# The Ignorant Observer*

Thibault Gajdos†        Feriel Kandil‡

First draft: February 2005.
This draft: April 2007.

### Abstract

We propose an extension of Harsanyi's Impartial Observer Theorem based on the representation of ignorance as the set of all possible probability distributions over individuals. We obtain a characterization of the observer's preferences that, under our most restrictive conditions, is a convex combination of Harsanyi's utilitarian and Rawls' egalitarian criteria. This representation is ethically meaningful, in the sense that individuals' utilities are cardinally measurable and fully comparable. This allows us to conclude that the impartiality requirement cannot be used to decide between Rawls' and Harsanyi's positions.

**Keywords:** Impartiality, Justice, Utilitarianism, Egalitarianism, Decision under ignorance.

**JEL Numbers:** D63, D81.

## 1 Introduction

According to a long tradition among moral philosophers, moral judgements have to be made from the point of view of a rational, impartial and sympathetic observer. This idea is at the core of two prominent economic models of justice, namely Harsanyi's (1953, 1977) "utilitarianism"[1] and Rawls' (1971) egalitarianism. The fundamental insight put forward by Harsanyi and Rawls (and, independently, by Vickrey (1945)), is that *impartiality* can be ensured if the observer is placed under appropriate conditions of *ignorance* (the "veil of ignorance", in Rawls' (1971) terms). In particular, the observer should "*not know in advance* what his own social position would be in each social situation" (Harsanyi, 1977, p. 49).[2] A strong link is hence established between the theory of morality and the theory of decision making under ignorance. However, although

[1]Here and in all the sequel, "utilitarianism" refers to the preference utilitarianism, as in Harsanyi's work, and not to the doctrine of classical utilitarians.

[2]See Mongin (2001) for a thorough comparison of Vickrey's and Harsanyi's approaches.

Rawls and Harsanyi agree on the idea that fair rules are those chosen by rational individuals from behind the veil of ignorance, they strongly disagree on what these rules should actually be. Starting from similar assumptions, they end up with opposite conclusions: according to Rawls, the impartiality requirement leads to the egalitarian (or maxmin) criterion, whereas according to Harsanyi, the same requirement leads to utilitarianism.

Our aim is to determine if the impartiality requirement, viewed as ignorance, implies Harsanyi's or Rawls' conclusion. There are two main difficulties here. First, whereas Harsanyi (1977) proposed a formal model of decision making from behind the veil of ignorance (his celebrated "Impartial Observer Model"), Rawls only proposed informal arguments.[3] We therefore need to build a model that can accommodate Rawls' views on impartial decisions. The second difficulty is related to a well-known weakness of Harsanyi's model: as shown by Sen (1976) and Weymark (1991), the weights attributed to individuals' utilities in Harsanyi's Impartial Observer Theorem are not meaningful if these utilities are not cardinally comparable. We extend Harsanyi's model so that $(i)$ Rawls' argument can be formalized and $(ii)$ the conclusions we obtain are meaningful, i.e., make use of cardinally measurable and fully comparable individual utility functions.[4]

To make our approach clear, let us briefly present Harsanyi's (1953, 1977) Impartial Observer Model. Let $N$ be the set of individuals and $X$ be the set of social alternatives (both finite). Each individual has a preference relation $\succeq_i$ on the set $Y$ of lotteries over $X$ (social-alternative lotteries). Furthermore, these preferences are assumed to obey the axioms of expected utility theory. The observer is assumed to be a rational individual and to be able to make judgements such as: "social-alternative lottery $y$ is better for individual $i$ than social-alternative lottery $z$ for individual $j$". To formalize this idea, Harsanyi (1977) assumes that the observer has preferences on the set $\Delta(X \times N)$ of probability distributions over $X \times N$. Elements of $\Delta(X \times N)$ are called extended lotteries. The observer's preferences on extended lotteries are assumed to satisfy the axioms of expected utility theory. Harsanyi then adds two axioms. The first, known as the acceptance principle, states that whenever the observer has to rank two extended lotteries in which he is the same individual for sure, he does it the same way as that individual ranks the corresponding social-alternative lotteries. This axiom is intended to capture the observer's sympathy towards individuals. The second axiom, which is intended to capture the observer's impartiality, states

---

[3]Nevertheless, Rawls clearly thought that his conclusions could be derived from formal arguments. For instance, in Rawls (1974b), he argues that Arrow and Hurwicz's (1972) model of decision making under ignorance (which is a combination of the maximin and the minimax utilities) can be viewed as a step in that direction. This argument is precisely formalized in Maskin (1979) (Maskin uses a multiprofiles approach).

[4]Our conclusions are thus meaningful in the sense that the decision rules we obtain cannot be manipulated. However, we do not claim that the utility functions that are used to *represent* individual preferences coincide with individual utilities (in the utilitarian sense), because we use the von Neumann-Morgenstern theory to represent individual preferences. In this respect, the Sen (1976) and Weymark (1991) analysis also applies to our results. See Weymark (2005) for an illuminating investigation of this point.

that the observer ranks two social-alternative lotteries as he would rank the extended lotteries in which there is an equal chance of being any individual and all individuals face the same social-alternative lottery. In other words, Harsanyi represents ignorance by equiprobability. As a result, he obtains that the observer's preferences over social-alternative lotteries can be represented by the arithmetic mean of some adequately chosen individual von Neumann-Morgenstern utility functions.

Harsanyi's theorem presents the following problem, raised by Sen (1976) and Weymark (1991): even if one assumes that individual von Neumann-Morgenstern utilities have a cardinal meaning (which is not the case in the standard expected utility theory), the choice of a *specific* representation of individual von Neumann-Morgenstern utilities implies that the weights that appear in Harsanyi's theorem are not meaningful. The reason for this is that in Harsanyi's model, individual utilities are not cardinally comparable.

Karni (1998) and Mongin (2001) proposed a nice solution to escape both problems: they assume that the observer's preferences conform to the *subjective* expected utility theory. Therefore, provided that one can identify the observer's subjective probabilities, the weights would be determined.[5] An important feature of these approaches is that they remain inside the Bayesian theory. But Rawls (1971) explicitly rejected such an assumption. Therefore, if we want to take into account Rawls' arguments, we need a model that does not assume Bayesianism from the outset.

Note that Harsanyi and Rawls agree that probabilities should be taken into account whenever they have some objective basis. This suggests that the decision maker's knowledge can be represented by a *set of probability distributions* that describes all probability distributions that are *possible* according to the decision maker's factual or logical knowledge. The first step in our reconstruction of Harsanyi's impartial observer model is therefore to provide an axiomatic foundation for the observer's preferences when her information about who she is to be in the society takes the form of a set of probability distributions. Several axiomatizations of such preferences have been recently proposed.[6] Among these models, the one considered by Gajdos, Hayashi, Tallon and Vergnaud (2007) is the closest to the one we provide here. However, their model cannot be used for the observer's preferences because it assumes state-independence (as do all models of this kind that we are aware of), which would force all individuals to have the same

---

[5]Karni (1998) obtains this result through a sophisticated construction that can be interpreted in terms of impartiality, whereas Mongin (2001) proposes axioms of an epistemic nature related to those introduced by Karni and Schmeidler (1981).

[6]The idea of modeling information as a set of probability distributions seems to have been first proposed by Jaffray (1989) in a model where preferences are defined over belief functions. Wang (2003) and Gajdos, Hayashi, Tallon and Vergnaud (2004) considered information as a set of probability distributions together with an "anchor", i.e., a probability distribution that has particular salience.

preferences.[7] Furthermore, the axiomatization proposed by Gajdos, Tallon and Vergnaud (2007) requires an infinite state space, which would be difficult to justify in the present framework. Finally, their article is mainly concerned with a formulation of uncertainty aversion directly related to comparisons of sets of information, instead of the classical formulation in terms of randomization. However, in the present context, randomization has a natural interpretation from an ethical point of view and we will therefore keep it explicitly in our model.

Viewing complete ignorance as equivalent to considering that *all* probability distributions are possible, we are then in position to reconstruct Harsanyi's impartial observer's theorem in an extended framework that does not assume Bayesianism from the outset. We obtain our "Ignorant Observer" Theorem which, in its most precise formulation, asserts that the observer's preferences on social-alternatives lotteries can be represented by:

$$V(y) = \theta \min_{i \in N} V_i(y) + (1 - \theta) \sum_{i \in N} \frac{1}{n} V_i(y)$$

for all $y \in Y$, where $\theta \in [0, 1]$ is uniquely determined (for a given observer) and the utility functions $V_i$ are cardinally measurable and fully comparable representations of individual preferences. More precisely, the functions $V_i$ ($i \in N$) are chosen such that $V_i(Y) = V_j(Y)$ for all $i, j \in N$. The above result can also be written as:

$$V(y) = \theta \min_{i \in N} \frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)} + (1 - \theta) \sum_{i \in N} \frac{1}{n} \frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}$$

where again $\theta \in [0, 1]$ is uniquely determined (for a given observer), but where the $V_i$ are arbitrarily chosen von Neumann-Morgenstern utility functions representing the individuals' preferences.[8]

We therefore conclude that the impartiality requirement is compatible with both Harsanyi's and Rawls' views, for Harsanyi's criterion is obtained for $\theta = 0$, whereas Rawls' criterion is obtained for $\theta = 1$. Considering under which conditions this model specializes into that of Harsanyi or Rawls, one sheds some light on the debate between them. This leads us to defend the view that a (strict) combination of Harsanyi's and Rawls' criteria (i.e., choosing $\theta \in (0, 1)$) leads to a reasonable criterion for social decision making.

This article is organized as follows. First, we extend Harsanyi's framework by considering sets of probability distributions instead of lotteries on individual identities (Section 2). We then provide an axiomatic characterization of the observer's preferences in this extended framework (Section 3). In Section 4, we formalize the impartiality requirement as complete ignorance, in

---

[7]We are grateful to Edi Karni for having drawn our attention to this point.

[8]This result seems, at first sight, very similar to the one obtained by Karni (1998). There are, however, important differences between Karni's approach and ours. We elaborate on this point in Sections 3 and 4.

the sense that all lotteries on individual identities are considered as possible. We then reconstruct Harsanyi's impartial observer theorem under these hypotheses, assuming that individual preferences satisfy the axioms of expected utility theory, and state our Ignorant Observer Theorems. Finally, in Section 5, we defend the view that both Harsanyi's and Rawls' solutions are unsatisfactory, whereas a mix of the two (i.e., with $\theta \in (0,1)$) is a reasonable criterion for social decision making.

## 2 Modeling ignorance

We consider a society made up of a finite number of agents $N = \{1, \ldots, n\}$. Let $X$ be a non-empty finite set of social alternatives (or consequences) and $Y$ be the set of probability distributions over $X$ (social-alternative lotteries). Following Harsanyi (1953, 1977), individuals are assumed to have preferences on $Y$. These preferences are denoted $\succeq_i$ ($i \in N$). As is customary, we denote by $\sim_i$ and $\succ_i$ the symmetric and asymmetric components of $\succeq_i$.

An *observer* is someone able to make social judgements of the following kind: "social-alternative lottery $y$ is better for individual $i$ than social-alternative lottery $z$ for individual $j$". In order to make such a statement formally, Harsanyi (1977) assumed that the observer has preferences over the set of all *extended lotteries*, i.e., lotteries on $X \times N$. We will denote by $\mathcal{E}$ the set of such lotteries. An element of $\mathcal{E}$ is thus a function $\rho : X \times N \to [0,1]$ such that $\sum_{x \in X} \sum_{i \in N} \rho(x,i) = 1$, where $\rho(x,i)$ is the probability of being individual $i$ and getting $x$.

Karni and Weymark (1998) proposed the following illuminating interpretation of an extended lottery. Such a lottery can be viewed as a two stage lottery, where a first lottery on $N$ determines which individual the observer is to be and a second lottery on $X$ then determines what the social state is. Formally, Karni and Weymark (1998) defined a "personal identity lottery" as a probability distribution $p$ on $N$ and an "allocation" $f$ as an assignment of a lottery on $X$ to each individual. Let $\Delta(N)$ be the set of all probability distributions on $N$ and $\mathcal{A}$ be the set of allocations, i.e., the set of all functions from $N$ to $Y$. Let $\mathcal{A}^c$ be the set of *constant* allocations, i.e., allocations $f$ such that $f(i) = f(j)$ for all $i, j \in N$. As noted by Mongin and d'Aspremont (1998) and Karni and Weymark (1998), interpreting individuals as states of the nature, an allocation is an act in the Anscombe-Aumann (1963) model.

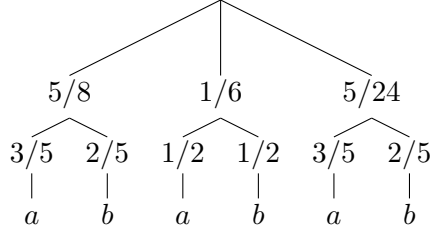The following example illustrates the correspondence between $\mathcal{E}$ and $\mathcal{A} \times \Delta(N)$. Assume that $N = \{1, 2, 3\}$, $X = \{a, b\}$ and consider the following extended lottery $\rho$:

|           | 1    | 2     | 3     |
|-----------|------|-------|-------|
| $a$       | 3/8  | 1/12  | 1/8   |
| $b$       | 1/4  | 1/12  | 1/12  |
| $p(\rho)(i)$ | 5/8  | 1/6   | 5/24  |

5

$y_i(\rho)$ is then:

|   | 1 | 2 | 3 |
|---|---|---|---|
| $a$ | 3/5 | 1/2 | 3/5 |
| $b$ | 2/5 | 1/2 | 2/5 |

Finally, $(f, p)$ can be represented as follows:



Our formulation of the observer's preferences will be based on this observation. Following Karni and Weymark (1998), we will identify an extended lottery $\rho$ with a couple $(f, p) \in \mathcal{A} \times \Delta(N)$ as follows. Let $p(\rho)(i) = \sum_{x \in X} \rho(x, i)$, for all $i \in N$, and whenever $p(\rho)(i) > 0$, let $y_i(\rho)(x) = \frac{\rho(x,i)}{\sum_{z \in X} \rho(z,i)}$ for all $x \in X$, with $f(\rho)(i) = y_i(\rho)$. If $p(\rho)(i) = 0$, let $y_i(\rho)$ be an arbitrary element of $Y$. Finally, define $p(\rho) \otimes f(\rho)$ by $(p(\rho) \otimes f(\rho))(x, i) = p(\rho)(i)y_i(\rho)(x)$. Clearly, $\rho = p(\rho) \otimes f(\rho)$.

Let $\mathbb{P}$ be the set of all non-empty, compact and convex sets of probability distributions on $N$, where compactness is defined with respect to the Euclidean space $\mathbb{R}^{|N|}$. A generic element of $\mathbb{P}$ will be denoted by $\mathcal{P}$. Finally, $\delta_i$ is the probability distribution on $N$ defined by $\delta_i(i) = 1$.

The observer's preferences $\succeq$ will be defined on the product $\mathcal{A} \times \mathbb{P}$ ($\sim$ and $\succ$ will, as usual, denote the symmetric and asymmetric components of $\succeq$, respectively). The couple $(f, \mathcal{P})$ can be interpreted as follows: the observer knows that the allocation is given by $f$ and he also knows that the probability distribution on $N$ according to which his identity will be chosen is in the set $\mathcal{P}$, but has no further information on the process assigning identities. It is important to note that a set of probability distribution $\mathcal{P} \in \mathbb{P}$ is here thought of as *objective* data of the decision problem in hand.

Harsanyi assumes that the observer is Bayesian *and* that being completely ignorant about the probability distribution that governs the individual lottery is equivalent to knowing for sure that the individual lottery has a uniform distribution. Translated into our framework, these assumptions are: $(i)$ for all $f \in \mathcal{A}$ and all $\mathcal{P} \in \mathbb{P}$, there exists $p \in \mathcal{P}$ such that $(f, \mathcal{P}) \sim (f, \{p\})$, and $(ii)$ for all $f \in \mathcal{A}$, $(f, \Delta(N)) \sim (f, \{\mu\})$, where $\mu$ is the uniform distribution on $N$. The first assumption is a version of Bayesianism, whereas the second is what Harsanyi (1977) calls the "Equal Chance Principle". These assumptions seem highly dubious and are actually related to the decision maker's attitude towards imprecise information and not to his supposed rationality: they have a psychological meaning (namely: neutrality towards uncertainty). Note that, in this

respect, Rawls' approach is no more convincing, since he assumes from the outset that the decision maker has an extreme aversion towards uncertainty. Our aim is therefore to propose a general decision model on $\mathcal{A} \times \mathbb{P}$ that leaves open the decision maker's attitude towards information imprecision.

A crucial property of our framework is that ignorance is not assumed from the outset, as it is the case in most theories of individual decision under uncertainty. This might seem strange, insofar as the *impartial* observer eventually only compares allocations under complete ignorance. In other words, the domain of the observer's preference might seem to be in some sense too large.[9] This choice was motivated by two reasons. First, because following Harsanyi and Rawls, we view moral decisions as rational decisions under some special circumstances (namely, ignorance), it is important to be sure that the decision rules we characterize are consistent with "rational"[10] decision rules on the full domain. If this theory cannot be extended (in a consistent way) so as to deal with all loteries, it is not sure that it would be thought of as a "good theory". The second reason for insisting on working with the large domain is more specific to the problem we deal with. We said that both Rawls and Harsanyi viewed moral decisions as *rational decisions* under *appropriate conditions of ignorance.* Thus there are two questions: $(i)$ which decision rule is rational when one has some possibly incomplete information, and $(ii)$ what is the information we should assume the observer has in order to implement impartial decisions. We believe that it is conceptually important to separate these two issues, as only the second has a moral content. Making such a distinction is only possible if one considers a large domain.[11] Moreover, we believe that statements like: "the observer prefers being individual $i$ for sure than being individual $j$ for sure when the social allocation is $f$", or "the observer prefers being individual $i$ for sure than being totally ignorant about his identity when the social allocation is $f$" are relevant. Such statements cannot be made if one restricts the domain to social allocations under complete ignorance. Thus, our approach radically differs from the literature on decision making under complete ignorance that imposes no probabilistic structure on the set of consequences (not even a state space), and often allows for incomplete preference, a route we totally ignore here.[12]

---

[9]A similar objection can be made to Harsanyi: The observer's preferences in Harsanyi's Theorem are defined over *all* extended lotteries, although only constant impartial lotteries are eventually considered by the impartial observer. Restricting the domain raises some technical difficulties, that were pointed out and solved by Karni and Weymark (1998) who provided an "informationally parsimonious" version of Harsanyi's Impartial Observer Theorem.

[10]Here, "rational" means "satisfy some well identified and acceptable axioms"

[11]Maybe part of the misunderstanding around the Rawls-Harsanyi debate can actually be explained by a confusion between these two issues.

[12]Barberà, Bossert and Pattanaik (2004) provides an extensive survey on that literature, including application to social choice.

# 3 The Observer's preferences

We now turn to the observer's preferences on $\mathcal{A} \times \mathbb{P}$. Our model is similar in spirit to the one axiomatized by Gajdos, Hayashi, Tallon and Vergnaud (2007). However, their model is state-independent, which would, in our framework lead to uniform individual preferences. We therefore need to relax the state-independence assumption. Furthermore, the above mentioned paper aimed at defining an imprecision aversion concept directly related to comparisons of sets of information. But, as we will show, the classical definition of uncertainty aversion through preference for randomization can be easily interpreted in the social choice framework. We will therefore keep such a definition of uncertainty aversion. Finally, the Gajdos, Hayashi, Tallon and Vergnaud (2007) axiomatization relies on operations on the state space that are difficult to interpret in this framework, where states are individuals. We will therefore avoid them. Let us note that Gilboa and Schmeidler's (1989) maxmin model cannot be used here, for two reasons. First, because it is state-independent, it would lead to uniform individual preferences. Second, this model does not permit taking into account objective information. Indeed, in Gilboa and Schmeidler's (1989) model, objects of choice are acts (i.e., elements of $\mathcal{A}$). There is therefore no way to say, for instance, that the decision maker prefers the act $f$ together with an information set $\mathcal{P}$ over an act $g$ together with an information set $\mathcal{Q}$. The (unique) set of probability distributions the decision maker uses to evaluate acts in Gilboa and Schmeidler's (1989) model is fixed and of purely subjective nature: it *only* depends on the decision maker's preferences.

We start by three quite standard axioms, that require the preference relation $\succeq$ on $\mathcal{A} \times \mathbb{P}$ to be complete, transitive, non-degenerate and continuous. As usual, convex combinations in $\mathcal{A}$ are performed pointwise: for $f, g \in \mathcal{A}$ and $\alpha \in [0, 1]$, $\alpha f + (1-\alpha)g = h$ where $h(i) = \alpha f(i) + (1-\alpha)g(i)$ for all $i \in N$.

**Axiom 1** (Ordering). *$\succeq$ is a reflexive, complete and transitive binary relation on $\mathcal{A} \times \mathbb{P}$.*

**Axiom 2** (Non-degeneracy). *For every $\mathcal{P} \in \mathbb{P}$, there exist $f, g \in \mathcal{A}$ such that $(f, \mathcal{P}) \succ (g, \mathcal{P})$.*

**Axiom 3** (Act Continuity). *For all $f$, $g$, $h \in \mathcal{A}$ and all $\mathcal{P} \in \mathbb{P}$, if $(f, \mathcal{P}) \succ (g, \mathcal{P}) \succ (h, \mathcal{P})$, then there exists an $\alpha$ in $(0, 1)$ such that :*

$$(\alpha f + (1 - \alpha)h, \mathcal{P}) \sim (g, \mathcal{P}).$$

The following notion of mixture of *sets* of probability distributions will be extensively used in the sequel.

**Notation 1.** *For all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ and all $\alpha \in [0, 1]$, the $\alpha-$mixture of $\mathcal{P}$ and $\mathcal{Q}$ is defined by:*

$$\alpha \mathcal{P} + (1 - \alpha)\mathcal{Q} = \{p \in \Delta(N) | p = \alpha p_1 + (1 - \alpha)p_2, \ p_1 \in \mathcal{P}, \ p_2 \in \mathcal{Q}\}.$$

An element $(f, \alpha\mathcal{P} + (1-\alpha)\mathcal{Q})$ of $\mathcal{A} \times \mathbb{P}$ can be interpreted as a compound lottery in which in the first stage $(\mathcal{P}, f)$ and $(\mathcal{Q}, f)$ are obtained with probabilities $\alpha$ and $(1-\alpha)$, respectively.

Let us consider $(f, \mathcal{P}_1)$, $(g, \mathcal{Q}_1)$, $(f, \mathcal{P}_2)$ and $(g, \mathcal{Q}_2)$ such that the Observer prefers $(f, \mathcal{P}_1)$ to $(g, \mathcal{Q}_1)$ and $(f, \mathcal{P}_2)$ to $(g, \mathcal{Q}_2)$. Assume, now, that the observer faces a choice between $(f, \alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2)$ and $(g, \alpha\mathcal{Q}_1 + (1-\alpha)\mathcal{Q}_2)$. He might reason as follows: with probability $\alpha$, I would obtain $(f, \mathcal{P}_1)$ if I have chosen $(f, \alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2)$ and $(g, \mathcal{Q}_1)$ if I have chosen $(g, \alpha\mathcal{Q}_1 + (1-\alpha)\mathcal{Q}_2)$. Since I prefer $(f, \mathcal{P}_1)$ over $(g, \mathcal{Q}_1)$, it is better for me to choose $(f, \alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2)$, conditional on the realization of the event whose probability is $\alpha$. Similarly, since I prefer $(f, \mathcal{P}_2)$ over $(g, \mathcal{Q}_2)$, it is better for me to choose $(f, \alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2)$, conditional on the realization of the event whose probability is $(1-\alpha)$. Thus, I prefer *unconditionally* $(f, \alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2)$ over $(g, \alpha\mathcal{Q}_1 + (1-\alpha)\mathcal{Q}_2)$. This leads us to the following axiom, which is a mere extension of the "Constrained Independence Axiom" proposed by Karni and Safra (2000).[13]

**Axiom 4** (Set-Mixture Independence). *For all $\mathcal{P}_1, \mathcal{Q}_1, \mathcal{P}_2, \mathcal{Q}_2 \in \mathbb{P}$, all $\alpha \in [0,1]$ and for all $f, g \in \mathcal{A}$,*

$$\left. \begin{array}{l} (f, \mathcal{P}_1) \succeq (\succ)(g, \mathcal{Q}_1) \\ (f, \mathcal{P}_2) \succeq (g, \mathcal{Q}_2) \end{array} \right\} \Rightarrow (f, \alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2) \succeq (\succ)(g, \alpha\mathcal{Q}_1 + (1-\alpha)\mathcal{Q}_2).$$

The next axiom concerns comparisons of information sets. It states that if an allocation $f$ is judged better than another allocation $g$ according to any probability distribution in $\mathcal{P}$, then $(f, \mathcal{P})$ is judged better than $(g, \mathcal{P})$.

**Axiom 5** (Dominance). *For all $\mathcal{P} \in \mathbb{P}$, if for all $p \in \mathcal{P}$, we have $(f, \{p\}) \succeq (g, \{p\})$ $((f, \{p\}) \succ (g, \{p\}))$ then $(f, \mathcal{P}) \succeq (g, \mathcal{P})$ $((f, \mathcal{P}) \succeq (g, \mathcal{P}))$.*

Now, because we do not want to impose state-independence (that would lead in our framework to the conclusion that all individuals' preferences on $Y$ are identical), we need to construct allocations that would play the role that constant allocations usually play. To do so, we define the set $\mathcal{A}^{cv}$ of constant-valued allocations. These allocations are characterized by the fact that the observer is indifferent between being individual $i$ or individual $j$ for sure, for all pairs $(i, j)$. Formally,

$$\mathcal{A}^{cv} = \{f \in \mathcal{A} | (f, \{\delta_i\}) \sim (f, \{\delta_j\}), \forall i, j \in N\}.$$

Notions similar to that of constant-valued acts have appeared in Drèze (1987), Karni (1993, 2007) and Skiadas (1997a, 1997b).

The next axiom is a classical boundedness requirement with respect to the set of constant-valued allocations. In particular, this axiom guarantees that $\mathcal{A}^{cv}$ is not empty.

---

[13]The above interpretation is exactly the one proposed by Karni and Safra (2000), where sets of probability distributions are considered instead of probability distributions.

**Axiom 6** (Boundedness). *For all $\mathcal{P} \in \mathbb{P}$ and $f \in \mathcal{A}$, there exist $\bar{f}$ and $\underline{f}$ in $\mathcal{A}^{cv}$ such that:*

$$(\bar{f}, \mathcal{P}) \succeq (f, \mathcal{P}) \succeq (\underline{f}, \mathcal{P}).$$

A similar axiom can be found, for instance, in Luce and Krantz's (1971) axiomatization of state-dependent expected utility. Essentially, it amounts to assuming that from the observer's point of view, the range of utility over social lotteries conditionally on being individual $i$ for sure is the same as the range of utility over social lotteries conditionally on being individual $j$ for sure, for all $i, j \in N$. In other words, from the observer's point of view (i.e., from behind the veil of ignorance), all individuals face the same opportunity set in terms of well-being. Put differently, this axioms implies that there is not *a priori* reason for the observer to prefer being an individual or another. It thus conveys a notion of impartiality. In that sense, this axiom has a strong ethical meaning. A similar idea can be found in Karni (1998). A natural objection to Axiom 6 would be the following. Assume that individuals $i$ and $j$ are identical in all respects, except that individual $i$ is disabled, whereas individual $j$ is not. In this case, one may reasonably think that individual $j$ can achieve a greater level of well-being than individual $i$. This is precisely a case where one might *not* want to be impartial between $i$ and $j$: The observer might want to favor an individual from the outset. Now, assume that the set of social alternatives includes a state in which individual $i$ *is not* disabled. Then, it is reasonable to say that individuals $i$ and $j$ can achieve the same level of well-being (observe that Axiom 6 doesn't require that all individuals *actually* achieve the same level of well-being). Thus Axiom 6 can be justified if one adopts a restrictive definition of preferences and an extensive definition of social alternatives. Finally, from a technical point of view, it is probably possible to weaken Axiom 6, for instance only assuming that there are non-trivial overlap of the ranges of conditional utilities (with some obvious change in the results). This would, however, add to our analysis complications that are, in our opinion, not worth the cost.

The next axiom is the analogue of the $C-$independence Axiom of Gilboa and Schmeidler (1989), where the set of constant-valued allocations replaces the set of constant allocations. It states that if $(f, \mathcal{P})$ is judged better than $(g, \mathcal{Q})$, then this relation is preserved if one mixes $f$ and $g$ with some constant-valued allocation $h$.[14]

**Axiom 7** ($\mathcal{A}^{cv}$-Independence). *For all $f, g \in \mathcal{A}$, $h \in \mathcal{A}^{cv}$, $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ and all $\alpha \in (0, 1)$,*

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow (\alpha f + (1 - \alpha)h, \mathcal{P}) \succeq (\alpha g + (1 - \alpha)h, \mathcal{Q}).$$

The next axiom is a reduction axiom. Assume that $f \in \mathcal{A}$ and $h_i \in \mathcal{A}^{cv}$ $(i \in N)$ are such that $(f, \{\delta_i\}) \sim (h_i, \{\delta_k\})$, where $k \in N$ is fixed. Let $p \in \Delta(N)$ be given. The pair $(f, \{p\})$ can be

---

[14]A similar axiom appears in Karni (2006).

viewed as a lottery over the product $Y \times N$, where $p(i)$ is the probability of being individual $i$ and getting $f(i)$, for all $i \in N$. Let us denote this lottery by $\ell = ((f(1), 1), p(1); \ldots ; (f(n), n), p(n))$, where $(f(i), i)$ means "being individual $i$ and getting $f(i)$".

Now, consider the pair $(\sum_{i \in N} p(i) h_i, \{\delta_k\})$. This pair can be viewed as a two-stage lottery $\tilde{\ell} = ((\ell_1, 1), 0; \ldots ; (\ell_k, k), 1; \ldots ; (\ell_n, n), 0)$, where $\ell_i = (h_1(i), p(1); \ldots ; h_n(i), p(n))$ are lotteries over $Y$. If one accepts the reduction of compound lottery axiom, one should be indifferent between $\tilde{\ell}$ and $\ell^* = ((h_1(k), k), p(1); \ldots ; (h_n(k), k), p(n))$. Because $h_i \in \mathcal{A}^{cv}$ for all $i \in N$, $(h_i, \{\delta_i\}) \sim (h_i, \{\delta_k\})$. Thus, the decision maker is indifferent between being individual $i$ and getting $h_i(i)$ and being individual $k$ and getting $h_i(k)$. But $(h_i, \{\delta_k\}) \sim (f, \{\delta_i\})$ and therefore the decision maker is indifferent between being individual $k$ and getting $h_i(k)$ and being individual $i$ and getting $f(i)$. Hence, for each $i$, $\ell$ and $\ell^*$ have, with probability $p(i)$, consequences among which the decision maker is indifferent. It is thus reasonable to assume that he is indifferent between them. Thus, he should also be indifferent between $\tilde{\ell}$ and $\ell$ and, therefore, between $(\sum_{i \in N} p(i) h_i, \{\delta_k\})$ and $(f, \{p\})$. This leads us to the following axiom.

**Axiom 8** (Reduction). *For all $f, g \in \mathcal{A}$, $p \in \Delta(N)$, and $k \in N$, if there exist $h_1, \ldots, h_n \in \mathcal{A}^{cv}$ such that:*

$$\begin{cases} g = \sum_i p(i) h_i \\ (h_i, \{\delta_k\}) \sim (f, \{\delta_i\}) \end{cases}$$

*then $(f, \{p\}) \sim (g, \{\delta_k\})$.*

Now, consider a constant-valued allocation $h$. Axiom 8 (with $h_i = h$ for all $i$) implies that for all $p \in \Delta(N)$ and any $k \in N$, $(h, \{p\}) \sim (h, \{\delta_k\})$. Thus, for any $p, q \in \Delta(N)$, $(h, \{p\}) \sim (h, \{q\})$. It is therefore reasonable to assume that the information set does not matter when the decision maker faces constant-valued allocations. This is precisely the meaning of the next axiom.

**Axiom 9** (Information indifference on $\mathcal{A}^{cv}$). *For all $h \in \mathcal{A}^{cv}$ and all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$, $(h, \mathcal{P}) \sim (h, \mathcal{Q})$.*

We will also assume that only the part of the allocation on which there is a positive probability for some probability distribution in the information set matters. Formally, let $S(\mathcal{P})$ be the subset of $N$ defined by $S(\mathcal{P}) = \{i \in N | \exists p \in \mathcal{P} \text{ s.t. } p(i) > 0\}$. In other words, $S(\mathcal{P})$ is the union of the supports of all probability distributions in the information set. For any subset $E$ of $N$ and pair of allocations $(f, g)$, we define the allocation $f_E g$ as $(f_E g)(i) = f(i)$ if $i \in E$ and $(f_E g)(i) = g(i)$ is $i \in N \setminus E$.

**Axiom 10** (Equivalence). *For all $f, g \in \mathcal{A}$, $\mathcal{P} \in \mathbb{P}$, $(f, \mathcal{P}) \sim (f_{S(\mathcal{P})} g, \mathcal{P})$.*

In particular, Axiom 10 implies that for all $f \in \mathcal{A}$ and $i \in N$, $(f, \{\delta_i\}) \sim (h, \{\delta_i\})$ for all $h$ such that $h(i) = f(i)$.

Finally, our last axiom is a version of the uncertainty aversion axiom of Schmeidler (1984), Chateauneuf (1991) and Gilboa and Schmeidler (1989). In the framework of decision making under uncertainty, it simply stipulates that the decision maker exhibits a (weak) preference for hedging. This axiom may also be interpreted from an ethical point of view. We defer this discussion to Section 5.

**Axiom 11** (Uncertainty aversion). *For all $f, g \in \mathcal{A}$, $\mathcal{P} \in \mathbb{P}$, $\alpha \in (0,1)$, $(f, \mathcal{P}) \sim (g, \mathcal{P})$ implies $(\alpha f + (1-\alpha)g, \mathcal{P}) \succeq (f, \mathcal{P})$.*

Let $\hat{\succeq}_i$ be the restriction of $\succeq$ to $\mathcal{A} \times \{\delta_i\}$. We say that a function $\hat{V}_i : Y \to \mathbb{R}$ *represents* $\hat{\succeq}_i$ if for all $f, g \in \mathcal{A}$:

$$(f, \{\delta_i\}) \succeq (g, \{\delta_i\}) \Leftrightarrow \hat{V}_i(f(i)) \geq \hat{V}_i(g(i)).$$

We can now state the following Theorem.

**Theorem 1.** *Axioms 1 to 11 hold if, and only if, (a) there exist affine functions $\hat{V}_i : Y \to \mathbb{R}$, $i \in N$, representing $\hat{\succeq}_i$ such that $\hat{V}_i(Y) = \hat{V}_j(Y)$, for all $i, j \in N$ and (b) there exists a function $\mathcal{F} : \mathbb{P} \to \mathbb{P}$ satisfying, for all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$:*

*1. $\mathcal{F}(\mathcal{P}) \subseteq \mathcal{P}$*

*2. For all $\alpha \in [0,1]$, $\mathcal{F}(\alpha \mathcal{P} + (1-\alpha)\mathcal{Q}) = \alpha \mathcal{F}(\mathcal{P}) + (1-\alpha)\mathcal{F}(\mathcal{Q})$*

*for which for all $(f, \mathcal{P}), (g, \mathcal{Q}) \in \mathcal{A} \times \mathbb{P}$,*

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(g(i)).$$

*Furthermore, $\mathcal{F}$ is unique and the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation.*

*Proof.* See the Appendix. $\qquad\square$

To better understand the meaning of Theorem 1, it might be useful to compare and contrast it with Gilboa and Schmeidler's (1989) maxmin expected utility. This theory is stated in the standard Anscombe-Aumann framework, i.e., *without* any explicit information concerning probabilities. They thus only provide a representation theorem for preferences on $\mathcal{A}$. Let $\unrhd$ be a binary relation on $\mathcal{A}$, with $f \unrhd g$ meaning "$g$ is not strictly preferred to $f$". According to the maxmin expected utility model, there exists a unique compact convex set of probability measures

$\mathcal{C} \in \mathbb{P}$ and a linear utility function $V$ (unique up to a strictly increasing affine transformation) such that for all $f, g \in \mathcal{A}$:

$$f \trianglerighteq g \Leftrightarrow \min_{p \in \mathcal{C}} \sum_i p(i) V(f(i)) \geq \min_{p \in \mathcal{C}} \sum_i p(i) V(g(i)).$$

This rule can be interpreted as follows: The decision maker has a *subjective* set of priors $\mathcal{C}$, and evaluates any allocation by computing its expected value with respect to the worst probability distribution in that set. A key feature of that model is that the set of priors $\mathcal{C}$ is totally subjective, and absolutely not related to any objective information. Now, consider Theorem 1. The decision rule characterized in that theorem looks quite similar in spirit to Gilboa and Schmeidler's representation result. However, the set of priors used to compute the minimum of the expected utility is now *explicitly related to the available information*.

Observe, first of all, that because $\mathcal{F}(\mathcal{P}) \subseteq \mathcal{P}$ for all $\mathcal{P} \in \mathbb{P}$, it is the case that $\mathcal{F}(\{p\}) = \{p\}$ for all $p \in \Delta(N)$. Thus $V_i(f(i))$ may be interpreted as the utility obtained by the observer when she is sure to be individual $i$ and the allocation is $f$. In the particular case where the information sets are reduced to singletons, Theorem 1 reduces to:

$$(f, \{p\}) \succeq (g, \{q\}) \Leftrightarrow \sum_i p(i) V_i(f(i)) \geq \sum_i q(i) V_i(g(i)),$$

for all $p, q \in \Delta(N)$ and all $f, g \in \mathcal{A}$. In other words, when the information available to the observer takes the form of a single probability distribution, she evaluates an allocation by computing its expected value with respect to that probability distribution. Now, consider what happens when the observer has to evaluate an allocation with a non-degenerated information set.[15]

First consider the case where $\mathcal{F}(\mathcal{P}) = \mathcal{P}$ for all $\mathcal{P} \in \mathbb{P}$. Such a selection function obviously satisfies all the conditions of Theorem 1, and we get:

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \min_{p \in \mathcal{P}} \sum_i p(i) V_i(f(i)) \geq \min_{p \in \mathcal{Q}} \sum_i p(i) V_i(g(i)),$$

for all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ and $f, g \in \mathcal{A}$. Such a rule can be interpreted as follows: The observer evaluates each allocation by computing its expected value for the worst probability distribution among the possible ones. Of course, such a behavior is quite pessimistic.

Now, consider a Bayesian observer. Facing an information set $\mathcal{P}$, she will reduce it to a single probability distribution, i.e., $\mathcal{F}(\mathcal{P}) \in \Delta(N)$ for all $\mathcal{P} \in \mathbb{P}$. For instance, $\mathcal{F}(\mathcal{P})$ could be the

---

[15]Of course, we are going far beyond what the theorem actually delivers. The theorem only says that everything is "as if" the observer actually thinks along the lines we describe here.

Steiner point (which is a generalization of the notion of "center" for arbitrary convex sets) of $\mathcal{P}$, which will be denoted as $St(\mathcal{P})$.[16] Theorem 1 will then give us:

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \sum_i St(\mathcal{P})(i)V_i(f(i)) \geq q \sum_i St(\mathcal{Q})(i)V_i(g(i)),$$

for all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ and $f, g \in \mathcal{A}$. Thus, the observer evaluates an allocation by computing its expected value with respect to the Steiner point of the information set. Observe that, in that case, the observer is neutral towards uncertainty: for all $\mathcal{P} \in \mathbb{P}$, $f, g \in \mathcal{A}$, and all $\alpha \in (0, 1)$, $(f, \mathcal{P}) \sim (g, \mathcal{P})$ implies $(\alpha f + (1 - \alpha)g, \mathcal{P}) \sim (f, \mathcal{P})$.

Finally, another natural candidate for $\mathcal{F}$ is a combination of the two selection functions given above, namely: $\mathcal{F}(\mathcal{P}) = \theta \mathcal{P} + (1 - \theta)St(\mathcal{P})$. This is precisely the selection function we characterize in what follows, on a restricted domain. Before stating our next result, we need some additional notation. For any $S \subseteq N$, let $\Delta(S)$ be the set of all probability distributions with support in $S$. Let:

$$\mathbb{B} = \{\mathcal{P} \in \mathbb{P} | \exists (\alpha_t)_{t=1,\ldots,r} \in [0, 1] \text{ with } \sum_t \alpha_t = 1 \text{ and } (S_t)_{t=1,\ldots,r} \subseteq N \text{ s.t. } \mathcal{P} = \sum_t \alpha_t \Delta(S_t)\}.$$

Hence, an element of $\mathbb{B}$ is a convex combination of simplices with supports in $N$. Furthermore, for $\mathcal{P} = \sum_t \alpha_t \Delta(S_t)$, let $c(\mathcal{P}) = \sum_t \alpha_t c(\Delta(S_t))$, where $c(\Delta(S_t))$ is the probability distribution defined by $c(\Delta(S_t))(s) = \frac{1}{|S_t|}$ (hence, $c(\Delta(S_t))$ is the uniform distribution on $S_t$).[17] Actually, it is easily shown that for all $\mathcal{P} \in \mathbb{B}$, $c(\mathcal{P}) = St(\mathcal{P})$, and that moreover $c(\mathcal{P})$ coincides with the Shapley value of the cooperative game whose core is $\mathcal{P}$.

In what follows, we will consider the restriction of $\succeq$ on $\mathcal{A} \times \mathbb{B}$. We propose two additional axioms on that restricted domain. The first may be interpreted as an anonymity requirement. The idea is that the only things that matter for the observer, besides the information set, are the utility *levels* he obtains, conditionally on being each individual. For any $f \in \mathcal{A}$ and any permutation $\varphi : N \to N$, define $\mathcal{A}(f^\varphi) = \{g \in \mathcal{A} \,|\, (g, \{\delta_i\}) \sim (f, \{\delta_{\varphi^{-1}(i)}\}), \forall i \in n\}$. Hence, for all $g \in \mathcal{A}(f^\varphi)$, individual $i$ gets the same utility level as individual $\varphi^{-1}(i)$ with the allocation $f$. Now, for any $\mathcal{P} \in \mathbb{B}$, let $\mathcal{P}^\varphi$ be defined by $\mathcal{P}^\varphi = \{p^\varphi | p \in \mathcal{P}\}$, where for all $p \in \mathcal{P}$, $p^\varphi(i) = p(\varphi^{-1}(i))$ for all $i \in N$. Essentially, $(f, \mathcal{P})$ and $(g, \mathcal{P}^\varphi)$ (with $\varphi$ a permutation and $g \in \mathcal{A}(f^\varphi)$) only differ by the names of the states. The following axiom states that the observer is indifferent between them.

---

[16]to be precise, let $e = (\frac{1}{|N|}, \cdots, \frac{1}{|N|})$ and $V = \{v \in \mathbb{R}^{|N|} : \langle v, e \rangle = 0, \|v\| = 1\}$ be the $|N| - 2$ dimensional unit sphere orthogonal to $e$. For $P \in \mathbb{P}$, its Steiner point is defined by $St(P) = \int_V \arg\max_{p \in P} \langle p, v \rangle \nu(dv)$, where $\nu$ is the uniform distribution over $V$. It satisfies the condition of Theorem 1. The Steiner point was introduced in decision theory by Gajdos, Hayashi, Tallon and Vergnaud (2007).

[17]The set $\mathbb{B}$ is known in decision theory as the set of cores of belief functions. For more details on beliefs functions see, e.g., Dempster (1967), Shafer (1976) and Jaffray (1989).

**Axiom 12** (Anonymity)**.** *For all $(f, \mathcal{P}) \in \mathcal{A} \times \mathbb{B}$, all permutations $\varphi : N \to N$ and all $g \in \mathcal{A}(f^\varphi)$, $(f, \mathcal{P}) \sim (g, \mathcal{P}^\varphi)$.*

The next axiom states that whenever two allocations share the same worst probability distribution in $\mathcal{P}$, a mixture of these allocations will not reduce the degree of uncertainty and therefore will not lead to an improvement, whenever $\mathcal{P}$ is the available information set. Recall that in Axiom 11, we interpreted preference for hedging (or mixture) as uncertainty aversion. With this in mind, indifference for hedging can be interpreted as *neutrality* towards uncertainty. The following axiom can therefore be viewed as a restricted neutrality towards uncertainty.[18]

**Axiom 13** (Restricted Mixture Neutrality)**.** *For all $\mathcal{P} \in \mathbb{B}$ and all $f, g \in \mathcal{A}$, if there exists $p^* \in \mathcal{P}$ such that $(f, \{p\}) \succeq (f, \{p^*\})$ and $(g, \{p\}) \succeq (g, \{p^*\})$ for all $p \in \mathcal{P}$, then for all $\alpha \in [0, 1]$,*

$$(f, \mathcal{P}) \sim (g, \mathcal{P}) \Leftrightarrow (\alpha f + (1 - \alpha)g, \mathcal{P}) \sim (g, \mathcal{P}).$$

We then obtain the following representation theorem when the information belongs to $\mathbb{B}$.[19]

**Theorem 2.** *Under the assumptions of Theorem 1 restricted to $\mathcal{A} \times \mathbb{B}$, Axioms 12 and 13 hold if, and only if, (a) there exist affine functions $\hat{V}_i : Y \to \mathbb{R}$, $i \in N$, representing $\hat{\succeq}_i$ such that $\hat{V}_i(Y) = \hat{V}_j(Y)$, for all $i, j \in N$ and (b) there exists $\theta \in [0, 1]$ such that for all $\mathcal{P}, \mathcal{Q} \in \mathbb{B}$ and all $f, g \in \mathcal{A}$,*

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \theta \min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f(i)) + (1 - \theta) \sum_i c(\mathcal{P})(i)\hat{V}_i(f(i))$$

$$\geq \theta \min_{p \in \mathcal{Q}} \sum_i p(i)\hat{V}_i(g(i)) + (1 - \theta) \sum_i c(\mathcal{Q})(i)\hat{V}_i(g(i)).$$

*Furthermore, $\theta$ is unique and the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation.*

*Proof.* See the Appendix. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Theorem 2 is a special case of Theorem 1, where the subjective set of priors is obtained by (*i*) solving for the "mean value" $c(\mathcal{P})$ of the information set and (*ii*) shrinking the information

---

[18]Mixture neutrality can be considered as the analogue of the *betweenness* property for preferences under risk (on the betweenness property, see, e.g., Chew (1983, 1993) and Dekel (1986)). Attitudes towards mixtures are considered in detail in Klibanoff (2001).

[19]Actually, the following representation theorem, which is only given when the information takes the form of a core of a belief function, can be extended to convex combinations of symmetric polytopes. However, for the problem under consideration, such an extension would not be a great improvement.

set toward the mean value according to a degree $(1-\theta)$ given by the decision maker's preference. The set of priors is then $\mathcal{F}(\mathcal{P}) = \theta\mathcal{P} + (1-\theta)\{c(\mathcal{P})\}$. Because

$$\min_{p \in \theta\mathcal{P}+(1-\theta)\{c(\mathcal{P})\}} \sum_{i \in N} p(i)\hat{V}_i(f(i)) = \theta \min_{p \in \mathcal{P}} \sum_{i \in N} p(i)\hat{V}_i(f(i)) + (1-\theta) \sum_{i \in N} c(\mathcal{P})(i)\hat{V}_i(f(i)),$$

one obtains the decision rule that appears in Theorem 2. Observe that the case $\theta = 0$ corresponds to a Bayesian decision maker who uses exclusively the "mean value" $c(\mathcal{P})$, whereas the case $\theta = 1$ corresponds to an extremely pessimistic decision maker, who keeps the whole set of information.[20]

## 4   The Ignorant Observer Theorem

We now define individual preferences $\succeq_i$ on $Y$. Following Harsanyi, we assume that individuals obey the von Neumann and Morgenstern axioms. It should be noted that this assumption is by no means in conflict with Rawls' views. Indeed, Rawls only rejected the use of the Bayesian doctrine from behind the veil of ignorance. Formally, these axioms are as follows.

**Axiom 14** (Ordering). $\succeq_i$ *is a reflexive, complete, transitive and nondegenerate binary relation on $Y$.*

**Axiom 15** (Continuity). *For all $w, y, z \in Y$ such that $w \succ_i y \succ_i z$, there exists an $\alpha \in (0,1)$ such that $\alpha w + (1-\alpha)z \sim_i y$.*

**Axiom 16** (Independence). *For all $w, y, z \in Y$ and all $\alpha \in [0,1]$, $y \succeq_i z$ implies $\alpha y + (1-\alpha)w \succeq_i \alpha z + (1-\alpha)w$.*

As is well known (see, e.g., Fishburn, 1970), a preference relation satisfies these three axioms if, and only if, it can be represented by an Expected Utility functional. This is stated formally in the following theorem.

**Theorem 3.** *Axiom 14 to 16 hold if, and only if, there exists an affine real-valued function $U_i$ on $Y$ such that for all $y, z \in Y$, $y \succeq_i z$ if, and only if, $U_i(y) \geq U_i(z)$. Furthermore, such a representation is unique up to a positive affine transformation.*

Our aim is to deduce from the individual preferences $\succeq_i$ on $Y$ and the observer's preferences $\succeq$ on $\mathcal{A} \times \mathbb{P}$, a "social preference" $\succeq^*$ on $Y$. In order to do so, one needs to specify how these preferences interact.

---

[20] A similar decision rule has been axiomatized in Gajdos, Hayashi, Tallon and Vergnaud (2007) on the domain $\mathcal{A} \times \mathbb{P}$. The axioms used there are very different from the ones we use (they involve a strong notion of invariance with respect to some joint transformations acts and information and continuity with respect to the information). In particular, they do not involve any restriction concerning attitude towards mixture of acts. Finally, it was stated in a state-independent framework.

The preferences of the individuals and the observer are linked by the so-called "acceptance principle" (see Harsanyi, 1977) which states that if the observer is *sure* to be $i$, her choices should be the same as those of $i$. The acceptance axiom can be restated as follows in our framework.

**Axiom 17** (Acceptance). *For all $i \in N$ and all $f, g \in \mathcal{A}$, $(f, \{\delta_i\}) \succeq (g, \{\delta_i\})$ if and only if $f(i) \succeq_i g(i)$.*

We now turn to the link between social preferences and the observer's preferences. The fundamental idea of the veil of ignorance is that (fair) social preferences are those of an observer who is *totally ignorant* about the position he will eventually get in the society. There are two issues here. First, social preferences are defined over social alternatives in $Y$, whereas the observer's preferences are defined on the product $\mathcal{A} \times \mathbb{P}$ of allocations and information sets. For all $y \in Y$, let $k_y \in \mathcal{A}$ be the allocation defined by $k_y(i) = y$ for all $i \in N$. It is natural to define the observer's preferences $\succeq_Y$ on $Y \times \mathbb{P}$ as follows: $(y, \mathcal{P}) \succeq_Y (z, \mathcal{Q})$ iff $(k_y, \mathcal{P}) \succeq (k_z, \mathcal{Q})$. Indeed, consider the social lottery $y$ together with the information set $\mathcal{P}$. One may interpret this couple as a two stages process. First, there is a probability distribution in $\mathcal{P}$ according to which his identity will be chosen. Then, she will face the social alternative $y$, *whatever* her identity is. But this is precisely how we interpret $(k_y, \mathcal{P})$.

Now, we should formalize the idea that the observer is totally ignorant about his position. In our framework, this is captured by the fact that the information set is $\Delta(N)$, the set of all probability distributions over the set $N$ of individuals. This leads us to the following axiom.

**Axiom 18** (Ignorance). *For all $y, z \in Y$, $y \succeq^* z$ if, and only if, $(k_y, \Delta(N)) \succeq (k_z, \Delta(N))$.*

We can now state a first representation Theorem.

**Theorem 4.** *Assume that the observer's preferences satisfy all of the axioms of Theorem 1 and that the individual preferences satisfy all of the axioms of Theorem 3. Take $\mathcal{F} : \mathbb{P} \to \mathbb{P}$ from Theorem 1 and affine functions $V_i : Y \to \mathbb{R}$, $i \in N$, from Theorem 3. Then for all $y, z \in Y$,*

$$y \succeq^* z \Leftrightarrow \min_{p \in \mathcal{F}(\Delta(N))} \sum_i p(i) \frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}$$

$$\geq \min_{p \in \mathcal{F}(\Delta(N))} \sum_i p(i) \frac{V_i(z) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}.$$

*Proof.* This theorem is a straightforward corollary of Theorem 1. Axiom 17 implies $\succeq_i = \hat{\succeq}_i$. Let $V_i$ be an affine representation of $\succeq_i$. By Theorem 3, $V_i$ is unique up to a positive affine transformation. Therefore, $V_i^*(y) = \frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}$ is also an affine representation of $\succeq_i$. Furthermore, $V_i^*(Y) = V_j^*(Y)$ for all $i, j \in N$. Define $\tilde{V}_i(f) = V_i^*(f(i))$ for all $f \in \mathcal{A}$.

Clearly, $\tilde{V}_i$ are affine representations of $\hat{\succeq}_i$ and satisfy $\tilde{V}_i(Y) = \tilde{V}_j(Y)$ for all $i, j \in N$. The result then follows from Theorem 1. $\qquad\square$

Observe that the exact form of $\mathcal{F}(\Delta(N))$ will actually depend on the decision maker's *uncertainty aversion*. A more precise form may be obtained if we use Axioms 12 and 13. Indeed, Theorem 2 leads to the following result.

**Theorem 5.** *Assume that the observer's preferences satisfy all of the axioms of Theorem 2 and that the individual preferences satisfy all of the axioms of Theorem 3. Take affine functions $V_i$, $i \in N$, from Theorem 3. Then Axioms 17 and 18 hold if, and only if for $y, z \in Y$,*

$$y \succeq^* z \Leftrightarrow \theta \min_{i \in N} \frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)} + (1 - \theta) \sum_{i \in N} \frac{1}{n} \frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}$$

$$\geq \theta \min_{i \in N} \frac{V_i(z) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)} + (1 - \theta) \sum_{i \in N} \frac{1}{n} \frac{V_i(z) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}.$$

*Furthermore, $\theta$ is unique.*

*Proof.* Let $V_i^*$ be defined as in the proof of Theorem 4. Reasoning as in the proof of Theorem 4, Theorem 2 implies that there exist affine functions $V_i : Y \to \mathbb{R}$, $i \in N$, representing $\succeq_i$ such that for all $y, z \in Y$,

$$y \succeq^* z \Leftrightarrow \theta \min_{p \in \Delta(N)} \sum_i p(i) V_i^*(y) + (1 - \theta) \sum_i \frac{1}{n} V_i^*(y)$$

$$\geq \theta \min_{p \in \Delta(N)} \sum_i p(i) V_i^*(z) + (1 - \theta) \sum_i \frac{1}{n} V_i^*(z).$$

Theorem 5 follows by noting that $\min_{p \in \Delta(N)} \sum_i p(i) V_i^*(x) = \min_{i \in N} V_i^*(x)$ for all $x \in Y$. $\qquad\square$

Such a criterion can be easily interpreted: it is a weighted average of Harsanyi's utilitarian criterion and Rawls' maxmin criterion. It should be noted that Axioms 12 and 13 are used to obtain this specific functional form. Axiom 12 plays a transparent role: it ensures that the symmetry of the set of probability distributions that represents the available information will be preserved in the decision rule. Axiom 13 forces the set of probability distributions used in the decision rule to have a similar shape as the set of probability distributions that represents the available information.

Let us note that, in both Theorems 4 and 5, the individual utility functions $\frac{V_i(y) - \min_{w \in Y} V_i(w)}{\max_{w \in Y} V_i(w) - \min_{w \in Y} V_i(w)}$ are cardinally measurable and fully comparable. Therefore, the weights assigned to these functions cannot be manipulated and are meaningful. As mentioned above, the normalization of individual utility functions that we have obtained is very similar to that obtained by Karni

(1998), Dhillon and Mertens (1999), Segal (2000) and Moreno-Ternero and Roemer (2005).[21] In all of these articles, the weight for each individual depends on the diameter of the range of her utility function.[22] Because Karni (1998) is the paper to which we are the closest, let us emphasize a significant difference between Karni's approach and the one that we have followed here. Karni assumes that the observer's preferences are defined on extended lotteries, in Harsanyi's sense.[23] This assumption implies that he faces the well-known problem of the determination of the weights the observer attaches to each individual. This is where his "Impartiality Axiom" plays a key role. Our approach is rather different. Because the weights that appear in our representation theorem are based on objective information (i.e., because probability distributions on individual lotteries are part of the model), we do not face the problem of their determination. Hence, the normalization of individual utilities can be seen as coming from a strictly epistemic axiom (namely, Axiom 6), which should not be interpreted in terms of impartiality. Our impartiality requirement actually lies in the nature of the information the observer can use to make her decisions.

## 5   Arguments for a compromise

According to our Ignorant Observer theorem, there is not one, but a plurality of decision rules complying with the impartiality requirement (as formalized by the veil of ignorance). These include both Harsanyi's utilitarian criterion and Rawls' maximin principle.

In particular, given the assumptions of Theorem 4, the observer would be Rawlsian if, and only if, she obeys the following axiom.

**Axiom 19** (Extreme aversion towards uncertainty). *For all $f \in \mathcal{A}$, $\mathcal{P} \in \mathbb{P}$ and $p \in \mathcal{P}$, $(f, \{p\}) \succeq (f, \mathcal{P})$.*

This axiom requires the observer to (weakly) prefer *any* lottery on the set of individuals to complete ignorance. In particular, she must (weakly) prefer to be the worst-off individual for sure, rather than facing complete ignorance concerning her identity. Such a requirement seems very unlikely from a decision theoretic point of view.

On the other hand, assuming that the ignorant observer is an expected utility maximizer, one would obviously obtain the utilitarian rule.[24] This requirement would take, in our framework, the following form, which is a strengthening of Axiom 11:

---

[21] Although they consider a very restrictive case (i.e., when individual preferences are risk-isomorphic).

[22] See also Karni (2003) for a related approach.

[23] He uses the Anscombe and Aumann (1963) formalism.

[24] Again, in the preference utilitarian sense, as in Harsanyi's work, and not as it was understood by classical utilitarians

**Axiom 20** (Neutrality towards uncertainty). *For all $f, g \in \mathcal{A}$ and all $\mathcal{P} \in \mathbb{P}$, $(f, \mathcal{P}) \sim (g, \mathcal{P})$ implies, for all $\alpha \in (0, 1)$, $(\alpha f + (1 - \alpha)g, \mathcal{P}) \sim (f, \mathcal{P})$.*

This axiom is relatively easy to interpret from an ethical point of view, because it is directly related to a kind of neutrality towards inequalities. To grasp the intuition behind the axiom, let us consider the following simple example.

Assume that the society is composed of two individuals, 1 and 2 and let $\mathcal{P} = \Delta(\{1, 2\})$. Let $f$ and $g$ be two allocations whose outcomes (in terms of utilities) are as follows:

|  | 1 | 2 |
|---|---|---|
| $\hat{V}_i(f(i))$ | 1 | 0 |
| $\hat{V}_i(g(i))$ | 0 | 1 |
| $\hat{V}_i(h(i))$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

Assume that the ignorant observer is indifferent between $(f, \Delta(\{1, 2\}))$ and $(g, \Delta(\{1, 2\}))$ (this would be the case, in particular, if Axiom 12 holds). Observe that $f$ and $g$ are highly unequal. Now, consider $h$, which is simply defined as $\frac{1}{2}f + \frac{1}{2}g$. Obviously, $h$ is less unequal than $f$ and $g$. It makes sense, therefore, to assume that the ignorant observer will prefer $h$ to both $f$ and $g$. However, Axiom 20 forces the observer to be *indifferent* between $f$ and $h$. Actually, this is the main point made in Diamond's (1967) famous objection to Harsanyi.[25]

Axiom 11 allows us to accommodate for fairness, and thus to escape from Diamond's critique, since we relax the independence axiom at the observer's level. A similar argument applied to Harsanyi's Aggregation Theorem can be found in Epstein and Segal (1992). Observe that there are other ways to relax the independence axiom, and thus to answer Diamond's critique. For instance, Grant, Kajii, Polak and Safra (2006) assume that the Impartial Observer's preferences are defined over $\Delta(N) \times \Delta(X)$ (instead of $\Delta(N \times X)$ as in Harsanyi (1953, 1977). This allows them to assume independence over outcome lotteries (i.e., elements of $\Delta(X)$) and independence with respect to identity lotteries (i.e., elements of $\Delta(N)$) without assuming that a randomization between individual lotteries is equivalent from the observer's point of view to randomization between outcome lotteries. In that framework, they prove that if the observer actually prefers randomization between outcome lotteries to randomization over identity lotteries, then her preferences can be represented by a "generalized utilitarian representation" of the form $V(y, p) = \sum_i p(i)\phi_i(U_i(y))$, where $(y, p) \in \Delta(X) \times \Delta(N)$, and the $\phi_i$ functions are concave. We admit that Grant, Kajii, Polak and Safra's answer to Diamond is deeper than ours, insofar as they show that it is not really the independence assumption that is at the heart of Diamond's critique, but the fact that lotteries over outcomes and lotteries over identities are considered

---

[25]Actually, Diamond's critique concerned Harsanyi's (1955) Social Aggregation Theorem. However, it readily translates into the Impartial Observer framework.

as equivalent. However, our aim was *not* to improve Harsanyi's model so as to answer Diamond's critique, but to provide a framework in which Rawls' and Harsanyi's arguments could be compared and evaluated (because they represent ignorance by individual lotteries, Grant, Kajii, Polak and Safra's (2006) approach cannot be related to Rawls' veil of ignorance).

However, we would like, here, to give some arguments in favor of the compromise suggested by Theorem 5. Once one is convinced that both Rawls' and Harsanyi's criteria are unappealing, adoption of the Axioms of Theorem 5 would indeed lead to a strict compromise between Rawls' egalitarianism and Harsanyi's utilitarianism.

The two key axioms that allow us to obtain Theorem 5 are Axioms 12 and 13. The first one is a standard anonymity assumption that simply states that individuals' names are irrelevant when one compares two allocations. Axiom 13 is less usual and can actually be viewed as a restriction of Axiom 20 that escapes Diamond's critique. Indeed, consider again Diamond's example. It is clearly not the case that $f$ and $g$ share the same worst probability distribution in $\Delta(\{1,2\})$, since the worst case if $f$ is chosen is to be individual 2 for sure, whereas the worst case if $g$ is chosen is to be individual 1 for sure. Therefore, one can not conclude that $h \sim f$. Now, consider the following variation of Diamond's example, with 3 individuals.

|  | 1 | 2 | 3 |
|---|---|---|---|
| $\hat{V}_i(f_1(i))$ | 1 | 0 | 0 |
| $\hat{V}_i(g_1(i))$ | 1 | 1 | 0 |
| $\hat{V}_i(h_1(i))$ | 1 | $\frac{1}{2}$ | 0 |

Assume that the observer is indifferent between $(f_1, \Delta(\{1,2,3\}))$ and $(g_1, \Delta(\{1,2,3\}))$. Since $\delta_3$ is the worst probability distribution for both $f_1$ and $g_1$, Axiom 13 leads to the conclusion that $(h_1, \Delta(\{1,2,3\})) \sim (f_1, \Delta(\{1,2,3\})) \sim (g_1, \Delta(\{1,2,3\}))$. Is this reasonable? Yes, insofar as the observer's indifference between $(f_1, \Delta(\{1,2,3\}))$ and $(g_1, \Delta(\{1,2,3\}))$ indicates that he cares a lot about the worst-off individual because this individual is the same whatever allocation is chosen (namely, 3) and has always the same utility.

Naturally, Axiom 13 is by no means indisputable. In particular, one may exhibit a particular version of Diamond's critique. Consider the following example, with the same notation as above:

|  | 1 | 2 | 3 |
|---|---|---|---|
| $\hat{V}_i(f_2(i))$ | 1 | 0 | 0 |
| $\hat{V}_i(g_2(i))$ | 0 | 1 | 0 |
| $\hat{V}_i(h_2(i))$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 |

Assume, again, that $(f_2, \Delta(\{1,2,3\})) \sim (g_2, \Delta(\{1,2,3\}))$. Since $\delta_3$ is the worst probability distribution for both $f$ and $g$, Axiom 13 leads to the conclusion that $(h_2, \Delta(\{1,2,3\})) \sim (f_2, \Delta(\{1,2,3\})) \sim (g_2, \Delta(\{1,2,3\}))$. But observe that if one considers only individuals 1 and

2, $h_2$ is strictly more equal than $f_2$ and $g_2$, whereas individual 3 gets the same outcome in $f_2$, $g_2$ and $h_2$. Thus one may consider that the observer should prefer $(h_2, \Delta(\{1, 2, 3\}))$ over $(f_2, \Delta(\{1, 2, 3\}))$. In other words, Axiom 13 does not allow compensating the situation of the worst-off individual by a decrease of the inequality among better-off individuals (only a Pareto-improvement can lead to such a compensation). Thus, in this case, we are essentially back to Diamond's critique. Thus, Axiom 13 might be seen as an axiom restricting the set of situations where Diamond's critique applies (as compared to Harsanyi's utilitarianism), without totally eliminating them.

This leads us to believe that the allocation rule proposed in Theorem 5 (with $\theta \in (0, 1)$) is a reasonable one.[26] Of course, this rule is in sharp conflict with Harsanyi's view, since it is not compatible with utilitarianism as soon as $\theta > 0$. On the other hand, it is not certain that Rawls would have been displeased with it. Although after *A Theory of Justice* was published, he wrote some articles to defend maximin (see, e.g., Rawls, 1974a, 1974b), his arguments were often conflicting and even contradictory with his own interpretation of the veil of ignorance. We take these contradictions as evidence that his main purpose was not to defend a specific criterion of rational decision under ignorance. Since his project was primarily to propose a theory of social justice alternative to Utilitarianism, his main purpose with the veil of ignorance model was to acknowledge a solution excluding Utilitarianism. As himself wrote:

> "But I do not wish to overemphasize this criterion: a deeper investigation (...) may show that some other conception of justice is more reasonable."

> Rawls, (1974a, p. 145)

---

[26]Yet, since as we have seen, Axiom 13 is not indisputable, we do not claim that this rule is the only one that should be considered. One may want to start from the more general representation of Theorem 4 and impose other axioms that would lead to different rules. We leave this investigation for future research.

# Appendix

## Proof of Theorem 1

The necessity part of the Theorem is easily checked. We therefore only prove the sufficiency part. The proof goes through several claims. Although not explicitly stated in the claims, all of the assumptions of Theorem 1 are made throughout this subsection.

**Claim 1.** $\mathcal{A}^{cv}$ *is convex.*

*Proof.* Let $f, g \in \mathcal{A}^{cv}$ and $\alpha \in [0, 1]$. By the definition of $\mathcal{A}^{cv}$, $(f, \{\delta_i\}) \sim (f, \{\delta_j\})$ for all $i, j \in N$. Therefore, by Axiom 7, $(\alpha f + (1 - \alpha)g, \{\delta_i\}) \sim (\alpha f + (1 - \alpha)g, \{\delta_j\})$ for all $i, j \in N$. Hence, $\alpha f + (1 - \alpha)g \in \mathcal{A}^{cv}$, which proves that $\mathcal{A}^{cv}$ is convex. □

**Claim 2.** *For all $i \in N$, all $f, g, h \in \mathcal{A}$ and all $\alpha \in (0, 1)$,*

$$(f, \{\delta_i\}) \succeq (g, \{\delta_i\}) \Leftrightarrow (\alpha f + (1 - \alpha)h, \{\delta_i\}) \succeq (\alpha g + (1 - \alpha)h, \{\delta_i\}).$$

*Proof.* Let $f, g, h \in \mathcal{A}$. By Axiom 6, there exist $\bar{h}, \underline{h} \in \mathcal{A}^{cv}$ such that $(\bar{h}, \{\delta_i\}) \succeq (h, \{\delta_i\}) \succeq (\underline{h}, \{\delta_i\})$. Hence, by Axiom 3, there exists $\theta \in [0, 1]$ such that $(h, \{\delta_i\}) \sim (\theta \bar{h} + (1 - \theta)\underline{h}, \{\delta_i\})$. Let $\hat{h} = \theta \bar{h} + (1 - \theta)\underline{h}$. By Claim 1, $\hat{h} \in \mathcal{A}^{cv}$. Next, let $\tilde{h}$ be defined by $\tilde{h}(i) = h(i)$ and $\tilde{h}(j) = \hat{h}(j)$ for all $j \neq i$. By Axiom 10, we then have $(\tilde{h}, \{\delta_i\}) \sim (h, \{\delta_i\}) \sim (\hat{h}, \{\delta_i\})$. Furthermore, for all $j \neq i$, $(\tilde{h}, \{\delta_j\}) \sim (\tilde{h}, \{\delta_i\})$ by construction. Therefore, $\tilde{h} \in \mathcal{A}^{cv}$. By Axiom 7, for all $\alpha \in (0, 1)$,

$$(f, \{\delta_i\}) \succeq (g, \{\delta_i\}) \Leftrightarrow (\alpha f + (1 - \alpha)\tilde{h}, \{\delta_i\}) \succeq (\alpha g + (1 - \alpha)\tilde{h}, \{\delta_i\}).$$

But, by Axiom 10, $(\alpha f + (1 - \alpha)\tilde{h}, \{\delta_i\}) \sim (\alpha f + (1 - \alpha)h, \{\delta_i\})$ and $(\alpha g + (1 - \alpha)\tilde{h}, \{\delta_i\}) \sim (\alpha g + (1 - \alpha)h, \{\delta_i\})$. Therefore,

$$(f, \{\delta_i\}) \succeq (g, \{\delta_i\}) \Leftrightarrow (\alpha f + (1 - \alpha)h, \{\delta_i\}) \succeq (\alpha g + (1 - \alpha)h, \{\delta_i\}),$$

the desired result. □

**Claim 3.** *For all $i$, there exits an affine function $\hat{V}_i : Y \to \mathbb{R}$ such that for all $f, g \in \mathcal{A}$,*

$$(f, \{\delta_i\}) \succeq (g, \{\delta_i\}) \Leftrightarrow \hat{V}_i(f(i)) \geq \hat{V}_i(g(i)).$$

*Furthermore, $\hat{V}_i$ is unique up to a positive affine transformation.*

*Proof.* Let $i$ be fixed in $N$ and $\hat{\succeq}_i$ be the restriction of $\succeq$ to $\mathcal{A} \times \{\delta_i\}$. By Axioms 1, 3 and Claim 2, $\hat{\succeq}_i$ satisfies the von Neumann-Morgenstern axioms (observe that by Axiom 2 the restriction

of $\hat{\succeq}_i$ is nondegenerate). Therefore, there exists an affine function $U_i : \mathcal{A} \to \mathbb{R}$, unique up to a positive affine transformation, such that for all $f, g \in \mathcal{A}$,

$$(f, \{\delta_i\}) \succeq (g, \{\delta_i\}) \Leftrightarrow U_i(f) \geq U_i(g).$$

By Axiom 10, for all $f, f' \in \mathcal{A}$ such that $f(i) = f'(i)$, $(f, \{\delta_i\}) \sim (f', \{\delta_i\})$. Therefore, defining $\hat{V}_i : Y \to \mathbb{R}$ by $\hat{V}_i(f(i)) = U_i(f')$ for all $f' \in \mathcal{A}$ such that $f'(i) = f(i)$, one obtains the desired result. $\qquad\square$

In the sequel, we will make the following slight abuse of notation: we will denote by $\hat{V}_i$ both the function $\hat{V}_i : Y \to \mathbb{R}$ defined in Claim 3 and the function $\tilde{V}_i : \mathcal{A} \to \mathbb{R}$ defined by $\tilde{V}_i(f) = \hat{V}_i(f(i))$.

**Claim 4.** *For all $\mathcal{P} \in \mathbb{P}$, $f, g \in \mathcal{A}$,*

$$\begin{cases} (f, \{\delta_i\}) \succ (g, \{\delta_i\}), \, \forall i \in N \Rightarrow (f, \mathcal{P}) \succ (g, \mathcal{P}) \\ (f, \{\delta_i\}) \succeq (g, \{\delta_i\}), \, \forall i \in N \Rightarrow (f, \mathcal{P}) \succeq (g, \mathcal{P}). \end{cases}$$

*Proof.* Let $f, g \in \mathcal{A}$ be such that $(f, \{\delta_i\}) \succ (g, \{\delta_i\})$ for all $i \in N$. By Axiom 4, for all $(\alpha_1, \ldots, \alpha_n)$ such that $\alpha_i \geq 0$ for all $i$ and $\sum_{i \in N} \alpha_i = 1$, we have $(f, \sum_{i \in N} \alpha_i \{\delta_i\}) \succ (g, \sum_{i \in N} \alpha_i \{\delta_i\})$. Therefore, for all $p \in \Delta(N)$, $(f, \{p\}) \succ (g, \{p\})$. Hence, by Axiom 5, for all $\mathcal{P} \in \mathbb{P}$, $(f, \mathcal{P}) \succ (g, \mathcal{P})$. The second part of the Claim is proved using the same argument. $\qquad\square$

**Claim 5.** *There exist $f$ and $g$ in $\mathcal{A}^{cv}$ such that for all $i \in N$, $(f, \{\delta_i\}) \succ (g, \{\delta_i\})$.*

*Proof.* Let $\mathcal{P} \in \mathbb{P}$ be fixed. By Axiom 2, there exists $\hat{f}, \hat{g} \in \mathcal{A}$ such that $(\hat{f}, \mathcal{P}) \succ (\hat{g}, \mathcal{P})$. By Axiom 6, there exist $f, g \in \mathcal{A}^{cv}$ such that $(f, \mathcal{P}) \succeq (\hat{f}, \mathcal{P})$ and $(\hat{g}, \mathcal{P}) \succeq (g, \mathcal{P})$. Therefore, $(f, \mathcal{P}) \succ (g, \mathcal{P})$. Since $f$ and $g$ belong to $\mathcal{A}^{cv}$, Axiom 9 implies that $(f, \{\delta_i\}) \succ (g, \{\delta_i\})$ for all $i \in N$. $\qquad\square$

We say that a function $V : \mathcal{A} \to \mathbb{R}$ is $\mathcal{A}^{cv}$-*affine* iff for all $f \in \mathcal{A}$, $g \in \mathcal{A}^{cv}$ and $\alpha \in [0, 1]$, $V(\alpha f + (1 - \alpha)g) = \alpha V(f) + (1 - \alpha)V(g)$.

**Claim 6.** *For all $\mathcal{P} \in \mathbb{P}$, there exists an $\mathcal{A}^{cv}-$affine functional $V_{\mathcal{P}} : \mathcal{A} \to \mathbb{R}$ such that for all $f, g \in \mathcal{A}$:*

$$(f, \mathcal{P}) \succeq (g, \mathcal{P}) \Leftrightarrow V_{\mathcal{P}}(f) \geq V_{\mathcal{P}}(g).$$

*Furthermore, $V_{\mathcal{P}}$ is unique up to a positive affine transformation and $V_{\mathcal{P}}(\mathcal{A}) = V_{\mathcal{P}}(\mathcal{A}^{cv})$.*

*Proof.* This result follows from Claim 1, Axioms 1, 3, 6, 7 and Corollary 2 in Castagnoli, Marinacci and Maccherroni (2003). $\qquad\square$

For all $h \in \mathcal{A} \setminus \mathcal{A}^{cv}$, let $\mathcal{A}_h = co\{h, \mathcal{A}^{cv}\}$.

**Claim 7.** *There exist affine representations $\hat{V}_i$ ($i \in N$) of $\succeq$ on $\mathcal{A} \times \{\delta_i\}$ satisfying $\hat{V}_i(\mathcal{A}) = \hat{V}_j(\mathcal{A})$, for all $i, j \in N$, such that, for all $\mathcal{P} \in \mathbb{P}$, $h \in \mathcal{A} \setminus \mathcal{A}^{cv}$, there exists an $\mathcal{A}^{cv}-$affine representation $V_\mathcal{P}$ of $\succeq$ on $\mathcal{A} \times \{\mathcal{P}\}$ and non-negative numbers $\lambda_1(h, \mathcal{P}), \ldots, \lambda_n(h, \mathcal{P})$, not all equal to zero and summing up to one, such that, for all $f \in \mathcal{A}_h$,*

$$V_\mathcal{P}(f) = \sum_{i \in N} \lambda_i(h, \mathcal{P}) \hat{V}_i(f).$$

*Moreover, the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation and coincide on $\mathcal{A}^{cv}$.*

*Proof.* For all $i$, let $f_i^*$ be such that $f_i^* \hat{\succeq}_i f$ for all $f \in \mathcal{A}$ and $f_{*i}$ be such that $f \hat{\succeq}_i f_{*i}$ for all $f \in \mathcal{A}$. These allocations are well defined since $\mathcal{A}$ is a compact set. Define $f^*$ and $f_*$ by $f^*(i) = f_i^*(i)$ and $f_*(i) = f_{*i}(i)$ for all $i \in N$.

By Axiom 10, $f^* \hat{\sim}_i f_i^*$ and $f_* \hat{\sim}_i f_{*i}$ for all $i \in N$. Therefore, by Claim 4, $f^* \hat{\succeq}_i f \hat{\succeq}_i f_*$ for all $f \in \mathcal{A}$ and all $i \in N$. Let $\hat{V}_i$ ($i \in N$) be affine representations of $\hat{\succeq}_i$, as defined in Claim 3 and $V_\mathcal{P}$ be an $\mathcal{A}^{cv}-$affine representation of $\succeq$ on $\mathcal{A} \times \{\mathcal{P}\}$, as defined in Claim 6. Without loss of generality, since the $\hat{V}_i$ and $V_\mathcal{P}$ are defined up to a positive affine transformation, we can choose $\hat{V}_i$ such that $\hat{V}_i(f^*) = V_\mathcal{P}(f^*) = 1$ and $\hat{V}_i(f_*) = V_\mathcal{P}(f_*) = -1$ for all $i \in N$. We thus have $\hat{V}_i(\mathcal{A}) = [-1, 1]$ for all $i \in N$.

We now show that for all $h \in \mathcal{A}^{cv}$ and all $i, j \in N$, $\hat{V}_i(h) = \hat{V}_j(h)$. Let $h \in \mathcal{A}^{cv}$. Because $f^*$ and $f_*$ are such that $f^* \hat{\succeq}_i f \hat{\succeq}_i f_*$ for all $f \in \mathcal{A}$ and all $i \in N$, we have in particular $f^* \hat{\succeq}_i h \hat{\succeq}_i f_*$, for all $i \in N$. By Axioms 1 and 3, for all $\mathcal{P} \in \mathbb{P}$, there exists $\alpha_\mathcal{P}$ such that $(h, \mathcal{P}) \sim (\alpha_\mathcal{P} f^* + (1 - \alpha_\mathcal{P}) f_*, \mathcal{P})$. By Claim 1, $\alpha_\mathcal{P} f^* + (1 - \alpha_\mathcal{P}) f_* \in \mathcal{A}^{cv}$. Therefore, by Axiom 9, $(\alpha_\mathcal{P} f^* + (1 - \alpha_\mathcal{P}) f_*, \mathcal{P}) \sim (\alpha_\mathcal{P} f^* + (1 - \alpha_\mathcal{P}) f_*, \{\delta_i\})$ and $(h, \mathcal{P}) \sim (h, \{\delta_i\})$. Thus $\alpha_\mathcal{P} f^* + (1 - \alpha_\mathcal{P}) f_* \hat{\sim}_i h$ for all $i \in N$. Hence, for all $i \in N$, $\hat{V}_i(h) = \alpha_\mathcal{P} \hat{V}_i(f^*) + (1 - \alpha_\mathcal{P}) \hat{V}_i(f_*) = 2\alpha_\mathcal{P} - 1$, which proves that $\hat{V}_i(h) = \hat{V}_j(h)$ for all $i, j \in N$.

Assume now that $\{\tilde{V}_i\}_{i \in N}$ is another set of affine functions representing $\{\hat{\succeq}_i\}_{i \in N}$ and satisfying $\tilde{V}_i(\mathcal{A}) = \tilde{V}_j(\mathcal{A})$ for all $i, j \in N$. By the same argument as above, we must have $\tilde{V}_i(h) = \tilde{V}_j(h)$ for all $i, j \in N$ and all $h \in \mathcal{A}^{cv}$. Because the $\tilde{V}_i$ are unique up to an affine transformation, there must exist $a_1, \ldots, a_n > 0$ and $b_1, \ldots, b_n \in \mathbb{R}$ such that for all $i$, $\tilde{V}_i = a_i \hat{V}_i + b_i$. But as we have shown, one must also have: $\tilde{V}_i(f) = \tilde{V}_j(f)$ for all $f \in \mathcal{A}^{cv}$. Hence, $a_i \hat{V}_i(f) + b_i = a_j \hat{V}_j(f) + b_j$ for all $i, j \in N$, $f \in \mathcal{A}^{cv}$. Since $\hat{V}_i(f) = \hat{V}_j(f)$ for all $f \in \mathcal{A}^{cv}$ and all $i, j \in N$ and $\hat{V}_i$ is not constant on $\mathcal{A}^{cv}$, this implies $a_i = a_j$ and $b_i = b_j$.

Let $F : \mathcal{A} \to \mathbb{R}^{n+1}$ be defined by $F(f) = (V_\mathcal{P}(f), \hat{V}_1(f), \ldots, \hat{V}_n(f))$. For any $h \in \mathcal{A} \setminus \mathcal{A}^{cv}$, let $K_h = F(\mathcal{A}_h)$. By Claims 1, 3 and 6, $K_h$ is convex. Therefore, by Claims 4 and 5, we

can apply Proposition 2 in De Meyer and Mongin (1995) : there exist non-negative numbers $\lambda_1(h, \mathcal{P}), \ldots, \lambda_n(h, \mathcal{P})$ not all zero, a nonnegative number $\kappa(h, \mathcal{P})$ and a real number $\mu(h, \mathcal{P})$ such that for all $f \in \mathcal{A}_h$,

$$\kappa(h, \mathcal{P}) V_{\mathcal{P}}(f) = \sum_{i \in N} \lambda_i(h, \mathcal{P}) \hat{V}_i(f) + \mu(h, \mathcal{P}).$$

Because $\hat{V}_i$ is not constant on $\mathcal{A}^{cv}$, $\kappa(h, \mathcal{P}) \neq 0$ and, without loss of generality, can be set equal to 1. We hence have:

$$\begin{cases} V_{\mathcal{P}}(f^*) = 1 = \sum_{i \in N} \lambda_i(h, \mathcal{P}) + \mu(h, \mathcal{P}) \\ V_{\mathcal{P}}(f_*) = -1 = -\sum_{i \in N} \lambda_i(h, \mathcal{P}) + \mu(h, \mathcal{P}). \end{cases}$$

Therefore, $\sum_i \lambda_i(h, \mathcal{P}) + \mu(h, \mathcal{P}) = \sum_{i \in N} \lambda_i(h, \mathcal{P}) - \mu(h, \mathcal{P})$, which implies $\mu(h, \mathcal{P}) = 0$ and therefore, $\sum_{i \in N} \lambda_i(h, \mathcal{P}) = 1$.

$\square$

**Claim 8.** *For all $\mathcal{P} \in \mathbb{P}$, there exists a unique compact and convex set $\mathcal{F}(\mathcal{P}) \in \mathbb{P}$ such that, for all $f, g \in \mathcal{A}$,*

$$(f, \mathcal{P}) \succeq (g, \mathcal{P}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i) \hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i) \hat{V}_i(f(i)),$$

*where the $\hat{V}_i$ are affine representations of $\hat{\succeq}_i$ such that $\hat{V}_i(\mathcal{A}) = \hat{V}_j(\mathcal{A})$ for all $i, j \in N$. Moreover, the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation and coincide on $\mathcal{A}^{cv}$.*

*Proof.* Let $\mathcal{P} \in \mathbb{P}$ be fixed and $V_{\mathcal{P}}, \hat{V}_i$ ($i \in N$) be defined as in Claim 7.

Let $\tilde{B} = \{\varphi : N \to [-1, 1]^n | \exists f \in \mathcal{A} \text{ s.t. } \varphi(i) = \hat{V}_i(f), \forall i \in N\}$ and $B(N, V_{\mathcal{P}}(\mathcal{A}))$ be the set of functions from $N$ to $V_{\mathcal{P}}(\mathcal{A})$.

We first prove that $\tilde{B} \subseteq B(N, V_{\mathcal{P}}(\mathcal{A}))$. Let $\varphi \in \tilde{B}$. By definition, there exists $f \in \mathcal{A}$ such that $\varphi(i) = \hat{V}_i(f)$ for all $i \in N$. By Axiom 6, for all $i \in N$ there exist $\bar{h}_i, \underline{h}_i \in \mathcal{A}^{cv}$ such that $(\bar{h}_i, \{\delta_i\}) \succeq (f, \{\delta_i\}) \succeq (\underline{h}_i, \{\delta_i\})$. By Axiom 3, there exists $\theta_i \in [0, 1]$ such that $(f, \{\delta_i\}) \sim (\theta_i \bar{h}_i + (1 - \theta_i) \underline{h}_i, \{\delta_i\})$. Let $h_i = \theta_i \bar{h}_i + (1 - \theta_i) \underline{h}_i$. Note that because $(h_i, \{\delta_i\}) \sim (f, \{\delta_i\})$, $\hat{V}_i(f) = \hat{V}_i(h_i)$. By Claim 1, $h_i \in \mathcal{A}^{cv}$. Therefore, as shown in the proof of Claim 7, $\hat{V}_i(h_i) = \hat{V}_j(h_i)$ for all $j \in N \setminus \{i\}$. Thus, by the displayed equation in Claim 7, $V_{\mathcal{P}}(h_i) = \hat{V}_i(h_i)$ and therefore $V_{\mathcal{P}}(h_i) = \hat{V}_i(f)$. Now, define $\psi \in B(N, V_{\mathcal{P}}(\mathcal{A}))$ by $\psi(i) = V_{\mathcal{P}}(h_i)$. We have $\psi(i) = \varphi(i)$ for all $i \in N$, proving that $\varphi \in B(N, V_{\mathcal{P}}(\mathcal{A}))$. Thus $\tilde{B} \subseteq B(N, V_{\mathcal{P}}(\mathcal{A}))$.

We show now that $B(N, V_{\mathcal{P}}) \subseteq \tilde{B}$. Let $\varphi : N \to V_{\mathcal{P}}(\mathcal{A})$. By Claim 6, $V_{\mathcal{P}}(\mathcal{A}) = V_{\mathcal{P}}(\mathcal{A}^{cv})$. Therefore, for all $i \in N$, there exist $h_i \in \mathcal{A}^{cv}$ such that $\varphi(i) = V_{\mathcal{P}}(h_i)$. Let $f \in \mathcal{A}$ be defined by $f(i) = h_i(i)$ for all $i \in N$. By Axiom 10 we then have $\hat{V}_i(f) = \hat{V}_i(h_i)$. Because $h_i \in \mathcal{A}^{cv}$,

26

the proof of Claim 7 shows that $\hat{V}_i(h_i) = V_{\mathcal{P}}(h_i)$. Therefore, $(\hat{V}_1(f), \ldots, \hat{V}_n(f)) = \varphi$. Thus $B(N, V_{\mathcal{P}}) \subseteq \tilde{B}$.

Let $I : B(N, V_{\mathcal{P}}(\mathcal{A})) \to \mathbb{R}$ be defined by:

$$I(\varphi) = V_{\mathcal{P}}(f) \text{ if } \varphi(i) = \hat{V}_i(f) \,\forall i \in N.$$

Observe that because $\tilde{B} = B(N, V_{\mathcal{P}}(\mathcal{A}))$, $I$ is well defined. By Claim 4, $I$ is monotone and by Claim 7, $I(0) = 0$, $I(1) = 1$ and $I$ is $V_{\mathcal{P}}(\mathcal{A}^{cv})$-affine, i.e., for all $f \in \mathcal{A}$, all $h \in \mathcal{A}^{cv}$ and all $\alpha \in [0,1]$, if $\phi, \psi \in B(N, V_{\mathcal{P}}(\mathcal{A}))$ are such that $\phi(i) = \hat{V}_i(f)$ and $\psi(i) = \hat{V}_i(h)$ for all $i \in N$, then $I(\alpha\phi + (1-\alpha)\psi) = \alpha I(\phi) + (1-\alpha)I(\psi)$. By Claim 6, $V_{\mathcal{P}}(\mathcal{A}) = V_{\mathcal{P}}(\mathcal{A}^{cv})$. Thus, $I$ is homogeneous of degree 1. Furthermore, Axiom 11 implies that $I$ is concave. Therefore, its homogeneous of degree 1 extension $J$ to $B(N)$, the set of all functions from $N$ to $\mathbb{R}$, is monotone, concave and such that $J(\varphi + k) = J(\varphi) + k$ for all $k \in \mathbb{R}$. Because $I$ is concave and homogeneous of degree 1, it is also superadditive. Therefore, by a classical result (see, e.g., the "Fundamental Lemma" in Chateauneuf, 1991, and Lemma 3.5 in Gilboa and Schmeidler, 1989), there exists a unique compact convex set $\mathcal{F}(\mathcal{P}) \in \mathbb{P}$ such that for all $\varphi \in B(N)$:

$$J(\varphi) = \min_{p \in \mathcal{F}(\mathcal{P})} \sum_{i \in N} p(i)\varphi(i).$$

Therefore, for all $f \in \mathcal{A}$,

$$(f, \mathcal{P}) \succeq (g, \mathcal{P}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_{i \in N} p(i)\hat{V}_i(f) \geq \min_{p \in \mathcal{F}(\mathcal{P})} \sum_{i \in N} p(i)\hat{V}_i(g)$$

which is equivalent to

$$(f, \mathcal{P}) \succeq (g, \mathcal{P}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_{i \in N} p(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{P})} \sum_{i \in N} p(i)\hat{V}_i(g(i)).$$

Finally, that the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation follows from Claim 7.

$\square$

**Claim 9.** *For all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$, there exist unique compact and convex sets $\mathcal{F}(\mathcal{P}), \mathcal{F}(\mathcal{Q}) \in \mathbb{P}$ such that, for all $f, g \in \mathcal{A}$,*

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(f(i)),$$

*where the $\hat{V}_i$ are affine representations of $\hat{\succeq}_i$ such that $\hat{V}_i(\mathcal{A}) = \hat{V}_j(\mathcal{A})$ for all $i, j \in N$. Moreover, the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation and coincide on $\mathcal{A}^{cv}$.*

*Proof.* Let $f, g \in \mathcal{A}$ and $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$. Let $\mathcal{F}(\mathcal{P})$ and $\mathcal{F}(\mathcal{Q})$ be defined as in Claim 8. Assume that $(f, \mathcal{P}) \succeq (g, \mathcal{Q})$. By Axiom 6, there exist $\bar{f}, \underline{f}, \bar{g}, \underline{g} \in \mathcal{A}^{cv}$ such that:

$$\begin{cases} (\bar{f}, \mathcal{P}) \succeq (f, \mathcal{P}) \succeq (\underline{f}, \mathcal{P}) \\ (\bar{g}, \mathcal{Q}) \succeq (g, \mathcal{Q}) \succeq (\underline{g}, \mathcal{Q}). \end{cases}$$

By Axiom 9, $(\underline{g}, \mathcal{P}) \sim (\underline{g}, \mathcal{Q})$ and $(\bar{f}, \mathcal{P}) \sim (\bar{f}, \mathcal{Q})$. Therefore,

$$\begin{cases} (\bar{f}, \mathcal{P}) \succeq (f, \mathcal{P}) \succeq (\underline{g}, \mathcal{P}) \\ (\bar{f}, \mathcal{Q}) \succeq (g, \mathcal{Q}) \succeq (\underline{g}, \mathcal{Q}). \end{cases}$$

Hence, by Axioms 1 and 3, there exist $\lambda, \mu \in [0, 1]$ such that $(f, \mathcal{P}) \sim (\lambda \bar{f} + (1 - \lambda)\underline{g}, \mathcal{P})$ and $(g, \mathcal{Q}) \sim (\mu \bar{f} + (1 - \mu)\underline{g}, \mathcal{Q})$. Hence, $(\lambda \bar{f} + (1 - \lambda)\underline{g}, \mathcal{P}) \succeq (\mu \bar{f} + (1 - \mu)\underline{g}, \mathcal{Q})$. Observe that by Claim 1, $\lambda \bar{f} + (1 - \lambda)\underline{g} \in \mathcal{A}^{cv}$ and $\mu \bar{f} + (1 - \mu)\underline{g} \in \mathcal{A}^{cv}$. Therefore, by Axiom 9, $(\lambda \bar{f} + (1 - \lambda)\underline{g}, \{\delta_1\}) \succeq (\mu \bar{f} + (1 - \mu)\underline{g}, \{\delta_1\})$ and therefore, $\hat{V}_1(\lambda \bar{f}(1) + (1 - \lambda)\underline{g}(1)) \geq \hat{V}_1(\mu \bar{f}(1) + (1 - \mu)\underline{g}(1))$.

Since $(\lambda \bar{f} + (1 - \lambda)\underline{g}, \mathcal{P}) \sim (f, \mathcal{P})$, we have by Claim 8:

$$\begin{aligned} \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) &= \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(\lambda \bar{f}(i) + (1 - \lambda)\underline{g}(i)) \\ &= \hat{V}_1(\lambda \bar{f}(1) + (1 - \lambda)\underline{g}(1)). \end{aligned}$$

Similarly,

$$\begin{aligned} \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(g(i)) &= \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(\mu \bar{f}(i) + (1 - \mu)\underline{g}(i)) \\ &= \hat{V}_1(\mu \bar{f}(1) + (1 - \mu)\underline{g}(1)). \end{aligned}$$

Since $\hat{V}_1(\lambda \bar{f}(1) + (1 - \lambda)\underline{g}(1)) \geq \hat{V}_1(\mu \bar{f}(1) + (1 - \mu)\underline{g}(1))$, we finally obtain:

$$\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(g(i)),$$

the desired result. $\square$

Hence, there exists a unique function $\mathcal{F} : \mathbb{P} \to \mathbb{P}$ such that $\mathcal{F}(\mathcal{P})$ is compact and convex for all sets $\mathcal{P} \in \mathbb{P}$ for which for all $f, g \in \mathcal{A}$, $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$,

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(f(i)),$$

where the $\hat{V}_i$ are affine representations of $\hat{\succeq}_i$ such that $\hat{V}_i(\mathcal{A}) = \hat{V}_j(\mathcal{A})$, for all $i, j \in N$. Moreover, $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation.

It remains to show that $\mathcal{F}(\mathcal{P}) \subseteq \mathcal{P}$ for all $\mathcal{P} \in \mathbb{P}$ and that for all $\alpha \in (0, 1)$, all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$, $\mathcal{F}(\alpha \mathcal{P} + (1 - \alpha)\mathcal{Q}) = \alpha \mathcal{F}(\mathcal{P}) + (1 - \alpha)\mathcal{F}(\mathcal{Q})$. This is done in the two following claims.

**Claim 10.** *For all $\mathcal{P} \in \mathbb{P}$, $\mathcal{F}(\mathcal{P}) \subseteq \mathcal{P}$.*

*Proof.* First, note that for all $\mathcal{P} \in \mathbb{P}$ and all $p \in \mathcal{F}(\mathcal{P})$, $p(S(\mathcal{P})) = 1$. Indeed, assume on the contrary that there exists $\tilde{p} \in \mathcal{F}(\mathcal{P})$ such that $p(S(\mathcal{P})) = q < 1$. Let $f$ be defined by $f(i) = f^*(i)$ for all $i \in S(\mathcal{P})$ and $f(i) = f_*(i)$ for all $i \in N \setminus S(\mathcal{P})$, where $f^*$ and $f_*$ are defined as in Claim 7. Let $g$ be defined by $g(i) = f^*(i)$ for all $i \in N$. We then have:

$$
\begin{aligned}
\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) &\leq \sum_i \tilde{p}(i)\hat{V}_i(f(i)) \\
&= q\hat{V}_i(f^*(i)) + (1-q)\hat{V}_i(f_*(i)) \\
&< \hat{V}_i(f^*(i)) \\
&= \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(g(i))
\end{aligned}
$$

Hence, $(g, \mathcal{P}) \succ (f, \mathcal{P})$, which contradicts Axiom 10. Therefore, $\mathcal{F}(\{\delta_i\}) = \{\delta_i\}$ for all $i \in N$.

Let $f \in \mathcal{A}$ and $p \in \Delta(N)$. By Axioms 3, 6 and Claim 1, for all $i \in N$ there exist $h_i \in \mathcal{A}^{cv}$ such that $(f, \{\delta_i\}) \sim (h_i, \{\delta_i\})$. Because $h_i \in \mathcal{A}^{cv}$, $(h_i, \{\delta_i\}) \sim (h_i, \{\delta_k\})$ and $\mathcal{F}(\{\delta_i\}) = \{\delta_i\}$, Claim 9 implies $\hat{V}_i(h_i(i)) = \hat{V}_k(h_i(k))$ for all $i \in N$.

Let $h = \sum_{i \in N} p(i)h_i$. By Axiom 8, $(h, \{\delta_k\}) \sim (f, \{p\})$. Thus,

$$
\begin{aligned}
\min_{p \in \mathcal{F}(\{p\})} \sum_{i \in N} p(i)\hat{V}_i(f(i)) &= \min_{p \in \mathcal{F}(\{\delta_k\})} \sum_{i \in N} p(i)\hat{V}_i(h(i)) && \text{(by Claim 9)} \\
&= \hat{V}_k(h(k)) && \text{(because } \mathcal{F}(\{\delta_k\}) = \{\delta_k\}) \\
&= \hat{V}_k\left(\sum_i p(i)h_i(k)\right) && \text{(by definition of } h) \\
&= \sum_i p(i)\hat{V}_k(h_i(k)) && \text{(because } \hat{V}_k \text{ is affine)} \\
&= \sum_i p(i)\hat{V}_i(h_i(i)) && \text{(because } (h_i, \{\delta_k\}) \sim (h_i, \{\delta_i\}) \text{ for all } i \in N).
\end{aligned}
$$

Hence, for all $f, g \in \mathcal{A}$, $p, q \in \Delta(N)$,

$$
(f, \{p\}) \succeq (g, \{q\}) \Leftrightarrow \sum_i p(i)\hat{V}_i(f(i)) \geq \sum_i q(i)\hat{V}_i(g(i)).
$$

Assume that there exists $\mathcal{P} \in \mathbb{P}$ such that $\mathcal{F}(\mathcal{P}) \nsubseteq \mathcal{P}$. Then, there exists $p^* \in \mathcal{F}(\mathcal{P})$ such that $p^* \notin \mathcal{P}$. Since $\mathcal{P}$ and $\mathcal{F}(\mathcal{P})$ are closed and convex sets, a separation argument implies that there exists a function $\phi : N \to \mathbb{R}$ such that $\sum_i p^*(i)\phi(i) < min_{p \in \mathcal{P}} \sum_i p(i)\phi(i)$. There exist numbers $a, b$ with $a > 0$ such that $a\phi(i) + b \in \hat{V}_i(\mathcal{A})$ for all $i$ (choosing $a$ sufficiently close to zero will ensure that $a_j\phi_j(i) + b_j \in [-1, 1]$ for all $i \in N$). Hence, for all $i$, there exists $y_i \in Y$ such that $a\phi(i) + b = \hat{V}_i(y_i)$. Define $f$ by $f(i) = y_i$ for all $i \in N$. Note that because

$\mathcal{A}^{cv}$ is convex (see Claim 1), $\hat{V}_i$ is continuous and $\hat{V}_i(g) = \hat{V}_j(g)$ for all $i, j \in N$ and $g \in \mathcal{A}^{cv}$, $\min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f(i)) \in \hat{V}_j(\mathcal{A}^{cv})$ for all $j \in N$. Therefore, there exists $\hat{h} \in \mathcal{A}^{cv}$ such that $\hat{V}_j(\hat{h}(j)) = \min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f(i))$ for all $j \in N$. We have, for all $q \in \mathcal{P}$ and all $j \in N$:

$$\sum_i q(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f(i)) = \hat{V}_j(\hat{h}(j)).$$

But we have shown that $\mathcal{F}(\{p\}) = \{p\}$ for all $p \in \Delta(N)$. Therefore, by Claim 9, for all $q \in \mathcal{P}$ and all $j \in N$, $(f, \{q\}) \succeq (\hat{h}, \{\delta_j\})$. Because $\hat{h} \in \mathcal{A}^{cv}$, by Axiom 9, $(\hat{h}, \{\delta_j\}) \sim (\hat{h}, \{q\})$ for all $j \in N$ and all $q \in \Delta(N)$. Thus $(f, \{q\}) \succeq (\hat{h}, \{q\})$ for all $q \in \mathcal{P}$. Hence, by Axiom 5, $(f, \mathcal{P}) \succeq (\hat{h}, \mathcal{P})$. But:

$$
\begin{aligned}
\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) &\leq \sum_i p^*(i)\hat{V}_i(f(i)) && \text{(because } p^* \in \mathcal{F}(\mathcal{P})\text{)} \\
&< \min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f(i)) \\
&= \hat{V}_j(\hat{h}(j)), \text{ for all } j \in N && \text{(by definition of } h\text{).}
\end{aligned}
$$

Hence, by Claim 9 and Axiom 9, $(\hat{h}, \mathcal{P}) \succ (f, \mathcal{P})$, a contradiction. $\qquad\square$

**Claim 11.** *For all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$, all $\alpha \in (0, 1)$, $\mathcal{F}(\alpha\mathcal{P} + (1-\alpha)\mathcal{Q}) = \alpha\mathcal{F}(\mathcal{P}) + (1-\alpha)\mathcal{F}(\mathcal{Q})$.*

*Proof.* Let $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ and $\alpha \in (0, 1)$. For all $f \in \mathcal{A}$, let $p^*(f) \in \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i))$ and $q^*(f) \in \arg\min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(f(i))$. Because by Claim 9, $\mathcal{F}(\mathcal{P})$ and $\mathcal{F}(\mathcal{Q})$ are compact, $p^*(f)$ and $q^*(f)$ are well-defined. By Claims 9 and 10, $(f, \mathcal{P}) \sim (f, \{p^*(f)\})$ and $(f, \mathcal{Q}) \sim (f, \{q^*(f)\})$. By Axiom 4, this implies:

$$(f, \alpha\mathcal{P} + (1-\alpha)\mathcal{Q}) \sim (f, \alpha\{p^*(f)\} + (1-\alpha)\{q^*(f)\}).$$

By Claim 10, $\mathcal{F}(\{\alpha p^*(f) + (1-\alpha)q^*(f)\}) = \{\alpha p^*f + (1-\alpha)q^*(f)\} = \alpha p^*(f) + (1-\alpha)\{q^*(f)\}$. Hence, by Claims 9 and 10:

$$
\begin{aligned}
\min_{p \in \mathcal{F}(\alpha\mathcal{P} + (1-\alpha)\mathcal{Q})} \sum_i p(i)\hat{V}_i(f(i)) &= \min_{p \in \mathcal{F}(\alpha\{p^*(f)\} + (1-\alpha)\{q^*(f)\})} \sum_i p(i)\hat{V}_i(f(i)) \\
&= \sum_i (\alpha p^*(f)(i) + (1-\alpha)q^*(f)(i))\hat{V}_i(f(i)) \\
&= \alpha\sum_i p^*(f)(i)\hat{V}_i(f(i)) + (1-\alpha)\sum_i q^*(f)(i)\hat{V}_i(f(i)) \\
&= \alpha\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) + (1-\alpha)\min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(f(i)) \\
&= \min_{p \in \alpha\mathcal{F}(\mathcal{P}) + (1-\alpha)\mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(f(i)).
\end{aligned}
$$

This holds for all $f \in \mathcal{A}$. Observe that by Claim 7, the set $\{(\hat{V}_1(f(1)), \ldots, \hat{V}_n(f_n)) | f \in \mathcal{A}\} = [-1, 1]^n$ and therefore includes a neighborhood of the origin. Thus, uniqueness of $\mathcal{F}$ (see Claim 9) implies that $\mathcal{F}(\alpha\mathcal{P} + (1 - \alpha)\mathcal{Q}) = \alpha(\mathcal{F})\mathcal{P} + (1 - \alpha)\mathcal{F}(\mathcal{Q})$, the desired result. $\square$

## Proof of Theorem 2

The necessity part of the Theorem is easily checked. We therefore only prove the sufficiency part.

We will use the following notation. For all subsets $T$ of $N$, let $\Delta(T)$ be the simplex over $T$. Let $c_T \in \Delta(T)$ be defined by $c_T(s) = \frac{1}{|T|}$ for all $s \in T$. Finally, let $\mathcal{H} : \mathbb{P} \times \Delta(N) \times [0, 1] \to \mathbb{P}$ be defined by :

$$\mathcal{H}(\mathcal{Q}, c, \theta) = \{p \in \Delta(N) | \exists q \in \mathcal{Q} \text{ s.t. } p = \theta q + (1 - \theta)c\}.$$

One can easily check that the proof of Theorem 1 is unaffected if one restrict the domain of $\succeq$ to $\mathcal{A} \times \mathbb{B}$. Thus, under the assumptions of Theorem 1 restricted to $\mathcal{A} \times \mathbb{B}$, there exist affine functions $\hat{V}_i : Y \to \mathbb{R}$, $i \in N$, representing $\hat{\succeq}_i$ such that $\hat{V}_i(Y) = \hat{V}_j(Y)$, for all $i, j \in N$ and a function $\mathcal{F} : \mathbb{B} \to \mathbb{P}$ satisfying for all $\mathcal{P}, \mathcal{Q} \in \mathbb{B}$,

1. $\mathcal{F}(\mathcal{P}) \subseteq \mathcal{P}$

2. For all $\alpha \in [0, 1]$, $\mathcal{F}(\alpha\mathcal{P} + (1 - \alpha)\mathcal{Q}) = \alpha\mathcal{F}(\mathcal{P}) + (1 - \alpha)\mathcal{F}(\mathcal{Q})$

for which for all $\mathcal{P}, \mathcal{Q} \in \mathbb{B}$ and all $f, g \in \mathcal{A}$,

$$(f, \mathcal{P}) \succeq (g, \mathcal{Q}) \Leftrightarrow \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f(i)) \geq \min_{p \in \mathcal{F}(\mathcal{Q})} \sum_i p(i)\hat{V}_i(g(i)).$$

Furthermore, $\mathcal{F}$ is unique and the functions $\{\hat{V}_i\}_{i \in N}$ are unique up to a common positive affine transformation and coincide on $\mathcal{A}^{cv}$.

As in the proof of Theorem 1, we will use the following slight abuse of notation: for all $i \in N$ and $f \in \mathcal{A}$, we define $\hat{V}_i(f)$ as $\hat{V}_i(f(i))$.

The proof goes through several claims. Although not explicitly stated in the claims, all the assumptions of Theorem 2 are made throughout this subsection.

**Claim 12.** *For all $\mathcal{P} \in \mathbb{B}$ and all bijections $\varphi : N \to N$, $\mathcal{F}(\mathcal{P}^\varphi) = (\mathcal{F}(\mathcal{P}))^\varphi$.*

*Proof.* Let $\mathcal{P} \in \mathbb{B}$ and $\varphi$ be a permutation on $N$. We will prove that $\mathcal{F}(\mathcal{P})^\varphi \subseteq \mathcal{F}(\mathcal{P}^\varphi)$. Assume that such is not the case, i.e., there exists $p^* \in \mathcal{F}(\mathcal{P})^\varphi$ such that $p^* \notin \mathcal{F}(\mathcal{P}^\varphi)$. Because $\mathcal{F}(\mathcal{P}^\varphi)$ and $\mathcal{F}(\mathcal{P})^\varphi$ are closed convex sets, a standard separation argument implies that there exists a function $\phi : N \to \mathbb{R}$ such that: $\sum_i p^*(i)\phi(i) < \min_{p \in \mathcal{F}(\mathcal{P}^\varphi)} \sum_i p(i)\phi(i)$. There exist numbers $a, b$ with $a > 0$ such that $a\phi(i) + b \in \hat{V}_i(\mathcal{A})$ for all $i$. Hence, for all $i$, there exists $y_i \in Y$ such that $a\phi(i) + b =$

$\hat{V}_i(y_i)$ (choosing $a$ sufficiently close to zero will ensure that $a_j\phi_j(i) + b_j \in [-1,1]$ for all $i \in N$). Define $f$ by $f(i) = y_i$ for all $i \in N$. We then have: $\sum_i p^*(i)\hat{V}_i(f) < \min_{p \in \mathcal{F}(\mathcal{P}^\varphi)} \sum_i p(i)\hat{V}_i(f)$. Axiom 6 and Claim 1 in the proof of Theorem 1 imply that $\mathcal{A}(h^\psi) \neq \emptyset$, for all permutations $\psi : N \to N$ and all $h \in \mathcal{A}$. Let $g \in \mathcal{A}(f^{\varphi^{-1}})$. For all $p \in \mathcal{F}(\mathcal{P})$, $\sum_i p(i)\hat{V}_i(g) = \sum_i p^\varphi(i)\hat{V}_i(f)$. Therefore, $\min_{p \in \mathcal{F}(\mathcal{P})^\varphi} \sum_i p(i)\hat{V}_i(f) = \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(g)$. Hence,

$$\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(g) = \min_{p \in \mathcal{F}(\mathcal{P})^\varphi} \sum_i p(i)\hat{V}_i(f) \leq \sum_i p^*(i)\hat{V}_i(f) < \min_{p \in \mathcal{F}(\mathcal{P}^\varphi)} \sum_i p(i)\hat{V}_i(f).$$

Thus, $(f, \mathcal{P}^\varphi) \succ (g, \mathcal{P})$, a contradiction with Axiom 12. The inclusion $\mathcal{F}(\mathcal{P}^\varphi) \subseteq \mathcal{F}(\mathcal{P})^\varphi$ can be proved using a similar argument. Therefore, $\mathcal{F}(\mathcal{P}^\varphi) = \mathcal{F}(\mathcal{P})^\varphi$. $\qquad \square$

**Claim 13.** *Let* $\mathcal{P} \in \mathbb{B}$, $I$ *a subset of* $N$ *and* $f_k \in \mathcal{A}$ ($k \in I$). *Then:*

$$\bigcap_{k \in I} \left( \arg\min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f_k) \right) \neq \emptyset \Rightarrow \bigcap_{k \in I} \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_k) \right) \neq \emptyset.$$

*Proof.* Let $\mathcal{P} \in \mathbb{B}$, $I = \{1, \ldots, m\}$ with $m \geq 2$ and $f_k \in \mathcal{A}$ ($k \in I$) be such that:

$$\bigcap_{k \in I} \left( \arg\min_{p \in \mathcal{P}} \sum_i \hat{V}_i(f_k) \right) \neq \emptyset.$$

Assume, without loss of generality, that $(f_1, \mathcal{P}) \succeq (f_k, \mathcal{P})$, for all $k \in I$. By Axiom 6, there exists $\bar{f} \in \mathcal{A}^{cv}$ such that $(\bar{f}, \mathcal{P}) \succeq (f_1, \mathcal{P})$. Hence, for all $k \in I \setminus \{1\}$, $(\bar{f}, \mathcal{P}) \succeq (f_1, \mathcal{P}) \succeq (f_k, \mathcal{P})$. Hence, by Axioms 1 and 3, for all $k \neq 1$, there exist $\alpha_k \in [0,1]$ such that $(\alpha_k \bar{f} + (1 - \alpha_k)f_k, \mathcal{P}) \sim (f_1, \mathcal{P})$. But observe that since $\bar{f} \in \mathcal{A}^{cv}$, $\arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(\alpha_k \bar{f} + (1 - \alpha_k)f_k) = \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_k)$ if $a_k \neq 1$. Therefore, there is no loss in generality if we assume $(f_1, \mathcal{P}) \sim (f_k, \mathcal{P})$, for all $k \in I$, an assumption we maintain throughout this proof.

We now proceed by induction. For all $r \leq m$ let

$$\mathbf{P}(r) = \left\{ \bigcap_{k \in I} \left( \arg\min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f_k) \right) \neq \emptyset \Rightarrow \bigcap_{k \in \{1, \ldots, r\}} \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_k) \right) \neq \emptyset \right\}.$$

We first prove that $\mathbf{P}(2)$ is true. Assume that:

$$\left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_1) \right) \cap \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i \hat{V}_i(f_2) \right) = \emptyset.$$

Let $p^* \in \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_1)$ and $\hat{p} \in \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_2)$. By assumption,

$$\sum_i \hat{p}(i)\hat{V}_i(f_2) < \sum_i p^*(i)\hat{V}_i(f_2),$$

and by Theorem 1:

$$\sum_i p^*(i)\hat{V}_i(f_1) = \sum_i \hat{p}(i)\hat{V}_i(f_2).$$

By assumption, there exists $\bar{p} \in \mathcal{P}$ such that

$$\bar{p} \in \left( \arg\min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f_1(i)) \right) \cap \left( \arg\min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f_2(i)) \right).$$

By Theorem 1, $\mathcal{F}(\{p\}) = \{p\}$ for all $p \in \Delta(N)$ and, hence, $(f_1, \{p\}) \succeq (f_1, \{\bar{p}\})$ and $(f_2, \{p\}) \succeq (f_2, \{\bar{p}\})$ for all $p \in \mathcal{P}$. Thus, by Axiom 13, for all $\alpha \in (0,1)$, $(\alpha f_1 + (1-\alpha)f_2, \mathcal{P}) \sim (f_1, \mathcal{P})$. Therefore, by Theorem 1,

$$\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(\alpha f_1 + (1-\alpha)f_2) = \min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_1) = \sum_i p^*(i)\hat{V}_i(f_1).$$

Hence, there exists $\tilde{p} \in \mathcal{F}(\mathcal{P})$ such that:

$$\sum_i \tilde{p}(i)\hat{V}_i(\alpha f_1 + (1-\alpha)f_2) = \sum_i p^*(i)\hat{V}_i(f_1),$$

i.e.,

$$\alpha \sum_i \tilde{p}(i)\hat{V}_i(f_1) + (1-\alpha) \sum_i \tilde{p}(i)\hat{V}_i(f_2) = \sum_i p^*(i)\hat{V}_i(f_1).$$

Therefore, because $\sum_i p^*(i)\hat{V}_i(f_1) = \sum_i \hat{p}(i)\hat{V}_i(f_2)$ :

$$\alpha \sum_i \tilde{p}(i)\hat{V}_i(f_1) + (1-\alpha) \sum_i \tilde{p}(i)\hat{V}_i(f_2) = \alpha \sum_i p^*(i)\hat{V}_i(f_1) + (1-\alpha) \sum_i \hat{p}(i)\hat{V}_i(f_2).$$

Hence, because $\tilde{p} \in \mathcal{F}(\mathcal{P})$, it follows from the definition of $\hat{p}$ and $p^*$ that

$$\tilde{p} \in \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_1) \right) \cap \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_2) \right),$$

a contradiction. Therefore, $\mathbf{P}(2)$ is true.

We now assume that $\mathbf{P}(r-1)$ is true, with $r-1 < m$ and prove that then $\mathbf{P}(r)$ is also true. Assume that $\bigcap_{k \in \{1,\dots,r\}} \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_k) \right) = \emptyset$. By the induction assumption, there exists $p^* \in \bigcap_{k \in \{1,\dots,r-1\}} \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_k) \right)$. By assumption, for all $\hat{p} \in \left( \arg\min_{p \in \mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_r) \right)$,

$$\sum_i \hat{p}(i)\hat{V}_i(f_r) < \sum_i p^*(i)\hat{V}_i(f_r)$$

and, by Theorem 1,

$$\sum_i \hat{p}(i)\hat{V}_i(f_r) = \sum_i p^*(i)\hat{V}_i(f_k), \forall k \in \{1,\dots,r-1\}.$$

Let $h = \sum_{k=1}^{r-1} \frac{1}{r-1} f_k$. We then have:

$$\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h) \;=\; \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i\left(\sum_{k=1}^{r-1} \frac{1}{r-1} f_k\right)$$

$$=\; \arg\min_{p\in\mathcal{P}} \sum_i p(i)\left(\sum_{k=1}^{r-1} \hat{V}_i(f_k)\right),$$

the last equality being implied by the fact that the $\hat{V}_i$ are affine.

Let $p_1 \in \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h)$ and assume that $p_1 \notin \bigcap_{k=1}^{r-1}\left(\arg\min_{p\in\mathcal{P}} \sum_i \hat{V}_i(f_k)\right)$. Then, for all $p_2 \in \bigcap_{k=1}^{r-1}\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)\right)$ and all $k \in \{1,\ldots,r-1\}$, $\sum_i p_2(i)\hat{V}_i(f_k) \leq \sum_i p_1(i)\hat{V}_i(f_k)$, with a strict inequality for some $j \in \{1,\ldots,r-1\}$. But, then, $\sum_i p_2(i)\hat{V}_i(h) < \sum_i p_1(i)\hat{V}_i(h)$, a contradiction with the fact that $p_1 \in \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h)$. Therefore, $\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h) \subseteq \bigcap_{k=1}^{r-1}\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)\right)$. Conversely, let

$$q_1 \in \bigcap_{i=1}^{r-1}\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)\right).$$

Then, by definition, $q_1 \in \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)$, for all $k \in \{1,\ldots,r-1\}$. Hence, $q_1 \in \arg\min_{p\in\mathcal{P}} \sum_i p(i)\left(\sum_{k=1}^{r-1} \hat{V}_i(f_k)\right) = \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h)$. Therefore,

$$\bigcap_{k=1}^{r-1}\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)\right) \subseteq \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h),$$

which implies that $\bigcap_{k=1}^{r-1}\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)\right) = \arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h)$. The same reasoning show that $\bigcap_{k=1}^{r-1}\left(\arg\min_{p\in\mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(f_k)\right) = \arg\min_{p\in\mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(h)$.

Hence, because, by assumption, $\bigcap_{k\in I}\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_k)\right) \neq \emptyset$, we have

$$\left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(h)\right) \cap \left(\arg\min_{p\in\mathcal{P}} \sum_i p(i)\hat{V}_i(f_r)\right) \neq \emptyset.$$

Furthermore, by assumption, $p^* \in \arg\min_{p\in\mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(h) = \arg\min_{p\in\mathcal{F}(\mathcal{P})} \sum_i p(i)\left(\sum_{k=1}^{r-1} \frac{1}{r-1}\hat{V}_i(f_k)\right)$.

Reasoning as in the $r=2$ case, Axiom 13 implies that for all $\alpha \in (0,1)$, $(\alpha h + (1-\alpha)f_r, \mathcal{P}) \sim (h, \mathcal{P})$. Therefore, by Theorem 1,

$$\min_{p\in\mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(\alpha h + (1-\alpha)f_r) = \min_{p\in\mathcal{F}(\mathcal{P})} \sum_i p(i)\hat{V}_i(h) = \sum_i p^*(i)\hat{V}_i(h).$$

Hence, there exists $\tilde{p} \in \mathcal{F}(\mathcal{P})$ such that:

$$\sum_i \tilde{p}(i)\hat{V}_i(\alpha h + (1-\alpha)f_r) = \sum_i p^*(i)\hat{V}_i(h).$$

34

Thus, because $\sum p^*(i)\hat{V}_i(h) = \sum_i \hat{p}(i)\hat{V}_i(f_r)$,

$$\alpha \sum_i \tilde{p}(i)\hat{V}_i(h) + (1-\alpha)\sum_i \tilde{p}(i)\hat{V}_i(f_r) = \alpha\sum_i p^*(i)\hat{V}_i(h) + (1-\alpha)\sum_i \hat{p}(i)\hat{V}_i(f_r),$$

which implies that $\tilde{p} \in \left(\arg\min_{p\in\mathcal{F}(\mathcal{P})}\sum_i p(i)\hat{V}_i(h)\right) \cap \left(\arg\min_{p\in\mathcal{F}(\mathcal{P})}\sum_i p(i)\hat{V}_i(f_r)\right)$. But:

$$\left(\arg\min_{p\in\mathcal{F}(\mathcal{P})}\sum_i p(i)\hat{V}_i(h)\right) = \bigcap_{k\in\{1,\dots,r-1\}}\left(\arg\min_{p\in\mathcal{F}(\mathcal{P})}\sum_i p(i)\hat{V}_i(f_k)\right).$$

Therefore, $\tilde{p} \in \bigcap_{l\in\{1,\dots,r\}}\left(\arg\min_{p\in\mathcal{F}(\mathcal{P})}\sum_i p(i)\hat{V}_i(f_k)\right)$, a contradiction. $\square$

**Claim 14.** *Let $S_1, S_2 \subset N$ such that $S_1 \neq \emptyset$, $S_2 \neq \emptyset$. Then, for all $\alpha \in [0,1]$, there exists $\theta \in [0,1]$ such that $\mathcal{F}(\alpha\Delta(S_1) + (1-\alpha)\Delta(S_2)) = \mathcal{H}(\alpha\Delta(S_1) + (1-\alpha)\Delta(S_2), \alpha c_{S_1} + (1-\alpha)c_{S_2}, \theta)$.*

*Proof.* Let $\alpha \in [0,1]$ and $S_1, S_2 \subset N$ such that $S_1 \neq \emptyset$, $S_2 \neq \emptyset$. To simplify notation, let $c_1 = c_{S_1}$ and $c_2 = c_{S_2}$, $\Delta_1 = \Delta(S_1)$, $\Delta_2 = \Delta(S_2)$ and $\Delta = \alpha\Delta_1 + (1-\alpha)\Delta_2$. We first consider the case $|S_1| > 1$ and $|S_2| > 1$.

We know by Theorem 1 that $\mathcal{F}(\Delta) = \alpha\mathcal{F}(\Delta_1) + (1-\alpha)\mathcal{F}(\Delta_2)$. Furthermore, because $\Delta_i^\varphi = \Delta_i$, for all permutation $\varphi: N \to N$ such that $\varphi(S_i) = S_i$, Claim 12 implies that $\mathcal{F}(\Delta_i)^\varphi = \mathcal{F}(\Delta_i)$ for all such permutations. Hence, because $\mathcal{F}(\Delta_i)$ is convex (by Theorem 1), $c_i \in \mathcal{F}(\Delta_i)$ and, therefore, $\alpha c_1 + (1-\alpha)c_2 \in \alpha\mathcal{F}(\Delta_1) + (1-\alpha)\mathcal{F}(\Delta_2) = \mathcal{F}(\Delta)$. Let $c = \alpha c_1 + (1-\alpha)c_2$

Because $\Delta$ is a polytope, it has a finite number of facets and vertices. Let $\{F_1, \dots, F_K\}$ be the set of all facets of $\Delta$ and $\Pi = \{\pi_1^*, \dots, \pi_M^*\}$ be the set of its vertices. Furthermore, for all $m \in \{1, \dots, M\}$, let $J_m$ be a subset of $\{1, \dots, K\}$ such that $\cap_{j\in J_m}F_j = \{\pi_m^*\}$.

For each facet $F_j$, let $q_j$ be in the (relative) interior of $F_j$. Let $\Psi_j$ be the (unique) hyperplane supporting $\Delta$ at $q_j$, defined by $\Psi_j = \{p\,|\,\phi_j(p) = \mu_j\}$, with $\mu_j \in \mathbb{R}$ and $\phi_j$ a linear function (observe that $F_j \subset \Psi_j$). Because $\Psi_j$ is a supporting hyperplane for $\Delta$, $\phi_j$ can be chosen such that $\phi_j(p) \geq \mu_j$ for all $p \in \Delta$. Let $\phi_j(p) = \sum_i \phi_j(i)p(i)$. There exist numbers $a_j$ and $b_j$, with $a_j > 0$ such that $a_j\phi_j(i) + b_j \in \hat{V}_i(\mathcal{A})$ for all $i \in N$ (choosing $a$ sufficiently close to zero will ensure that $a_j\phi_j(i) + b_j \in [-1,1]$ for all $i \in N$). Therefore, there exist $y_i^j \in Y$ such that $a_j\phi_j(i) + b_j = \hat{V}_i(y_i^j)$ for all $i \in N$. Define $f_j$ by $f_j(i) = y_i^j$ for all $i \in N$.

Observe that $\mathcal{H}(\Psi_j \cap \Delta, c, \theta)$ are the sets for which $\sum_i p(i)\hat{V}_i(f_j)$ is constant and smaller than $\sum_i c(i)\hat{V}_i(f_j)$. Furthermore, for all $p \in \mathcal{H}(\Psi_j \cap \Delta, c, \theta)$, $p' \in \mathcal{H}(\Psi_j, c, \theta')$, $\theta < \theta'$ if, and only if, $\sum_i p(i)\hat{V}_i(f_j) > \sum_i p'(i)\hat{V}_i(f_j)$. Therefore, it is the case that, for some $\hat{\theta}_j$,

$$\arg\min_{p\in\mathcal{F}(\Delta)}\sum_i p(i)\hat{V}_i(f) \subseteq \mathcal{H}(\Psi_j \cap \Delta, c, \hat{\theta}_j).$$

Now, assume that there are two facets $F_r$ and $F_t$ such that $\hat{\theta}_r \neq \hat{\theta}_t$. This implies that there are two facets $F_j$ and $F_\ell$ such that $\hat{\theta}_j \neq \hat{\theta}_\ell$ and $F_\ell \cap F_j \neq \emptyset$ and $\hat{\theta}_\ell > \hat{\theta}_j$.[27]

Finally, consider $\pi \in F_\ell \cap F_j$. There exist numbers $\eta_1, \eta_2, \lambda_1, \lambda_2$ with $\eta_1$ and $\eta_2$ both positive such that $\eta_1\phi_\ell(i) + \lambda_1 \in \hat{V}_i(\mathcal{A})$ for all $i \in N$, $\eta_2\phi_j(i) + \lambda_2 \in \hat{V}_i(\mathcal{A})$ for all $i \in N$ and

$$\begin{cases} \sum_i \pi(i)(\eta_1\phi_\ell(i) + \lambda_1) = \sum_i \pi(i)(\eta_2\phi_j(i) + \lambda_2) \\ \sum_i c(i)(\eta_1\phi_\ell(i) + \lambda_1) = \sum_i c(i)(\eta_2\phi_j(i) + \lambda_2). \end{cases}$$

(Again, choosing these numbers sufficiently close to zero will ensure that $\eta_1\phi_\ell(i) + \lambda_1$ and $\eta_2\phi_j(i) + \lambda_2$ belong to $[-1, 1]$.) Therefore, there exist $\tilde{y}_i^j$ and $\tilde{y}_i^\ell$ in $Y$ such that $\hat{V}_i(\tilde{y}_i^\ell) = \eta_1\phi_\ell(i) + \lambda_1$ and $\hat{V}_i(\tilde{y}_i^j) = \eta_2\phi_j(i) + \lambda_2$ for all $i \in N$. Let $g_\ell$ be defined by $g_\ell(i) = \tilde{y}_i^\ell$ for all $i$ and $g_j$ be defined by $g_j(i) = \tilde{y}_i^j$ for all $i$. Observe that:

$$\arg\min_{p \in \Delta} \sum_i p(i)\hat{V}_i(g_j) = \arg\min_{p \in \Delta} \sum_i p(i)\hat{V}_i(f_j)$$

and

$$\arg\min_{p \in \Delta} \sum_i p(i)\hat{V}_i(g_\ell) = \arg\min_{p \in \Delta} \sum_i p(i)\hat{V}_i(f_\ell).$$

Therefore,

$$\arg\min_{p \in \Delta} \sum_i p(i)\hat{V}_i(g_j) \cap \arg\min_{p \in \Delta} \sum_i p(i)\hat{V}_i(g_\ell) \neq \emptyset.$$

Hence, by Axiom 13, $(g_j, \Delta) \sim (g_\ell, \Delta)$, i.e.,

$$\min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(g_j) = \min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(g_\ell).$$

But we also have

$$\min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(g_j) = \sum_i q_j(i)\hat{V}_i(g_j)$$

for any $q_j \in \mathcal{H}(\Psi_j \cap \Delta, c, \hat{\theta}_j)$ and similarly,

$$\min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(g_\ell) = \sum_i q_\ell(i)\hat{V}_i(g_\ell)$$

for any $q_\ell \in \mathcal{H}(\Psi_\ell \cap \Delta, c, \hat{\theta}_\ell)$. So, in particular, we have

$$\min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(g_j) = \sum_i \mathcal{H}(\pi, c, \hat{\theta}_j)(i)\hat{V}_i(g_j)$$

and

$$\min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(g_\ell) = \sum_i \mathcal{H}(\pi, c, \hat{\theta}_\ell)(i)\hat{V}_i(g_\ell).$$

---

[27]Assume $F_r \cap F_t = \emptyset$. Because $\Delta$ is a polytope, there is as sequence of adjacent facets $(F_{r_1}, F_{r_2}, \ldots, F_{r_s})$ such that $F_{r_1}$ is adjacent to $F_r$ and $F_{r_s}$ is adjacent to $F_t$. Thus, it must be the case that $\hat{\theta}_{r_k} \neq \hat{\theta}_{r_{k+1}}$ for some $k \in \{1, \ldots, s-1\}$.

Therefore, we must have:

$$\sum_i \mathcal{H}(\pi, c, \hat{\theta}_j)(i)\hat{V}_i(g_j) = \sum_i \mathcal{H}(\pi, c, \hat{\theta}_\ell)(i)\hat{V}_i(g_\ell)$$

$$\sum_i (\hat{\theta}_j \pi(i) + (1 - \hat{\theta}_j)c(i))\hat{V}_i(g_j) = \sum_i (\hat{\theta}_\ell \pi(i) + (1 - \hat{\theta}_\ell)c(i))\hat{V}_i(g_\ell)$$

$$\hat{\theta}_j \sum_i \pi(i)\hat{V}_i(g_j) + (1 - \hat{\theta}_j) \sum_i c(i)\hat{V}_i(g_j) = \hat{\theta}_\ell \sum_i \pi(i)\hat{V}_i(g_\ell) + (1 - \hat{\theta}_\ell) \sum_i c(i)\hat{V}_i(g_\ell).$$

But $\sum_i \pi(i)\hat{V}_i(g_j) = \sum_i \pi(i)\hat{V}_i(g_\ell)$, $\sum_i c(i)\hat{V}_i(g_j) = \sum_i c(i)\hat{V}_i(g_\ell)$, $\sum_i \pi(i)\hat{V}_i(g_\ell) < \sum_i c(i)\hat{V}_i(g_\ell)$ and $\sum_i \pi(i)\hat{V}_i(g_j) < \sum_i c(i)\hat{V}_i(g_j)$. Therefore $\hat{\theta}_\ell > \hat{\theta}_j$ implies

$$\hat{\theta}_j \sum_i \pi(i)\hat{V}_i(g_j) + (1 - \hat{\theta}_j) \sum_i c(i)\hat{V}_i(g_j) > \hat{\theta}_\ell \sum_i \pi(i)\hat{V}_i(g_\ell) + (1 - \hat{\theta}_\ell) \sum_i c(i)\hat{V}_i(g_\ell),$$

a contradiction. Therefore, $\hat{\theta}_j = \hat{\theta}_\ell$. Let $\hat{\theta} = \hat{\theta}_k$ for all $k \in \{1, \ldots, K\}$. Observe that because $\mathcal{H}(\Psi_j, c, \hat{\theta})$ are supporting hyperplanes of $\mathcal{F}(\Delta)$ for all $j \in \{1, \ldots, K\}$, $\mathcal{F}(\Delta) \subseteq \mathcal{H}(\Delta, c, \hat{\theta})$.

Now, consider any vertex $\pi_m^*$ of $\Delta$. Because $\pi_m^* \in \bigcap_{k \in J_m} F_k$,

$$\bigcap_{k \in J_m} \left( \arg \min_{p \in \Delta} \sum_i p(i)\hat{V}_i(f_k) \right) \neq \emptyset.$$

Then, Claim 13 implies:

$$\bigcap_{k \in J_m} \left( \arg \min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(f_k) \right) \neq \emptyset.$$

But:

$$\bigcap_{k \in J_m} \mathcal{H}(F_k, c, \hat{\theta}) = \mathcal{H}(\pi_m^*, c, \hat{\theta}).$$

We thus have:

$$\emptyset \neq \bigcap_{k \in J_m} \left( \arg \min_{p \in \mathcal{F}(\Delta)} \sum_i p(i)\hat{V}_i(f_k) \right) \subseteq \bigcap_{k \in J_m} \mathcal{H}(F_k, c, \hat{\theta}) = \mathcal{H}(\pi_m^*, c, \hat{\theta}).$$

Therefore, $\mathcal{H}(\pi_m^*, c, \hat{\theta}) \in \mathcal{F}(\Delta)$. Now consider any other vertex $\pi_r^*$ of $\Delta$. Then there exists a permutation $\varphi : N \to N$ (that depends of $\pi_r^*$) satisfying $\varphi(S_1) = S_1$ and $\varphi(S_2) = S_2$ such that $\pi_r^* = (\pi_m^*)^\varphi$. Because $\Delta = \Delta^\varphi$ for any such permutation, Claim 12 implies that:

$$\left( \mathcal{H}(\pi_m^*, c, \hat{\theta}) \right)^\varphi = \mathcal{H}((\pi_m^*)^\varphi, c, \hat{\theta}) = \mathcal{H}(\pi_r^*, c, \hat{\theta}) \in \mathcal{F}(\Delta).$$

Thus, for any vertex $\pi^*$ of $\Delta$, $\mathcal{H}(\pi^*, c, \hat{\theta}) \in \mathcal{F}(\Delta)$. Because $\Delta$ is polyhedral, $\mathcal{H}(\Delta) = co\{\mathcal{H}(\pi_m^*, c, \hat{\theta}) | \pi_m^* \in \Pi\}$. Therefore $\mathcal{H}(\Delta, c, \hat{\theta}) \subseteq \mathcal{F}(\Delta)$. Because we proved that $\mathcal{F}(\Delta) \subseteq \mathcal{H}(\Delta, c, \hat{\theta})$, we finally obtain $\mathcal{H}(\Delta, c, \hat{\theta}) = \mathcal{F}(\Delta)$.

It remains to consider the case $|S_i| = 1$ for some $i \in \{1, 2\}$ (if $|S_1| = |S_2| = 1$, the result follows trivially from Theorem 1). Assume without loss of generality, that $|S_1| = 1$ and $|S_2| > 1$. Then $\Delta = \alpha \delta_k + (1 - \alpha)\Delta_2$, for some $k \in N$, where $\delta_k$ is the probability distribution on $N$ defined by $\delta_k(k) = 1$. By Theorem 1, $\mathcal{F}(\Delta) = \alpha \delta_k + (1 - \alpha)\mathcal{F}(\Delta_2)$. But by the above result, $\mathcal{F}(\Delta_2) = \mathcal{H}(\Delta_2, c_2, \hat{\theta})$. Hence, $\mathcal{F}(\Delta) = \alpha \delta_k + (1 - \alpha)\mathcal{H}(\Delta_2, c_2, \hat{\theta}) = \mathcal{H}(\alpha \delta_k + (1 - \alpha)\Delta_2, \alpha \delta_k + (1 - \alpha)c_2, \hat{\theta})$, which proves the desired result. $\qquad \square$

**Claim 15.** *There exists $\theta \in [0, 1]$ such that for all subsets $S$ of $N$, $\mathcal{F}(\Delta(S)) = \mathcal{H}(\Delta(S), c_S, \theta)$.*

*Proof.* Let $S_1$ and $S_2$ be two subsets of $N$. We use the same notation as in Claim 14. By Claim 14, we know that there exist $\theta_1, \theta_2 \in [0, 1]$ such that $\mathcal{F}(\Delta_i) = \mathcal{H}(\Delta_i, c_i, \theta_i)$. What remains to be proved is that $\theta_1 = \theta_2$. Let $\alpha \in (0, 1)$ and $\Delta = \alpha \Delta_1 + (1 - \alpha)\Delta_2$. By Claim 14, we also know that there exists $\theta_3 \in [0, 1]$ such that:

$$\mathcal{F}(\Delta) = \mathcal{H}(\Delta, \alpha c_1 + (1 - \alpha)c_2, \theta_3) = \alpha \mathcal{H}(\Delta_1, c_1, \theta_3) + (1 - \alpha)\mathcal{H}(\Delta_2, c_2, \theta_3).$$

Finally, by Theorem 1,

$$\mathcal{F}(\Delta) = \alpha \mathcal{F}(\Delta_1) + (1 - \alpha)\mathcal{F}(\Delta_2).$$

Therefore,

$$\alpha \mathcal{H}(\Delta_1, c_1, \theta_1) + (1 - \alpha)\mathcal{H}(\Delta_2, c_2, \theta_2) = \alpha \mathcal{H}(\Delta_1, c_1, \theta_3) + (1 - \alpha)\mathcal{H}(\Delta_2, c_2, \theta_3)$$

which implies that $\theta_1 = \theta_2 = \theta_3$, the desired result. $\qquad \square$

Finally, let $\mathcal{P} = \sum_t \alpha_t(\Delta(S_t))$ with $\alpha_t \in [0, 1]$, $\sum_t \alpha_t = 1$ and $S_t \subseteq N$ for all $t$. By Claims 14 and 15, we know that

$$
\begin{aligned}
\mathcal{F}(\mathcal{P}) &= \mathcal{H}(\mathcal{P}, c(\mathcal{P}), \theta) \\
&= \{p \in \Delta(N) \mid \exists q \in \mathcal{P} \text{ s.t. } p = \theta q + (1 - \theta)c(\mathcal{P})\}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\min_{p \in \mathcal{F}(\mathcal{P})} p(i)\hat{V}_i(f(i)) &= \min_{q \in \mathcal{P}} \sum_i [\theta q(i) + (1 - \theta)c(\mathcal{P})(i)] \hat{V}_i(f(i)) \\
&= \theta \min_{p \in \mathcal{P}} \sum_i p(i)\hat{V}_i(f(i)) + (1 - \theta) \sum_i c(\mathcal{P})(i)\hat{V}_i(f(i)),
\end{aligned}
$$

which completes the proof of Theorem 2.

# References

Anscombe, F., and R. Aumann (1963): "A definition of subjective probability," *Annals of Mathematical Statistics*, 34, 199–205.

Arrow, K., and L. Hurwicz (1972): "An optimality criterion for decision making under ignorance," in *Uncertainty and Expectations in Economics*, ed. by C. Carter, and J. Ford, pp. 1–11. B. Blackwell, Oxford.

Barberà, S., W. Bossert, and K. Pattanaik (2004): "Ranking sets of objects," in *Handbook of Utility Theory. Volume II Extensions*, ed. by S. Barberà, P. Hammond, and C. Seidl, pp. 893–977. Kluwer Academic Publishers, Dordrecht.

Castagnoli, E., F. Maccheroni, and M. Marinacci (2003): "Expected Utility with Multiple Priors," *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, pp. 121–132.

Chateauneuf, A. (1991): "On the use of capacities in modeling uncertainty aversion and risk aversion," *Journal of Mathematical Economics*, 20, 343–369.

Chew, S. (1983): "A Generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox," *Econometrica*, 51, 1065–1092.

——— (1989): "Axiomatic utility theories with the betweenness property," *Annals of Operations Research*, 19, 273–298.

De Meyer, B., and P. Mongin (1995): "A Note on affine aggregation," *Economics Letters*, 47(2), 177–183.

Dekel, E. (1986): "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 40, 304–318.

Dempster, A. P. (1967): "probabilities induced by a multi-valued mapping," *Annals of Mathematical Statistics*, 38, 325–339.

Dhillon, A., and J. Mertens (1999): "Relative utilitarianism: An improved axiomatization," *Econometrica*, 67, 471–498.

Diamond, P. (1967): "Cardinal welfare, individualistic ethics, and interpersonal comparison of utility : Comment," *Journal of Political Economy*, 75, 765–766.

Drèze, J. (1987): "Decision theory with moral hazard and state-dependent preferences," in *Essays on Economic Decisions under Uncertainty*, ed. by J. Drèze, pp. 23–89. Cambridge University Press, Cambridge.

Epstein, L., and U. Segal (1992): "Quadratic social welfare functions," *Journal of Political Economy*, 99, 263–286.

Fishburn, P. (1970): *Utility Theory for Decision Making*. Wiley, New York.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2007): "Attitude toward imprecise information," *mimeo*, available http://eurequa.univ-paris1.fr/membres/tallon/GHTV-final.pdf.

Gajdos, T., J.-M. Tallon, and J.-C. Vergnaud (2004): "Decision making with imprecise probabilistic information," *Journal of Mathematical Economics*, 40, 647–681.

Gilboa, I., and D. Schmeidler (1989): "Maximin expected utility with a non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Grant, S., A. Kajii, B. Polak, and Z. Safra (2006): "A Generalized Utilitarianism and Harsanyi's Impartial Observer Theorem," Cowles Foundation Discussion Paper N$^o$. 1578.

Harsanyi, J. (1953): "Cardinal utility in welfare economics and in the theory of risk-taking," *Journal of Political Economy*, 61, 434–435.

——— (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons if Utility," *Journal of Political Economy*, 63, 309–321.

Harsanyi, J. C. (1977): *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press, Cambridge.

Jaffray, J.-Y. (1989): "Linear utility for belief functions," *Operations Research Letters*, 8, 107–112.

Karni, E. (1993): "Subjective expected utility with state-dependent preferences," *Journal of Economic Theory*, 60, 428–438.

——— (1998): "Impartiality: Definition and representation," *Econometrica*, 66, 1405–1515.

——— (2003): "Impartiality and interpersonal comparisons of variations in well-being," *Social Choice and Welfare*, 21, 95–111.

——— (2006): "Agency theory with maxmin expected utility players," mimeo, Johns Hopkins University, available at http://www.econ.jhu.edu/People/Karni/.

——— (2007): "Foundations of Bayesian theory," *Journal of Economic Theory*, 132, 167–188, Forthcoming.

Karni, E., and Z. Safra (2000): "An extension of a theorem of von Neumann and Morgenstern with an application to social choice theory," *Journal of Mathematical Economics*, 34, 315–327.

Karni, E., and D. Schmeidler (1981): "An expected utility theory for state-dependent preferences," Working paper 48-40 Foerder Institute for Economic Research, Tel Aviv University, Tel Aviv.

Karni, E., and J. A. Weymark (1998): "An informationally parsimonious impartial observer theorem," *Social Choice and Welfare*, 15, 321–332.

Klibanoff, P. (2001): "Characterizing uncertainty aversion through preference for mixtures," *Social Choice and Welfare*, 18, 289–301.

Luce, R. D., and D. Krantz (1971): "Conditional expected utility," *Econometrica*, 39, 253–271.

Maskin, E. (1979): "Decision-making under ignorance with implications for social choice," *Theory and Decision*, 11, 319–337.

Mongin, P. (2001): "The impartial observer theorem of social ethics," *Economics and Philosophy*, 17, 147–179.

Mongin, P., and C. d'Aspremont (1998): "Utility theory and ethics," in *Handbook of Utility Theory, I*, ed. by S. Barberà, P. Hammond, and C. Seidl, pp. 233–289. Kluwer, Dordrecht.

Moreno-Ternero, J., and J. Roemer (2005): "Objectivity, priority and the veil of ignorance," Core discussion paper 2005-81.

Rawls, J. (1971): *A Theory of Justice*. Harvard University Press, Cambridge, Mass.

——— (1974a): "Concepts of distributional equity: Some reasons for the maximin criterion," *American Economic Review*, 64, 141–146.

——— (1974b): "Reply to Alexander and Musgrave," *Quartely Journal of Economics*, 88, 633–655.

Schmeidler, D. (1984): "Subjective Probability and Expected Utility Without Additivity," IMA preprint series.

Segal, U. (2000): "Let's agree that all dictatorship are equally bad," *Journal of Political Economy*, 108, 569–589.

Sen, A. (1976): "Welfare inequalities and Rawlsian axiomatics," *Theory and Decision*, 7, 243–262.

Shafer, G. (1976): *A methematical theory of evidence*. Princeton University Press, Princeton, New Jersey.

Skiadas, C. (1997a): "Conditioning and aggregation of preferences," *Econometrica*, 65, 242–271.

——— (1997b): "Subjective probability under additive aggregation of conditional preferences," *Journal of Economic Theory*, 76, 347–367.

Vickrey, W. (1945): "Measuring marginal utility by reaction to risk," *Econometrica*, 13, 319–333.

Wang, T. (2003): "Two classes of multiple priors," Discussion paper, Department of Finance, University British Columbia,.

Weymark, J. A. (1991): "A reconsideration of the Harsanyi-Sen debate on utilitarianism," in *Interpersonnal Comparisons of Well-Being*, ed. by J. Elster, and J. Roemer, pp. 255–320, Cambridge. Cambridge University Press.

——— (2005): "Measurement theory and the foundations of utilitarianism," *Social Choice and Welfare*, 25, 527–555.