over menus of lotteries, Dekel and Lipman (2011) point out that a stochastic version of Gul and Pesendorfer (2001)'s temptation model is indistinguishable from a random Strotz model.

[9] By contrast, Wakker (1990) discusses optimism and pessimism as phenomena that can be accommodated in models such as rank-dependent expected utility (RDU), but cannot in expected utility. Wakker *defines* pessimism through behavior (similarly to uncertainty aversion) and shows that within RDU, pessimism (optimism) holds if greater decision weights are given to worse (better) ranks. Unlike in our model, fixing the ranks, changes in outcomes do not affect changes in decision weights in Wakker's model.

[10] Grant and Karni (2005) argue that there are situations in which Savage's notion of subjective probabilities (which is based on the convention that the utilities of consequences are state-independent) is inadequate for the study of incentive contracts. For example, in a principal-agent framework, misconstrued probabilities and utilities may lead the principal to offer the agent a contract that is acceptable yet incentive incompatible.

Chatterjee and Krishna (2009) show that a preference with a Gul and Pesendorfer (2001) representation also has a representation where there is a menu-dependent probability that the choice is made by the tempted (the "alter-ego") self, and otherwise the choice is made by the untempted self. Spiegler (2008) extends Brunnermeier and Parker's (2005) model of optimal expectations by adding a preliminary stage to the decision process, in which the DM chooses a signal from a set of feasible signals. Spiegler establishes that the DM's behavior throughout the two-stage decision problem, and particularly his choices between signals in the first stage, is indistinguishable from those of a standard DM who tries to maximize the expectation of some state-dependent utility function over actions. In the context of preferences over acts, Strzalecki (2011) shows that for the class of multiplier preferences, there is no way of disentangling risk aversion from concern about model misspecification. Consequentially, he points out that "...policy recommendations based on such a model would depend on a somewhat arbitrary choice of the representation. Different representations of the same preferences could lead to different welfare assessments and policy choices, but such choices would not be based on observable data." Some of the papers above suggest an additional choice data that is sufficient to distinguish between the models. For example, in Dekel and Lipman (2011) as well as in Chatterjee and Krishna (2009), the two indistinguishable models predict different choices from menus, suggesting that data on second-stage choice is needed.

Can our model be distinguished from Savage's? In the spirit of the papers above, can we find a domain of choice in which optimism and pessimism may be distinguished both from each other and from a simple expected utility maximizer? One might think that a person with pessimistic (or optimistic) beliefs could be identified on the basis of choice by confronting the DM with choices over acts whose prizes are based on objective probabilities; for example, by soliciting enough responses to questions of the sort "Which do you prefer: a lottery that gives you \$100 with probability .5 and \$0 with probability .5 or getting \$40 for sure?" The responses to such questions would allow one to elicit the DM's utility function, and once this is known, one can determine whether the DM has stake-dependent probabilities over subjective events. (See the discussion in Section 4). This argument, however, presumes that the DM takes these probabilities at face value. Her choices will depend on the likelihoods in her mind of getting the various prizes. She may well think "I'm very unlikely to get the \$100 if I take the gamble - I never win anything." There is no compelling reason to believe that a pessimistic or optimistic DM's mental assessment of the likelihood of an event can be controlled by arguing what the DM "should" believe. This is reminiscent of a point in Pesendorfer (2006), who questions whether it is reasonable in the face of a failure of one aspect of the standard model to assume that the other aspects continue to apply.

Another approach for attempting to distinguish the models would be to study their behavior after receiving additional information, such as a signal about the true state. This involves specifying a somewhat arbitrary updating rule for optimists and pessimists. Since the DM is distorting the subjective probability distribution p, the most standard choice that comes to mind is to distort the Bayesian updated subjective probability distribution (this is similar to the choice of Spiegler (2008) in extending Brunnermeier and Parker's (2005) model to include signals). The updated beliefs of the optimist or pessimist would lead to choices that would, once again, be indistinguishable from an expected utility maximizer using Bayesian updating. Other choices of updating rules (for example, performing a Bayesian update on the distortion) may well lead to different behaviors. In general, it is not immediately clear why an individual who views the likelihood of events as dependent upon their consequences would employ standard Bayesian updating.

A reduced form interpretation. The pessimistic representation may capture a DM who suffers from performance anxiety. Consider a DM who is about to perform a task that yields high stakes upon succeeding. But the high stakes may increase his anxiety and stress, which, in turn, reduces the probability of success.[11] If the DM can foresee his performance anxiety and were to bet on his success, then he will bias down that probability.

Alternatively, suppose that the probability of success depends monotonically on the amount of effort the DM exerts. In a stylized principal-agent model, the principal, who likes to induce the agent to exert high effort, will need to incentivise him by increasing the stakes. If the agent responds to these incentives and puts higher effort, then the probability of success is indeed higher. More generally, one can think of a situation in which the agent simply believes that he can affect the probabilities by putting more effort. Our notion of optimism captures such a situation, without explicitly mentioning the effort dependence; the higher the stakes are, the more likely the DM believes the good state occurs.[12]

[11] Ariely (2010) found that potentially large bonuses may lower employee performance due to stress.

[12] A related idea appears in Drèze (1987). Drèze develops a decision theory with state-dependent preferences and moral hazard, based on the Anscombe and Aumann model, in which the reversal of order assumption (which requires indifference between ex-ante and ex-post randomization) is relaxed. The DM might display preferences for ex-ante randomization (that is, preferences to knowing the outcome of a lottery before the state of nature becomes known) if he believes that he can influence the outcome by his (unobserved) actions. Drèze derives a representation that entails the maximization of subjective expected utility over a convex set of subjective probability measures.

Appendix

Proof of Proposition 1

Using $\frac{\partial f}{\partial z}(a) := \frac{\partial f}{\partial z}|_{z=a}$, and assuming that x_2 is a point such that $\frac{\partial f}{\partial z}(u(x_2)) \in (0, \infty)$:

$$\frac{\partial P_1(x_1, x_2; p_1)}{\partial x_1} \leq 0 \text{ for all } x_1, x_2 \text{ and for fixed } p_1 \in (0,1)$$

if and only if

$$\frac{(f(u(x_1)) - f(u(x_2)))}{\frac{\partial f}{\partial z}(u(x_1))} \leq \frac{(f(u(x_2) + p(u(x_1) - u(x_2))) - f(u(x_2)))}{\frac{\partial f}{\partial z}(v(x_1, x_2))\, p}.$$

Divide both sides by $u(x_1) - u(x_2)$. If $u(x_1) - u(x_2) > 0$ then

$$\frac{(f(u(x_1)) - f(u(x_2)))}{\frac{\partial f}{\partial z}(u(x_1))(u(x_1) - u(x_2))} \leq \frac{(f(u(x_2) + p(u(x_1) - u(x_2))) - f(u(x_2)))}{\frac{\partial f}{\partial z}(pu(x_1) + (1-p)u(x_2))(u(x_1) - u(x_2))\, p}. \tag{4}$$

A sufficient condition is that the right hand side (4) is decreasing in p (so that it is minimized at $p = 1$). Take $f(z) = z^r$ with $r > 1$. Note that the right hand side (4) is then

$$\frac{(u(x_2) + p(u(x_1) - u(x_2)))^r - (u(x_2))^r}{p(u(x_1) - u(x_2))\, r\, (u(x_2) + p(u(x_1) - u(x_2)))^{r-1}} = \frac{1}{r}\left[\sum_{j=0}^{r-1}\left(\frac{u(x_2)}{u(x_2) + p(u(x_1) - u(x_2))}\right)^j\right], \tag{5}$$

since the term $(u(x_2) + p(u(x_1) - u(x_2)))^r - (u(x_2))^r$ can be factored as

$$p(u(x_1) - u(x_2))\left[\sum_{j=0}^{r-1}(u(x_2) + p(u(x_1) - u(x_2)))^{r-1-j}(u(x_2))^j\right].$$

Since all the summands in the right hand side of (5) are decreasing with p (recall that $(u(x_1) - u(x_2)) > 0$ and $u(x_2) > 0$), the whole term is decreasing with p.

If $u(x_1) - u(x_2) < 0$ then

$$\frac{(f(u(x_1)) - f(u(x_2)))}{\frac{\partial f}{\partial z}(u(x_1))(u(x_1) - u(x_2))} \geq \frac{(f(u(x_2) + p(u(x_1) - u(x_2))) - f(u(x_2)))}{\frac{\partial f}{\partial z}(pu(x_1) + (1-p)u(x_2))(u(x_1) - u(x_2))\, p}. \tag{6}$$

Again, a sufficient condition would be that the right hand side of (6) is increasing with p. But, for $f(z) = z^r$ this is indeed the case since the factorization of the right hand side is as above, but now the coefficient of p in the denominator is $(u(x_1) - u(x_2)) < 0$.

For the optimist, we would like the following condition to hold:

$$\frac{\partial P_1(x_1,x_2;p_1)}{\partial x_1} \geq 0 \text{ (for all } (x_1,x_2) \text{ and for fixed } p \in (0,1)\text{).}$$

Mimicking the calculations in the pessimism case, the condition is (for $u(x_1) - u(x_2) > 0$)

$$\frac{(f(u(x_1)) - f(u(x_2)))}{\frac{\partial f}{\partial z}(u(x_1))(u(x_1) - u(x_2))} \geq \frac{(f(u(x_2) + p(u(x_1) - u(x_2))) - f(u(x_2)))}{\frac{\partial f}{\partial z}(pu(x_1) + (1-p)u(x_2))(u(x_1) - u(x_2))p}. \tag{7}$$

A sufficient condition is that the right hand side (7) is increasing in p (so that it is maximized at $p = 1$). Take, for example, $f(z) = z^{\frac{1}{r}}$ with $r > 1$. By factoring and rearranging, note that the right hand side (7) is then,

$$\begin{aligned}
& \frac{(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1}{r}} - u(x_2)^{\frac{1}{r}}}{(u(x_1) - u(x_2))p\frac{1}{r}(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1-r}{r}}} \\
= & \frac{(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1}{r}} - u(x_2)^{\frac{1}{r}}}{(u(x_2) + p(u(x_1) - u(x_2)) - u(x_2))\frac{1}{r}(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1-r}{r}}} \\
= & \frac{(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1}{r}} - u(x_2)^{\frac{1}{r}}}{\left(\left((u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1}{r}}\right)^r - \left(u(x_2)^{\frac{1}{r}}\right)^r\right)\frac{1}{r}(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1-r}{r}}} \\
= & \frac{1}{\left[\sum_{j=0}^{r-1}(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{r-1-j}{r}}u(x_2)^{\frac{j}{r}}\right]\frac{1}{r}(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{1-r}{r}}} \\
= & \frac{r(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{r-1}{r}}}{\left[\sum_{j=0}^{r-1}(u(x_2) + p(u(x_1) - u(x_2)))^{\frac{r-1-j}{r}}u(x_2)^{\frac{j}{r}}\right]} \\
= & \frac{r}{\left(\frac{\sum_{j=0}^{r-1}(u(x_2)+p(u(x_1)-u(x_2)))^{\frac{r-1-j}{r}}u(x_2)^{\frac{j}{r}}}{(u(x_2)+p(u(x_1)-u(x_2)))^{\frac{r-1}{r}}}\right)} \\
= & \frac{r}{\left(\sum_{j=0}^{r-1}\left(\frac{u(x_2)}{(u(x_2)+p(u(x_1)-u(x_2)))}\right)^{\frac{j}{r}}\right)}.
\end{aligned}$$

Since $u(x_1) > u(x_2)$, each summand is decreasing with p and hence the entire expression is increasing with p. Similarly, if $u(x_1) < u(x_2)$ the expression is decreasing with p (and the $\geq$ sign in (7) is reversed). ∎

References

Anscombe, F.J. and R.J. Aumann (1963), "A Definition of Subjective Probability," *Annals of Mathematical Statistics,* 34: 199-205.

Ariely, Dan (2010), "The Upside of Irrationality: The Unexpected Benefits of Defying Logic at Work and at Home," Harper (publisher).

Brunnermeier , Markus K. and Jonathan Parker (2005), "Optimal Expectations," *American Economic Review,* 95(4): 1092-1118.

Chatterjee, Kalyan and R. Vijay Krishna (2009), " A 'Dual Self' Representation for Stochastic Temptation," *American Economic Journal: Microeconomics,* 1: 148-167.

Chew, Soo Hong and Jacob Sagi (2008), "Small Worlds: Modeling Attitudes Toward Sources of Uncertainty," *Journal of Economic Theory,* 139: 1-24.

Dekel, Eddie and Bart Lipman (2010), "How (Not) to Do Decision Theory," *Annual Review of Economics* 2: 257-82.

Dekel, Eddie and Bart Lipman (2011), "Costly Self Control and Random Self Indulgence", mimeo.

Drèze, Jacques H. (1987), "Decision Theory with Moral Hazard and State-Dependent Preferences," in *Drèze, J. H. Essays on Economic Decisions Under Uncertainty.* Cambridge University Press, Cambridge.

Gilboa, Itzhak, Andrew Postlewaite and David Schmeidler (2009), "Is It Always Rational to Satisfy Savage's Axioms?," *Economics and Philosophy*, 25:285-296.

Gilboa, Itzhak and Massimo Marinacci (2011), "Ambiguity and the Bayesian Paradigm," mimeo.

Grant, Simon and Edi Karni (2005), "Why Does It Matter that Beliefs and Valuations be Correctly Represented?" *International Economic Review* 46: 917-934.

Grant, Simon, Ben Polak and Tomasz Strzalecki (2009), "Second-Order Expected Utility," *mimeo.*

Gul, Faruk (1992), "Savage's Theorem With a Finite Number of States," *Journal of Economic Theory 57(1).*

Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and Self-Control," *Econometrica* 69: 1403–1436.

Hey, John (1984), "The Economics of Optimism and Pessimism," *Kyklos* 37: 181-205.

Karni, Edi (2009), "A Theory of Medical Decision Making Under Uncertainty," *Journal of Risk and Uncertainty* 39: 1-16.

Karni, Edi (2011), "Subjective Probabilities on a State Space," *American Economic Journal: Microeconomics*, forthcoming.

Pesendorfer, Wolfgang (2006), "Behavioral Economics Comes of Age: A Review Essay on Advances in Behavioral Economics," *Journal of Economic Literature* 44: 712-721.

Savage, L. J. (1954), *The Foundation of Statistics*, Wiley, New York.

Spiegler Rani (2008), "On Two Points of View Regarding Revealed Preferences and Behavioral Economics," *The Foundations of Positive and Normative Economics*, OUP: 95-115.

Strzalecki, Tomasz (2011), "Axiomatic Foundations of Multiplier Preferences," *Econometrica* 79: 47-73.

von Neumann, John (1955), 'Method in the Physical Sciences', *Collected Works*, vol. 6, 491-498; originally in L. Leary (Ed.): *The Unity of Knowledge*, New York: Doubleday, 157-164.

Wakker, Peter (1984), "Cardinal Coordinate Independence for Expected Utility," *Journal of Mathematical Psychology* 28: 110-117.

Wakker, Peter (1990), "Characterizing Optimism and Pessimism Directly Through Comonotonicity,"*Journal of Economic Theory* 52: 453-463.

Wakker, Peter (2008), "Lessons Learned by (from?) an Economist Working in Medical Decision Making," *Medical Decision Making* 28: 690-698.

Wakker, Peter (2010), "Prospect Theory For Risk and Uncertainty," Cambridge University Press, Cambridge, UK.

Supplement to "Optimism and Pessimism with Expected Utility" by DPR

In this supplement, we (1) prove the comparative statics for the example of an optimistic distortion described in the text of Section 3, and (2) provide an alternate construction of a probability distortion and study the corresponding comparative statics.

For ease of exposition, we will assume that $x_1 > \cdots > x_n$; it is straightforward to handle the general case at the cost of more complicated notation.

Comparative statics for the optimism example

Proposition 2. *Consider $f(z) = z^r$ for $r \in (0,1)$. Let $P_\lambda(x) = \lambda \cdot (1,0,0,...,0) + (1-\lambda) \cdot p$, where λ solves $c(x; f(u), P_\lambda(x)) = c(x; u, p)$. Then the probability placed by $P_\lambda(x)$ on the best outcome increases when the worst outcome becomes worse.*

Proof. The λ solving $c(x; f(u), P_\lambda(x)) = c(x; u, p)$ is given by

$$\lambda = \frac{f\left(\sum_{i=1}^n p_i u(x_i)\right) - \sum_{i=1}^n p_i f\left(u\left(x_i\right)\right)}{f\left(u\left(x_1\right)\right) - \sum_{i=1}^n p_i f\left(u\left(x_i\right)\right)}. \tag{8}$$

By construction of $P_\lambda(x)$, it suffices to show that $\frac{\partial \lambda}{\partial x_n} < 0$ when $f(z) = z^r$ for $r \in (0,1)$. Taking the derivative of (8) and rearranging, this is true if and only if

$$\frac{f(u(x_1)) - \sum_{i=1}^n p_i f(u(x_i))}{f'(u(x_n))} < \frac{f(u(x_1)) - f(\sum_{i=1}^n p_i u(x_i))}{f'(\sum_{i=1}^n p_i u(x_i))}. \tag{9}$$

For fixed x, define

$$F(p) = \frac{u(x_1)^r}{u(x_n)^{r-1}} - \frac{1}{u(x_n)^{r-1}} \sum_{i=1}^n p_i u(x_i)^r - \frac{u(x_1)^r}{(\sum_{i=1}^n p_i u(x_i))^{r-1}} + \sum_{i=1}^n p_i u(x_i).$$

Then (9) is equivalent to $F(p) < 0$. We proceed in two steps. First, we show that $F(p)$ is convex in p. Let $\zeta(x) := (1-r)r(\sum_{i=1}^n p_i u(x_i))^{-r-1} > 0$. Then the Hessian of F is given by

$$H_F(p) = \zeta(x) \begin{pmatrix} (u(x_1))^2 & (u(x_1))(u(x_2)) & \cdots & (u(x_1))(u(x_n)) \\ (u(x_1))(u(x_2)) & (u(x_2))^2 & \cdots & (u(x_2))(u(x_n)) \\ \vdots & \cdots & \ddots & \vdots \\ (u(x_1))(u(x_n)) & (u(x_2))(u(x_n)) & \cdots & (u(x_n))^2 \end{pmatrix},$$

where the i, j-th entry is the cross-partial derivative of F with respect to p_i and p_j. H_F is a positive constant times a Gramian matrix, hence is positive semi-definite.

Second, since F is a convex function, it achieves its maximum on the vertices of the probability simplex. Let e^i be an $n-$dimensional vector such that $e^i_j = 0$ for $j \neq i$ and $e^i_i = 1$. Then for any $p = (p_1, p_2, ..., p_n)$, $F(p) = F\left(\sum_i p_i e^i\right) \leq p_i \sum_i F\left(e^i\right)$. Since $x_1 > x_i$ for $i = 2, .., n$, note after some rearrangement that

$$F(e_i) = \begin{cases} 0 & \text{if } i = 1 \\ 0 & \text{if } i = n \\ (u(x_1)^r - u(x_i)^r)(\frac{1}{u(x_n)^{r-1}} - \frac{1}{u(x_i)^{r-1}}) < 0 & \text{otherwise.} \end{cases}$$

Hence $F(p)$ is weakly negative, and strictly so whenever we are not in the two state case. ■

An alternate construction

Here we demonstrate one (of the many) additional ways to construct a pessimistic distortion. Let $P_i(x; p)$ be the probability of the state that yields the ith-best outcome among all n states. To fix a distribution, we normalize the ratio of the probabilities of any pair of states (except of the state giving the worst prize) to be equal to the ratio of the Savage probabilities corresponding to these two states. That is,

$$\frac{P_j(x; p)}{P_{j+1}(x; p)} = \frac{p_j}{p_{j+1}} \text{ for } j = 1, 2, ..., n - 2. \tag{10}$$

Substitute (10) and $P_n(x; p) = 1 - \sum_{i=1}^{n-1} P_i(x; p)$ in (3) and solving for $P_1(x; p)$ yields

$$P_1(x; p) = p_1 \frac{f(v(x)) - f(u(x_n))}{\sum_{i=1}^{n} p_i f(u(x_i)) - f(u(x_n))}. \tag{11}$$

When f is convex, Jensen's inequality says that $\sum_{i=1}^{n} p_i f(u(x_i)) \geq f(v(x))$. Thus

$$P_1(x; p) = p_1 \frac{f(v(x)) - f(u(x_n))}{\sum_{i=1}^{n} p_i f(u(x_i)) - f(u(x_n))} \leq p_1.$$

Furthermore, for $j = 2, ..., n - 1$, $P_j(x; p) = \frac{p_j}{p_1} P_1(x; p) \leq p_j$. Hence

$$P_n(x; p) = 1 - \sum_{i=1}^{n-1} P_i(x; p) \geq 1 - \sum_{i=1}^{n-1} p_i = p_n.$$

We have now proved the first part of the upcoming result.

Proposition 3. *For a convex transformation function f, the probability distribution constructed above is first-order stochastically dominated by the original Savage distribution. Moreover, if $f(z) = z^r$ for $r > 1$, an increase in the best prize causes a decrease in all probabilities except that of the worst outcome.*

Proof. It is enough to show that $\frac{\partial P_1}{\partial x_1}(x;p) \leq 0$, since $P_j(x;p) = \frac{p_j}{p_1} P_1(x;p)$ for $j \leq n-2$. Taking derivatives and rearranging, if

$$\frac{\left(\sum_{i=1}^n p_i \left(u\left(x_i\right)\right)^r - \left(u\left(x_n\right)\right)^r\right)}{\left(u\left(x_1\right)\right)^{r-1}} \leq \sum\nolimits_{i=1}^n p_i u\left(x_i\right) - \frac{\left(u\left(x_n\right)\right)^r}{\left(\sum_{i=1}^n p_i u\left(x_i\right)\right)^{r-1}}.$$

Let $p = (p_1, ..., p_n)$ and define, for fixed consequences x,

$$F(p) = \frac{\left(\sum_{i=1}^n p_i \left(u\left(x_i\right)\right)^r - \left(u\left(x_n\right)\right)^r\right)}{\left(u\left(x_1\right)\right)^{r-1}} - \sum\nolimits_{i=1}^n p_i u\left(x_i\right) + \frac{\left(u\left(x_n\right)\right)^r}{\left(\sum_{i=1}^n p_i u\left(x_i\right)\right)^{r-1}}. \tag{12}$$

We proceed from here in two steps. First, we show that $F(p)$ is convex. Let $\zeta(x) := -r(1-r)\left(u\left(x_n\right)\right)^r \left(\sum_{i=1}^n p_i u\left(x_i\right)\right)^{-r-1} > 0$. The Hessian matrix of F is

$$H_F(p) = \zeta(x) \begin{pmatrix} \left(u\left(x_1\right)\right)^2 & \left(u\left(x_1\right)\right)\left(u\left(x_2\right)\right) & \cdots & \left(u\left(x_1\right)\right)\left(u\left(x_n\right)\right) \\ \left(u\left(x_1\right)\right)\left(u\left(x_2\right)\right) & \left(u\left(x_2\right)\right)^2 & \cdots & \left(u\left(x_2\right)\right)\left(u\left(x_n\right)\right) \\ \vdots & \cdots & \ddots & \vdots \\ \left(u\left(x_1\right)\right)\left(u\left(x_n\right)\right) & \left(u\left(x_2\right)\right)\left(u\left(x_n\right)\right) & \cdots & \left(u\left(x_n\right)\right)^2 \end{pmatrix},$$

where the i,j-th entry is the cross-partial derivative of F with respect to p_i and p_j. H_F is a positive constant times a Gramian matrix, hence is positive semi-definite.

Second, since F is a convex function, it achieves its maximum on the vertices of the probability simplex. Since $x_1 > x_i$ for $i = 2, 3, .., n$, note that

$$F\left(e^i\right) = \frac{\left(u\left(x_i\right)\right)^r - \left(u\left(x_n\right)\right)^r}{\left(u\left(x_1\right)\right)^{r-1}} - \frac{\left(u\left(x_i\right)\right)^r - \left(u\left(x_n\right)\right)^r}{\left(u\left(x_i\right)\right)^{r-1}} \begin{cases} = 0 & \text{if } i = 1 \\ = 0 & \text{if } i = n \\ < 0 & \text{otherwise.} \end{cases}$$

We conclude that $F(p)$ is weakly negative (and strictly negative whenever we are not in the two-state case), and thus $\frac{\partial P_1}{\partial x_1}(x;p) \leq 0$. ■