EIEF WORKING PAPER SERIES

Einaudi Institute for Economics and Finance

EIEF Working Paper 04/08
November 2008

Relative Minimax

by

Daniele Terlizzese
(EIEF and Bank of Italy)

Relative Minimax*

Daniele Terlizzese
EIEF[†] and Bank of Italy

May 2008

Abstract

To achieve robustness, a decision criterion that recently has been widely adopted is Wald's minimax, after Gilboa and Schmeidler (1989) showed that (one generalisation of) it can be given an axiomatic foundation with a behavioural interpretation. Yet minimax has known drawbacks. A better alternative is Savage's minimax regret, recently axiomatized by Stoye (2006). A related alternative is relative minimax, known as competitive ratio in the computer science literature, which is appealingly unit free. This paper provides an axiomatisation with behavioural content for relative minimax.

JEL Classification Numbers: D81, B41, C19

Keywords: Robust decisions, Minimax, Minimax Regret.

*The author thanks Fernando Alvarez, Luca Anderlini, David Andolfatto, Francesco Lippi, Ramon Marimon, Karl Schlag and Jörg Stoye for comments and useful discussions. The views are personal and do not involve the institutions with which he is affiliated. The usual disclaimer applies.

[†]Einaudi Institute for Economics and Finance, EIEF, via Due Macelli 73, 00187 Rome, Italy. E-mail: daniele.terlizzese@eief.it, daniele.terlizzese@bancaditalia.it

1 Introduction

The interest in robust control techniques has recently been on the rise. Confronted with a considerable amount of uncertainty concerning the correct specification of the model representing the economy and the parameters of any given specification, economists and policy makers seem to find increasing appeal in the notion that choices (policy choices, but also microeconomic choices like for example pricing decisions) should be robust, meaning that their consequences should remain relatively good irrespective of the model of the economy that turns out to be the best approximation of reality.

The want for robustness has been met in different ways. A natural tack is provided by the standard approach to decision under uncertainty: specify a prior probability distribution over a set of alternative representations of economic reality (models) and find the policy whose expected utility is maximal; a flat (uniform) prior over the models is sometimes adopted in this context, to capture the state of high uncertainty under which the choice is made (see for example Levin, Wieland and Williams, 2003, or Onatski and Williams, 2003). However, a skepticism on the possibility to express a prior over the alternative models (often justified invoking a notion of "Knightian uncertainty", whereby certain events would not be amenable to a probabilistic assessment) has led many studies to focus on decision criteria that do not require the use of prior probabilities. In particular, the decision criterion that is often adopted is Wald's minimax (Wald, 1950), which selects the policy whose minimal utility across states of the world (here, models) is maximal.[1] Indeed, most of the recent papers (see Hansen and Sargent, 2001, Onatski and Williams, 2003, or Giannoni, 2002) justify the decision criterion adopted with reference to a generalisation of Wald's original approach, provided by Gilboa and Schmeidler (1989; GS); in the latter paper a *set* of

[1] Strictly speaking, the decision criterion as described in the text should be called maximin, while Wald's formulation requires to minimise the maximum of the negative utility. This terminologic distinction will be ignored in the following.

priors is allowed, and the choice is that which maximises the expected utility taken with respect to the "least favourable prior". GS "maxmin expected utility" reduces to standard minimax if the set of priors is large enough (in particular when it includes the degenerate priors that put all the probability mass on each single state), and therefore makes in practice little difference as to the selection of the putative robust policy. It has however the distinctive advantage of an explicit axiomatic foundation.

Yet, there are circumstances in which neither minimax nor its GS generalisation provide a satisfactory decision criterion. Consider a simple example in which there are only two possible actions and two states of the world: action 1, yielding a utility of 1 in state 1 and of 10 in state 2, and action 2 yielding a utility of 0.99 in state 1 and of 40 in state 2 (see Table 1).

Table 1 (*utiles*)

	state 1	state 2
action 1	1	10
action 2	0.99	40

The minimax solution is clearly action 1 (whether or not randomisation between the two actions, or mixing, is allowed). To compute GS maxmin expected utility, consider a set of probability distributions over the two states, which in this case is represented by the set $P = \{(p, 1-p) | 0 \leq \underline{p} \leq p \leq \overline{p} \leq 1\}$, for some values $\underline{p}$ and $\overline{p}$, where p represents the probability of state 1. The two parameters $\underline{p}$ and $\overline{p}$ capture the degree of uncertainty that characterizes the problem; a condition of Knightian uncertainty is probably best interpreted as corresponding to the largest possible set of priors, leading to the natural boundaries $\underline{p} = 0$ and $\overline{p} = 1$. Whatever the values of the boundaries, it is easily verified that the lowest expected utility of each action is achieved for $p = \overline{p}$. If $\overline{p}$ is close enough to 1 (in particular, if $\overline{p} > \frac{3000}{3001}$) then the action that maximises the minimal expected utility is again action 1 (randomisation is never optimal with maxmin expected utility).

There is something unsatisfactory about action 1 being chosen. While its lowest utility is indeed largest, it obtains in one state where the best that can be achieved is only marginally better than the alternative, but yields a utility of 10 in a state in which a much better outcome could be achieved. This latter feature is totally neglected by the minimax criterion, that only focuses on the lowest utility, and is given little weight by maxmin expected utility, at least as long as the assumption of Knightian uncertainty is interpreted as leading to a large enough set of priors.

The problem with these decision criteria, in the case at hand, is however deeper than that. It can be shown that they are insensitive (totally, in the case of minimax; almost completely, for maxmin expected utility) to the availability of information concerning the state, no matter how precise (as long as it is not perfect). This point, with reference to minimax, was originally made by Savage (1954). Imagine that, in the above example, it is available, free of charge, a noisy signal of the state: if state 1 (2) is true, the signal says so with probability q (h), and these probabilities are known to the decision maker, who can select action 1 or 2 depending on the signal (obviously, action 1 if the signal suggests that state 1 is true, action 2 otherwise).[2] This new action, call it action 3, has (expected) utility equal to $0.99 + 0.01q$ in state 1, $10 + 30h$ in state 2. It can easily be checked that minimax will never select action 3, however large are h or q (as long as they are less than 1), and that it will always select action 1 (irrespective to the possibility of randomisation), being thus totally unaffected by the availability of the signal. This is clearly a serious drawback (one that makes minimax, according to Savage (1954), *"utterly untenable for statistics"*). As to maxmin

[2] For definitiveness, assume that h and q are at least 0.5, i.e. that the signal is positively correlated with the state; otherwise, it would be rational to define action 3 as dictating action 1 if the signal suggests that state 2 is true, and action 2 if the signal suggest that state 1 is true, and the results in the text would still go through.

expected utility, a little algebra suffices to show that action 1 would always be chosen if $\overline{p} = 1$, neglecting altogether the signal independently of its precision (as long as h or q are less than 1).[3] There are thus circumstances in which GS decision criterion is insensitive to the availability of (noisy but very precise) information about the state. Similarly to what claimed for standard minimax, this seems a serious drawback.

These difficulties arise whenever, across all actions, the largest utility in one state is smaller than the smallest utility in the other state.[4] Technically, this implies that to identify the minimax action only the utilities in the "low" state needs to be compared (no randomisation is ever required); hence, the availability of a new action that exploits the signal (whose utilities in each state are convex combinations of the original utilities in the same state) would make no difference. A similar story holds for maxmin expected utility, as long as the set of priors includes a weight on the "low" state close enough to 1.

In intuitive terms, these two decision criteria run into problems since they make no (or very little) difference as to whether a given low utility is obtained in one state in which *all* utilities (associated to *all* available actions) tend to be small, or rather in one state where utilities associated to some other actions are large. The problem would not occur if the utilities were "normalised" in each state, on the basis of the largest utility that could be achieved in that state. This corresponds to interpret a putative robust action as one that should not lead in any state to too large a departure from the best. To put it differently, robustness seems best interpreted as a *relative* concept: a robust action is one that never does too badly, *relative to the best*

[3] It should be aknowledged that in this case action 3 might be chosen if $\overline{p}$ is allowed to be strictly smaller than 1, and in particular if it falls in the interval $[\frac{30(1-h)}{30(1-h)+0.01q}, \frac{30h}{30h+0.01(1-q)}]$. At any rate, provided that h and q are smaller than 1, this interval is very narrow: for example, for h and q not larger than 0.95, its length is at most 0.006.

[4] Puppe and Schlag (2006) show that when there is no overlap among the consequences of the various actions in different states the axioms provided by Milnor (1954) for minimax and minimax regret no longer suffice to characterize these decision criteria.

that could be done.

To achieve this, divide the utilities of all actions in a given state by the state-dependent largest utility, and apply the minimax criterion to the "relative utility" thus obtained, i.e. select the action whose smallest "relative utility" is maximal. In Table 2 the original problem is transformed by dividing the entries in each column of Table 1 by the maximum of the column.

Table 2 (*utiles/max(utiles)*)

	state 1	state 2
action 1	1	0.25
action 2	0.99	1

Applying the minimax criterion to the transformed problem in Table 2 leads to action 2, if no mixing is allowed (or, in case of mixing, to a mixture that gives probability $\frac{1}{76}$ to action 1 and $\frac{75}{76}$ to action 2). This seems a more sensible choice. If a signal about the state were available, as before, action 3 would be chosen, alone (if $0.99 + 0.01q \leq 0.25 + 0.75h$) or as a part of a mixture (if $0.99 + 0.01q > 0.25 + 0.75h$, in which case action 3 would be chosen with probability $\frac{0.01}{0.01q+0.75(1-h)}$, and action 2 otherwise). Therefore, the choice would not neglect the availability of information about the states. This decision criterion, known in the computer science literature as *competitive ratio*, will be called in this paper *relative minimax* (to avoid the possible confusion resulting from the term competitive, which in the economic literature has a different meaning). There is a tight connection between relative minimax and the so called *minimax regret*, which will be further discussed in the next Section. Both decision criteria introduce a normalisation based on the state-dependent largest utility, and in this way they implicitly take into account the fact that the same level of utility in different states might mean something different. Even with minimax regret the choice would be affected by the availability of information about the states. It is interesting that the notion of a state-dependent best option providing a benchmark against which to assess the available choices is starting (?) to attract attention in the literature on behavioural finance [quotes; complete].

However, relative minimax lacks so far an axiomatic foundation with an explicit behavioural interpretation, like that provided by Gilboa and Schmeidler (1989) for minimax; Stoye (2006) presents axiomatic foundations for minimax regret (see the following Section for a brief discussion). This paper fills the gap, by presenting a set of axioms concerning the preferences of an agent that rigorously justify the adoption of relative minimax as a decision criterion. In particular, the axioms will be seen to imply the existence of a utility function that, in each state, assigns a real value to the consequences of each action, unique up to a positive affine transformation, such that one action is preferred to another if and only if the smallest "standardised relative utility" of the first is larger than that of the second. The standardised relative utility in each state, whose precise definition will be given in Section 5, is always in $[0, 1]$, and is invariant to any positive affine transformation of the utility. Moreover, an affine transformation of the utility can be chosen that makes the standardised relative utility coincide with the normalisation appearing in the relative minimax.

In the next Section a brief discussion of related work will be presented. In Section 3 a motivating example, taken from Altissimo, Siviero and Terlizzese (2005), will be shown. In Section 4 the axioms will be introduced, and in Section 5 the representation theorem will be proved (all proofs are collected in the Appendix). Section 6 briefly concludes.

2 Related work

What is known in the literature as (Savage's) minimax regret is the interpretation that Savage (1954) gave of Wald's (1950) original formulation of the minimax criterion. Indeed, Savage attributed to Wald the idea to minimise the maximal *difference* from the highest achievable utility in each state of the world, a difference that became known with the term "regret" (Savage himself called it "loss", and objected to the use of the term regret, but the latter became nevertheless entrenched). From a practical point of view, the difference between relative minimax and minimax regret is that the latter considers the *difference* with the utility of the best (state contingent) consequence, while the former considers the *ratio*. In terms of the example presented before, the minimax regret criterion would lead to consider the following transformation of Table 1, where the entries are the difference between the column maximum and the original value:

Table 3 (*max(utiles)* $-$ *utiles*)

	state 1	state 2
action 1	0	30
action 2	0.01	0

The largest regret of action 1 is 30, while the largest regret of action 2 is 0.01, and the latter action would then be chosen according to the minimax regret criterion (if no mixing is considered; otherwise action 2 would be chosen with probability $\frac{3000}{3001}$). In this example this is the same choice that would result from relative minimax (neglecting mixtures). However, it is not necessarily the case that the two criteria yield the same choice. Consider for example the following choice problem:

Table 4 (*utiles*)

	state 1	state 2
action 1	1	9
action 2	0.1	10

where it is easily checked that (neglecting mixtures) relative minimax (as well as standard minimax) would select action 1, while minimax regret would select action 2.[5] The reason why relative minimax would lead to action 1 is that, *in relative terms*, action 2 is 10 times worse than action 1 in state 1, while it is only slightly better in state 2. The large difference in the level of utility in the two states, however, makes the *absolute* advantage of action 1 in state 1 (0.9 utiles) smaller than the absolute advantage of action 2 in state 2 (1 utile), and this is what matters for minimax regret. It might be argued that the normalisation of utilities in terms of ratios is more appropriate than the normalisation in terms of differences, as the former is "scale free" while the latter remains dependent on large differences in the level of utility across states. However, whether one finds more compelling a comparison based on ratios or on differences might depend on circumstances and it is largely a matter of taste. Spelling out the axioms that justify the different choice criteria is a way to provide guidance to any such debate.[6]

The relative minimax criterion has been proposed in the economic literature only very recently (see Altissimo, Siviero and Terlizzese, 2005)[7]. However, as mentioned before, it is extensively used in the theoretical computer science literature (in particular in the optimisation of the on-line algorithms), where it is known as *competitive ratio.* Brafman and Tennenholtz (1999) provide an axiomatic foundation for minimax, that turns out to support at the

[5]Considering mixtures, action 1 would be selected by relative minimax with a probability of $\frac{9}{10}$, by minimax regret with a probability of $\frac{9}{19}$ (standard minimax would select action 1 for sure).

[6]Clearly, a "scale free" comparison could also be obtained by taking the logarithm of the entries in Table 4 and then comparing the largest regret computed on the transformed values, as the differences among the (transformed) levels carry the same information about the preference as the ratios among the original levels. More generally, by using a log transformation, the axiomatisation of relative minimax that is provided in this paper could be straightforwardly adapted to represent minimax regret as well. However the resulting representation would no longer be invariant to affine transformations of the utility. Moreover, the behavioural content of some of the axioms will be seen to hinge upon a comparison in relative terms, so that it is relative minimax which is their "natural" implication.

[7]Interestingly, the recent literature assessing the robustness of various policy rules often uses the "relative normalisation" *as a diagnostic criterion.* Taking for example Levin, Wieland and Williams (2003), the performance of a given monetary policy rule in the various models considered is assessed by showing its loss divided by the optimized loss for each particular model. This "relative loss", however, is not used to *design* a robust rule.

same time also relative minimax and minimax regret, through appropriate transformations of the function that assigns a value to each state-action pair (what they term the “value function”). More in detail, they define as *policy* a function that specifies which action would be chosen in each of the possible subsets of the set of states of the world (each subset represents the states that are deemed possible given a particular state of the information). Brafman and Tennenholtz’s result is that if the policy satisfies three appropriate axioms, then it is possible to assign values to each state-action pair in such a way that applying minimax to these values results in the same choices that would be dictated by the policy. Moreover, applying minimax regret or relative minimax to appropriate transformations of the “value function” would still result in the same policy. Therefore, all three decision criteria are shown to yield exactly the *same choices*, provided that the values assigned to the state-action pairs are appropriately defined. The latter are in no way restricted by the proposed axioms. This serves well Brafman and Tennenholtz very practical goal of representing a given policy in a way that minimises the use of memory space in the computer. However, it is not satisfactory as a characterisation of behaviour, since all the differences among the three criteria are hidden in the choice of the “value function”, and the latter is defined residually as that which supports the decision criterion at hand. The axioms proposed in this paper, instead, will tightly restrict the utility function consistent with given preferences.

Milnor (1954) provides an axiomatisation of both minimax and minimax regret (as well as a number of other non probabilistic decision criteria, but not of relative minimax). In Milnor’s paper, however, the consequences of the various choices are directly measured in utiles, and the formulation of the axioms takes advantage of this fact so that, for example, it is possible to consider “adding a constant” to the consequences of one action, or checking that a new action has all its consequences “smaller than the corresponding consequences of all other actions”; in particular, one of the axioms uses utility differences, thus introducing the presumption of cardinally measurable utility. In this paper, the axioms can be given a direct behavioural interpretation. Stoye (2006), similarly to Milnor, considers a range of decision criteria (that include minimax regret but not relative minimax), but avoids the presumption of cardinal utility and provides them with a behavioural axiomatization. The approach in Stoye (2006) is somewhat different from the one in this paper. With both minimax regret and relative minimax, changes in the menu of actions might change the preference, as will be discussed later. Here, this difficulty is faced by fixing the menu of actions, and stating all the axioms on the single preference relationship which pertains to the given menu.[8] In Stoye, some of the axioms establish consistency requirements among preference relationships associated to different menus.

3 An example

Consider a central bank (CB) that has the standard quadratic loss function defined on inflation (π) and output gap (x): $L_t = (1-\beta)E_t \sum_{\tau=0}^{\infty} \beta^{\tau}[(\pi_{t+\tau})^2 + \alpha x_{t+\tau}^2]$, where β is the discount factor and α is the relative weight assigned to output gap variability. The economy the CB faces is described by a basic New-Keynesian model with staggered price setting behaviour and some form of indexation, summarized by a Phillips curve $\pi_t = (1-\gamma)\beta E_t(\pi_{t+1}) + \gamma\beta\pi_{t-1} + \lambda x_t + e_t$, where e is a random shock and γ measures the degree of inflation inertia. For $\gamma = 1$ the economy is fully backward-looking (fully inertial), while for $\gamma = 0$ it is fully forward-looking. The CB is completely uncertain about the true value of γ, and would like to choose its policy, which in this simplified set-up amounts to the choice of the output-gap, in a "robust" fashion, i.e. in such a way that, whatever the true value of γ, the loss is not too big. In selecting its policy, the CB considers four alternatives:

- the policy that minimizes the expected loss, assuming for γ a uniform

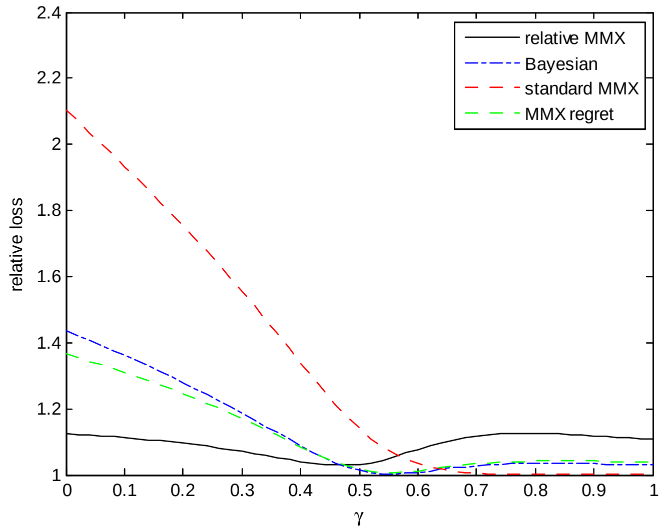
[8] There will be one exception to this. The axiom concerned, however, will not be needed to prove the basic representation theorem, but only to extend the result to a more general setting.

prior in $[0,1]$

- the policy that minimizes the maximal loss, as γ varies in $[0,1]$
- the policy that minimizes the maximal regret, as γ varies in $[0,1]$
- the policy that minimizes the maximal relative loss, as γ varies in $[0,1]$

Each line in the chart corresponds to one of these four policies and shows, as γ varies, the ratio between the associated loss and the minimal loss that could be achieved if the value of γ were known. A flat line at 1 would then correspond to a policy that always achieves the minimal loss, the quintessential robust policy. Conversely, if for a given γ the line associated to a policy is, say, 1.5, this means that, for that γ, the policy would lead to a loss which is 50% higher than the minimum achievable. The chart then shows, in a graphically convenient way, the trade-off between optimality and robustness: a line which is very close or at 1 for a subset of values of γ, but then departs substantially from 1 for other values of γ, is associated to a policy that hardly qualifies as robust.

The results presented in the chart, corresponding to a fairly standard calibration of the model and robust to alternative calibrations (see Altissimo, Siviero and Terlizzese (2005) for details), signal a clear advantage of the relative minimax policy over the alternatives: while never fully optimal (the closest it gets to the optimum is about 3% away), it never strays too far from the optimum either (at most, about 12% away). The example is clearly special, and no general conclusion should be drawn from it. However, it suggests that relative minimax policies can be of interest when robustness is sought, and motivates the effort in this paper to put this decision criterion on a firmer ground.



4 Axioms

Let A be the set of the n actions available in a particular choice problem. It is convenient, to define mixtures of actions, to arbitrarily fix one enumeration of the elements of A, so that the $i-$th element of A always refers to the same action. Let S be the set of m states of the world, denoted by $\{1, 2, ...m\}$. $m \geq 3$ will be required (this is needed for one of the axioms to bite). Each action in A yields a well defined consequence in each of the states in S. Let $\overline{C}(A, j)$ be the set whose $i-$th element is the consequence that obtains in state j when the $i-$th action in A is taken. Clearly, each of $\overline{C}(A, j)$ has n elements, not necessarily all distinct. It will be assumed the existence of a "best" consequence in $\overline{C}(A, j)$, denoted by b_j, for each j (in which sense b_j is the best consequence will be made precise by one of the axioms to follow). Also, it will be assumed the existence of at least one "worst" consequence overall, denoted by w (again, the precise meaning of this will be specified by one of the axioms below). As w is not necessarily included in any of the $\overline{C}(A, j)$, it is useful to define the sets $C(A, j) = \overline{C}(A, j) \cup \{w\}, j = 1, 2, ...m$. Later on, to assess how the results depend on the (largely arbitrary) choice of w, a more general $C^{+}(A,j) = \overline{C}(A,j) \cup \{w, w'\}, j = 1,2,...m$ will be considered.

Following Anscombe and Aumann (1963), and similarly to Gilboa and Schmeidler (1989), it will be assumed the availability of a randomising device, which allows objective lotteries to be formed with "prizes" being the consequences of the original actions and, possibly, w. Let in particular $Q(A,j)$ ($\overline{Q}(A,j)$) be the set of simple probability measures (i.e. probability measures with finite support) on $C(A,j)$ ($\overline{C}(A,j)$). Obviously, if p and $\widetilde{p}$ are in $Q(A,j)$ so is $\alpha p + (1-\alpha)\widetilde{p}$, for all $\alpha \in [0,1]$, where the convex combination is intended as a combination of probability distributions: if a given $c_i \in C(A,j)$ is assigned probability π_i^p under p and probability $\pi_i^{\widetilde{p}}$ under $\widetilde{p}$, then $\alpha\pi_i^p + (1-\alpha)\pi_i^{\widetilde{p}}$ is its probability under $\alpha p + (1-\alpha)\widetilde{p}$. The set $H_A = \prod_{j=1,2...m} Q(A,j)$, with typical element $h = (p_1, p_2, ...p_m)$, for $p_j \in Q(A,j)$, is the set of $m-$tuples of such simple probability measures (or lotteries), with the following interpretation (as in Anscombe and Aumann, 1963): according to which state occurs, the corresponding lottery in h determines the final outcome.[9] If $h = (p_j)_{j=1,2...m}$ and $k = (q_j)_{j=1,2...m}$ are in H_A, then $\alpha h + (1-\alpha)k = (\alpha p_j + (1-\alpha)q_j)_{j=1,2...m}$, $\alpha \in [0,1]$, is also in H_A.

Clearly each of the original actions can be identified with one element of H_A by appropriately choosing two degenerate probability distributions. Also, any mixture among the original actions can be represented as one element of H_A in which all components associate the same probability to the consequences of the same action. By extension, any element of H_A will be called an action. In the following, with a slight abuse of the notation, b_j, $j = 1,2...m$ and w will also denote degenerate distributions putting all the probability mass on, respectively, b_j, $j = 1,2...m$ and w. To simplify the notation, denote by (b^{-j}, p) the vector $(b_1, b_2, ..., b_{j-1}, p, b_{j+1}, ...b_m)$,

[9] We allow for the possibility that w is the only element common to the sets $C(A,j), j = 1,2...m$. In the Anscombe and Aumann setting there must be at least two elements (not indifferent to each other) common to all sets of (state-dependent) consequences (see Fishburn, 1970).

by (q^{-j}, p) the vector $(q_1, q_2, ..., q_{j-1}, p, q_{j+1}, ...q_m)$, by $(b^{-(j,i)}, p, q)$ the vector $(b_1, ..., b_{j-1}, p, b_{j+1}, ..., b_{i-1}, q, b_{i+1}, ..., b_m)$.

Finally, let $\succsim$ denote a binary relation on H_A, with $\succ$ and $\sim$ be defined in the usual way from $\succsim$.

A number of properties will be assumed for this binary relation. The first three are relatively standard.

*A*1. $\succsim$ is a preference relation (complete, reflexive, transitive).

With reference to this axiom, it is worth stressing that $\succsim$ can be interpreted as a single preference relation only with reference to the given set of actions, A (extended to allow for the possibly extraneous worst consequence w). Indeed, no axiomatization of relative minimax (or, for that matter, of minimax regret) can insist on a single preference ordering of all conceivable alternatives, since the ordering induced by the relative minimax criterion (as well as the one induced by minimax regret) can be different depending on the set of available alternatives. Consier for example the following decision problem:

Table 5

	state 1	state 2
h	12	13
k	14	12

$\longrightarrow$

	state 1	state 2
h	6/7	1
k	1	12/13

where the panel on the left presents the consequences (in utiles) of the two choices ($A = \{h, k\}$), and the panel on the right normalizes these utilities by the column maximum. According to relative maximum $k \succ h$.[10]. Suppose now we consider a third action, as in the following table:

Table 5′

[10] Allowing for mixtures the relative minimax is h with probability $\frac{7}{20}$, k with probability $\frac{13}{20}$. Note that $k \succ h$ also according to minimax regret.

	state 1	state 2
h	12	13
k	14	12
g	0	15

$\rightarrow$

	state 1	state 2
h	6/7	13/15
k	1	12/15
g	0	1

where again the left panel presents the consequences in utiles and the right one presents the normalization by the column maximum. Now $A = \{h, k, g\}$, and $h \succ k$.[11]. The preference reversal can occur when the addition of new alternatives modifies the best consequence in some state. The dependence on the choice menu[12] was already noted by Savage (1954), as a possible criticism of minimax regret (Savage tersely encapsulated the criticism by quipping: "*Fancy saying to the butcher, 'Seeing that you have geese, I'll take a duck instead of a chicken or a ham'.*"). While recognizing that this phenomenon makes it "*absurd to contend that the objectivistic minimax rule* (which amounts, in the context of this quotation, to minimax regret) *selects the best available act*", Savage himself countered the criticism by noting that the decision criterion was not intended to select the *best* available act (to which end he obviously advocated expected utility!) but rather as a "*...sometimes practical rule of thumb in contexts where the concept of "best" is impractical – impractical for the objectivist, where it amounts to the concept of personal probability, in which he does not believe at all; and for the personalist, where the difficulty of vagueness becomes overwhelming.*". Such a remark clearly anticipates the quest for robustness that underlies the current, renewed interest for decision criteria that do not involve the use of subjective probability. In-

[11] Allowing for mixtures, relative minimax is h with probability $\frac{21}{22}$, k with probability $\frac{1}{22}$. With minimax regret the preference reversal is even more extreme, as it leads to h with probability 1.

[12] The menu-dependence of the preferences is sometimes referred to a failure of postulate known as "independence of irrelevant alternatives" (IIA; see for example Stoye, 2006). However IIA is a condition that emerges in the context of social choice, not of individual choice, and strictly speaking refers to a somewhat different phenomenon. As Arrow put it, "the choice made by society should be independent of the very existence of alternatives outside the given set" (Arrow, 1984). Here the preferences might change when new alternatives are included in the given set. As long as they are not feasible, they do not affect preferences.

deed, the situations where a concern for robustness might arise, and therefore a decision criterion different from expected utility might become relevant, are likely to be those in which a single, consistent preference ordering over all possible pairs of actions is supposed to be too difficult to achieve (it is impractical, to use Savage's term). If such a single preference ordering were to exist, and if it satisfied Savage's sure thing principle, a subjective probability would essentially follow, and the expected utility criterion would be justified. If, conversely the single ordering were not to satisfy the sure thing principle, then it would be a poor basis for the decision (according to Savage (1954), in this case the ordering would be *"absurd as an expression of preference"*). In fact, standard minimax provides a single ordering (it is not subject to preference reversal), but one that violates the sure thing principle.

A2. given $p, q \in Q(A, j)$ such that $(b^{-j}, p) \succ (b^{-j}, q)$, then $(b^{-j}, \alpha p + (1 - \alpha)r) \succ (b^{-j}, \alpha q + (1 - \alpha)r)$, $\forall r \in Q(A, j), \alpha \in (0, 1], j \in S$.

This is a weaker version of the standard "independence" axiom (see for example Fishburn, 1970). In particular, the preference which is postulated to survive when a mixture with a common lottery is introduced is between a specific kind of actions and concerns a specific kind of mixture: actions whose components are, in all but one state, equal to the (state-dependent) best lottery and mixtures that only involve the state in which the component is not equal to the best lottery. Both restrictions to the standard independence property are essential, since a more general formulation would allow the sort of mixture that might make one of the components in the original pair of actions so "bad" that the subject becomes indifferent among the modified actions. This can be easily shown by way of examples (as before, the left panel presents the consequences of the various actions measured in utiles, the right one normalised by the column maximum).

Table 6

	state 1	state 2
h	9	8
k	10	6
g	20	0
l	0	40

$\rightarrow$

	state 1	state 2
h	0.45	0.20
k	0.5	0.15
g	1	0
l	0	1

Here, according to relative minimax, $h \succ k$, but neither h nor k have the form required by $A2$.Consider now $r = 30$ (included in the convex hull of the set of consequences achievable in state 2, as required by $A2$), and take $\alpha = 0.1$. Define $h' = (9, 8\alpha + (1-\alpha)30) = (9, 27.8); k' = (10, 6\alpha + (1-\alpha)30) = (10, 27.6)$. Therefore:

Table 6$'$

	state 1	state 2
h'	9	27.8
k'	10	27.6
g	20	0
l	0	40

$\rightarrow$

	state 1	state 2
h'	0.45	0.695
k'	0.5	0.69
g	1	0
l	0	1

Now, $k' \succ h'$. A similar failure to preserve the preference under mixtures could be shown starting from $h = (20, 8)$ and $k = (20, 6)$ (i.e. with a pair of actions that have the form required by $A2$) and considering a mixture in state 1, rather than in state 2, with $r = 0$, $\alpha = 0.1$, leading to $h' = (2, 8), k' = (2, 6)$, which are indifferent to each other under relative minimax, since they correspond, in term of normalised utilities, to $(0.1, 0.2)$ and $(0.1, 0.15)$.[13]

[13] It is worth noting that the independence axiom proposed in Stoye (2006) for minimax regret would not be satisfied by relative minimax. Stoye version of indipendence is as follows. Let $h, k \in M$, where M is a given menu of actions, and let c be a new action, not necessarily in M. Define a new menu, $\alpha M + (1-\alpha)c$, by replacing *each* element h in M with $\alpha h + (1-\alpha)c$. Let $\succsim_M$be the preference among the elements in M. The axiom then reads as follows:

$h \succsim_M k \Longleftrightarrow \alpha h + (1-\alpha)c \succsim_{\alpha M + (1-\alpha)c} \alpha k + (1-\alpha)c, \forall \alpha \in (0,1).$

It is easy to construct examples showing that relative minimax does not satisfy this axiom.

$A3$. given $h, k, l \in H_A$ such that $h \succ k \succ l$, then $\exists \alpha, \beta \in (0,1)$ such that $\alpha h + (1-\alpha) l \succ k \succ \beta h + (1-\beta) l$.

This is the standard "Archimedean" axiom (see again Fishburn, 1970).

The following two axioms are little more than definitions.

$A4$. $\forall p \in Q(A,j), q_i \in Q(A,i), i \neq j : (q^{-j}, b_j) \succsim (q^{-j}, p) \succsim (q^{-j}, w), \forall j, i \in S$.

This axiom specifies in which sense each of b_j is the best available consequence in state j, $j = 1, 2...m$, and w is the worst in each state.

$A5$. $(b^{-j}, b_j) \succ (b^{-j}, w), \forall j \in S$.

This axiom specifies that not all actions are indifferent to each other (in this sense, it is a "non triviality" axiom).

The following three axioms are those that are mostly responsible for the relative minimax representation.

The first is a weak form of the sure thing principle (Savage, 1954):

$A6$. given $p, q \in Q(A,j)$ such that $(b^{-j}, p) \succsim (b^{-j}, q)$, then $\forall s_i \in Q(A,i), i \neq j, (s^{-j}, p) \succsim (s^{-j}, q), j \in S$.

Note that, differently from Savage's stronger formulation, here the preference in the first part of the axiom is between actions whose structure, as in $A2$, is tightly restricted. Again, this restriction is essential, since it easy to show examples in which, under relative minimax, $(s^{-j}, p) \succsim (s^{-j}, q)$ but $(b^{-j}, q) \succ (b^{-j}, p)$. Note also that, again differently from Savage's formulation, even if the preference between the original pair of actions is strict (i.e. if $(b^{-j}, p) \succ (b^{-j}, q)$), the axiom only guarantees a weak preference between the modified actions. The reason for this is that the common components s^{-j} could be "bad" enough so as to make the modified actions indifferent to one another.

The second axiom imparts a form of symmetry to the preferences:

$A7$. Let $h \equiv (\alpha_1 b_1 + (1-\alpha_1) w, \alpha_2 b_2 + (1-\alpha_2) w ..., \alpha_m b_m + (1-\alpha_m) w)$, $\alpha_i \in [0,1]$. Let E and F be any two non empty, disjoint subsets of the set of states S, such that α_i is constant over E and F, with $\alpha_i = \alpha_E$ for $i \in E$, and $\alpha_j = \alpha_F$ for $j \in F$. Define $h' \equiv (\alpha'_1 b_1 + (1-\alpha'_1) w, \alpha'_2 b_2 + (1-\alpha'_2) w ..., \alpha'_m b_m +$

$(1 - \alpha'_m)w)$, where $\alpha'_i = \alpha_F$ for $i \in E$, $\alpha'_j = \alpha_E$ for $j \in F$ and $\alpha'_s = \alpha_s$ otherwise. Then, $h \sim h'$.

This axiom considers a special class of standardized actions, that yield in each state a simple lottery between the best (state-dependent) and the worst consequences (it will be shown later that under $A1 - A6$ all actions can be put in this form, so that the loss of generality is only apparent). Given a set of m lotteries, the axiom essentially states that it does not matter how the lotteries are coupled with the states: all standardized actions yielding these m lotteries are equivalent. This is consistent with the idea that the preference ordering is independent of the labelling of the states (this requirement is imposed by Milnor, 1954, and Stoye, 2006, for preference orderings that can be represented by minimax or minimax regret). However, $A7$ imposes something more than pure symmetry. The latter amounts, as in Milnor (1954), to the possibility of "switching columns", and requires consistency between the preference ordering relative to a given menu of actions and the preference ordering relative to the new menu obtained from the former by swapping, for all the actions, the consequences in some state with the consequences in a different state. $A7$ also requires a form of relativity of preferences: the preference is relative to what is (best) achievable, since what matters is how "close" one is to the best, not what the best actually is. This relativity amounts to require consistency between the preference ordering relative to the original set of standardized actions and the preference ordering relative to the set of standardized actions obtained from the former by swapping the best consequences.[14] It is possible to show that $A7$ is basically equivalent to symmetry and relativity, so defined. The advantage of $A7$, however, is that

[14] Taking for simplicity of notation the case of two states, let all standardized actions $(\alpha b_1 + (1-\alpha)w, \beta b_2 + (1-\beta)w)$ be denoted by $(\alpha b_1, \beta b_2)$. Let $\succsim_A$ indicates the preference ordering relative to the original set of actions. For each standardized action $(\alpha b_1, \beta b_2)$ define a new standardized action $(\alpha b_2, \beta b_1) = (\alpha b_2 + (1-\alpha)w, \beta b_1 + (1-\beta)w)$. Let A' be the set of actions so obtained, and $\succsim_{A'}$ the preference relation on this set. The consistency requirement mentioned in the text can now be stated as follows: $(\alpha b_1, \beta b_2) \succsim_A (\gamma b_1, \delta b_2) \Leftrightarrow (\alpha b_2, \beta b_1) \succsim_{A'} (\gamma b_2, \delta b_1)$

it does not involve different menus of actions. A stronger form of symmetry, which would be satisfied by standard minimax, would impose that $h \sim h'$ whenever $h = (p_1, p_2...p_m)$ and $h' = (p_{j1}, p_{j2}...p_{jm})$, where $(p_{j1}, p_{j2}...p_{jm})$ is any permutation of $(p_1, p_2...p_m)$. This would in general not be satisfied by relative minimax, nor by minimax regret: even when $p_{ji} \in Q(A, i)$ for all i, so that $h' \in H_A$, the relationship between p_i and b_i – which is what matters for both decision criteria – would in general be different than the relationship between p_{ji} and b_i (the "switching columns" kind of symmetry requires that also b_i be swapped with b_{ji}). Finally, note that minimax regret does not in general satisfy $A7$, since it does not satisfy the relativity requirement.

The third axiom is the standard ambiguity aversion (as in Gilboa and Schmeidler, 1989):

$A8$. Consider $h, k \in H_A$, such that $h \sim k$. Then, $\forall \alpha \in [0, 1]$, $\alpha h + (1 - \alpha)k \succsim h$

This axiom is the key ingredient for the "minimax part" of relative minimax, and captures the benefits of hedging: the worst outcome of the convex combination of the two actions might be better, and cannot be worse, than the worst outcome of the original actions. The full bite of this axiom requires the states to be at least 3.

The last axiom imposes a form of dominance.

$A9$. Let $p_j, q_j \in Q(A, j)$ such that $(b^{-j}, p_j) \succsim (b^{-j}, q_j)$, $\forall j \in S$. Then, $(p_1, p_2...p_m) \succsim (q_1, q_2...q_m)$. If $\succsim$ in the first part of the axiom is replaced by $\succ$, for all j, then $\succsim$ is replaced by $\succ$ in the second part.

This set of axioms is similar to the set of axioms that is presented in Kreps (1988) to justify standard minimax.[15] The main difference is the symmetry axiom, which in Kreps is stated in the strong form mentioned above (in the 2 states case considered by Kreps, this reads $(p, q) \sim (q, p)$ for all p and q, Kreps' axiom (b)), and in particular the relativity part of axiom $A7$, which

[15] In fact, Kreps (1988) only states the axioms, for the special case of 2 states and common outcome space, leaving it as an exercise to show that they justify Wald's minimax.

in Kreps formulation is missing.[16] A second notable difference is that instead of ambiguity aversion Kreps directly takes as an axiom the main implication of ambiguity aversion (i.e. that if $(b_1, p) \succsim (b_1, q)$, then $(q, p) \sim (q, q)$, which is Kreps' axiom (h)); this also implies that dominance is not required.

5 Representation

To prove that the set of axioms imply the relative minimax as a decision criterion, it will first be proved that each lottery component in each action can be replaced by an appropriate standardized lottery defined on the (state-dependent) best and worst available consequences (Lemmas 1-2). Then it will be shown that only the most favourable and the least favourable standardized lotteries are needed to describe any action (Lemmas 3-4), and further that only the least favourable is actually needed (Lemma 5). Then it will be shown that the preference relationship can be represented by the minimax criterion applied to a vector of appropriately defined functions (Lemma 6). It will then be shown that these latter functions can be interpreted as "normalised" utilities (Lemma 8). Theorem 1 summarises these results and the reverse implication, that the relative minimax decision criterion implies the axioms. All proofs are gathered in the appendix.

The first step is a relatively standard result.

Lemma 1

Under $A1$, $A2$, $A3$, the following hold

(a) if $(b^{-j}, p) \succ (b^{-j}, q)$, and $0 \leq \alpha < \beta \leq 1$, then $(b^{-j}, \beta p + (1-\beta)q) \succ (b^{-j}, \alpha p + (1-\alpha)q)$, $j = 1, 2...m$;

(b) if $(b^{-j}, p) \succsim (b^{-j}, q) \succsim (b^{-j}, r)$, $(b^{-j}, p) \succ (b^{-j}, r)$ then $\exists \alpha_j^* \in [0, 1]$,unique, such that $(b^{-j}, \alpha_j^* p + (1-\alpha_j^*)r) \sim (b^{-j}, q)$, $j = 1, 2...m$;

[16] In fact, Kreps assumes that the outcome space is the same for all states (in the notation of the present paper, $Q(A, i) = Q(A, j)$ for all i and j), and also that the best consequences are the same ($b_i = b_j$ for all i and j). As a consequence, the strong notion of symmetry is equivalent to the weaker "switching columns" form.

(c) if $(b^{-j}, p) \sim (b^{-j}, q)$, and $\alpha \in [0,1]$, then $(b^{-j}, \alpha p + (1-\alpha)r) \sim (b^{-j}, \alpha q + (1-\alpha)r), \forall r \in Q(A,j)$, $j = 1, 2...m$.

From $A4$, setting $q_i = b_i, \forall i \neq j$, it can be concluded that,$\forall p \in Q(A,j)$, $(b^{-j}, b_j) \succsim (b^{-j}, p) \succsim (b^{-j}, w)$; moreover, $(b^{-j}, b_j) \succ (b^{-j}, w)$, for all $j \in S$, from $A5$. Hence it is possible to apply part (b) of the Lemma 1, which guarantees the existence and uniqueness of a value $\alpha_j^* \in [0,1]$ such that $(b^{-j}, \alpha_j^* b_j + (1-\alpha_j^*)w) \sim (b^{-j}, p)$.Therefore, the following definition can be introduced:

Definition 1

For each $(p_1, p_2, ...p_m) \in H_A$ let $f : H_A \rightarrow [0,1]^m$ be defined by $f(p_1, p_2, ...p_m) = (f_1(p_1), f_2(p_2), ...f_m(p_m))$, where $f_j(p_j) \in [0,1]$ is the unique value such that $(b^{-j}, f_j(p)b_j + (1 - f_j(p_j))w) \sim (b^{-j}, p_j)$, $j = 1, 2...m$. Clearly, $f_j(b_j) = 1$ and $f_j(w) = 0$, $j = 1, 2...m$.

It is now possible to prove the following:

Lemma 2

If $A1 - A6$ hold, there exists a unique function $f : H_A \rightarrow [0,1]^m, f = (f_1, f_2, ...f_m)$,with $f_i(w) = 0, f_i(b_i) = 1, i = 1, 2...m$, such that for any $(p_1, p_2...p_m) \in H_A$, $(p_1, p_2...p_m) \sim (f_1(p_1)b_1 + (1 - f_1(p_1))w, f_2(p_2)b_2 + (1 - f_2(p_2))w..., f_m(p_m)b_m + (1 - f_m(p_m))w)$. Moreover, for all $r, s \in Q(A,i)$, and for any $\alpha \in [0,1]$, $f_i(\alpha r + (1-\alpha)s) = \alpha f_i(r) + (1-\alpha) f_i(s), i = 1, 2..., m$ (i.e., all components of f are affine).

Because of Lemma 2, to each $h = (p_1, p_2...p_m) \in H_A$ can be associated a vector $\alpha^h = (\alpha_1^h, \alpha_2^h, ...\alpha_m^h) \in [0,1]^m$, where $\alpha_j^h = f_j(p_j)$. Whenever there is no ambiguity, h will be identified with the vector α^h.

It is useful to rekord a straightforward implication of $A9$, that will also be referred to as "dominance" and will be repeatedly used in the proofs of results presented in the following. Consider $h = (p_1, p_2...p_m), k = (q_1, q_2...q_m) \in H_A$. Let $(\alpha_1^h, \alpha_2^h, ...\alpha_m^h)$ and $(\alpha_1^k, \alpha_2^k, ...\alpha_m^k)$ be the vectors in $[0,1]^m$ associated to h and k, respectively, and suppose that $\alpha_j^h \geq \alpha_j^k$ for all $j = 1, 2...m$. It is immediate to show that in this case $h \succsim k$ (if $\alpha_j^h > \alpha_j^k$ for all j, then $h \succ k$). Indeed, by definition $(b^{-j}, p_j) \sim (b^{-j}, \alpha_j^h b_j + (1 - \alpha_j^h)w)$, and $(b^{-j}, q_j) \sim (b^{-j}, \alpha_j^k b_j + (1 - \alpha_j^k)w)$, $j = 1, 2...m$. Because of Lemma 1 and transitivity, $(b^{-j}, p_j) \succsim (b^{-j}, q_j), j = 1, 2...m$. Axiom $A9$ now implies that $(p_1, p_2...p_m) \succsim (q_1, q_2...q_m)$ (if $\alpha_j^h > \alpha_j^k$ for all j, then $(p_1, p_2...p_m) \succ (q_1, q_2...q_m)$).

Let E be a non empty subset of S and denote by $\alpha_E\beta$, $\alpha, \beta \in [0, 1]$, the action $h = (\alpha_1^h, \alpha_2^h, ...\alpha_m^h)$, with $\alpha_j^h = \alpha$ if $j \in E$, $\alpha_j^h = \beta$ if $j \notin E$. In words, $\alpha_E\beta$ is a special kind of action that, loosely speaking, takes only two consequences, α in a subset of the states and β otherwise. Let F be another non empty subset of S and consider the action $\alpha_F\beta$, defined similarly. Then it is possible to establish the following:

Lemma 3

If $A1 - A7$ and $A9$ hold, $\alpha_E\beta \sim \alpha_F\beta$. Moreover, $\alpha_E\beta \sim \beta_E\alpha$.

Consider now any $h = (\alpha_1^h, \alpha_2^h, ...\alpha_m^h)$. Let $\overline{\alpha}^h = \max_{j \in S} \alpha_j^h$, and $\underline{\alpha}^h = \min_{j \in S} \alpha_j^h$ (if there is no ambiguity, the superscript h to $\overline{\alpha}$ and $\underline{\alpha}$will be dropped). To avoid trivial cases assume $\overline{\alpha} > \underline{\alpha}$. It is then easy to show the following:

Lemma 4

If $A1 - A7$ and $A9$ hold, $h \sim \underline{\alpha}_E\overline{\alpha}$, for any non empty subset of the states E.

This result implies that there is no loss of generality in considering actions h that only take the two values $\overline{\alpha}^h$ and $\underline{\alpha}^h$. In particular, given a partition of S in three sets, (F, K, M), any action $h = (\alpha_1^h, \alpha_2^h, ...\alpha_m^h)$ is, because of Lemmas 3 and 4, equivalent to actions that take value $\alpha_j^h = \overline{\alpha}^h$ or $\alpha_j^h = \underline{\alpha}^h$ for $j \in F$ or K or M. Denote these actions as (α, β, γ), where each of α, β, γ is equal to either $\underline{\alpha}$ or $\overline{\alpha}$. Then, what has been established so far is that $h \sim (\underline{\alpha}, \overline{\alpha}, \overline{\alpha}) \sim (\underline{\alpha}, \overline{\alpha}, \underline{\alpha}) \sim (\underline{\alpha}, \underline{\alpha}, \overline{\alpha})$. It is possible to show the following important:

Lemma 5

If $A1 - A9$ hold, $h \sim (\underline{\alpha}, \underline{\alpha}, \underline{\alpha})$.

It is now easy to establish the following Lemma, that provides a key characterisation of the preferences:

Lemma 6

If $A1 - A9$ hold, then there exists a unique function $f : H_A \to [0,1]^m, f = (f_1, f_2, ...f_m)$,with $f_i(w) = 0, f_i(b_i) = 1, i = 1, 2...m$, such that for any $(p_1, p_2, ..., p_m), (q_1, q_2, ..., q_m) \in H_A$:

$(p_1, p_2, ..., p_m) \succsim (q_1, q_2, ..., q_m)$ iff $\min(f_1(p_1), f_2(p_2), ...f_m(p_m)) \geq \min(f_1(q_1), f_2(q_2), ...f_m(q_m))$.

Moreover, all components of the f function are affine.

[Clearly, by the definition of the function f, its first component f_1 assigns a value to the first component of any $m-$tuple in H_A, its second component f_2 assigns a value to the second component, and so on. Thus, even if $\cap_{j \in S} Q(A, j) \neq \emptyset$ (a condition that might or migh not be true, depending on the choice problem under consideration), there is no ambiguity as to the value that would correspond to a $p \in \cap_{j \in S} Q(A, j)$: if p occurs when state j is true (i.e. if p is the $j-$th component of the $m-$tuple) then the value associated to p is $f_j(p)$. Indeed, $f_j(p)$ can be interpreted as a numerical value associated to p when p is assessed in terms of the benchmark provided by the two consequences b_j and w. There is then no reason to require $f_j(p)$ to be the same value that would be associated to p if p were assessed in terms of a different benchmark, as the one provided by the consequences b_i and w, $i \neq j$. However, if the decision problem happens to be symmetrical (with $b_i = b_j, \forall i, j$), so that the benchmark in each state is the same, it might be considered as reasonable, or even desirable, that $f_i(p) = f_j(p), \forall i, j$, for any $p \in \cap_{j \in S} Q(A, j)$. This is a property which is not warranted by axioms $A1 - A9$, but which is also not needed for the proof of the representation theorem provided below, and in general will not be imposed. It will however be shown in the next lemma that the property holds if the following axiom is added to $A1 - A9$.

$A10$. if $b_i = b_j, \forall i, j$, $(b^{-i}p) \sim (b^{-j}, p)$ $\forall p \in \cap_{j \in S} Q(A, j)$, where b indicates the common value of $b_j, j = 1, 2...m$.

The axiom states that in a symmetrical situation, where the best that can be achieved in all the states are the same, and if the same lottery is conceivable in all states, then it does not matter whether that lottery is associated to one or to another of the states. It needs to be stressed that this axiom imposes a consistency requirement that only makes sense if, in the *given* set A, the conditions described by the axiom were to apply. The axiom does not require A to vary, and therefore it does not open up the possibility of a preference reversal.

Adding $A10$ to the other axioms it is possible to establish the following:

Lemma 7
Let $A1-A10$ hold. Suppose $b_i = b_j, \forall i, j$, and suppose that $\cap_{j\in S} Q(A, j) \neq \emptyset$. Then, $f_i(p) = f_j(p) \forall p \in \cap_{s\in S} Q(A, s), \forall i, j.$]

The unicity of the function f follows from fixing the reference consequences b_j and w, $j = 1, 2...m$. These are not arbitrary, since they appear in the formulation of most of the axioms. It is however possible to represent the preferences through a different function, which takes as reference two pairs of arbitrary lotteries, much in the same way as in standard derivations of the expected utility criterion.

To this end, take any two p^1 and p^0 in $Q(A, j)$ such that $(b^{-j}, p^1) \succ (b^{-j}, p^0)$, $j = 1, 2...m$. $A5$ guarantees the existence of at least two elements in $Q(A, j)$ for which this is true (b_j and w), but in general there will be more. Consider any other $p \in Q(A, j)$. There are three possible cases:

(i) $(b^{-j}, p^1) \succsim (b^{-j}, p) \succsim (b^{-j}, p^0)$

(ii) $(b^{-j}, p) \succsim (b^{-j}, p^1) \succ (b^{-j}, p^0)$

(iii) $(b^{-j}, p^1) \succ (b^{-j}, p^0) \succsim (b^{-j}, p)$.

Lemma 1 (b) established the existence of unique values α, β and γ such that:

in case (i): $(b^{-j}, \alpha p^1 + (1-\alpha)p^0) \sim (b^{-j}, p)$;

in case (ii): $(b^{-j}, \beta p + (1-\beta)p^0) \sim (b^{-j}, p^1)$;

in case (iii):$(b^{-j}, \gamma p^1 + (1-\gamma)p) \sim (b^{-j}, p^0)$.

Correspondingly, a function $g_j : Q(A,j) \rightarrow \mathbb{R}$ can be defined as:

in case (i): $g_j(p) = \alpha$; hence, $g_j(p)$ is the value such that $(b^{-j}, g_j(p)p^1 + (1 - g_j(p))p^0) \sim (b^{-j}, p)$;

in case (ii): $g_j(p) = \frac{1}{\beta}$; hence, $g_j(p)$ is the value such that $(b^{-j}, \frac{1}{g_j(p)}p + (1 - \frac{1}{g_j(p)})p^0) \sim (b^{-j}, p^1)$;

in case (iii):$g_j(p) = \frac{\gamma}{\gamma-1}$; hence, $g_j(p)$ is the value such that $(b^{-j}, \frac{g_j(p)}{g_j(p)-1}p^1 + (1 - \frac{g_j(p)}{g_j(p)-1})p) \sim (b^{-j}, p^0)$.

Consider now the function $g : H_A \rightarrow \mathbb{R}^m$ defined by $g(p_1, p_2, ...p_m) = (g_1(p_1), g_2(p_2), ...g_m(p_m))$. It is possible to show that an appropriate transformation of this function plays the same role as the function f in representing the preferences. This is established by the following:

Lemma 8

If $A1 - A9$ hold, for any $(p_1, p_2, ..., p_m) \in H_A$,

$f_j(p_j) = \frac{g_j(p_j) - \min_{s \in Q(A,j)} g_j(s)}{\max_{s \in Q(A,j)} g_j(s) - \min_{s \in Q(A,j)} g_j(s)}, j = 1, 2...m.$

It is now possible to state the representation theorem:

Theorem 1

The preference relationship $\succsim$ on the set H_A satisfies $A1 - A9$ if and only if there exists a function $g : H_A \rightarrow \mathbb{R}^m, g = (g_1, g_2...g_m)$, such that for any $(p_1, p_2, ..., p_m), (q_1, q_2, ..., q_m) \in H_A$ it holds true that

$$(p_1, p_2, ..., p_m) \succsim (q_1, q_2, ..., q_m) \ iff \tag{RMX}$$
$$\min_j \left(\frac{g_j(p_j) - \min\limits_{s \in Q(A,j)} g_j(s)}{\max\limits_{s \in Q(A,j)} g_j(s) - \min\limits_{s \in Q(A,j)} g_j(s)} \right)_{j=1,2...m} \geq$$
$$\min_j \left(\frac{g_j(q_j) - \min\limits_{s \in Q(A,j)} g_j(s)}{\max\limits_{s \in Q(A,j)} g_j(s) - \min\limits_{s \in Q(A,j)} g_j(s)} \right)_{j=1,2...m}.$$

Moreover, for all $r, s \in Q(A,i)$, and for any $\alpha \in [0,1]$, $g_i(\alpha r + (1-\alpha)s) = \alpha g_i(r) + (1-\alpha)g_i(s), i = 1, 2...m$ (i.e., all components of g are affine). Also, each component of the g function is unique up to a positive affine transformation.

Finally, it can be shown that the function g is linear in the (objective) probabilities used in generating the lotteries in $Q(A,i), i = 1, 2...m$.

Lemma 9

If $A1 - A9$ hold, there exists a function $u : \prod_{j \in S} C(A,j) \rightarrow \mathbb{R}^m, u = (u_1, u_2...u_m)$ such that for any $(p_1, p_2...p_m) \in H_A$, it is true that, for all j, $g_j(p_j) = \sum_c u_j(c)p_j(c)$ where the summation is extended to all elements $c \in C(A,j)$ in the support of p_j, $p_j(c)$ are the probability that p_j assigns to a given $c \in C(A,j)$.

Within the class of functions that represent the same preference relationship $\succsim$, it is always possible to choose one whose components take all value 0 at w (as this just requires adding an appropriate constant to each component, and therefore only involves an affine transformation). This choice is not restrictive, since the same preferences obtain.

For this particular choice, the characterisation (RMX) takes the following, simplified form:

$$\begin{aligned} (p_1, p_2, ..., p_m) \; &\succsim \; (q_1, q_2, ..., q_m) \; iff \\ \min_j \left(\frac{g_j(p_j)}{\max\limits_{s \in Q(A,j)} g_j(s)} \right)_{j=1,2...m} \; &\geq \; \min_j \left(\frac{g_j(q_j)}{\max\limits_{s \in Q(A,j)} g_j(s)} \right)_{j=1,2...m}. \end{aligned} \tag{RMXS}$$

The characterisation (RMXS) corresponds precisely to the relative minimax (or competitive ratio) decision criterion. Therefore, axioms $A1 - A9$ provide the behavioural foundation for that decision criterion.

Moreover, as earlier mentioned, a simple by-product of the representation theorem is the possibility to show that the same set of axioms also supports minimax regret, via a logarithmic transformation of the function g.

To provide the details choose, in the class of functions $g : H_A \rightarrow \mathbb{R}^2, g = (g_1, g_2)$ that, according to Theorem 1 represent the preferences, any function whose components both take value greater or equal to 0 at w (to simplify the notation, in the following the choice that underlies the characterisation (RMXS) will be made, but the same conclusion would be reached for any other transformation of the g function that guarantees its non negativity). For such a function g, define a function $l = (l_1, l_2)$ whose $i - th$ component $(i = 1, 2)$ is equal to $\log(g_i)$ for the set $\{p : p \in Q(A, i), g_i(p) > 0\}$, and is equal to $-\infty$ for the set $\{p : p \in Q(A, i), g_i(p) = 0\}$. Clearly, if g_i is one-to-one so is l_i; if g_i is not one-to-one, the equivalence classes of g_i (i.e. the sets $G_\alpha = \{p : p \in Q(A, i), g_i(p) = \alpha\}$) coincide with the equivalence classes of l_i (i.e. the sets $L_\alpha = \{p : p \in Q(A, i), l_i(p) = \log(\alpha) \ if \ \alpha > 0, l_i(p) = -\infty \ if \ \alpha = 0\}$). Consider now any pair of actions $(p', q'), (p, q) \in H_A$, and the corresponding inequality:

$$\min\Big(\frac{g_1(p')}{\max\limits_{s\in Q(A,1)} g_1(s)}, \frac{g_2(q')}{\max\limits_{s\in Q(A,2)} g_2(s)}\Big) \gtreqless \min\Big(\frac{g_1(p)}{\max\limits_{s\in Q(A,1)} g_1(s)}, \frac{g_2(q)}{\max\limits_{s\in Q(A,2)} g_2(s)}\Big).$$

First of all note that, since each component of the l is a monotone increasing transformation of the corresponding component of the g, the same argument maximises both functions (provided a unique maximiser for g exists; if not, select arbitrarily one element in the two corresponding equivalence classes). Therefore, $\max\limits_{s\in Q(A,1)} l_1(s) = l_1(b_1)$ and $\max\limits_{s\in Q(A,2)} l_2(s) = l_2(b_2)$. Recalling that $\min(a, b) = -\max(-a, -b)$, it is then immediate to establish the following chain of equivalences:

$$\begin{aligned}
\min\Big(\frac{g_1(p')}{g_1(b_1)}, \frac{g_2(q')}{g_2(b_2)}\Big) &\gtreqless \min\Big(\frac{g_1(p)}{g_1(b_1)}, \frac{g_2(q)}{g_2(b_2)}\Big) \Longleftrightarrow \\
-\max\Big(-\frac{g_1(p')}{g_1(b_1)}, -\frac{g_2(q')}{g_2(b_2)}\Big) &\gtreqless -\max\Big(-\frac{g_1(p)}{g_1(b_1)}, -\frac{g_2(q)}{g_2(b_2)}\Big) \Longleftrightarrow \\
\max\Big(-\frac{g_1(p')}{g_1(b_1)}, -\frac{g_2(q')}{g_2(b_2)}\Big) &\lesseqgtr \max\Big(-\frac{g_1(p)}{g_1(b_1)}, -\frac{g_2(q)}{g_2(b_2)}\Big) \Longleftrightarrow \\
\max\Big(-\log\Big(\frac{g_1(p')}{g_1(b_1)}\Big), -\log\Big(\frac{g_2(q')}{g_2(b_2)}\Big)\Big) &\lesseqgtr \max\Big(-\log\Big(\frac{g_1(p)}{g_1(b_1)}\Big), -\log\Big(\frac{g_2(q)}{g_2(b_2)}\Big)\Big) \Longleftrightarrow \\
\max(l_1(b_1) - l_1(p'), l_2(b_2) - l_2(q')) &\lesseqgtr \max(l_1(b_1) - l_1(p), l_2(b_2) - l_2(q)) \Longleftrightarrow \\
\max\Big(\max_{s\in Q(A,1)} l_1(s) - l_1(p'), \max_{s\in Q(A,2)} l_2(s) - l_2(q')\Big) &\lesseqgtr \max\Big(\max_{s\in Q(A,1)} l_1(s) - l_1(p), \max_{s\in Q(A,2)} l_2(s) - l_2(q)\Big).
\end{aligned}$$

Therefore:

$$(p', q') \succsim (p, q) \ iff$$
$$\min(\frac{g_1(p')}{\max\limits_{s \in Q(A,1)} g_1(s)}, \frac{g_2(q')}{\max\limits_{s \in Q(A,2)} g_2(s)}) \geq \min(\frac{g_1(p)}{\max\limits_{s \in Q(A,1)} g_1(s)}, \frac{g_2(q)}{\max\limits_{s \in Q(A,2)} g_2(s)}) \ iff$$
$$\max(\max_{s \in Q(A,1)} l_1(s) - l_1(p'), \max_{s \in Q(A,2)} l_2(s) - l_2(q')) \leq \max(\max_{s \in Q(A,1)} l_1(s) - l_1(p), \max_{s \in Q(A,2)} l_2(s) - l_2(q)).$$

If the utility is measured by the logarithm of the g function, the LHS member in the last expression can be interpreted as the maximal regret associated to (p', q') (see Section 2), and the RHS as the maximal regret associated to (p, q). Hence, $A1 - A8$ can be seen to provide the axiomatic foundation of minimax regret, as well as that of relative minimax. It should however be noted that these axioms, and in particular $A7$, which is mostly responsible for the "relative normalisation" of the utility, are formulated so as to capture a relative preference: loosely speaking, the indifference between p and an $x\%$ chance of b naturally corresponds to the statement that the preference for p is only $x\%$ as strong as the preference for b or equivalently that, relative to the preference for b, the preference for p is only $x\%$. Therefore, the decision criterion which seems to be naturally implied by the axioms is relative minimax. To get minimax regret, the scale of the utility need being distorted, as it were, in order to have ratios represented as differences between levels, so that the relative comparisons explicitly required by relative minimax are implicitly hidden by the logarithmic transformation. This is the reason why the emphasis of the paper is on relative minimax. The following representation theorem remains nevertheless true (the proof is obvious, given the above established equivalence, and is not provided):

Theorem 2

The preference relationship $\succsim$ on the set H_A satisfies $A1 - A8$ if and only if there exists a function $l : H_A \to \widetilde{\mathbb{R}} \times \widetilde{\mathbb{R}}$ (where $\widetilde{\mathbb{R}}$ is the set of real numbers extended to include $-\infty$) such that

$$(p', q') \succsim (p, q) \; iff$$

$$\max(\max_{s \in Q(A,1)} l_1(s) - l_1(p'), \max_{s \in Q(A,2)} l_2(s) - l_2(q')) \leq \max(\max_{s \in Q(A,1)} l_1(s) - l_1(p), \max_{s \in Q(A,2)} l_2(s) - l_2(q)).$$

is true for all couples $(p', q'), (p, q) \in H_A$.

6 Conclusions

As Savage (1954) remarked long ago, "*the minimax rule founded on the negative of income* (Savage's term for the negative of utility) *seems altogether untenable*". More appealing decision criteria – still free from the need to specify prior probabilities over the states of the world – are obtained when the utility is "normalised", either dividing it by, or subtracting it from, the largest state dependent utility. These normalisations yield, respectively, relative minimax and minimax regret. Stoye (2006) has only recently provided an axiomatic foundation with an explicit behavioural content for the second decision criterion. This paper does the same for the first. The lack of behavioural axiomatic foundations probably explains why the two criteria have been almost neglected in the recent literature, in spite of a renewed and increasing interest on robustness. Interesting exceptions are Bergemann and Schlag (2005), who use regret to define robust monopoly pricing, Cozzi and Giordani (2005), who explore the implication of minimax regret (as well as standard minimax) for the investment in R&D in a Shumpeterian growth model, and Altissimo et al. (2005), who use relative minimax to assess robust monetary policy; an earlier application is provided by Linhart and Radner (1989), who investigate the use of minimax regret in mechanism design. With an explicit behavioural foundation for relative minimax and for minimax regret these decision criteria are now on firmer ground. It is hoped that more applications will follow.

7 Appendix

Proof of Lemma 1

The proof follows well known arguments (see for example Kreps, 1988), and is provided in detail only for completeness.

(a) Fix $j \in S$. Take first $\alpha = 0$. Then, using $A2$ with $r = q \in Q(A, j)$, for any $\beta \in (0, 1]$ we have that $(b^{-j}, \beta p + (1 - \beta)q) \succ (b^{-j}, \beta q + (1 - \beta)q) = (b^{-j}, q) = (b^{-j}, \alpha p + (1 - \alpha)q)$, where the last equality is true since $\alpha = 0$. Let now $\alpha > 0$ (which in turn implies that $0 < \frac{\alpha}{\beta} < 1$), and define for ease of notation $r = \beta p + (1 - \beta)q$. We have just established that $(b^{-j}, r) \succ (b^{-j}, q)$. Hence we can again apply $A2$ to establish that $(b^{-j}, (1 - \frac{\alpha}{\beta})r + \frac{\alpha}{\beta}r) \succ (b^{-j}, (1 - \frac{\alpha}{\beta})q + \frac{\alpha}{\beta}r)$ Now, $(1 - \frac{\alpha}{\beta})q + \frac{\alpha}{\beta}r = (1 - \frac{\alpha}{\beta})q + \frac{\alpha}{\beta}(\beta p + (1 - \beta)q) = \alpha p + (1 - \frac{\alpha}{\beta} + \frac{\alpha}{\beta} - \frac{\alpha}{\beta}\beta)q = \alpha p + (1 - \alpha)q$. Hence, $(b^{-j}, (1 - \frac{\alpha}{\beta})r + \frac{\alpha}{\beta}r) = (b^{-j}, r) = (b^{-j}, \beta p + (1 - \beta)q) \succ (b^{-j}, \alpha p + (1 - \alpha)q)$, which is what we wanted to show. The same argument can be repeated for all other $j \in S$.

(b) Fix $j \in S$. First note that if α_j^* exists it is unique, since part (a) of the lemma implies that any two couples $(b^{-j}, \alpha^* p + (1 - \alpha^*)r)$ and $(b^{-j}, \alpha^{**} p + (1 - \alpha^{**})r)$ could not be both indifferent to (b^{-j}, q). That such an α_j^* exists for the cases $(b^{-j}, p) \sim (b^{-j}, q)$ and $(b^{-j}, q) \sim (b^{-j}, r)$ is obvious: it would be $\alpha_j^* = 1$ in the first case and $\alpha_j^* = 0$ in the second. Consider then the case $(b^{-j}, p) \succ (b^{-j}, q) \succ (b^{-j}, r)$. Define the set $D = \{\alpha \in [0, 1] : (b^{-j}, q) \succsim (b^{-j}, \alpha p + (1 - \alpha)r)\}$, which is not empty since it includes $\alpha = 0$, and define $\alpha_j^* = \sup D$. If $\alpha > \alpha_j^*$, by the definition of α_j^* it will be the case that $(b^{-j}, \alpha p + (1 - \alpha)r) \succ (b^{-j}, q)$. If $\alpha < \alpha_j^*$, then $(b^{-j}, q) \succ (b^{-j}, \alpha p + (1 - \alpha)r)$. To prove this, note that there exists in this case an $\alpha < \alpha' < \alpha_j^*$, with $(b^{-j}, q) \succsim (b^{-j}, \alpha' p + (1 - \alpha')r) \succ (b^{-j}, \alpha p + (1 - \alpha)r)$, where the first relationship follows from the fact that α' is in D and the second from part (a) of the lemma. Now, there are three possible cases: (i) $\alpha_j^* \notin D$; (ii) $\alpha_j^* \in D$, but the preference is strict; (iii) $\alpha_j^* \in D$, and there is the desired indifference. Consider first (i). In formal terms this case amounts to $(b^{-j}, \alpha_j^* p + (1 - \alpha_j^*)r) \succ (b^{-j}, q) \succ (b^{-j}, r)$. By $A3$ there is $\alpha \in (0, 1)$ such that $(b^{-j}, \alpha(\alpha_j^* p + (1 -$

$\alpha_j^*)r) + (1-\alpha)r) \succ (b^{-j}, q)$. But $(b^{-j}, \alpha(\alpha_j^* p + (1-\alpha_j^*)r) + (1-\alpha)r) = (b^{-j}, \alpha\alpha_j^* p + (1-\alpha\alpha_j^*)r)$, and since $\alpha\alpha_j^* < \alpha_j^*$ it must be, as shown above, that $(b^{-j}, q) \succ (b^{-j}, \alpha\alpha_j^* p + (1-\alpha\alpha_j^*)r)$. Therefore we have a contradiction. Consider now case (ii). Then $(b^{-j}, p) \succ (b^{-j}, q) \succ (b^{-j}, \alpha_j^* p + (1-\alpha_j^*)r)$. By $A3$ there is $\beta \in (0,1)$ such that $(b^{-j}, q) \succ (b^{-j}, \beta p + (1-\beta)(\alpha_j^* p + (1-\alpha_j^*)r))$. But $(b^{-j}, \beta p + (1-\beta)(\alpha_j^* p + (1-\alpha_j^*)r)) = (b^{-j}, (\alpha_j^* + \beta(1-\alpha_j^*))p + (1 - \alpha_j^* - \beta(1-\alpha_j^*))r)$, and since $\alpha_j^* + \beta(1-\alpha_j^*) > \alpha_j^*$ it must be, as shown above, that $(b^{-j}, \beta p + (1-\beta)(\alpha_j^* p + (1-\alpha_j^*)r)) \succ (b^{-j}, q)$. Therefore we have a contradiction. Hence we are left with the only remaining possibility, which is the one we wanted to prove to be true. The same argument can be repeated for all other $j \in S$.

(c) Fix $j \in S$. the result is trivial if $(b^{-j}, p) \sim (b^{-j}, q)$ for all p and $q \in Q(A, j)$. Hence assume that there exists at least one couple $(b^{-j}, s) \succ (b^{-j}, p) \sim (b^{-j}, q)$. Suppose now that (c) is not true, and take $(b^{-j}, \alpha p + (1-\alpha)r) \succ (b^{-j}, \alpha q + (1-\alpha)r)$ (the other possibility can be handled in a similar way). From $A2$ we have $(b^{-j}, \beta s + (1-\beta)q) \succ (b^{-j}, \beta q + (1 - \beta)q) = (b^{-j}, q) \sim (b^{-j}, p)$, for any $\beta \in (0,1]$. We can again apply $A2$ to get $(b^{-j}, \alpha(\beta s + (1-\beta)q) + (1-\alpha)r) \succ (b^{-j}, \alpha p + (1-\alpha)r)$. Since we assumed that $(b^{-j}, \alpha p + (1-\alpha)r) \succ (b^{-j}, \alpha q + (1-\alpha)r)$, $A3$ ensures that there exists, for each β, a value $\alpha(\beta) \in (0,1)$ such that $(b^{-j}, \alpha p + (1 - \alpha)r) \succ (b^{-j}, \alpha(\beta)(\alpha(\beta s + (1-\beta)q) + (1-\alpha)r) + (1-\alpha(\beta))(\alpha q + (1-\alpha)r))$. However, $\alpha(\beta)(\alpha(\beta s + (1-\beta)q) + (1-\alpha)r) + (1-\alpha(\beta))(\alpha q + (1-\alpha)r) = \alpha(\beta)\alpha\beta s + (\alpha - \alpha(\beta)\alpha\beta)q + (1-\alpha)r = \alpha(\alpha(\beta)\beta s + (1-\alpha(\beta)\beta)q) + (1-\alpha)r$. Hence, $(b^{-j}, \alpha p + (1-\alpha)r) \succ (b^{-j}, \alpha(\alpha(\beta)\beta s + (1-\alpha(\beta)\beta)q) + (1-\alpha)r)$. However, $(b^{-j}, \alpha(\beta)\beta s + (1-\alpha(\beta)\beta)q) \succ (b^{-j}, p)$, as shown before (since $\alpha(\beta)\beta \in (0,1)$), and therefore we can apply $A2$ to get $(b^{-j}, \alpha(\alpha(\beta)\beta s + (1 - \alpha(\beta)\beta)q) + (1-\alpha)r) \succ (b^{-j}, \alpha p + (1-\alpha)r)$, which is a contradiction. The same argument can be repeated for all other $j \in S$. ∎

Proof of Lemma 2

Take any $(p_1, p_2 ..., p_m) \in H_A$. An alternative notation for this vector is

(p^{-j}, p_j), $j = 1, 2...m$, or more generally (p^{-E}, p_E), where E denotes any non empty subset of S. Consider the first component of the function in definition 1. By definition $(b^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w) \sim (b^{-1}, p_1)$. This implies both that $(b^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w) \succsim (b^{-1}, p_1)$ and $(b^{-1}, p_1) \succsim (b^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w)$. Using $A6$ we then have that $(p^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w) \succsim (p^{-1}, p_1)$ and $(p^{-1}, p_1) \succsim (p^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w)$. Therefore, $(p^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w) \sim (p^{-1}, p_1)$. In a similar way, starting from $(b^{-2}, p_2) \sim (b^{-2}, f_2(p_2)b_2 + (1 - f_2(p_2))w)$ we can prove, using again $A6$, that $(p^{-(1,2)}, f_1(p)b_1 + (1 - f_1(p))w, f_2(q)b_2 + (1 - f_2(q))w) \sim (, p^{-(1,2)}, f_1(p_1)b_1 + (1 - f_1(p_1))w, p_2)$ (note that the latter is identical to $(p^{-1}, f_1(p_1)b_1 + (1 - f_1(p_1))w))$. Using the transitivity of $\sim$, we then have $(p^{-1}, p_1) \sim (p^{-(1,2)}, f_1(p)b_1 + (1 - f_1(p))w, f_2(q)b_2 + (1 - f_2(q))w)$. The same procedure can be repeated to conclude that $(p_1, p_2...p_m) \sim (f_1(p_1)b_1 + (1 - f_1(p_1))w, f_2(p_2)b_2 + (1 - f_2(p_2))w..., f_m(p_m)b_m + (1 - f_m(p_m))w)$. It remain to show that all components of f are affine (the claims that $f_i(w) = 0, f_i(b_i) = 1, i = 1, 2...m$ and that they are unique are obvious from their definition). Take $p \in Q(A, j)$. By definition of $f_j(p_j)$, $(f_j(p_j)b_j + (1 - f_j(p_j))w, b^{-j}) \sim (p_j, b^{-j})$. By part (c) of Lemma 1 we have that $(\alpha p_j + (1 - \alpha)q_j, b^{-j}) \sim (\alpha(f_j(p_j)b_j + (1 - f_j(p_j))w) + (1 - \alpha)q_j, b^{-j})$ for any $q_j \in Q(A, j)$ and for any $\alpha \in [0, 1]$. Moreover, again by definition, $(f_j(q_j)b_j + (1 - f_j(q_j))w, b^{-j}) \sim (q_j, , b^{-j})$, and again by part (c) of Lemma 1 we have that

$$(\alpha(f_j(p_j)b_j + (1 - f_j(p_j))w) + (1 - \alpha)q_j, b^{-j}) \sim$$
$$(\alpha(f_j(p_j)b_j + (1 - f_j(p_j))w) + (1 - \alpha)(f_j(q_j)b_j + (1 - f_j(q_j))w, b^{-j}) =$$
$$((\alpha f_j(p_j) + (1 - \alpha)f_j(q_j))b_j + (1 - (\alpha f_j(p_j) + (1 - \alpha)f_j(q_j)))w, b^{-j})$$

Transitivity then implies that $(\alpha p_j + (1 - \alpha)q_j, b^{-j}) \sim ((\alpha f_j(p_j) + (1 - \alpha)f_j(q_j))b_j + (1 - (\alpha f_j(p_j) + (1 - \alpha)f_j(q_j)))w, b^{-j})$ or, equivalently, that $f_j(\alpha p_j + (1 - \alpha)q_j) = \alpha f_j(p_j) + (1 - \alpha)f_j(q_j)$, for all $p_j, q_j \in Q(A, j)$ and for any $\alpha \in [0, 1]$. ■

Proof of Lemma 3

To prove this, consider first the case $E \cap F \neq \emptyset$, $E \cap F^c \neq \emptyset$. Let $L = E \cap F$, $G = E \cap F^c$, $K = E^c \cap F$, $M = E^c \cap F^c$. Then $\alpha_E \beta = (\alpha_1, \alpha_2, ...\alpha_m)$, with $\alpha_j = \alpha$ if $j \in L$ or $j \in G$, $\alpha_j = \beta$ if $j \in K$ or $j \in M$. Using $A7$ (symmetry), with the role of E here taken by G and the role of F here taken by K (i.e. exchanging the values taken in G with those taken in K), we conclude that $\alpha_E \beta$ is indifferent to the action $(\alpha'_1, \alpha'_2, ...\alpha'_m)$, with $\alpha'_j = \alpha$ if $j \in L$ or K, $\alpha'_j = \beta$ if $j \in G$ or M. Clearly, $(\alpha'_1, \alpha'_2, ...\alpha'_m) = \alpha_F \beta$. Consider now the case $E \subset F$. Now, $\alpha_E \beta$ can be written as $(\alpha_1, \alpha_2, ...\alpha_m)$, with $\alpha_j = \alpha$ if $j \in E$, $\alpha_j = \beta$ if $j \in F \cap E^c$ or $j \in F^c$. Again, $A7$ can be invoked (twice), to establish first that $\alpha_E \beta$ is indifferent to the action $(\alpha'_1, \alpha'_2, ...\alpha'_m)$, with $\alpha'_j = \beta$ if $j \in E$ or $F \cap E^c$, $\alpha'_j = \alpha$ if $j \in F^c$ (i.e. exchanging the values taken in the two subsets E and F^c), and then that the latter, which amounts to $\beta_F \alpha$, is indifferent to the action $(\alpha''_1, \alpha''_2, ...\alpha''_m)$, with $\alpha''_j = \alpha$ if $j \in F$, $\alpha''_j = \beta$ if $j \in F^c$ (i.e.exchanging the values taken in the two subsets F and F^c), which clearly is $\alpha_F \beta$; transitivity then establishes the desired result. Finally, if $E \cap F = \emptyset$, $A7$ can be directly invoked with reference to the two sets E and F, to conclude that $\alpha_E \beta \sim \alpha_F \beta$. It is also immediate to establish that $\alpha_E \beta \sim \beta_E \alpha$, for any non empty set E in S, by using $A7$ with reference to the two sets E and E^c. ■

Proof of Lemma 4

To prove this, let $\underline{E}$ and $\overline{E}$ be the sets of states where $\underline{\alpha}$ and $\overline{\alpha}$ are, respectively, achieved. By dominance, we know that $\underline{\alpha}_{\underline{E}} \overline{\alpha} \succsim h \succsim \overline{\alpha}_{\overline{E}} \underline{\alpha}$. Because of Lemma 3, $\underline{\alpha}_{\underline{E}} \overline{\alpha} \sim \underline{\alpha}_{\overline{E}} \overline{\alpha} \sim \overline{\alpha}_{\overline{E}} \underline{\alpha}$. Moreover, for any non empty E in S, $\underline{\alpha}_{\underline{E}} \overline{\alpha} \sim \underline{\alpha}_E \overline{\alpha}$ and $\overline{\alpha}_{\overline{E}} \underline{\alpha} \sim \overline{\alpha}_E \underline{\alpha}$. Hence, by transitivity, we obtain the result. ■

Proof of Lemma 5

To prove this, we can apply $A8$ to $(\underline{\alpha}, \overline{\alpha}, \underline{\alpha}) \sim (\underline{\alpha}, \underline{\alpha}, \overline{\alpha})$ and conclude that $(\underline{\alpha}, \frac{\overline{\alpha}+\underline{\alpha}}{2}, \frac{\overline{\alpha}+\underline{\alpha}}{2}) \succsim (\underline{\alpha}, \overline{\alpha}, \underline{\alpha}) \sim (\underline{\alpha}, \overline{\alpha}, \overline{\alpha})$. Dominance implies that $(\underline{\alpha}, \overline{\alpha}, \overline{\alpha}) \succsim (\underline{\alpha}, \frac{\overline{\alpha}+\underline{\alpha}}{2}, \frac{\overline{\alpha}+\underline{\alpha}}{2})$. Therefore, $h \sim (\underline{\alpha}, \frac{\overline{\alpha}+\underline{\alpha}}{2}, \frac{\overline{\alpha}+\underline{\alpha}}{2})$. Because of Lemmas 3 and 4, $(\underline{\alpha}, \frac{\overline{\alpha}+\underline{\alpha}}{2}, \frac{\overline{\alpha}+\underline{\alpha}}{2}) \sim (\underline{\alpha}, \frac{\overline{\alpha}+\underline{\alpha}}{2}, \underline{\alpha}) \sim (\underline{\alpha}, \underline{\alpha}, \frac{\overline{\alpha}+\underline{\alpha}}{2})$. Using again $A8$ we have that $(\underline{\alpha}, \frac{\overline{\alpha}+3\underline{\alpha}}{4}, \frac{\overline{\alpha}+3\underline{\alpha}}{4}) \sim h$. Iterating in this way we have $(\underline{\alpha}, \underline{\alpha} + \frac{\overline{\alpha}-\underline{\alpha}}{2^n}, \underline{\alpha} + \frac{\overline{\alpha}-\underline{\alpha}}{2^n}) \sim h$, for all $n = 0, 1, 2....$ This implies, given that the set of points $\{\frac{1}{2^n}\}_{n=0,1,2...}$ is dense in $(0, 1]$, and using dominance and transitivity, that $(\underline{\alpha}, \underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha}), \underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha})) \sim h$, for all $\gamma \in (0, 1]$. Suppose now, by contradiction, that $(\underline{\alpha}, \overline{\alpha}, \overline{\alpha}) \succ (\underline{\alpha}, \underline{\alpha}, \underline{\alpha})$ (we can simplify the notation, in what follows, by representing actions of the form (α, β, β) as (α, β)). Consider first the case $\underline{\alpha} > 0$. Then, given that $(\underline{\alpha}, \underline{\alpha}+\gamma(\overline{\alpha}-\underline{\alpha})) \sim (\underline{\alpha}, \overline{\alpha})$, for all $\gamma \in (0, 1]$, the contradiction hypothesis and dominance imply $(\underline{\alpha}, \underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha})) \succ (\underline{\alpha}, \underline{\alpha}) \succ (0, 0)$. By $A3$, for all $\gamma \in (0, 1] \exists \delta \in (0, 1)$, which might in principle depend on γ and which we therefore denote $\delta(\gamma)$, such that $(\delta(\gamma)\underline{\alpha}, \delta(\gamma)[\underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha})]) \succ (\underline{\alpha}, \underline{\alpha})$. Consider the value $\widehat{\gamma} \equiv \frac{(1-\delta(\gamma))\underline{\alpha}}{\delta(\gamma)(\overline{\alpha}-\underline{\alpha})}$, which is positive given that $\delta(\gamma) > 0$, $\overline{\alpha} > \underline{\alpha} > 0$. For $0 < \gamma < \widehat{\gamma}$, $\delta(\gamma)[\underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha})] < \underline{\alpha}$. Therefore, dominance implies that, for all such γ, $(\underline{\alpha}, \underline{\alpha}) \succsim (\delta(\gamma)\underline{\alpha}, \delta(\gamma)[\underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha})])$, contradicting the claim that $(\delta(\gamma)\underline{\alpha}, \delta(\gamma)[\underline{\alpha} + \gamma(\overline{\alpha} - \underline{\alpha})]) \succ (\underline{\alpha}, \underline{\alpha})$, for all $\gamma \in (0, 1]$. Consider now the other case, $\underline{\alpha} = 0$. We therefore have, because of previous results, that $(0, \gamma\overline{\alpha}) \sim (0, \overline{\alpha})$, for all $\gamma \in (0, 1]$. Suppose, by contradiction, that $(0, \overline{\alpha}) \succ (0, 0)$. Because of $A5$ and of previous results we can write $(1, 1) \succ (0, 1)$. Because of dominace we can also write $(0, 1) \succsim (0, \overline{\alpha})$, and by transitivity we conclude $(1, 1) \succ (0, \overline{\alpha}) \sim (0, \gamma\overline{\alpha}) \succ (0, 0)$. Therefore, by $A3$, for all $\gamma \in (0, 1], \exists \delta(\gamma) \in (0, 1)$ such that $(0, \gamma\overline{\alpha}) \succ (\delta(\gamma), \delta(\gamma))$. Consider the value $\widehat{\gamma} \equiv \frac{\delta(\gamma)}{\overline{\alpha}}$, which is positive given that $\delta(\gamma) > 0$, $\overline{\alpha} > 0$. For $0 < \gamma < \widehat{\gamma}$, $\gamma\overline{\alpha} < \delta(\gamma)$. Dominance implies that, for all such γ, $(\delta(\gamma), \delta(\gamma)) \succ (0, \gamma\overline{\alpha})$, leading to a contradiction. ■

Proof of Lemma 6

Consider the function f whose existence and unicity was established in Lemma 2. It was proved there that the components of f are affine and that they take values 0 and 1 at, respectively, w and $b_i, i = 1, 2, ...m$. Consider generic $(p_1, p_2, ..., p_m)$ and $(q_1, q_2, ..., q_m) \in H_A$. Let $(\alpha_1, \alpha_2, ...\alpha_m)$ and $(\beta_1, \beta_2, ...\beta_m)$, with $\alpha_j = f_j(p_j)$ and $\beta_j = f_j(q_j), j = 1, 2...m$, denote the two actions that are equivalent to $(p_1, p_2, ..., p_m)$ and $(q_1, q_2, ..., q_m)$, respectively, according to that Lemma. Let moreover $\underline{\alpha} = \min_j \alpha_j$ and $\underline{\beta} = \min_j \beta_j$. Because of Lemma 5, we know that $(p_1, p_2, ..., p_m) \sim (\underline{\alpha}, \underline{\alpha}, \underline{\alpha})$

and $(q_1, q_2, ..., q_m) \sim (\underline{\beta}, \underline{\beta}, \underline{\beta})$, where the notation $(\underline{x}, \underline{x}, \underline{x})$ makes implicit reference to any arbitrary partition (F, K, M) of the set of states S. Suppose now that $\underline{\alpha} \geq \underline{\beta}$. Then dominance immediately implies that $(p_1, p_2, ..., p_m) \succsim (q_1, q_2, ..., q_m)$. Conversely, suppose that $(p_1, p_2, ..., p_m) \succsim (q_1, q_2, ..., q_m)$. Transitivity then implies that $(\underline{\alpha}, \underline{\alpha}, \underline{\alpha}) \succsim (\underline{\beta}, \underline{\beta}, \underline{\beta})$. If it were the case that $\underline{\alpha} < \underline{\beta}$, dominance would lead to $(\underline{\beta}, \underline{\beta}, \underline{\beta}) \succ (\underline{\alpha}, \underline{\alpha}, \underline{\alpha})$, a contradiction. Therefore it must be that $\underline{\alpha} \geq \underline{\beta}$. ■

Proof of Lemma 7

Consider a decision problem in which there exists a $p \in \cap_{j \in S} Q(A, j)$, and in which $b_i = b_j \forall i, j$, and denote by b their common value. Fix a j and an i. Now, $(b^{-j}p) \sim (b^{-j}, f_j(p)b + (1 - f_j(p))w) \sim (b^{-i}, f_j(p)b + (1 - f_j(p))w)$, the first relationship following from the definition of f_j and the second because of $A7$, by identifying the set E there with the singleton $\{j\}$, and the set F with the singleton $\{i\}$. Also, $(b^i, p) \sim (b^{-i}, f_i(p)b + (1 - f_i(p))w)$ by definition. It then follows, by transitivity and part (a) of Lemma 1, that $(b^j, p) \succ (b^i, p)$ or $(b^i, p) \succ (b^j, p)$ according to $f_j(p)$ being greater or smaller than $f_i(p)$. Since in both cases we have a contradiction with $A10$, we conclude that $f_i(p) = f_j(p)$. The same argument can be repeated for any other i and j. ■

Proof of Lemma 8

Fix $j \in S$, and p^1 and p^0 in $Q(A, j)$ such that $(b^{-j}, p^1) \succ (b^{-j}, p^0)$, and take any $p \in Q(A, j)$. Assume first that the chosen p falls in case (i). Then

$$(b^{-j}, p) \sim (b^{-j}, g_j(p)p^1 + (1 - g_j(p))p^0) \qquad (a)$$

by definition of g_j. Also,

$$(b^{-j}, p) \sim (b^{-j}, f_j(p)b_j + (1 - f_j(p))w), \qquad (b)$$

$$(b^{-j}, p^0) \sim (b^{-j}, f_j(p^0)b_j + (1 - f_j(p^0))w) \qquad (c)$$

and

$$(b^{-j}, p^1) \sim (b^{-j}, f_j(p^1)b_j + (1 - f_j(p^1))w), \qquad (d)$$

by definition of f. Since in this case $g_j(p) \in [0,1]$, part (c) of Lemma 1 can be applied to (d) to conclude that $(b^{-j}, g_j(p)p^1 + (1-g_j(p))p^0) \sim (b^{-j}, g_j(p)[f_j(p^1)b_j + (1-f_j(p^1))w] + (1-g_j(p))p^0)$. Moreover, in a similar way, from (c) we get:

$$\begin{gathered}(b^{-j}, (1-g_j(p))p^0 + g_j(p)[f_j(p^1)b_j + (1-f_j(p^1))w]) \sim \\ (b^{-j}, (1-g_j(p))[f_j(p^0)b_j + (1-f_j(p^0))w] + g_j(p)[f_j(p^1)b_j + (1-f_j(p^1))w]) = \\ (b^{-j}, [g_j(p)f_j(p^1) + (1-g_j(p))f_j(p^0)]b_j + [1-g_j(p)f_j(p^1) - (1-g_j(p))f_j(p^0)]w).\end{gathered}$$

From transitivity we have:

$$\begin{gathered}(b^{-j}, f_j(p)b_j + (1-f_j(p))w) \sim \\ (b^{-j}, [g_j(p)f_j(p^1) + (1-g_j(p))f_j(p^0)]b_j + [1-g_j(p)f_j(p^1) - (1-g_j(p))f_j(p^0)]w).\end{gathered}$$

Part (b) of Lemma 1 then imply:

$$f_j(p) = g_j(p)[f_j(p^1) - f_j(p^0)] + f_j(p^0). \tag{e}$$

Also, $(b^{-j}, p^1) \succ (b^{-j}, p^0) \succsim (b^{-j}, w)$, because of $A4$ (with $q^{-j} = b^{-j}$) Hence, w falls in case (iii), and by definition: $(b^{-j}, \frac{g_j(w)}{g_j(w)-1}p^1 + (1 - \frac{g_j(w)}{g_j(w)-1})w) \sim (b^{-j}, p^0)$. From (d), as before, we can conclude that

$$\begin{gathered}(b^{-j}, \frac{g_j(w)}{g_j(w)-1}p^1 + \frac{1}{1-g_j(w)}w) \sim \\ (b^{-j}, \frac{g_j(w)}{g_j(w)-1}[f_j(p^1)b_j + (1-f_j(p^1))w] + \frac{1}{1-g_j(w)}w) = \\ (b^{-j}, \frac{g_j(w)}{g_j(w)-1}f_j(p^1)b_j + [1 - \frac{g_j(w)}{g_j(w)-1}f_j(p^1)]w).\end{gathered}$$

Recalling (c) we conclude, using transitivity and part (b) of Lemma 1, that:

$$f_j(p^0) = \frac{g_j(w)}{g_j(w)-1}f_j(p^1). \tag{f}$$

Moreover, $(b^{-j}, b_j) \succsim (b^{-j}, p^1) \succ (b^{-j}, p^0)$, because of $A4$ (with $q^{-j} = b^{-j}$). Hence, b_j falls in case (ii), and by definition: $(b^{-j}, \frac{1}{g_j(b_j)}b_j + (1 - \frac{1}{g_j(b_j)})p^0) \sim$

(b^{-j}, p^1). From (c), as before, we conclude that:

$$
\begin{aligned}
&(b^{-j}, \frac{1}{g_j(b_j)} b_j + (1 - \frac{1}{g_j(b_j)}) p^0) \sim \\
&(b^{-j}, \frac{1}{g_j(b_j)} b_j + (1 - \frac{1}{g_j(b_j)})[f_j(p^0) b_j + (1 - f_j(p^0)) w]) = \\
&(b^{-j}, \frac{f_j(p^0)[g_j(b_j) - 1] + 1}{g_j(b_j)} b_j + (1 - \frac{f_j(p^0)[g_j(b_j) - 1] + 1}{g_j(b_j)}) w).
\end{aligned}
$$

Recalling (d) we conclude, using transitivity and part (b) of Lemma 1, that:

$$
f_j(p^1) = \frac{f_j(p^0)[g_j(b_j) - 1] + 1}{g_j(b_j)}. \tag{g}
$$

The system of equations $(e$-f-$g)$ can now be solved to obtain:

$$
f_j(p) = \frac{g_j(p) - g_j(w)}{g_j(b_j) - g_j(w)}. \tag{h}
$$

Following a very similar reasoning it can easily be checked that also when $p \in Q(A, j)$ falls in case (ii) or (iii) one still obtains equation (e). Combining this with equations $(f$-$g)$ it is still then true that equation (h) holds. We now show that $g_j(w) = \min_{s \in Q(A,j)} g_j(s)$ by contradiction. Since $(b^{-j}, p^1) \succ (b^{-j}, p^0) \succsim (b^{-j}, w)$, because of $A4$ (with $q^{-j} = b^{-j}$), $g_j(w) = \frac{\gamma_w}{\gamma_w - 1} < 0$, where $\gamma_w \in [0, 1]$ is such that $(b^{-j}, \gamma_w p^1 + (1 - \gamma_w) w) \sim (b^{-j}, p^0)$. Suppose there is a $s \in Q(A, j)$ such that $g_j(s) < g_j(w) < 0$. It must then be that s falls in case (iii) and $g_j(s) = \frac{\gamma_s}{\gamma_s - 1}$, where $\gamma_s \in [0, 1]$ is such that $(b^{-j}, \gamma_s p^1 + (1 - \gamma_s) s) \sim (b^{-j}, p^0)$ and $\gamma_s > \gamma_w$. Because of $A4$ (with $q^{-j} = b^{-j}$) we have $(b^{-j}, p^1) \succ (b^{-j}, p^0) \succsim (b^{-j}, s) \succsim (b^{-j}, w)$. There are two possibilities: either $(b^{-j}, s) \sim (b^{-j}, w)$, or $(b^{-j}, s) \succ (b^{-j}, w)$. In the first case we have:

$$
\begin{aligned}
(b^{-j}, p^0) \sim (b^{-j}, \gamma_s p^1 + (1 - \gamma_s) s) \sim (b^{-j}, \gamma_s p^1 + (1 - \gamma_s) w) \succ \\
(b^{-j}, \gamma_w p^1 + (1 - \gamma_w) w) \sim (b^{-j}, p^0),
\end{aligned}
$$

where the first relationship follows from the definition of $g_j(s)$, the second from part (c) of Lemma 1 applied to $(b^{-j}, s) \sim (b^{-j}, w)$, the third from part (a) of Lemma 1, given that $\gamma_s > \gamma_w$, the fourth from the definition of $g_j(w)$.

Since we get a contradiction, it cannot be that $(b^{-j}, s) \sim (b^{-j}, w)$ (and at the same time $g_j(s) < g_j(w)$). Consider then the alternative. In this case we have $(b^{-j}, p^1) \succ (b^{-j}, s) \succ (b^{-j}, w)$. Because of part (b) of Lemma 1 there exists a value $\alpha \in [0,1]$ such that $(b^{-j}, \alpha p^1 + (1-\alpha)w) \sim (b^{-j}, s)$. We then have:

$$(b^{-j}, p^0) \sim (b^{-j}, \gamma_s p^1 + (1-\gamma_s)s) \sim (b^{-j}, \gamma_s p^1 + (1-\gamma_s)[\alpha p^1 + (1-\alpha)w]) =$$
$$(b^{-j}, (\gamma_s + \alpha(1-\gamma_s))p^1 + (1-(\gamma_s + \alpha(1-\gamma_s)))w) \succ (b^{-j}, \gamma_w p^1 + (1-\gamma_w)w) \sim (b^{-j}, p^0),$$

where the second relationship follows from part (c) of Lemma 1 applied to $(b^{-j}, \alpha p^1 + (1-\alpha)w) \sim (b^{-j}, s)$, the third from part (a) of Lemma 1, given that $\gamma_s + \alpha(1-\gamma_s) \geq \gamma_s > \gamma_w$, and the others follows as before from definitions. Again, we get a contradiction, and we conclude that there cannot be any $s \in Q(A,1)$ such that $g_j(s) < g_j(w)$. A very similar argument can also be used to show that $g_j(b_j) = \max_{s \in Q(A,j)} g_j(s)$. Using these conclusions, for any $p \in Q(A,j)$ we can write equation (h) as: $f_j(p) = \frac{g_j(p) - \min_{s \in Q(A,j)} g_j(s)}{\max_{s \in Q(A,j)} g_j(s) - \min_{s \in Q(A,j)} g_j(s)}$. Clearly, the whole line of proof can be followed for any other $j \in S$.

■

Proof of Theorem 1

We first prove that the axioms imply the characterisation (RMX), together with the properties of the g function. That there exist a function such that (RMX) is true is obvious, given Lemmas 6 and 8. The conclusion that all components of this function are affine is also immediate, given that, according to Lemma 8, they are all an affine transformation of affine functions. To prove that all components are unique up to a positive affine transformation, suppose that $g = (g_1, g_2, ...g_m)$ satisfies (RMX). Clearly, $v = (\kappa_{01} + \kappa_{11} g_1, \kappa_{02} + \kappa_{12} g_2, ...\kappa_{0m} + \kappa_{1m} g_m)$, with $\kappa_{0j}, j = 1, 2...m$ arbitrary constants and $\kappa_{1j}, j = 1, 2...m$ positive (and otherwise arbitrary) constants, also satisfies (RMX). To go the other way, suppose that both $g = (g_1, g_2, ...g_m)$ and $v = (v_1, v_2, ...v_m)$ satisfy (RMX). Assume, for the moment, that both g_j and v_j are minimised at w and maximised at $b_j, j = 1, 2, ...m$. Then, for any $(p_1, p_2, ..., p_m) \in H_A$,

$h = (h_j)_{j=1,2...m} = (\frac{g_j(p_j)-\min_{s\in Q(A,j)} g_j(s)}{\max_{s\in Q(A,j)} g_j(s)-\min_{s\in Q(A,j)} g_j(s)})_{j=1,2...m}$ and $k = (k_1)_{j=1,2...m} = (\frac{v_j(p_j)-\min_{s\in Q(A,j)} v_j(s)}{\max_{s\in Q(A,j)} v_j(s)-\min_{s\in Q(A,j)} v_j(s)})_{j=1,2...m}$ define functions in $[0,1]^m$, with $x_i(w) = 0, x_i(b_i) = 1$, such that $(p_1, p_2, ..., p_m) \succsim (q_1, q_2, ..., q_m)$ iff $\min_{j\in S}(x_j(p_j))_{j=1,2...m} \geq \min(x_j(q_j))_{j=1,2...m}, x = h, k$ and $j = 1, 2...m$. Since lemma 6 ensures that there is only one such function, it must be that $h = k$. Hence $\frac{g_j(p_j)-\min_{s\in Q(A,j)} g_j(s)}{\max_{s\in Q(A,j)} g_j(s)-\min_{s\in Q(A,j)} g_j(s)} = \frac{v_j(p_j)-\min_{s\in Q(A,j)} v_j(s)}{\max_{s\in Q(A,j)} v_j(s)-\min_{s\in Q(A,j)} v_j(s)}, j = 1, 2...m$. Therefore $v_j(p_j) = g_j(p_j)\frac{\max_{s\in Q(A,j)} v_j(s)-\min_{s\in Q(A,j)} v_j(s)}{\max_{s\in Q(A,j)} g_j(s)-\min_{s\in Q(A,j)} g_j(s)} + \frac{\min_{s\in Q(A,j)} v_j(s)-\min_{s\in Q(A,j)} g_j(s)}{\max_{s\in Q(A,j)} g_j(s)-\min_{s\in Q(A,j)} g_j(s)} = \kappa_{1j}g_j + \kappa_{0j}$, $\kappa_{1j} > 0$, as desired. It remains to be shown that both g_j and v_j are minimised at w and maximised at $b_j, j = 1, 2...m$.Suppose instead that, for some j, $\min_{s\in Q(A,j)} x_j(s) \neq x_j(w)$, and that there exist $r \in Q(A, j)$ such that $x_j(r) < x_j(w)$, with $x = g, v$ Therefore, $\min(\frac{x_j(r)-\min_{s\in Q(A,j)} x_j(s)}{\max_{s\in Q(A,j)} x_j(s)-\min_{s\in Q(A,j)} x_j(s)}, 1) < \min(\frac{x_j(w)-\min_{s\in Q(A,j)} x_j(s)}{\max_{s\in Q(A,j)} x_j(s)-\min_{s\in Q(A,j)} x_j(s)}, 1)$. Since x satisfies (RMX) by hypothesis, $(q^{-j}, r) \prec (q^{-j}, w)$, with $q_i = \arg\max_{s\in Q(A,i)} x_i(s), i \neq j$. However, from $A4$ we have $(q^{-j}, r) \succsim (q^{-j}, w)$. The contradiction implies that both g_j and v_j are minimised at w, for all j. Also, suppose that, for some j, $\max_{s\in Q(A,j)} x_j(s) \neq x_j(b_j)$, and that there exist $r \in Q(A, j)$ such that $x_j(r) > x_j(b_j)$, with $x = g, v$ Therefore, $\min(1, 1) > \min(\frac{x_j(b_j)-\min_{s\in Q(A,j)} x_j(s)}{\max_{s\in Q(A,j)} x_j(s)-\min_{s\in Q(A,j)} x_j(s)}, 1)$. Since x satisfies (RMX) by hypothesis, $(q^{-j}, r) \succ (q^{-j}, b_j)$, with $q_i = \arg\max_{s\in Q(A,i)} x_i(s), i \neq j$. This however contradicts $A4$, and it shows that both g_j and v_j are maximised at b_j, for all j.

To prove thath the axioms are implied by the characterisation (RMX), together with the properties of the g function, is a lengthy but straightforward exercise. We will use, without explicitly mentioning, the following obvious facts: $\min_j(p_j)_{j=1,2..m} = \min(\min_j(p_j)_{j\in S_1}, \min_j(p_j)_{j\in S_2})$ where S_1 and S_2 is any partition of S; $\min(a, b) = \min(b, a)$. Let $b_i = \arg\max_{r\in Q(A,i)} g_i(r)$ and $w = \arg\min_{r\in Q(A,i)} g_i(r), i = 1, 2...m$. Clearly, $A1$ is directly implied by (RMX). As to $A2$, let $(b^{-j}, p) \succ (b^{-j}, q)$; hence, $\min(1, \frac{g_j(p)-g_j(w)}{g_j(b_j)-g_j(w)}) > \min(1, \frac{g_j(q)-g_j(w)}{g_j(b_j)-g_j(w)})$. This implies that $g_j(p) > g_j(q)$ (and $g_j(q) < g_j(b_j)$). Take any $r \in Q(A, j)$. Then, for all $\alpha \in (0, 1]$, $g_j(\alpha p + (1-\alpha)r) = \alpha g_j(p) + (1-\alpha)g_j(r) > \alpha g_j(q)+(1-\alpha)g_j(r) = g_j(\alpha q+(1-\alpha)r)$, where we exploited the fact that the function g_j is affine. Moreover, $\frac{\alpha g_j(q)+(1-\alpha)g_j(r)-g_j(w)}{g_j(b_j)-g_j(w)} < 1$ as long as $\alpha > 0$, given that $g_j(r) \leq g_j(b_j)$. Therefore, $\min(1, \frac{g_j(\alpha p+(1-\alpha)r)-g_j(w)}{g_j(b_j)-g_j(w)}) > \min(1, \frac{g_j(\alpha q+(1-\alpha)r)-g_j(w)}{g_j(b_j)-g_j(w)})$, which is equivalent to $(b^{-j}, \alpha p+(1-\alpha)r) \succ (b^{-j}, \alpha q+(1-\alpha)r)$; repeating the same argument for any other j proves $A2$. Moving to $A3$, take $(p_1, p_2, ...p_m), (q_1, q_2, ...q_m), (r_1, r_2, ...r_m) \in H_A$ such that $(p_1, p_2, ...p_m) \succ (q_1, q_2, ...q_m) \succ (r_1, r_2, ...r_m)$. To simplify the notation, let $u_j = \frac{g_j(p_j)-g_j(w)}{g_j(b_j)-g_j(w)}$, $v_j = \frac{g_j(q_j)-g_j(w)}{g_j(b_j)-g_j(w)}$, $z_j = \frac{g_j(p_j)-g_j(w)}{g_j(b_j)-g_j(w)}$. The assumed preference implies $\min_j(u_j)_{j=1,2...m} > \min_j(v_j)_{j=1,2...m} > \min_j(z_j)_{j=1,2...m}$. We need to show that $\exists \alpha, \beta \in (0,1)$ such that $\min_j(\alpha u_j + (1-\alpha)z_j)_{j=1,2...m} > \min_j(v_j)_{j=1,2...m} > \min_j(\beta u_j + (1-\beta)z_j)_{j=1,2...m}$; this, given that $g_j(\cdot)$ is affine, would immediately lead to $A3$. Let $v_s = \min_j(v_j)_{j=1,2...m}$ and denote as i_1 the state where the minimum of z_j is achieved, i.e. $z_{i_1} = \min_{j\in S}(z_j)_{j=1,2...m}$ (for simplicity assume there is a single minimum). It must then be that $u_j > v_s, \forall j \in S$, and $z_{i_1} < v_s$. It is then true that $u_{i_1} > v_s > z_{i_1}$. For notational convenience define the function $h(\alpha, i) = \alpha u_i + (1-\alpha)z_i$, which is continuous in α for each given i. The intermediate value theorem guarantees that there exist a value $\alpha' \in (0,1) : h(\alpha', i_1) = v_s$. Hence, for all $0 < \beta < \alpha'$ we have that $h(\beta, i_1) < v_s$, while for all $\alpha' < \alpha < 1$ we have that $h(\alpha, i_1) > v_s$. Since $\min_j h(\beta, j)_{j=1,2...m} \leq h(\beta, i_1) < v_s$ this proves the second half of the claim. To prove the first, there are 3 possibilities to consider: (a) $h(\alpha', i_1)$ is the unique minimum across the various j of $h(\alpha', j)$; (b) $h(\alpha', i_1)$ is the minimum across the various j of $h(\alpha', j)$, but there is at least another state, denote it as i_2, such that $h(\alpha', i_1) = h(\alpha', i_2)$; (c) $h(\alpha', i_1) > \min_j h(\alpha', j)_{j=1,2...m} = h(\alpha', i_2)$. In case (a) continuity ensures that, for small enough $\alpha > \alpha'$, $h(\alpha, i_1)$ is still the minimum across states; since for any $\alpha > \alpha', h(\alpha, i_1) > v_s$, the first part of the claim would be proved. Cases (b) and (c) are only possible if there is more than one state i such that $z_i < v_s$, since otherwise for each state $i \neq i_1$, $h(\alpha', i)$ would be a convex combination of two numbers both bigger than v_s, and $h(\alpha', i_1) = v_s$ would be the unique minimum. In addition to $i_1$there are at most $m-1$ states $i_2, i_3, ...i_m$ such that z_{i_j} could be smaller than v_s. Consider now case (b), and assume for simplicity that the minimum is achieved only at i_1 and i_2. Continuity implies that for small enough

$\alpha > \alpha'$ the unique minimum across states is reached either at i_1 or at i_2, and in both cases it is larger than v_s, proving the first part of the claim. In case (c) the intermediate value theorem guarantees the existence of $\alpha'' > \alpha'$ such that $h(\alpha'', i_2) = v_s$, and $h(\alpha, i_2) > v_s$ for $\alpha > \alpha''$. As with $h(\alpha', i_1)$ there are again 3 possibilities, and the same analysis can be repeated. In at most $m-1$ steps we would then find the state i_j such that $h(\widehat{\alpha}, i_j) = v_s$ is the unique minimum across states, and we could then use the argument presented for case (a) to prove the first part of the claim. To prove $A4$, take any $p \in Q(A,j)$ and $q_i \in Q(A,i), i \neq j$. Let $\alpha = \min_i(\frac{g_i(q_i)-g_i(w)}{g_i(b_i)-g_i(w)})_{i \neq j}$ and $\beta = \frac{g_j(p)-g_j(w)}{g_j(b_j)-g_j(w)}$. It is either $\alpha < \beta \leq 1$ or $1 \geq \alpha \geq \beta$. In the first case $\min(\alpha,1) = \min(\alpha,\beta)$. In the second case $\min(\alpha,1) \geq \min(\alpha,\beta)$. Therefore, $(q^{-j}, b_j) \succsim (q^{-j}, p)$. Moreover, $\min(\alpha,\beta) \geq \min(\alpha,0)$. Hence $(q^{-j}, p) \succsim (q^{-j}, w)$. Since the chosen j is generic, this proves $A4$. The proof of $A5$ is obvious. To prove $A6$ consider $p, q \in Q(A,j)$ such that $(b^{-j}, p) \succsim (b^{-j}, q)$. Let $\alpha = \frac{g_j(p)-g_j(w)}{g_j(b_j)-g_j(w)}$ and $\beta = \frac{g_j(q)-g_j(w)}{g_j(b_j)-g_j(w)}$. It must then be that $\min(1,\alpha) \geq \min(1,\beta)$, which in turn implies that $\alpha \geq \beta$. Consider now any $s_i \in Q(A,i), i \neq j$, and let $\gamma = \min_i(\frac{g_i(s_i)-g_i(w)}{g_i(b_i)-g_i(w)})_{i \neq j}$. It is either $\gamma \geq \beta$ or $\gamma < \beta$. In the first case, $\min(\gamma,\alpha) \geq \beta = \min(\gamma,\beta)$. In the second, $\min(\gamma,\alpha) = \gamma = \min(\gamma,\beta)$. Hence, $(s^{-j}, p) \succsim (s^{-j}, q)$. Since the chosen j is generic, this proves $A6$. To prove $A7$, consider the two actions h and h' defined in the statement of the axiom. Using the fact that all components of g are affine, we have $\frac{g_j(\alpha_j b_j + (1-\alpha_j)w) - g_j(w)}{g_j(b_j)-g_j(w)} = \alpha_j$, for all $j = 1, 2...m$. Since swapping any subset of α_j does not modify the value of $\min_j(\alpha_j)_{j=1,2...m}$, it clearly follows that $h \sim h'$. To prove $A8$, take $(p_1, p_2...p_m), (q_1, q_2...q_m) \in H_A$, with $(p_1, p_2...p_m) \sim (q_1, q_2...q_m)$. Let $\alpha_j = \frac{g_j(p_j)-g_j(w)}{g_j(b_j)-g_j(w)}$ and $\beta_j = \frac{g_j(q_j)-g_j(w)}{g_j(b_j)-g_j(w)}$. It must then be that $\min_j(\alpha_j)_{j=1,2...m} = \min_j(\beta_j)_{j=1,2...m}$. Denote by $\underline{\alpha}$ the common value of the minimum. Let $M_\alpha = \{j : \alpha_j = \underline{\alpha}\}$, and $M_\beta = \{j : \beta_j = \underline{\alpha}\}$. There are two possible cases. Either $M_\alpha = M_\beta$, or $M_\alpha \neq M_\beta$. In the first case the minimum of any convex combination of α_j and β_j is still $\underline{\alpha}$, in the second case it is strictly greater than $\underline{\alpha}$. Hence $A8$ is satisfied. To prove $A9$, let $\alpha_j = \frac{g_j(p_j)-g_j(w)}{g_j(b_j)-g_j(w)} \geq \beta_j = \frac{g_j(q_j)-g_j(w)}{g_j(b_j)-g_j(w)}, j = 1, 2...m$. Therefore $(b^{-j}, p_j) \succsim (b^{-j}, p_j)$ for all j. Clearly, $\min_j(\alpha_j)_{j=1,2...m} \geq \min_j(\beta_j)_{j=1,2...m}$. This implies that $(p_1, p_2...p_m) \succsim (q_1, q_2...q_m)$. ■

Proof of Lemma 9

The proof is by induction. Denote by δ_c the degenerate distribution whose mass is concentrated on $c \in C(A, j)$, and define $u_j(c) = g_j(\delta_c)$. If the support of $p \in Q(A, j)$ has only one element, $p = \delta_c$ for some $c \in C(A, j)$ and the conclusion $g_j(p) = \sum_c u_j(c)p(c)$ follows trivially. Suppose now that the same conclusion holds for all $p \in Q(A, j)$ whose support has size $n - 1$. Take any $p \in Q(A, j)$ whose support has size n and let c' be in the support of such p. Define q as follows: $q(c) = 0$ if $c = c'$; $q(c) = p(c)/(1 - p(c'))$ if $c \neq c'$. Hence the support of q has size $n - 1$ and $p = p(c')\delta_{c'} + (1 - p(c'))q$. Since g_j is affine we have that $g_j(p) = p(c')g_j(\delta_{c'}) + (1 - p(c'))g_j(q) = p(c')u_j(c') + (1 - p(c'))g_j(q)$. Moreover, the induction hypothesis on q allows us to write $g_j(p) = p(c')u_1(c') + (1 - p(c')) \sum_{c \neq c'} u_j(c)q(c) = \sum_c u_j(c)p(c)$. Since we are dealing with simple probability distributions (i.e. distributions with finite support), this concludes the proof for u_j. The other components are proved in the same way. ■

References

[1] Altissimo, F., Siviero, S. and D. Terlizzese, (2005), "On Robust Monetary Policy", in *Long-Run Growth and Short-Run Stabilization: Essays in Memory of Albert Ando*, L.R. Klein Ed., Edward Elgar, Cheltenham, UK. Forthcoming.

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

[3] Arrow, K.J., (1984), Collected Papers (Vol. 1), *Social Choice and Justice*, Basil Blackwell, Oxford.

[4] Bergemann D. and K. Schlag, (2005), "Robust Monopoly Pricing: The Case of Regret", European University Institute, mimeo.

[5] Brafman, R. I. and M. Tennenholtz, (2003), "An Axiomatic Treatment of Three Qualitative Decision Criteria", *Journal of the ACM.*

[6] Cozzi, G. and P. E. Giordani (2005), "Is Strong Uncertainty Harmful for Shumpeterian Growth?", University of Macerata and European University Institute, mimeo.

[7] Fishburn, P.C, (1970), *Utility Theory for Decision Making*, New York, Wiley.

[8] Giannoni, M.P., (2002), "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward- Looking Model," *Macroeconomic Dynamics*, 6(1), pp. 111–144.

[9] Gilboa, I. and D. Schmeidler, (1989), "Maxmin Expected Utility with Non-Unique Prior", *Journal of Mathematical Economics*, 18, pp. 141-153, North Holland.

[10] Hansen, L.P. and T.J. Sargent (2001), "Robust Control and Model Uncertainty", *American Economic Review Papers and Proceedings,* 91 (2), pp. 60-66.

[11] Kreps, D.M., (1988), *Notes on the Theory of Choice*, Underground Classics in Economics, Westview Press/Boulder and London.

[12] Levin, A.T., V. Wieland, and J.C. Williams, (2003), "The Performance of Forecast-Based Monetary Policy Rules under Model Uncertainty," *American Economic Review*, 93 (3), pp. 622–645.

[13] Linhart, P.B. and R. Radner, (1989), "Minimax - Regret Strategies for Bargaining over Several Variables", *Journal of Economic Theory* 48, pp. 152-178.

[14] Milnor, J. (1954), “Games Against Nature”, in *Decision Processes*, R.M. Thrall, C.H. Comb and R.L. Davis, Eds. New York, Wiley.

[15] Onatski, A. and N. Williams, (2003), “Modeling Model Uncertainty,” *Journal of the European Economic Association*, 1(5), pp. 1087-1122.

[16] Puppe, C. and K. Schlag, (2006), “Choice Under Uncertainty when Outcome Spaces are State-Dependent”, European University Institute, mimeo.

[17] Savage, L.J., (1954), *The Foundations of Statistics,* New York, Wiley.

[18] Stoye, J. (2006), Statistical Decisions under Ambiguity, New York University, mimeo.

[19] Wald A., (1950), *Statistical Decision Functions.* New York, Wiley.