GENERALIZED UTILITARIANISM AND HARSANYI'S IMPARTIAL OBSERVER THEOREM

By

Simon Grant, Atsushi Kajii, Ben Polak and Zvi Safra

September 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1578

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

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

September 11, 2006

Generalized Utilitarianism and Harsanyi's Impartial Observer Theorem*

Abstract

We provide an axiomatization of *generalized* utilitarian social welfare functions in the context of Harsanyi's impartial observer theorem. To do this, we reformulate Harsanyi's problem such that lotteries over identity (accidents of birth) and lotteries over outcomes (life chances) are independent. We show how to accommodate (first) Diamond's critique concerning fairness and (second) Pattanaik's critique concerning differing attitudes toward risk. In each case, we show what separates them from Harsanyi by showing what extra axioms return us to Harsanyi. Thus we provide two new axiomatizations of Harsanyi's utilitarianism..

Keywords: generalized utilitarianism, impartial observer, social welfare function, fairness, ex ante egalitarianism.

JEL Classification: D63, D71

Simon Grant
Department of Economics
Rice University

Atsushi Kajii
Insitute of Economic Research
Kyoto University

Ben Polak
Department of Economics & School of Management
Yale University

Zvi Safra
Tel Aviv University and College of Management

*We thank Jurgen Eichberger, Edi Karni, Bart Lipman, Stephen Morris, Heve Moulin, David Pearce, John Quiggin, John Roemer and John Weymark for many helpful comments. Zvi Safra thanks the Israel Science Foundation (grant no. 1299/05) and the Henry Crown Institute of Business Research for their support.

1 Introduction

This paper revisits Harsanyi's (1953, 1955) utilitarian impartial observer theorem. Consider a society of individuals $\mathcal{I}$. The society has to choose among different social policies, each of which induces a probability distribution or 'lottery' ℓ over a set of social outcomes $\mathcal{X}$. Each individual i has preferences $\succsim_i$ over these lotteries. These preferences are known, and they differ.

To help choose among social policies, Harsanyi proposed that each individual should imagine herself as an 'impartial observer' who does not know which person she will be. That is, the impartial observer faces not only the real lottery ℓ over the social outcomes in $\mathcal{X}$, but also a hypothetical lottery z over which identity in $\mathcal{I}$ she will assume. In forming preferences $\succsim$ over all such 'extended lotteries', the impartial observer is forced to make interpersonal comparisons; for example, she is forced to compare being person i in social state x with being person j in social state x'.

Harsanyi assumed that when the impartial observer imagines herself being person i she adopts person i's preferences over the outcome lotteries. He also assumed that all individuals are expected utility maximizers, and that they continue to be so in the role of the impartial observer. Harsanyi argued that these "Bayesian rationality" axioms force the impartial observer to be a (weighted) utilitarian. More formally, over all extended lotteries (z, ℓ) in which the identity lottery and the outcome lotteries are independently distributed, the impartial observer's preferences admit a representation of the form

$$V(z, \ell) = \sum_i z_i U_i(\ell)$$

where z_i is the probability of assuming person i's identity and $U_i(\ell) := \int_{\mathcal{X}} u_i(x)\, \ell(dx)$ is person i's von Neuman-Morgenstern expected utility for the outcome lottery ℓ.

Harsanyi's utilitarianism has attracted many criticisms.[1] We confront just two: one associated

[1] Sen(1970, 1977), Weymark (1991) and others have observed that Harsanyi's utilitarianism is not 'welfare' utilitarianism in the nineteenth-century sense. Harsanyi's representation is additive in individuals' von Neumann-Morgenstern utilities but not necessarily in individuals' 'welfares'. We will return to this issue in section 7.

There have been other criticisms beyond the scope of this paper. For example, it is unclear that placing different individuals in the role of the impartial observer will lead them to agree on the appropriate interpersonal comparisons. Even within Harsanyi's utilitarian form, the impartial observer has to decide among affine transformations which von Neumann-Morgenstern utility function to use for each individual, and which is the appropriate weighted identity lottery (see, for example, Mongin (2001)). It is also unclear that actual individuals will (or even should) feel bound with Diamond (1967) concerning fairness; and one associated with Pattanaik (1968) concerning different attitudes toward risk. To illustrate both criticisms, consider two individuals, i and j and two social outcomes x_i and x_j. Person i strictly prefers outcome x_i to outcome x_j, but person j strictly prefers x_j to x_i. Perhaps, there is some (possibly indivisible) good, and x_i is the state in which person i has the good while x_j is the state in which person j has it. Suppose that the impartial observer would be indifferent being person i in state x_i and being person j in state x_j; hence $u_i(x_i) = u_j(x_j) =: u^H$. She is also indifferent between being i in x_j and being j in x_i; hence $u_i(x_j) = u_j(x_i) =: u^L$. And she strictly prefers the first pair (having the good) to the second (not having the good); hence $u^H > u^L$.

To illustrate Diamond's criticism, consider the two extended lotteries illustrated in tables (a) and (b) in which rows are the people and columns are the outcomes.

	x_i	x_j
i	1/2	0
j	1/2	0

(a)

	x_i	x_j
i	1/4	1/4
j	1/4	1/4

(b)

In each, the impartial observer has a half chance of being person i or person j. But in table (a), the good is simply given outright to person i: outcome x_i has probability 1. In table (b), the good is allocated by tossing a coin: the outcomes x_i and x_j each have probability 1/2. Diamond argued that a fair-minded person might prefer the second allocation policy since it gives each person a "fair shake".[2] But Harsanyi's utilitarian impartial observer is indifferent to such considerations of fairness. Each policy (or its associated extended lottery) involves a half chance of getting the good and hence yields the impartial observer $\frac{1}{2}u^H + \frac{1}{2}u^L$. The impartial observer cares only about her total chance of getting the good, not how this chance is distributed between person i and j.

To illustrate Pattanaik's criticism, consider the two extended lotteries illustrated in tables (c)

by their evaluations in the role of the impartial observer once they 'resume' their role as real people (see, for example, Broome (1991)).

[2] Societies often use both simple lotteries and weighted lotteries to allocate goods (and bads), presumably for fairness considerations. Examples include the draft, kidney machines, oversubscribed events, schools, and public housing, and even whom should be thrown out of a lifeboat! For a long list and an enlightening discussion, see Elster (1989).

and (d).

	x_i	x_j
i	1/2	1/2
j	0	0

(c)

	x_i	x_j
i	0	0
j	1/2	1/2

(d)

In each, the impartial observer has a half chance of being in state x_i or state x_j, and hence a half chance of getting the good. But in (c), the impartial observer faces this risk as person i, while in (d), she faces the risk as person j. Pattanaik argued that if person i is more comfortable facing such a risk than is person j then the impartial observer might prefer to face the risk as person i. But Harsanyi's utilitarian impartial observer is indifferent to such considerations of risk attitude. Each of the extended lotteries (c) and (d) again yield $\frac{1}{2}u^H + \frac{1}{2}u^L$. Thus, the impartial does not care who faces this risk.

Most attempts to adapt Harsanyi's axioms to deal with these concerns have focussed on the independence axiom of expected utility theory. For example, Karni & Safra (2002) relax independence for the individual preferences, while Epstein & Segal (1992) relax independence for the impartial observer.[3] It is not clear, however that the independence axiom pe se is at the crux of the disagreements between Harsanyi and his critics.[4]

In his own response to Diamond, Harsanyi (1975) argued that, even if randomizations were of value for promoting 'fairness' (which he doubted), any explicit randomization is superfluous since 'the great lottery of (pre-)life' may be viewed as having already given each child an equal chance of being each individual. That is, for Harsanyi, it does not matter whether a good is allocated entirely by 'accidents of birth' (as in the extended lottery (a) above), or whether the good is allocated entirely by individuals' 'life chances' (as in the extended lottery (c)): for Harsanyi, they are equivalent. The dispute between Diamond and Harsanyi thus seems to rest on whether or not

[3] Strictly speaking, Epstein & Segal's paper is in the context of Harsanyi's (1955) aggregation theorem, not his impartial observer's theorem.

[4] Broome (1991) also argues that independence per se is not the key issue for Harsanyi's utilitarianism. He expands the outcome set to allow us to distinguish not just who gets the good but also the means of allocation. Debates about independence then become debates about "rational indifference": that is, which such *outcomes* should be viewed as equivalent. Somewhat analogously, we emphasize which *lotteries* should be viewed as equivalent. Broome's own critiques, however, both of the impartial observer theorem and of Diamond's notion of fairness (see, for example, Broome 1984) are on other grounds.

we think (imagined) lotteries over identities are indeed equivalent to (real) lotteries over outcomes. We will argue that such an equivalence is also at the heart of Pattanaik and Harsanyi's dispute.

Which of Harsanyi's axioms yield the equivalence between identity lotteries and outcome lotteries? In most formulations of Harsanyi's theorem, the impartial observer is assumed to form preferences not just over extended lotteries in which the identity lottery and outcome lottery are independently distributed but over the entire set of joint distributions $\triangle(\mathcal{I} \times \mathcal{X})$ over identities and outcomes.[5] In such a set up, it is hard even to distinguish the outcome lottery from the identity lottery since the resolution of identity can partially or fully resolve the outcome. For example, the impartial observer could face a joint distribution in which, if she becomes person i then society holds the outcome lottery ℓ, but if she becomes person j then social outcome x obtains for sure. Harsanyi's utilitarianism comes from later restricting the representation to the set of independent or 'product' lotteries.

Suppose instead that we restrict attention from the outset to product lotteries, $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$, That is, the impartial observer only forms preferences over extended lotteries in which the outcome lottery she faces is the same regardless of which identity she assumes. This setting seems closer to most informal accounts of Harsanyi's thought experiment. Suppose we then impose each of Harsanyi axioms in this simpler setting: in particular, each individual satisfies the independence axiom for outcome lotteries, that the impartial observer respects these individuals' preferences (and so inherits this independence), and that the impartial observer herself satisfies independence for identity lotteries. In this case, we are no longer forced to Harsanyi's utilitarianism. Instead (see theorem 2 below), we obtain a generalized (weighted) utilitarian representation:

$$V(z, \ell) = \sum_i z_i \phi_i(U_i(\ell))$$

where z_i is again the probability of assuming person i's identity and $U_i(\ell)$ is again person i's expected utility from the outcome lottery ℓ, but each $\phi_i(.)$ is a (possibly non-linear) transformation of person i's expected utility. Generalized utilitarianism is well known to applied welfare economists but has, till now, lacked axiomatic foundations.

[5] See, for example, Weymark (1991). An exception is Safra & Weissengrin (2002).

Generalized utilitarianism can accommodate both Diamond and Pattanaik. Diamond's concern for fairness can be accommodated if the ϕ_i-functions are concave.[6] Pattanaik's concern for different risk attitudes can be accommodated by allowing the ϕ_i-functions to differ in their degree of concavity or convexity.[7] Harsanyi's utilitarianism can be thought of as the special case where each ϕ_i is affine.

Since generalized utilitarianism nests Diamond's, Pattanaik's and Harsanyi's models, we proceed to axiomatize each in turn. We show formally that what separates Harsanyi from Diamond is that Harsanyi assumes that identity lotteries (accidents of birth) and outcome lotteries (life chances) are 'equivalent' in the sense of being indifferent. By contrast, Diamond assumes a preference for life chances; that is, in the example above, he prefers that the good be allocated by a real outcome lottery (as it is in (b),(c) and (d)) than by the imaginary chance of assuming the right identity (as it is in (a)).[8] What separates Harsanyi from Pattanaik is again that Harsanyi assumes that identity lotteries and outcome lotteries are 'equivalent' but this time in the sense that all axioms are symmetric across the two types of lottery. By contrast, Pattanaik imposes independence over identity lotteries directly on the impartial observer but allows her only to inherit independence over outcome lotteries indirectly from the individuals who will actually face those lotteries.

These two different notions of equivalence yield two different new axiomatizations of Harsanyi's utilitarianism. Each is built by adding an axiom to those that delivered generalized utilitarianism. More abstractly, we can also think of these as two new bi-linearity theorems for products of lottery spaces.

Although restricting attention to product lotteries seems natural and yields the results we want, it comes at a technical cost in that we can no longer rely on well-known results from decision

[6] In our story, we have $\phi_i(u_i(x_i)) = \phi_j(u_j(x_j)) > \phi_i(u_i(x_j)) = \phi_j(u_j(x_i))$. Thus, if the ϕ-functions are (strictly) concave, the impartial observer evaluatation of allocation policy (c) $\phi_i\left(\frac{1}{2}u_i(x_i) + \frac{1}{2}u_i(x_j)\right) > \frac{1}{2}\phi_i(u_i(x_i)) + \frac{1}{2}\phi_i(u_i(x_j)) = \frac{1}{2}\phi_i(u_i(x_i)) + \frac{1}{2}\phi_j(u_j(x_i))$, her evaluation of policy (a). The argument comparing (b) and (a) is similar.

[7] For example, if ϕ_i is strictly concave but ϕ_j is linear, then the impartial observer's evaluation of policy (c) $\phi_i\left(\frac{1}{2}u_i(x_i) + \frac{1}{2}u_i(x_j)\right) > \frac{1}{2}\phi_i(u_i(x_i)) + \frac{1}{2}\phi_i(u_i(x_j)) = \frac{1}{2}\phi_j(u_j(x_j)) + \frac{1}{2}\phi_j(u_j(x_i)) = \phi_j\left(\frac{1}{2}u_j(x_j) + \frac{1}{2}u_j(x_i)\right)$, her evaluation of policy (d).

[8] Indeed, Diamond's original example was between allocation policies like (a) and (c). He strictly preferred (c).

theory. First, the set of product lotteries $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$ is not a convex subset of $\triangle(\mathcal{I} \times \mathcal{X})$. In particular, we have to be careful that our independence axioms only involve mixtures that remain in the set of product lotteries. Fortunately, we can adapt some axioms developed by Fishburn (1982) to study mixed strategies in games.[9]

Second, the set of product lotteries does not have a nice recursive structure. With the full set of joint distributions, it is as if each individual i faces his own 'personal' outcome lottery. Each vector of personal lotteries induces a vector of individual utilities. In this setting, by changing person i's personal outcome lottery, holding fixed the other people's lotteries, we can induce a rich range of these individual-utility vectors. With only the set of product lotteries, however, each person faces the same outcome lottery (although their preferences over those lotteries may differ). This limits the set of individual-utility vectors we can induce. If this set is not rich enough then our axioms will lack bite.

The richness of this set depends on the degree to which different individuals differ in their ranking of outcome lotteries, and the degree to which different outcome lotteries lead the impartial observer to different welfare ranking of individuals. Most of the results below require relatively mild richness conditions: either that individuals do not all agree in their preference for one outcome over another or that the impartial observer does not always prefer to be one individual rather than another.[10] For one result, however, – our second axiomatization of Harsanyi's utilitarianism – we use a stronger condition that requires there to be three or more agents.

Section 2 sets up the product-lottery framework. Section 3 shows that, if we adapt the Harsanyi axioms to that framework, we obtain an axiomatization of generalized utilitarianism. Section 4 provides an additional axiom to accommodate Diamond's concern for fairness. It shows that forcing the impartial observer to ignore these concerns corresponds to being indifferent between identity and outcome lotteries. This yields our first new axiomatization of Harsanyi's utilitarianism. Section 5 shows how to accommodate Paittanaik's concerns for different attitudes toward risk. It shows that forcing the impartial observer to ignore these concerns corresponds to imposing

[9] Axioms of this form were also used by Safra & Weissengrin (2002).

[10] But, in most cases, these conditions are essential to the results. We provide counter-examples in appendix A.

a weak form of independence axiom directly on the impartial observer. Going further and imposing all axioms symmetrically on identity and outcome lotteries (given a rich environment) yields our second new axiomatization of Harsanyi. There is a very large literature discussing Harsanyi, and we do not attempt to summarize it, but 6 discusses some related technical papers. In particular, it considers what happens if we impose less structure and what happens if, like Harsanyi, we impose more. Section 7 (following, for example, Weymark (1991)) introduces an explicit notion of comparable welfare, and uses it to interpret some of our representation results. Appendix A provides counter-examples to show that our axioms are essential. Appendix B provides those proofs not in the text.

2 Set up and Notation

Let society consist of a finite set of individuals $\mathcal{I} = \{1, \ldots, I\}$, $I \geq 2$, with generic elements i and j. The set of final outcomes or social states is denoted by $\mathcal{X}$ with generic element x. The set $\mathcal{X}$ is assumed to have more than one element and to be a compact metrizable space and associated with it is the set of events $\mathcal{E}$, which is taken to be the Borel sigma-algebra of $\mathcal{X}$. Let $\triangle(\mathcal{X})$ (with generic element ℓ) denote the set of *outcome lotteries*; that is the set of probability measures on $(\mathcal{X}, \mathcal{E})$ endowed with the weak convergence topology. We will sometimes refer to these lotteries over outcomes as *life chances*: they represent the real risks faced by each individual in their real lives. With slight abuse of notation, we will let x or sometimes $[x]$ denote the degenerate outcome lottery that assigns probability weight 1 to social state x.

Each individual i in $\mathcal{I}$, is endowed with a preference relation $\succsim_i$ defined over the set of life-chances $\triangle(\mathcal{X})$. We assume throughout that for each i in $\mathcal{I}$, the preference relation $\succsim_i$ is a complete, transitive and continuous binary relation on $\triangle(\mathcal{X})$, and that its asymmetric part $\succ_i$ is non-empty. Hence for each $\succsim_i$ there exists a non-constant function $V_i : \triangle(\mathcal{X}) \to \mathbb{R}$, satisfying for any ℓ and ℓ' in $\triangle(\mathcal{X})$, $V_i(\ell) \geq V_i(\ell')$ if and only if $\ell \succsim_i \ell'$. In summary, a society may be characterized by the tuple $\langle \mathcal{X}, \mathcal{E}, \mathcal{I}, \{\succsim_i\}_{i \in \mathcal{I}} \rangle$.

In Harsanyi's story, the impartial observer imagines herself behind a veil of ignorance, uncertain about which identity she will assume in the given society. Let $\triangle(\mathcal{I})$ denote the set of *identity* *lotteries* on I. Let z denote the typical element of $\triangle(\mathcal{I})$, and let z_i denote the probability assigned by the identity lottery z to individual i. We will sometimes refer to these lotteries over identity as *accidents of birth*: they represent the imaginary risks in the mind of the impartial observer of being born as someone else. With slight abuse of notation, we will let i or sometimes $[i]$ denote the degenerate identity lottery that assigns probability weight 1 to the impartial observer's assuming the identity of individual i.

As discussed above, we assume that the outcome and identity lotteries faced by the impartial observer are independently distributed; that is, she faces a product lottery $(z, \ell) \in \triangle(\mathcal{I}) \times \triangle(\mathcal{X})$. We shall sometimes refer to this as a product *identity-outcome lottery* or (where no confusion arises) simply as an product lottery.

The impartial observer is endowed with a preference relation $\succsim$ defined over $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$. We assume throughout that $\succsim$ is complete, transitive and continuous, and that its asymmetric part $\succ$ is non-empty, and so it admits a (non-trivial) continuous representation $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$. That is, for any pair of product lotteries, (z, ℓ) and (z', ℓ'), $(z, \ell) \succsim (z', \ell')$ if and only if $V(z, \ell) \geq V(z', \ell')$.

3 Generalized Utilitarianism

In this section, we adapt the axioms from Harsanyi's impartial observer theorem to apply to the product-lottery framework, add a richness condition that there is some disagreement in the underlying individual preferences over policies, and hence provide an axiomatization of generalized utilitarianism.

The first axiom is Harsanyi's acceptance principle. In degenerate product lotteries of the form (i, ℓ) or (i, ℓ'), the impartial observer knows she will assume identity i for sure. The acceptance principle requires that, in this case, the impartial observer's preferences $\succsim$ must coincide with that individual's preferences $\succsim_i$ over life chances.

The Acceptance Principle. *For all i in $\mathcal{I}$ and all $\ell, \ell' \in \triangle(\mathcal{X})$, $\ell \succsim_i \ell'$ if and only if $(i, \ell) \succsim (i, \ell')$.*

Second, following Harsanyi, we assume that each individual i's preferences satisfy the independence axiom for the lotteries he faces, that is over the set of outcome lotteries $\triangle(\mathcal{X})$. We state this axiom in a slightly non-standard form.

Independence over Outcome Lotteries (for Individual i). *Suppose ℓ, $\ell' \in \triangle(\mathcal{X})$ are such that $\ell \sim_i \ell'$. Then, for all $\tilde{\ell}$, $\tilde{\ell}' \in \triangle(\mathcal{X})$, $\tilde{\ell} \succsim_i \tilde{\ell}'$ if and only if $\alpha\tilde{\ell}+(1-\alpha)\,\ell \succcurlyeq_i \alpha\tilde{\ell}'+(1-\alpha)\,\ell'$ for all α in $(0,1]$.*

Notice that the two outcome lotteries, ℓ and ℓ' that are "mixed in" with weight $(1-\alpha)$ to $\tilde{\ell}$ and $\tilde{\ell}'$ are themselves indifferent. The axiom states that 'mixing in' two indifferent lotteries (with equal weight) preserves the original preference order between $\tilde{\ell}$ and $\tilde{\ell}'$ prior to mixing.

The standard version of the independence axiom states that for all $\tilde{\ell}, \tilde{\ell}', \tilde{\ell}''$ in $\triangle(\mathcal{X})$, $\tilde{\ell} \succsim_i \tilde{\ell}'$ if and only if $\alpha\tilde{\ell} + (1-\alpha)\,\tilde{\ell}'' \succsim_i \alpha\tilde{\ell}' + (1-\alpha)\,\tilde{\ell}''$ for all α in $(0,1]$. That is, in its standard form, the *same* outcome lottery $\tilde{\ell}''$ is 'mixed-in" with weight $(1-\alpha)$ to $\tilde{\ell}$ and $\tilde{\ell}'$. It is a simple exercise to show that these two versions of independence are equivalent.[11] We use the form above to emphasize the symmetry with the next axiom.

Third, following Harsanyi, we assume that the impartial observer's preferences also satisfy independence. Here, however, we need to be careful. First, the set of product lotteries $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$ is not a convex subset of $\triangle(\mathcal{I} \times \mathcal{X})$ and hence not all probability mixtures of product lotteries are well defined. Second, the impartial observer faces two types of lottery, over outcomes and over identities. The former risks are faced directly by real people, but are only faced indirectly by the impartial observer once she assumes the identity of a real person. Once we impose the independence axiom on each individual's preferences, the acceptance principle already ensures that the impartial observer respects those individual preferences (and hence independence) over outcome lotteries. Identity lotteries, however, are not faced by real people, but only faced by the impartial observer in her thought experiment. Thus, to get independence over identity lotteries, we need to impose it directly on the impartial observer's preferences. The following axiom achieves this.[12]

[11] In particular, the current form immediately implies the standard form. For the other direction, the standard form implies that $\ell \sim_i \ell'$ if and only if $\alpha\tilde{\ell}' + (1-\alpha)\,\ell \sim_i \alpha\tilde{\ell}' + (1-\alpha)\,\ell'$.

[12] This axiom is based on Fishburn's (1982, p.88) and Safra & Weissengrin's (2003) substitution axioms for

Independence over Identity Lotteries (for the Impartial Observer). *Suppose* (z,ℓ), $(z',\ell') \in \triangle(\mathcal{I}) \times \triangle(\mathcal{X})$ *are such that* $(z,\ell) \sim (z',\ell')$. *Then, for all* $\tilde{z}$, $\tilde{z}' \in \triangle(\mathcal{I})$: $(\tilde{z},\ell) \succsim (\tilde{z}',\ell')$ *if and only if* $(\alpha\tilde{z} + (1-\alpha)z, \ell) \succsim (\alpha\tilde{z}' + (1-\alpha)z', \ell')$ *for all* α *in* $(0,1]$.

To understand this axiom, first notice that the two mixtures on the right side of the implication are identical to $\alpha(\tilde{z},\ell)+(1-\alpha)(z,\ell)$ and $\alpha(\tilde{z}',\ell')+(1-\alpha)(z',\ell')$ respectively. These two mixtures of product lotteries are well defined: they mix identity lotteries holding the outcome lottery fixed. Second, notice that the two product lotteries, (z,ℓ) and (z',ℓ'), that are 'mixed in' with weight $(1-\alpha)$ are themselves indifferent. The axiom states that 'mixing in' two indifferent lotteries (with equal weight) preserves the the original preference order between $(\tilde{z},\ell)$ and $(\tilde{z}',\ell')$ prior to mixing.[13] Finally, notice that this axiom only applies to mixtures of identity lotteries holding the outcome lotteries fixed, not to the opposite case: mixtures of outcome lotteries holding the identity lotteries fixed. We will discuss this 'opposite' axiom in section 5 below.

How do these axioms relate to the discussion in the introduction? Given acceptance, the impartial observer inherits her preferences over outcome lotteries from the preferences of the individuals who will face those lotteries and whose identities she will assume. In particular, the impartial observer inherits independence over outcome lotteries indirectly from individuals' preferences. By contrast, we can think of Harsanyi imposing such independence directly. We will show in section 5 that this distinction allows us to accommodate Pattanaik's concern about different individuals' different attitudes toward risk. None the axioms above say anything about how the impartial observer compares identity and outcome lotteries. In particular, unlike Harsanyi, we do not implicitly assume that she is indifferent between accidents of birth and life chances. We

product lottery spaces. Their axioms, however, apply whereever probability mixtures are well defined in this space. For example, in our context, their axioms would apply to mixtures of outcome lotteries. We only allow mixtures of identity lotteries. In this respect, our axiom is similar to Karni & Safra's (2000) 'constrained independence' axiom, but their axiom applies to all joint distributions over identities and outcomes, not just to product lotteries.

[13] One technical remark might interest some readers. In the axiom, we allow the mixing of identity lotteries to occur at two different outcome lotteries; that is, we do not restrict ℓ to equal ℓ'. We could define a weaker axiom – call in *conditional independence* – that simply imposes independence over identity lotteries at each *fixed* outcome lottery $\bar{\ell}$. That is, for all $\bar{\ell} \in \triangle(\mathcal{X})$, if $z, z' \in \triangle(\mathcal{I})$ are such that $(z,\bar{\ell}) \sim (z',\bar{\ell})$ then for all $\tilde{z}, \tilde{z}' \in \triangle(\mathcal{I})$, $(\tilde{z},\bar{\ell}) \succsim (\tilde{z}',\bar{\ell})$ if and only if $(\alpha\tilde{z} + (1-\alpha)z, \bar{\ell}) \succsim (\alpha\tilde{z}' + (1-\alpha)z', \bar{\ell})$ for all α in $(0,1]$. Our stronger axiom is necessary for the representation results that follow. To show this, example 2 in appendix A shows that preferences can satisfy the acceptance principle, independence over outcome lotteries for individuals, and *conditional* independence over identity lotteries for the impartial observer but not satisfy the (unconditional) independence axiom over identity lotteries defined above.

will show in section 4 that this allows us accommodate Diamond's concerns about fairness.

To obtain our representation results, we work with a richness condition on the domain of individual preferences: we assume that none of the outcome lotteries under consideration are Pareto dominated.

Absence of Unanimity *For all ℓ, $\ell' \in \triangle(\mathcal{X})$ if $\ell \succ_i \ell'$ for some i in $\mathcal{I}$ then there exists j in $\mathcal{I}$ such that $\ell' \succ_j \ell$.*

This condition is perhaps a natural restriction in the context of Harsanyi's thought experiment. That exercise is motivated by the need to make social choices when agents disagree. We do not need to imagine ourselves as an impartial observer facing a identity lottery to rule out social alternatives that are Pareto dominated.[14]

The following lemma does yet not impose independence over outcome lotteries on individuals and hence yields a more general representation. The idea for this lemma comes from Karni & Safra (2000) but they work with the full set of joint distributions $\triangle(\mathcal{I} \times \mathcal{X})$ whereas we are restricted to the set of product lotteries $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$.

Lemma 1 *Suppose absence of unanimity applies. Then the impartial observer satisfies the acceptance principle and independence over identity lotteries if and only if there exist a continuous function $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ that represents $\succsim$, and, for each individual i in $\mathcal{I}$, a function $V_i : \triangle(\mathcal{X}) \to \mathbb{R}$, that represents $\succsim_i$, such that for all (z, ℓ) in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$,*

$$V(z, \ell) = \sum_{i=1}^{I} z_i V_i(\ell).$$

Moreover the functions V_i are unique up to common affine transformations.

The proof is in the appendix but a sketch is as follows. The first step follows Karni & Safra (2000).[15] Fix some outcome lottery ℓ^1. Notice that, by independence, there exist two individual

[14] For example, if absence of unanimity fails but individuals satisfy independence, we could first discard all those outcomes that are Pareto dominated by other outcome lotteries and then carry out the Harsanyi thought experiment on the set of lotteries over the remaining undominated outcomes. Given independence, the set of undominated lotteries over the original outcomes is equal to the set of lotteries over the undominated outcomes, and absence of unanimity will hold for the new domain.

[15] An alternate strategy would be to prove it as a special case of Theorem 8.

i^1 and i_1 such that $\left(i^1,\ell^1\right) \succsim \left(z,\ell^1\right) \succsim \left(i_1,\ell^1\right)$ for all z; that is, i^1 is a best identity and i_1 is a worst identity to assume given that the impartial observer will then face the outcome lottery ℓ^1. Next, construct a representation for all product lotteries (z,ℓ) such that $\left(i^1,\ell^1\right) \succsim (z,\ell) \succsim \left(i_1,\ell^1\right)$ by finding the weight β in $[0,1]$ such that the identity lottery $\beta\left[i^1\right]+(1-\beta)\left[i_1\right]$ facing the outcome lottery ℓ^1 is indifferent to the identity lottery z facing the outcome lottery ℓ. Set $V(z,\ell) := \beta$. Independence over identity lotteries ensures that this representation is unique and affine.[16]

Up to this point the argument resembles a standard proof of the von Neumann-Morgenstern theorem except that (so far) we have only constructed an affine representation for those identity-outcome lotteries (z,ℓ) such that $\left(i^1,\ell^1\right) \succsim (z,\ell) \succsim \left(i_1,\ell^1\right)$; that is, those (z,ℓ) that are indifferent to $\left(z',\ell^1\right)$ for some identity lottery z' at the particular fixed outcome lottery ℓ^1. Loosely speaking, we have only represented an 'interval' of the impartial observer's preferences. To go further, Karni & Safra exploit the fact that (for them) each individual faces a different outcome lottery. Instead, we rely on our richness condition.

Lemma 9 in the appendix shows that, given absence of unanimity, we need at most two 'fixed' outcome lotteries (i.e., at most two 'intervals') to cover the entire range of the impartial observer's preferences. To keep the notation consistent with that in the appendix, let these two outcome lotteries be denoted ℓ^1 and ℓ_2. That is, there exists two outcome lotteries ℓ^1 and ℓ_2 such that for all product lotteries (z,ℓ) either $(z,\ell) \sim \left(z',\ell^1\right)$ for some z', or $(z,\ell) \sim (z'',\ell_2)$ for some z'' or both. Moreover we can choose ℓ^1 and ℓ_2 such that their 'intervals' overlap. With this step in hand, standard arguments ensure that the affine representations are consistent on the two 'intervals', and satisfy the usual uniqueness condition.

Finally, applying affinity implies that the representation takes the form $\sum_i z_i V(i,\ell)$, and, by acceptance, we can set $V(i,\cdot) := V_i(\cdot)$ to complete the proof.

In section 6, we show that without absence of unanimity, we can still obtain a representation similar to that in Lemma 1 but it will lack the uniqueness properties.

[16] For uniqueness: strictly speaking, we need $\left(i^1,\ell^1\right) \succ \left(i_1,\ell^1\right)$. For affinity, this step is where the weaker conditional independence discussed in footnote X would not be sufficient: the product lotteries (z,ℓ) we are representing contain outcome lotteries other than just ℓ^1.

The representation in Lemma 1 puts no restriction on the V_i-functions. But, if we now add the assumption that each individual satisfies independence over outcome lotteries then it follows immediately that each V_i-function must be a strictly increasing transformation of a von Neumann-Morgenstern expected-utility representation. Thus, we obtain a generalized utilitarian representation.

Theorem 2 (Generalized Utilitarianism) *Suppose that absence of unanimity applies. Then the following are equivalent:*

(a) The impartial observer satisfies the acceptance principle and independence over identity lotteries, and each individual satisfies independence over outcome lotteries

(b) There exist a continuous function $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ that represents $\succsim$, and, for each individual i in $\mathcal{I}$, a von Neumann-Morgenstern function $U_i : \triangle(\mathcal{X}) \to \mathbb{R}$ that represents $\succsim_i$ and a continuous, strictly increasing function $\phi_i : \mathbb{R} \to \mathbb{R}$, such that, for all (z, ℓ) in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$,

$$V(z, \ell) = \sum_{i=1}^{I} z_i \phi_i [U_i(\ell)] .$$

where, for each i, $U_i(\ell) = \int_{\mathcal{X}} U_i(x)\, \ell(dx)$. Moreover the functions U_i are unique up to affine transformations, and the composite functions $\phi_i \circ U_i$ are unique up to a common affine transformation.

Notice that, while the representation of each individual's preferences U_i is affine in outcome lotteries, in general, the representation of the impartial observer's preferences V is not.

4 Accommodating Fairness: Harsanyi vs. Diamond.

In this section, we first introduce a new axiom on the impartial observer's preferences to ensure that the generalized utilitarian representation is concave and hence accommodates Diamond. We then show that tightening this axiom yields Harsanyi's utilitarianism.

So far we have placed no restriction on the shape of the ϕ_i-functions except that they are increasing. An analogy may help to see why we want concavity. In a standard utilitarian social welfare function, each u_i-function maps individual i's income to an individual utility. These incomes differ across people, and concavity is associated with income egalitarianism. In a generalized utilitarian social welfare function, each ϕ_i-function maps individual i's expected utility $U_i(\ell)$ to a utility of the impartial observer. These expected utilities differ across people, and concavity is associated with expected-utility egalitarianism.[17]

It is easy to show that, if the ϕ_i-functions are concave then the impartial observer will respect Diamond's preferences in our initial example.[18] But having preferences respect Diamond's choice in this particular example is not enough to ensure in general that the ϕ_i-functions are concave: for example, the underlying social choice problem may not contain two outcomes and two people with $(i, x_i) \sim (j, x_j)$ and $(i, x_j) \sim (j, x_i)$. We need to generalize the idea of the example.

The example involved two indifference sets of the impartial observer, that containing (i, x_i) and (j, x_j) and that containing (i, x_j) and (j, x_i). Diamond preferred a randomization between these indifference sets in outcome lotteries (i.e., real life chances) to a randomization in identity lotteries (i.e., imaginary accidents of birth). To generalize, suppose the impartial observer is indifferent between (z, ℓ') and (z', ℓ), and consider the product lottery (z, ℓ) that (in general) lies in a different indifference set. There are two ways to randomize between these indifference sets while remaining in the set of product lotteries. The product lottery $(z, \alpha\ell + (1-\alpha)\ell')$ randomizes between these indifference sets in outcome lotteries (i.e., real life chances); while the product lottery $(\alpha z + (1-\alpha) z', \ell)$ randomizes between these indifference sets in identity lotteries (i.e., imaginary accidents of birth). The example suggests that Diamond prefers the former.

Preference for Life Chances. *For any pair of identity lotteries z and z' in $\triangle(\mathcal{I})$, and any pair of outcome lotteries ℓ and ℓ' in $\triangle(\mathcal{X})$, $(z, \ell') \sim (z', \ell)$ then $(z, \alpha\ell + (1-\alpha)\ell') \succsim (\alpha z + (1-\alpha) z', \ell)$ for all α in $(0,1)$.*

If the preference sign is reversed in the above axiom, we say that the impartial observer exhibits preference for accidents of birth. If both apply, we say that the impartial observer is indifferent between accidents of birth and life chances.

[17] This is sometimes called 'ex ante egalitarianism'. See for example, Broome (1984), Myerson (1981), Hammond (1981, 1982) and Meyer (1991). In our context, it is perhaps better to call this 'interim' egalitarianism since it refers to distributions 'after' the resolution of the identity lottery but 'before' the resolution of the outcome lottery.

[18] See footnote 6. To get Diamond's strict preference, we require strict concavity.

If we add this axiom to the conditions of Theorem 2, then we obtain concave generalized utilitarianism.

Proposition 3 (Concavity) *Suppose that absence of unanimity and all the axioms of Theorem 2 apply, so that* $V(z,\ell) = \sum_{i=1}^{I} z_i \phi_i [U_i(\ell)]$ *is a generalized utilitarian representation. Then the impartial observer exhibits preference for life chances if and only if each of the* ϕ_i*-functions is concave.*

To show that concavity is sufficient, recall that V is affine in identity lotteries and each U_i is affine in outcome lotteries. Thus, if we set $V_i(\ell) := \phi_i [U_i(\ell)]$ for all ℓ, then V_i is concave if and only if ϕ_i is concave. Imposing concavity, we obtain $V(z, \alpha\ell + (1-\alpha)\ell') = \sum_{i=1}^{I} z_i V_i(\alpha\ell + (1-\alpha)\ell') \geq \sum_{i=1}^{I} z_i [\alpha V_i(\ell) + (1-\alpha) V_i(\ell')] = \alpha V(z,\ell) + (1-\alpha) V(z,\ell')$. Using the fact that $(z,\ell') \sim (z',\ell)$, the last expression is equal to $\alpha V(z,\ell) + (1-\alpha) V(z',\ell) = V(\alpha z + (1-\alpha) z', \ell)$. Hence the impartial observer exhibits a preference for life chances.

The proof that preference for life chances implies concavity is in the appendix but a discussion follows. At first glance, the representation in Lemma 1 resembles a recursive expected-utility model such as $\sum_i z_i v(\ell_i)$ in which z_i is the probability of being faced by the outcome lottery ℓ_i. In that setting, an analog of preference for life chances implies that the function v is concave.[19] But there are two ways in which the current model differs from this recursive one. First, in place of a single v, each V_i represents a different individual's preferences. Second, in place of a vector of ℓ_i's, each individual faces the same outcome lottery ℓ, thus the set of product lotteries (our objects of choice) are not isomorphic to the set of compound lotteries. In a sense, however, the first problem alleviates the second.

The role of absence of unanimity in the proof is to ensure that there is enough variation in the individual preferences to make up for the lack of variation in the outcome lottery. Consider the map that takes each outcome lottery ℓ to its corresponding vector of utilities $(V_1(\ell), \ldots, V_I(\ell))$. Absence of unanimity ensures that the range of this map is rich enough for our axioms to bite. And absence of unanimity is essential: example 3 in appendix A shows that without this richness

[19] See, for example, Kreps & Porteus (1979) or Grant, Kajii & Polak (1998).

condition on the underlying individual preferences, the V_i's (and hence the ϕ_i's) need not be concave.

In contrast to Diamond, Harsanyi implicitly imposes indifference between life chances and accidents of birth. If we impose this indifference as an explicit axiom then, as a corollary of Proposition 3, we obtain that each ϕ_i-function must be affine. In this case, if we let $\hat{U}_i := \phi_i \circ U_i$, then $\hat{U}_i$ is itself a von Neumann-Morgenstern expected-utility representation of $\succsim_i$. Thus, we immediately obtain our first new axiomatization of Harsanyi's utilitarian representation.

Theorem 4 (Utilitarianism I) *Suppose that absence of unanimity applies. Then the following are equivalent:*

(a) The impartial observer satisfies the acceptance principle and independence over identity lotteries, each individual satisfies independence over outcome lotteries, and the impartial observer is indifferent between life chances and accidents of birth.

(b) There exist a continuous function $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ *that represents* $\succsim$*, and, for each individual* i *in* $\mathcal{I}$*, a function* $\hat{U}_i : \triangle(\mathcal{X}) \to \mathbb{R}$ *that is a von Neumann-Morgenstern expected-utility representation of* $\succsim_i$*, such that for all* (z, ℓ) *in* $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$*,*

$$V(z, \ell) = \sum_{i=1}^{I} z_i \hat{U}_i(\ell)$$

where, for each i*,* $\hat{U}_i(\ell) = \int_{\mathcal{X}} \hat{U}_i(x)\, \ell(dx)$*. Moreover the functions* $\hat{U}_i$ *are unique up to common affine transformation.*

To summarize: if we start from the axioms that gave us a generalized utilitarianism and add Diamond's preference for life chances then we obtain concave generalized utilitarianism. But if we assume that life chances and accidents of birth are equivalent in the sense of indifference we are forced back to Harsanyi's utilitarianism.

5 Different risk attitudes: Harsanyi vs. Pattanaik

In this section, we first show how generalized utilitarianism can accommodate different risk attitudes. More interestingly, we then show that if we impose a weak form of independence over outcome lotteries directly on the impartial observer (rather than just allow her to inherit outcome-lottery independence via the acceptance principle), then she is forced to ignore different risk attitudes. Finally, we will show that strengthening this axiom – so that identity and outcome lotteries are treated symmetrically in terms of the axioms – (almost) forces us again back to Harsanyi's utilitarianism.

Recall that Pattanaik's critique concerned different risk attitudes of different individuals. The impartial observer's interpersonal welfare comparisons might rank $(i, x_i) \sim (j, x_j)$ and $(i, x_j) \sim (j, x_i)$, but if person i is more comfortable facing risk than person j, she might rank $\left(i, \frac{1}{2}[x_i] + \frac{1}{2}[x_j]\right) \succ \left(j, \frac{1}{2}[x_i] + \frac{1}{2}[x_j]\right)$. Harsanyi's utilitarianism rules this out. An analogy might be useful. In the standard representative-agent model of consumption over time, each time period is assigned one utility function. This utility function must reflect both risk aversion in that period and substitutions between periods. Once utilities are scaled for inter-temporal welfare comparisons, there is limited scope to accommodate different risk attitudes across periods. Harsanyi's utilitarian impartial observer assigns one utility function per person. This utility function must reflect both risk aversion of that person and substitutions between people. Once utilities are scaled for interpersonal welfare comparisons, there is limited scope to accommodate different risk attitudes across people.

Given the analogy, it is not surprising that generalized utilitarianism can accommodate Pattanaik. Each person is now assigned two functions, ϕ_i and u_i, so we can separate interpersonal welfare comparison from risk aversion. To be more specific, we generalize Pattanaik's example. Suppose the impartial observer ranks $(i, \ell) \sim (j, \ell')$ and $(i, \tilde{\ell}) \sim (j, \tilde{\ell}')$. Then, for all α in $(0, 1)$, we say that the two outcome lotteries $\alpha\tilde{\ell} + (1-\alpha)\,\ell$ and $\alpha\tilde{\ell}' + (1-\alpha)\,\ell'$ are *similar risks* for individuals i and j respectively. Suppose that the generalized utilitarian impartial observer always prefers to face similar risks as person i than as person j. In this case, loosely speaking, we require the function ϕ_i to be a concave transformation of ϕ_j on the 'relevant domain'. The next proposition makes this precise.

Proposition 5 (Different Risk Attitudes.) *Suppose that absence of unanimity and all the axioms of Theorem 2 apply, so that* $V(z, \ell) = \sum_{i=1}^{I} z_i \phi_i [U_i(\ell)]$ *is a generalized utilitarian represen-*

tation. Then the impartial observer always (weakly) prefers to face similar risks as individual i than as individual j if and only if the composite function $\phi_i^{-1} \circ \phi_j$ is convex on the the domain $\mathcal{U}_{ji} := \{u \in \mathbb{R} :$ there exists $\ell, \ell' \in \triangle(\mathcal{X})$ with $(i, \ell) \sim (j, \ell')$ and $U_j(\ell') = u\}$.

The proof is in the appendix but a discussion follows. Recall that we say that agent j is more (income) risk averse than agent i if the function that maps income to agent j's von Neumann-Morgenstern utility is a concave transformation of that for agent i. For each i, the function ϕ_i maps agent i's von Neumann-Morgenstern utility to utilities of the impartial observer. Thus, to say that $\phi_i^{-1} \circ \phi_j$ is convex is to say that the function that maps the impartial observer's utility to agent j's von Neumann-Morgenstern utility (i.e., ϕ_j^{-1}) is a concave transformation of that of agent i. We return to this discussion in section 7 when we introduce a cardinally measurable and comparable welfare.

In contrast to Pattanaik, Harsanyi implicitly imposes indifference as to which person should face similar risks; that is, he ignores different risk attitudes.[20] Harsanyi makes this assumption when he imposes independence over outcome lotteries directly on the impartial observer, rather than just allowing such independence to be inherited from individual preferences via the acceptance principle. In fact, even the following weak independence suffices.

Weak Independence over Outcome Lotteries (for the impartial observer). *Suppose (i, ℓ), $(j, \ell') \in \mathcal{I} \times \triangle(\mathcal{X})$ are such that $(i, \ell) \sim (j, \ell')$. Then, for all $\tilde{\ell}$, $\tilde{\ell}' \in \triangle(X)$: $(i, \tilde{\ell}) \succsim (j, \tilde{\ell}')$ if and only if $(i, \alpha\tilde{\ell} + (1-\alpha)\ell) \succsim (j, \alpha\tilde{\ell}' + (1-\alpha)\ell')$ for all α in $(0, 1]$*

Notice first that this axiom is almost symmetric to independence over identity lotteries for the impartial observer: there the mixing involves identity lotteries holding the outcome lotteries fixed; here the mixing involves outcome lotteries holding the identity lotteries fixed. But this axiom is weak in that it restricts the (fixed) identity lotteries to be degenerate. Second, given acceptance, imposing weak independence over outcome lotteries directly on the impartial observer implies independence over outcome lotteries for each individual i, but the converse is not true. Third, if

[20] We know from section 4 that indifference as to which agent should face similar risks is implied by indifference between accidents of birth and life chances. But the converse is not true: indifference as to which agent should face similar risks is weaker.

we look at the case where both $(i, \ell) \sim (j, \ell')$ and $(i, \tilde{\ell}) \sim (j, \tilde{\ell}')$, the new axiom immediately forces the impartial observer to be indifferent between facing similar risks as person i and person j. Thus (given Proposition 5), the ϕ_i and ϕ_j-functions must be identical up to affine transformations.

To obtain a tighter result, we introduce a second richness condition that is the symmetric analog of absence of unanimity.

Redistributive Scope For all z, z' in $\triangle(\mathcal{I})$, if $(z, x) \succ (z', x)$ for some x in $\mathcal{X}$ then there exists x' in $\mathcal{X}$ such that $(z', x') \succ (z, x')$.

Given acceptance, absence of unanimity implies: for all ℓ, ℓ' in $\triangle(\mathcal{X})$, if $(i, \ell) \succ (i, \ell')$ for some i in $\mathcal{I}$ then there exists j in $\mathcal{I}$ such that $(j, \ell') \succ (j, \ell)$. With absence of unanimity, there are no Pareto dominated outcome lotteries. With redistributive scope, there are no 'dominated' identity lotteries. In particular, for each pair of individuals i and j, if there is an outcome at which the impartial observer would prefer to be individual i then there is another outcome at which she would prefer to be individual j. Intuitively, there is some other policy outcome in which either person i has been made sufficiently worse off or person j has been made sufficiently better off (or both) such that their ranking has been reversed. Despite the formal symmetry between these two richness conditions, redistributive scope is perhaps more restrictive in practice: for example, there may be policy settings in which, under every policy under consideration, one agent is always better off than an other. It will apply however, in standard private-good allocation problems with (ex ante) symmetric agents.

With redistributive scope, imposing weak independence over outcome lotteries directly on the impartial observer yields a generalized utilitarian representation in which there is a common ϕ-function.

Proposition 6 (Common ϕ-Function) *Suppose that absence of unanimity and redistributive scope both apply. Then the impartial observer satisfies the acceptance principle, independence over identity lotteries and weak independence over outcome lotteries if and only if there exist a continuous function $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ that represents $\succsim$, for each individual i in $\mathcal{I}$, a von Neumann-Morgenstern function $\hat{U}_i : \triangle(\mathcal{X}) \to \mathbb{R}$ that represents $\succsim_i$, and a (*common*) continuous,* *strictly increasing function $\phi : \mathbb{R} \to \mathbb{R}$, such that, for all (z, ℓ) in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$,*

$$V(z, \ell) = \sum_{i=1}^{I} z_i \phi \left[\hat{U}_i(\ell) \right].$$

where, for each i, $\hat{U}_i(\ell) = \int_{\mathcal{X}} \hat{U}_i(x) \ell(dx)$. Moreover the functions $\hat{U}_i$ are unique up to a common affine transformations, and the composite functions $\phi \circ \hat{U}_i$ are unique up to a common affine transformation.

The proof is in the appendix but a sketch follows. Given acceptance, weak independence over outcome lotteries for the impartial observer implies independence over outcome lotteries for each individual. Hence, theorem 2 implies that preferences admit a generalized utilitarian representation. Weak independence over outcome lotteries also implies that the impartial observer is indifferent between facing similar risks as person i or person j. Proposition 5 then tells us that $\phi_i^{-1} \circ \phi_j$ is affine on the relevant interval, $\mathcal{U}_{ji}$.

Our redistributive scope condition ensures that there exist two individuals, call them i^1 and i_2, such that for all individuals j either $\mathcal{U}_{ji^1}$ or $\mathcal{U}_{ji_2}$ is not trivial. Thus, loosely speaking, all the ϕ_j-functions are affine transformations of one another. The proof then constructs a common ϕ-function (applying appropriate affine transformations to each von Neumann-Morgenstern utility function U_i to form $\hat{U}_i$). Without the redistributive scope condition we could still construct a representation with a common ϕ-function – a remark in the appendix gives an example – but we would lose the tight uniqueness conditions.

Recall that when we assumed that the impartial observer was indifferent between identity and outcome lotteries – that is, she ignored Diamond's concerns – we were forced back to Harsanyi's utilitarianism. If we assume that the impartial observer is indifferent as to who should face similar risks – that is, she ignores Pattanaik's concerns — we are not forced back to Harsanyi, but only to a common ϕ-function. Nevertheless, once we introduce weak independence over outcome lotteries directly on the impartial observer, it seems natural to ask what happens if we impose strong independence; that is, if we treat identity and outcome lotteries symmetrically in terms of the axioms.

Strong Independence over Outcome Lotteries (for the Impartial Observer). *Suppose* (z, ℓ), $(z', \ell') \in \triangle(\mathcal{I}) \times \triangle(\mathcal{X})$ *are such that* $(z, \ell) \sim (z', \ell')$. *Then for all* $\tilde{\ell}$, $\tilde{\ell}' \in \triangle(\mathcal{X})$: $(z, \tilde{\ell}) \succsim (z', \tilde{\ell}')$ *if and only if* $(z, \alpha\tilde{\ell} + (1-a)\ell) \succsim (z', \alpha\tilde{\ell}' + (1-a)\ell')$ *for all* α *in* $(0, 1]$.

This axiom is the symmetric analog of our independence over identity lotteries for the impartial observer reversing the roles of identity lotteries and outcome lotteries. It is stronger than weak independence over outcome lotteries in that it allows the (fixed) identity lotteries z and z' to be non-degenerate.

One might conjecture that treating identity and outcome lotteries symmetrically in terms of the axioms would force us back to Harsanyi: more precisely, given the acceptance principle, if we symmetrically impose absence of unanimity and redistributive scope, and we symmetrically impose (strong) independence both over outcome lotteries and over identity lotteries on the impartial observer then we would again obtain Harsanyi's utilitarianism. After all, we know from lemma 1 that absence of unanimity and independence over identity lotteries gives us a representation that is affine in identity lotteries. Symmetrically, redistributive scope and (strong) independence over outcome lotteries give us a representation that is affine in outcome lotteries. This suggests that having both sets of properties gives us a representation that is affine in both identity and outcome lotteries; i.e., utilitarianism. But there is a flaw in this argument: the representation that is affine in identity lotteries need not be the same representation as that which is affine in outcome lotteries. The following example illustrates what can go wrong.

For the purpose of the example, let $\mathcal{I} = \{1, 2\}$ and $\mathcal{X} = \{x_1, x_2\}$. To simplify notation, for each $z \in \triangle(\mathcal{I})$, let $q = z_2$; and for each $\ell \in \triangle(\mathcal{X})$ let $p := \ell(x_2)$. With slight abuse of notation, we will write $(q, p) \succsim (q', p')$ for $(z, \ell) \succsim (z', \ell')$, and write $V(q, p)$ for $V(z, \ell)$.

Example 1 *Let agent* 1*'s preferences be given by* $U_1(p) = (1 - 2p)$, *and let agent* 2*'s preferences be given by* $U_2(p) = (2p - 1)$. *Let the impartial observer's preferences be given by* $V(q, p) := (1-q)\phi[U_1(p)] + q\phi[U_2(p)]$, *where the (common)* ϕ*-function is given by:*

$$\phi[u] = \begin{cases} u^k & \text{for } u \geq 0 \\ -(-u)^k & \text{for } u < 0 \end{cases}, \text{ for some } k > 0$$

These preferences are not utilitarian unless $k = 1$. Nevertheless, it is clear that absence of unanimity and redistributive scope both apply in this example. And, by proposition 6 (the common ϕ result), this impartial observer satisfies acceptance and independence over identity lotteries. It remains to show that she also satisfies strong independence over outcome lotteries.

Consider the inverse function $\phi^{-1}(u) = u^{1/k}$ for $u \geq 0$ and $\phi^{-1}(u) = -(-u)^{1/k}$ for $u < 0$. This is a strictly increasing function. Therefore, the function $\phi^{-1}[V(\cdot,\cdot)]$ represents the same preferences as $V(\cdot,\cdot)$. Some simple algebra (provided in the appendix) shows that we can write

$$\phi^{-1}[V(q,p)] = (1-p)\,\phi^{-1}[(1-2q)] + p\phi^{-1}[(2q-1)]\,.$$

This alternative representation is symmetric to the original representation $V(\cdot,\cdot)$ with the p's and q's reversed and ϕ^{-1} replacing ϕ. Since the alternative representation is affine in p, preferences must satisfy strong independence over outcome lotteries.

Notice that the underlying preferences in this example resemble those in the example in the introduction. We could think of x_1 as the outcome in which person 1 gets some indivisible good, and x_2 as the outcome in which person 2 gets it. As advocated by Harsanyi, if the impartial observer thinks she is equally likely to be either person, she is indifferent as to whom is given the good. And if the impartial observer thinks the good is equally likely to be given to either person, she is indifferent as to whom she is when she faces that risk. Nevertheless, her preferences are not utilitarian for more complicated randomizations.

Although this example is special, some aspects of it are quite general. For any generalized utilitarian representation V with common ϕ, the associated function $\phi^{-1} \circ V$ will always be an alternative representation of the same preferences. Moreover, lemma 11 in the appendix shows that, if we start from the conditions of proposition 6 (the common ϕ result) but replace weak with strong independence over outcome lotteries then this alternative representation $\phi^{-1} \circ V$ is always affine in outcome lotteries.

Our alternative representation takes the form:

$$\phi^{-1}[V(z,\ell)] := \phi^{-1}\left(\sum_{i=1}^{I} z_i \phi_i\left[\hat{U}_i(\ell)\right]\right)$$

We can think of this as a "generalized mean" of the distribution of utilities induced by (z, ℓ).[21] In particular, given each individual's (expected) utility function $\hat{U}_i$, we can again think of each outcome lottery ℓ as inducing a vector of individual (expected) utilities $\left(\hat{U}_1(\ell), \ldots, \hat{U}_I(\ell)\right)$. Thus, each product lottery (z, ℓ) induces a distribution of individual utilities where each $\hat{U}_i(\ell)$ is assigned probability z_i. Since $\phi^{-1} \circ V$ and each $\hat{U}_i$ function are affine on outcome lotteries, the generalized mean must be affine on the *induced* set of utility vectors. If the generalized mean were affine over *all* utility vectors (or over a sufficiently rich set) then it would have to be the ordinary arithmetic mean; that is, ϕ would have to be affine and our representation would reduce to Harsanyi's utilitarianism.[22] But since, in our product-lottery setting, every individual faces the same outcome lottery, the set of utility vectors that are induced by outcome lotteries need not be rich. In example 1, the induced utility vectors all lie in the line segment from $(1, -1)$ to $(-1, 1)$. Once again, we need a rich set of underlying outcomes and/or preferences to induce a sufficiently rich set of utility lotteries to be forced to utilitarianism.

The example shows that the two richness conditions we have used so far, absence of unanimity and redistributive scope, are not enough. But they are close. The ϕ-function in the example is a homogenous function. Again, this is general: lemma 12 in the appendix shows that if start from the conditions of proposition 6 but replace weak with strong independence over outcome lotteries then the common ϕ-function is always homogenous; that is homogeneity is necessary.[23] But homogeneity is not sufficient except in very special cases. Notice in the example that the point of inflection of the homogenous ϕ-function (the "zero") occurs exactly at outcome lottery $p = 1/2$ where the impartial observer is indifferent over which identity lottery she faces. This is a very "knife-edge" property, and it can be ruled out in a number of ways. The following extra richness condition suffices.

Three-Player Richness For all outcomes x, y in X, and all α in $[0, 1]$, there exist individuals i

[21] See, for example, Hardy, Littlewood and Polya (1934) ch.3.

[22] Strictly speaking, an affine transformation of the arithmetic mean. For a proof, see Hardy, Littlewood and Polya (1934) p.86.

[23] Strictly speaking, it must be an affine transformation of a homogenous function since our representation is only unique up to affine transformations.

and j in $\mathcal{I}$ such that $(i, \alpha[x] + (1-\alpha)[y]) \succ (j, \alpha[x] + (1-\alpha)[y])$.

Given redistributive scope, *three-player richness* implies that there must be at least three individuals. In words, it says that there is no outcome lottery involving just two outcomes at which the impartial observer is indifferent over all the possible identities she could assume. In example 1, the condition was violated at $p = 1/2$, since the impartial observer was indifferent between being either person there. If we add a third person to the example, then the condition would be met provided that either the outcome lottery at which the impartial observer is indifferent between being person 1 or person 3 or the outcome lottery at which she is indifferent between being person 2 or person 3 is not exactly equal to 1/2.[24]

With this extra condition in place (and hence troublesome examples like example 1 ruled out), the symmetric richness conditions and symmetric independence axioms over identity and outcome lotteries yield Harsanyi's utilitarianism.

Theorem 7 (Utilitarianism II) *Suppose that absence of unanimity, redistributive scope and three-player richness all apply. Then the following are equivalent:*

(a) The impartial observer satisfies the acceptance principle, independence over identity lotteries and (strong) independence over outcome lotteries

(b) There exist a continuous function $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ that represents $\succsim$ and, for each i in $\mathcal{I}$, a function $\hat{U}_i : \triangle(\mathcal{X}) \to \mathbb{R}$ that is a von Neumann-Morgenstern expected-utility representation of $\succsim_i$, such that for all (z, ℓ) in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$,

$$V(z, \ell) = \sum_{i=1}^{I} z_i \hat{U}_i(\ell)$$

where, for each i, $\hat{U}_i(\ell) = \int_{\mathcal{X}} \hat{U}_i(x)\, \ell(dx)$. Moreover the functions $\hat{U}_i$ are unique up to common affine transformation.

[24] Notice that, if there are three or more possible outcomes, the condition still only places restrictions for lotteries involving just two. In particular, there still could be some lottery on the interior of the simplex where the impartial observer is indifferent as to identity. The condition would, however, be violated if there were divisible and disposable private goods and (hence) some outcome that equalized the welfare of all individuals. In that setting, however, since the outcome set is itself very rich, we can anyway induce a sufficiently rich set of utility lotteries.

The proof is in the appendix. Example 1 shows that three-player richness is essential. Example 4 in appendix A shows that redistributive scope is also essential.

To summarize: if we start from the axioms that gave us generalized utilitarianism then it is easy to accommodate Pattanaik's concerns about different attitudes toward risk. Forcing the impartial observer to be indifferent as to who faces risk, does not force us to Harsanyi's utilitarianism but only to a common ϕ-function. Such indifference however, is equivalent to imposing a weak form of independence over outcome lotteries directly on the impartial observer. If we go further and assume that identity and outcome lotteries are equivalent in the sense that all axioms are symmetric across the two types of lottery, then (provided the underlying problem is rich) we are again forced to Harsanyi's utilitarianism.

6 Assuming less and assuming more: related literature.

In this section, we first ask what happens to our representations if we assume less. In particular, we consider dropping our richness considerations altogether. Then we switch around and compare our results to those in the literature that assume more. In particular, we show how Harsanyi's axioms (imposed on preferences over all joint distributions over identities and outcomes) imply all of the axioms (imposed just on preferences over product lotteries) of each of our utilitarian theorems.

Assuming less. If we are not worried about uniqueness, we can obtain a representation of the form $\sum_{i=1}^{I} z_i V_i(\ell)$ without imposing absence of unanimity. Recall our sketch proof of lemma 1. Absence of unanimity ensured there exists two outcome lotteries ℓ^1 and ℓ_2 such that for all product lotteries (z, ℓ) either $(z, \ell) \sim \left(z', \ell^1\right)$ for some z', or $(z, \ell) \sim (z'', \ell_2)$ for some z'' or both. Moreover we can choose ℓ^1 and ℓ_2 such that the 'intervals' of such indifferent lotteries 'overlap'. If we do not impose absence of unanimity then two complications arise. First we may require many more than two 'intervals' to cover the indifference sets of the impartial observer. The main step in the proof of theorem 8 is to show that we can always find a countable number of 'intervals' that cover these indifference sets, and to construct a representation using these intervals. The second complication is that the intervals may not overlap. Without such overlapping, we do not obtain uniqueness. Absence of unanimity is sufficient but not necessary to obtain uniqueness. The precise condition is given in case one of the proof of theorem 8. This theorem provides the most general representation result in the paper.

Theorem 8 *The following are equivalent:*

(a) *The impartial observer's preferences $\succsim$ satisfy the acceptance principle and independence over identity lotteries.*

(b) *There exist a continuous function $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ and, for each i in $\mathcal{I}$, a function $V_i : \triangle(\mathcal{X}) \to \mathbb{R}$, such that V represents $\succsim$; for each i, V_i represents $\succsim_i$; and for all (z, ℓ) in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$,*

$$V(z, \ell) = \sum_{i=1}^{I} z_i V_i(\ell).$$

Assuming more. At a technical level, the papers closest to ours are Karni & Safra (2000), Fishburn (1982, ch 7) and Safra & Weissengrin (2002). Karni & Safra produce a representation similar to lemma 1. The key difference is that their axioms apply to the full set of joint distributions so they can apply recursive arguments. Both Fishburn and Safra & Weissengrin work with product lottery spaces like ours. Fishburn provides axioms on product spaces of mixture sets to obtain multi-linear representations. His context was games in which opponents' mixed strategies are independent. Safra & Weissengrin adapt this approach to derive Harsanyi's utilitarianism in a setting where the impartial observer faces only product lotteries. The key difference between their result and our two utilitarianism theorems (theorem 4 and theorem 7) is that Safra & Weissengrin directly impose independence on the impartial observer for all mixtures that are well defined in the space of product lotteries. Implicitly, therefore, they not only impose both independence over identity lotteries and strong independence over outcome lotteries, but also a third independence axiom over hybrids of the other two:

Independence over Hybrid Lotteries (for the Impartial Observer). *Suppose (z, ℓ), $(z', \ell') \in \triangle(\mathcal{I}) \times \triangle(\mathcal{X})$ are such that $(z, \ell) \sim (z', \ell')$. Then for all $\in \triangle(\mathcal{I})$ and all $\tilde{\ell}' \in \triangle(\mathcal{X})$: $(\tilde{z}, \ell) \succsim$*

$(z', \tilde{\ell}')$ *if and only if* $(\alpha \tilde{z} + (1-\alpha) z, \ell) \succsim (z', \alpha \tilde{\ell}' + (1-a) \ell')$ *for all* α *in* $(0,1]$.

This axiom is similar to the other independence axioms for the impartial observer except that the lotteries being mixed on the left are identity lotteries (holding outcome lotteries fixed), while the lotteries being mixed on the right are outcome lotteries (holding identity lotteries fixed). This hybrid independence axiom is quite strong: in particular, it can be shown that it implies indifference between life chances and accidents of birth. Thus, we can think of Safra & Weissengrin as assuming the union of our axioms from theorem 4 and theorem 7, our two utilitarianism results.

The Safra & Weissengrin theorem helps explain how Harsanyi comes implicitly to assume both indifference between life chances and accidents of birth and indifference over who should face similar risks; and hence conflict with Diamond and Pattanaik. Recall that Harsanyi works with the full set of joint distributions $\triangle(\mathcal{I} \times \mathcal{X})$. He imposes independence directly on the impartial observer for all mixtures defined on that space. If we then restrict these preferences from the larger set $\triangle(\mathcal{I} \times \mathcal{X})$ to just the product lotteries $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$ then independence applies to any mixture that is still well defined. That is, all three of Safra & Weissengrin's independence axioms apply. But, as we have just argued, the third of the Safra & Weissengrin axioms (independence over hybrid lotteries for the impartial observer) implies indifference between life chances and accidents of birth; in conflict with Diamond. And, as we argued in section 5, the second of the Safra & Weissengrin axioms (strong independence over outcome lotteries for the impartial observer) implies indifference over who should face similar risks; in conflict with Pattanaik.

Notice that these conflicts are not over the idea of independence per se. We can assume that each individual satisfies independence over the lotteries he faces, namely outcome lotteries; that the impartial observer respects these individual preferences over outcome lotteries; and that she satisfies independence over the lottery she faces, namely identity lotteries, but this does not imply Harsanyi's utilitarianism: it only implies generalized utilitarianism.

7 Welfare Inequality and Risk

In this section, we introduce an explicit notion of comparable welfare, and use it to interpret some of our representation results. Let $w_i : \triangle(\mathcal{X}) \to \mathbb{R}$ be agent i's welfare function, and let

$w : \triangle(\mathcal{I}) \times \triangle(\mathcal{X}) \to \mathbb{R}$ be the impartial observer's welfare function. These w_i's are functions of life chances rather than final outcomes, so think of them as *interim* welfares. Similarly, we can think of the impartial observer's welfare function w as *ex ante* welfare. Let us assume that these welfare functions guide choice. That is,

Congruence For each individual i in $\mathcal{I}$ and for ℓ, ℓ' in $\triangle(\mathcal{X})$, $\ell \succsim_i \ell'$ if and only if $w_i(\ell) \geq w_i(\ell')$. For the impartial observer, for all $(z, \ell), (z', \ell')$ in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$, $(z, \ell) \succsim (z', \ell')$ if and only if $w(z, \ell) \geq w(z', \ell')$.

Following Weymark (1991), let us further assume that the impartial observer adopts the welfare of agent i when she puts herself in the shoes of agent i.

Principle of Welfare Identity For each individual i in $\mathcal{I}$ and for ℓ in $\triangle(\mathcal{X})$, $w_i(\ell) = w(i, \ell)$.

For the remainder of this section, we assume that both of these axioms apply. As Weymark (1991) notes, taken together, congruence and the principle of welfare identity imply acceptance. Furthermore, they entail that for any pair of individuals i and j, and any pair of life-chances ℓ and ℓ', the ranking between (i, ℓ) and (j, ℓ') is completely determined by the ranking between $w_i(\ell)$ and $w_j(\ell')$. That is, the welfare functions $(w_1(.), \ldots, w_I(.))$ are at least *ordinally measurable* and *fully comparable*.

To relate these welfare measures to the generalized utilitarian representation obtained in theorem 2, define for each individual i, the function $g_i : \mathbb{R} \to \mathbb{R}$ that maps individual i's interim welfare to his von Neumann-Morgenstern expected utility. That is, for each individual i, and for all ℓ in $\triangle(\mathcal{X})$:

$$g_i(w_i(\ell)) = U_i(\ell)$$

Similarly, let $g : \mathbb{R} \to \mathbb{R}$ denote the mapping from the impartial observer's ex ante welfare to her von Neumann-Morgenstern utility. Thus, if the conditions of theorem 2 apply, then for all (z, ℓ) in $\triangle(\mathcal{I}) \times \triangle(\mathcal{X})$:

$$g(w(z, \ell)) = \sum_{i=1}^{I} z_i \phi_i \circ U_i(\ell)$$

Given this, we can now re-interpret the functions ϕ_i in terms of welfare. Applying the principle of welfare identity, we get $\phi_i \circ U_i(\ell) = g[w(i,\ell)] = g[w_i(\ell)] = g\left[g_i^{-1}(U_i(\ell))\right]$. Thus, for each individual i, the function ϕ_i is given by the function $g \circ g_i^{-1}$.

We can then re-express the generalized utilitarian social welfare function from theorem 2 in terms of our g-functions and welfare to yield

$$w(z,\ell) = g^{-1}\left(\sum_{i=1}^{I} z_i g_i(w_i(\ell))\right).$$

With only *ordinally measurable* welfares, the shape and hence degree of curvature of g can vary as one considers different (common) monotonic transformations of $(w_1(.), \ldots, w_I(.))$. In this case (following Sen (1977)), we can no more interpret the shape of g than can Harsanyi interpret his social welfare function as being linear in welfare. We know from theorem 2, however, that ϕ_i (which is equal to $g \circ g_i^{-1}$) is invariant to any common increasing transformations of the welfare functions $(w_1(\cdot), \ldots, w_I(\cdot))$. That is, if we take $(\hat{w}_1(\cdot), \ldots, \hat{w}_I(\cdot))$, where $\hat{w}_i = h \circ w_i$ and $h : \mathbb{R} \to \mathbb{R}$ is an increasing function, then the functions that map the transformed welfares to the von Neumann-Morgenstern utilities are now given by $\hat{g}_i = g_i \circ h^{-1}$ and $\hat{g} = g \circ h^{-1}$ And so, $\hat{\phi}_i \equiv \hat{g} \circ \hat{g}_i^{-1} = g \circ h^{-1} \circ h \circ g_i^{-1} = g \circ g_i^{-1}$. This provides some intuition why the ϕ_i-functions are unique up to positive affine transformations.

Suppose we go further and however assume that welfares are cardinally measurable. In this case, we can give more interpretation to the g-functions. In particular, for a fixed identity lottery z (for example, a uniform lottery), we can associate the representation $w(z,\ell)$ above with a Bergson-Samuelson social welfare function that maps the induced vectors of individual interim welfares, $(w_1(\ell), \ldots, w_I(\ell))$, to 'social welfare' (that is, the impartial observer's ex ante welfare, before she knows whom she will become).[25]

In such an interpretation, since we have imposed cardinal measurability, g is uniquely-defined up to positive affine transformations. In this case, the degree of concavity of g may be interpreted as measuring the degree of the impartial observer's aversion to interim welfare inequality or her attitudes toward the risks embodied in accidents of birth.

[25] Analogous to the discussion in sections 4 and 5, we need a rich set of underlying preferences to induce a rich set of individual interim welfares.

This is exactly what we should expect. Had we started from the viewpoint of a Bergson-Samuelson social welfare function, then we would immediately have interpreted Diamond's notion of fairness as aversion to interim welfare inequality. This in turn would have led us to a social welfare function which is concave in individual interim utilities. Instead we started from the viewpoint of representing an impartial observer's preferences, and replaced Harsanyi's implicit assumption that the impartial observer is indifferent between life chances and accidents of birth with Diamond's preference for life chances. If we now impose cardinally measurable welfare, we arrive at the same point.

An explicit notion of welfare also helps us interpret the different attitudes to risk in Proposition 5. Recall that the impartial observer is more willing to take on similar risks in the identity of person i than that of person j if and only if ϕ_i is a concave transformation of ϕ_j. If welfare is cardinally measurable then ϕ_i is a concave transformation of ϕ_j if and only if g_i is a convex transformation of g_j. This corresponds to our usual notion of income risk aversion except that instead of being risk averse over income, our individuals are risk averse over welfares: individual i is less welfare risk averse than individual j. In other words, each function g_i captures individual i's attitudes toward the welfare risk embodied in her life chances. In this setting, imposing either that the impartial observer is indifferent between life chances and accidents of birth or imposing directly that she respect even weak independence over outcome lotteries forces all people to have the same welfare risk aversion.

Once we allow our social welfare function to take into account that different agents may have different degrees of welfare risk aversion, we may in fact no longer wish to accept Diamond's fairness axiom. There may be cases where the impartial observer may actually prefer accidents of birth to life chances. Suppose society contains people who are extremely welfare risk averse, but suppose that our impartial observer is only mildly (interim) welfare inequality averse. In this case, the functions g_i might be more concave than the function g, and hence the functions $\phi_i = g \circ g_i^{-1}$ would be convex. The impartial observer, anticipating the discomfort that real-life uncertainty would cause real people, prefers to absorb the risk in the imaginary world of her thought experiment.

Although the case of convex ϕ_i functions may seem odd, it corresponds to an argument sometimes used by conservatives to defend caste-like societies. Preferring accidents of birth to life chances corresponds to preferring risks that resolve early. In a different context, Grant, Kajii & Polak (1998) argue that a preference for early resolution corresponds to an intrinsic preference for information: anxious agents may prefer to know their fate soon. In the context of the impartial observer, if individuals are highly risk averse over their welfares, they may prefer for uncertainty to have been resolved by the time they are born. They might prefer "to know their place".

Appendix A: Examples

For each of the following examples, let $\mathcal{I} = \{1, 2\}$ and $\mathcal{X} = \{x_1, x_2\}$. To simplify notation, for each $z \in \triangle(\mathcal{I})$, let $q = z_2$; and for each $\ell \in \triangle(\mathcal{X})$ let $p := \ell(x_2)$. Then, with slight abuse of notation, we write $(q, p) \succsim (q', p')$ for $(z, \ell) \succsim (z', \ell')$, and write $V(q, p)$ for $V(z, \ell)$.

Example 1 is introduced and discussed in the text. It shows that absence of unanimity and redistributive scope can apply and the impartial observer can satisfy acceptance and (strong) independence over both identity lotteries and outcome lotteries but that the impartial observer need not be utilitarian in Harsanyi's sense; in particular, the ϕ-function need not be affine. That is, three-person richness is essential for theorem 7 (Utilitarianism II).

Here we just complete the argument to show we can write:

$$\phi^{-1}[V(q, p)] = (1 - p)\,\phi^{-1}[(1 - 2q)] + p\phi^{-1}[(2q - 1)]\,.$$

To see this, it is instructive to rewrite $V(q,p)$ as follows:

$$
\begin{aligned}
V(q,p) &= \begin{cases} (1-2q)(1-2p)^k & \text{for } p<1/2 \\ (2q-1)(2p-1)^k & \text{for } p>1/2 \end{cases} \\
&= \begin{cases} (1-2q)(1-2p)^k & \text{for } q<1/2,\ p<1/2 \text{ (and } V(q,p)>0) \\ -(2q-1)(1-2p)^k & \text{for } q>1/2,\ p<1/2 \text{ (and } V(q,p)<0) \\ 0 & \text{for } (2q-1)(2p-1)=0 \\ -(1-2q)(2p-1)^k & \text{for } q<1/2,\ p>1/2 \text{ (and } V(q,p)<0) \\ (2q-1)(2p-1)^k & \text{for } q>1/2,\ p>1/2 \text{ (and } V(q,p)>0) \end{cases} .
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\phi^{-1}\circ V(q,p) &= \begin{cases} (1-2q)^{1/k}(1-2p) & \text{for } q<1/2,\ p<1/2 \\ -(2q-1)^{1/k}(1-2p) & \text{for } q>1/2,\ p<1/2 \\ 0 & \text{for } (2q-1)(2p-1)=0 \\ -(1-2q)^{1/k}(2p-1) & \text{for } q<1/2,\ p>1/2 \\ (2q-1)^{1/k}(2p-1) & \text{for } q>1/2,\ p>1/2 \end{cases} \\
&= \begin{cases} (1-p)\left[(1-2q)^{1/k}\right]+p\left[-\left[-(2q-1)\right]^{1/k}\right] & \text{for } q<1/2 \\ 0 & \text{for } q=1/2 \\ (1-p)\left[-\left[-(1-2q)\right]^{1/k}\right]+p\left[(2q-1)^{1/k}\right] & \text{for } q>1/2 \end{cases} \\
&= (1-p)\,\phi^{-1}\left[(1-2q)\right]+p\phi^{-1}\left[(2q-1)\right]
\end{aligned}
$$

which equals $(1-p)\,\phi^{-1}[(1-2q)]+p\phi^{-1}[(2q-1)]$ as desired. $\square$

Example 2 shows that preferences can satisfy the acceptance principle, independence (over outcome lotteries) for individuals, and *conditional* independence over identity lotteries for the impartial observer (as defined in footnote 13 but not satisfy the (unconditional) independence axiom over identity lotteries.[26]

[26] Karni & Safra (2000, p.324) provide an example of preferences defined on $\triangle(\mathcal{I})\times\mathbf{X}$ that satisfy the analog of conditional independence but not the analog of unconditional independence. This example extends the idea to $\triangle(\mathcal{I})\times\triangle(\mathbf{X})$.

Example 2 *Let agent 1's preferences be given by* $U_1(p) = -(2p-1)/4$*, and let agent 2's preferences be given by* $U_2(p) = 3(2p-1)/4$*. Notice that these individual preferences satisfy independence. Let the impartial observer's preferences be given by* $V(q,p) = (2p-1)\left(q^2 - 1/4\right)$*.*

By construction, these preferences satisfy the acceptance principle. To show that they satisfy conditional independence over identity lotteries, notice that for each fixed $\bar{p}$, the function $V(q,\bar{p})$ is monotone in q. If $\bar{p} = 1/2$, then the impartial observer is indifferent over all q and conditional independence follows trivially. If $\bar{p} > 1/2$, then $(\tilde{q},\bar{p}) \succsim (\tilde{q}',\bar{p})$ if and only if $\tilde{q} \geq \tilde{q}'$. Thus, for all $\alpha \in (0,1]$, $(\alpha\tilde{q} + (1-\alpha)q, \bar{p}) \succsim (\alpha\tilde{q}' + (1-\alpha)q, \bar{p})$ if and only if $\tilde{q} \geq \tilde{q}'$. The case for $\bar{p} < 1/2$ is similar.

To show that these preferences violate (unconditional) independence over identity lotteries, let $p := 0$ and let $p' := 1$. Let $q = q' := 1/2$ so that $V(q,p) = V(q',p') = 0$. Let $\tilde{q} := 0$ and let $\tilde{q}' = 1/\sqrt{2}$ so that $V(\tilde{q},p) = V(\tilde{q}',p') = 1/4$. Let $\alpha := 1/2$. Then $V(\alpha\tilde{q} + (1-\alpha)q, p) = -\left((1/4)^2 - 1/4\right) = 3/16$. But $V(\alpha\tilde{q}' + (1-\alpha)q', p') = \left(1/4 + 1/\left(2\sqrt{2}\right)\right)^2 - 1/4 = (1/16 + 1/8 + 1/\left(4\sqrt{2}\right) - 1/4) = \left(3/16 + \left(1-\sqrt{2}\right)/\left(4\sqrt{2}\right)\right) < 3/16$, violating (unconditional) independence. □

Example 3 shows that the impartial observer's preferences can satisfy all the conditions of proposition 3 (the concavity result) except absence of unanimity and yet the functions ϕ_i need not be concave. That is, absence of unanimity is essential.

Example 3 *Let the individual's preferences be given by* $U_1(p) = U_2(p) = p$*, and let the impartial observer's preferences be given by* $V(q,p) := (1-q)\phi_1[U_1(p)] + q\phi_2[U_2(p)]$ *where*

$$\phi_1(u) := \begin{cases} 1/4 + u/2 & \text{for } u \leq 1/2 \\ u & \text{for } u > 1/2 \end{cases}$$

$$\phi_2(u) := \begin{cases} u & \text{for } u \leq 1/2 \\ 2u - 1/2 & \text{for } u > 1/2 \end{cases}$$

Since $U_1 = U_2$, both individuals have the same ranking over outcome lotteries and so the impartial observer's preferences violate absence of unanimity. Clearly, the functions $\phi_1(.)$ and $\phi_2(.)$ are not concave. To see that the impartial observer satisfies preference for life chances, without loss of generality let $p \leq p'$ and notice that $(q,p') \sim (q',p)$ implies either $p \leq p' \leq 1/2$ or $p' \geq p \geq 1/2$. But in either case, the functions ϕ_1 and ϕ_2 are concave (in fact, affine) on the domain $[p,p']$ and hence $V(\alpha q + (1-\alpha)q', p) \leq$ (in fact, $=$) $V(q', \alpha p + (1-\alpha)p')$, as desired.

Notice that redistributive scope cannot replace absence of unanimity in this proposition. Indeed, the preferences above satisfy redistributive scope. $\square$

Example 4 shows that preferences can satisfy all the conditions of Theorem 7 except redistributive scope yet the ϕ_i's are not affine. That is, redistributive scope is essential.

Example 4 *Let agent 1's preferences be given by $\hat{U}_1(p) = 1 - bp$, and let agent 2's preferences be given by $\hat{U}_2(p) = -1 + bp = -\hat{U}_1(p)$, for some $b \in (0,1)$. Fix $k > 0$, $k \neq 1$ and set $\phi(u) := u^k$, for u in $[0,1]$ and set $\phi(u) = -(-u)^k$, for u in $[-1,0)$. Suppose the impartial observer's preference relation $\succsim$ is represented by the following function:*

$$\begin{aligned} V(q,p) &= (1-q)\,\phi\left(\hat{U}_1(p)\right) + q\phi\left(\hat{U}_2(p)\right) \\ &= (1-2q)(1-bp)^k \end{aligned}$$

The impartial observer's preference relation, $\succsim$, generated by $V(q,p)$ clearly fails to satisfy redistributive scope, since for all p in $[0,1)$,

$$\phi\left(\hat{U}_1(p)\right) - \phi\left(\hat{U}_2(p)\right) = 2(1-bp)^k > 0.$$

That is, $(1,\ell) \succ (2,\ell)$ for all ℓ. For the same reason, three-player richness holds trivially: V is decreasing in q.

Absence of unanimity holds, since

$$\left[\hat{U}_1(p) - \hat{U}_1(p')\right]\left[\hat{U}_2(p) - \hat{U}_2(p')\right] = -b^2(p-p')^2 < 0$$

for all $p \neq p'$.

To see that strong independence over outcome lotteries holds, notice that

$$\phi^{-1} \circ V(q,p) = \begin{cases} (1-2q)^{1/k}(1-bp) & \text{if } q \leq 1/2 \\ (2q-1)^{1/k}(-1+bp) & \text{if } q > 1/2 \end{cases}$$

which for given q is affine in p. $\square$

Appendix B: Proofs

Proof of Lemma 1. Since the representation is affine in identity lotteries, it is immediate that the represented preferences satisfy the axioms. We will show that the axioms imply the representation.

Let outcome lotteries ℓ^1, ℓ_2 (not necessarily distinct) and identity lotteries z^1, z_2 (not necessarily distinct) be such that $(z^1, \ell^1) \succ (z_2, \ell_2)$ and such that $(z^1, \ell^1) \succsim (z, \ell) \succsim (z_2, \ell_2)$ for all product lotteries (z, ℓ). That is, the product lottery (z^1, ℓ^1) is weakly better than all other product lotteries, and the product lottery (z_2, ℓ_2) is weakly worse than all other product lotteries. And let the identity lotteries z_1 and z^2 (not necessarily distinct) be such that $(z^1, \ell^1) \succsim (z, \ell^1) \succsim (z_1, \ell^1)$ for all product lotteries (z, ℓ^1), and $(z^2, \ell_2) \succsim (z, \ell_2) \succsim (z_2, \ell_2)$ for all product lotteries (z, ℓ_2). That is, given outcome lottery ℓ^1, the identity lottery z_2 is (weakly) worse than all other identity lotteries; and, given outcome lottery ℓ_2, the identity lottery z^2 is (weakly) better than all other identity lotteries. Their existence of these special lotteries follows from continuity of $\succsim$, non-emptyness of $\succ$, and the compactness of $\Delta(\mathcal{I}) \times \Delta(\mathcal{X})$. Moreover, by independence over identity lotteries, we can take z^1, z_1, z^2, and z_2 each to be a degenerate identity lottery. Let these be i^1, i_1, i^2, and i_2 respectively.

The following lemma is helpful.

Lemma 9 *Assume absence of unanimity applies and that the impartial observer satisfies acceptance and independence over identity lotteries. Let $i^1, i_1, i^2, i_2, \ell^1$, and ℓ_2 be defined as above. Then (a) either $(i_1, \ell^1) \sim (i_2, \ell_2)$, or $(i^2, \ell_2) \sim (i^1, \ell^1)$, or $(i^2, \ell_2) \succ (i_1, \ell^1)$. And (b), for all product lotteries (z, ℓ), either $(i^1, \ell^1) \succsim (z, \ell) \succsim (i_1, \ell^1)$ or $(i^2, \ell_2) \succsim (z, \ell) \succsim (i_2, \ell_2)$ or both.*

Proof. (a) If $\ell^1 = \ell_2$, then the first two cases both hold. Otherwise, suppose that the first two cases do not hold; that is, $(i_1, \ell^1) \succ (i_2, \ell_2)$ and $(i^1, \ell^1) \succ (i^2, \ell_2)$. By the definition of i_1, we know that $(i_2, \ell^1) \succsim (i_1, \ell^1)$, and hence $(i_2, \ell^1) \succ (i_2, \ell_2)$. Using absence of unanimity and acceptance , there must exist another individual $\hat{\imath} \neq i_2$ such that $(\hat{\imath}, \ell_2) \succ (\hat{\imath}, \ell^1)$. Again by the definition of i_1, we know that $(\hat{\imath}, \ell^1) \succsim (i_1, \ell^1)$, and hence $(\hat{\imath}, \ell_2) \succ (i_1, \ell^1)$. By the definition of i^2, we know that $(i^2, \ell_2) \succsim (\hat{\imath}, \ell_2)$, and hence $(i^2, \ell_2) \succ (i_1, \ell^1)$, as desired. Part (b) follows immediately from (a). $\square$

Given continuity, an immediate consequence of the lemma is that there exists two outcome lotteries ℓ^1 and ℓ_2 such that for all product lotteries (z, ℓ) either $(z, \ell) \sim (z', \ell^1)$ for some z', or $(z, \ell) \sim (z'', \ell_2)$ for some z'' or both. Moreover, we can choose the z' such that its support only contains individuals i^1 and i_1. And similarly for z'' with respect to i^2 and i_2.

The proof of lemma now proceeds with two cases.

Case (1)[27] The easiest case to consider is where the lotteries ℓ^1 and ℓ_2, defined above are equal. In this case, for the individuals i^1 and i_1 defined in lemma 9, $\left(i^1, \ell^1\right) \succ \left(i_1, \ell^1\right)$, and $\left(i^1, \ell^1\right) \succsim (z, \ell) \succsim \left(i_1, \ell^1\right)$, for all (z, ℓ). Then, for each (z, ℓ), let $V(z, \ell)$ be defined by

$$\left(V(z,\ell)\left[i^1\right] + (1 - V(z,\ell))\left[i_1\right], \ell^1\right) \sim (z, \ell).$$

By continuity and independence over identity lotteries, such a $V(z, \ell)$ exists and is unique.

To show that this representation is affine, notice that if $\left(V(z,\ell)\left[i^1\right] + (1 - V(z,\ell))\left[i_1\right], \ell^1\right) \sim (z, \ell)$ and $\left(V(z',\ell)\left[i^1\right] + (1 - V(z',\ell))\left[i_1\right], \ell^1\right) \sim (z', \ell)$ then independence over identity lotteries implies $\left(\left[\alpha V(z,\ell) + (1-\alpha)V(z',\ell)\right]\left[i^1\right] + \left[1 - \alpha V(z,\ell) - (1-\alpha)V(z',\ell)\right]\left[i_1\right], \ell^1\right) \sim (\alpha z + (1-\alpha) z', \ell)$. Hence $\alpha V(z,\ell) + (1-\alpha)V(z',\ell) = V(\alpha z + (1-\alpha) z', \ell)$.

Since any identity lottery z in $\Delta(\mathcal{I})$ can be written as $z = \sum_i z_i [i]$, proceeding sequentially on $\mathcal{I}$, affinity implies $V(z,\ell) = \sum_i z_i V(i,\ell)$. Finally, by acceptance, $V(i,\cdot)$ agrees with $\succsim_i$ on $\Delta(\mathcal{X})$ Hence, if we define $V_i : \Delta(\mathcal{X}) \to \mathbb{R}$ by $V_i(\ell) = V(i,\ell)$, then V_i represents individual i's preferences. The uniqueness argument is standard: see for example, Karni & Safra (2000, p.321).

Case (2). Let outcome lotteries ℓ^1, ℓ_2 be defined as above, and let individuals i^1, i_1, i^2, i_2 be defined as in lemma 9 and its proof. If $\left(i_1, \ell^1\right) \sim (i_2, \ell_2)$ then $\left(i^1, \ell^1\right) \succsim (z,\ell) \succsim \left(i_1, \ell^1\right)$ for all (z, ℓ) and hence case (1) applies. Similarly, if $\left(i^2, \ell_2\right) \sim \left(i^1, \ell^1\right)$ then $\left(i^2, \ell_2\right) \succsim (z,\ell) \succsim (i_2, \ell_2)$ for all (z,ℓ), and again case (1) applies (with ℓ_2 in place of ℓ^1). Hence suppose that $\left(i^1, \ell^1\right) \succ \left(i^2, \ell_2\right)$ and that $\left(i_1, \ell^1\right) \succ (i_2, \ell_2)$. Then, by lemma 9, $\left(i^1, \ell^1\right) \succ \left(i^2, \ell_2\right) \succ \left(i_1, \ell^1\right) \succ (i_2, \ell_2)$; that is, we have two overlapping intervals that 'span' the entire range of the impartial observer's preferences.

Then, just as in case (1), we can construct an affine function $V^1(\cdot,\cdot)$ to represent the impartial observer's preferences $\succsim$ restricted to those (z,ℓ) such that $\left(i^1, \ell^1\right) \succsim (z,\ell) \succsim \left(i_1, \ell^1\right)$, and we can construct an affine function $V^2(\cdot,\cdot)$ to represent $\succsim$ restricted to those (z,ℓ) such that $\left(i^2, \ell_2\right) \succsim$

[27] This case is similar to case (1) in Safra & Weisengrin (2003, p.184). This case is also analogous to case (1) of Karni & Safra (2000, p.320) except that, in their setting, the analog of ℓ^1 is a vector of outcome lotteries, with a different outcome lottery for each agent. One implication of this is that, in their setting, if case (1) does not apply, then there must exist an agent i and two (vectors of) outcome lotteries ℓ' and ℓ'' such that $(i,\ell') \succsim (z,\ell) \succsim (i,\ell'')$ for all (z,ℓ). This is not true here.

$(z, \ell) \succsim (i_2, \ell_2)$. We can then apply an affine re-normalization of either V_1 or V_2 such the (re-normalized) representations agree on the 'overlap' $\left(i^2, \ell_2\right) \succsim (z, \ell) \succsim \left(i_1, \ell^1\right)$. Since $V_1(\cdot, \cdot)$ and $V_2(\cdot, \cdot)$ are affine, the re-normalized representation is affine, and induction on I (plus acceptance) gives us $V(z, \ell) = \sum_i z_i V_i(\ell)$ as before. Again, uniqueness follows from standard arguments. ■

Proof of Theorem 2 (Generalized Utilitarianism): In the text.

Proof of Proposition 3 (Concavity) The proof of sufficiency is in the text. For necessity, again define $V_i(\ell) := \phi_i \circ U_i(\ell)$. We need to show that for all i and all ℓ, $\ell' \in \triangle(\mathcal{X})$, $V_i(\alpha\ell + (1-\alpha)\ell') \geq \alpha V_i(\ell) + (1-\alpha) V_i(\ell')$ for all α in $[0,1]$. By acceptance, it is enough to show that $V(i, \alpha\ell + (1-\alpha)\ell') \geq \alpha V(i, \ell) + (1-\alpha) V(i, \ell')$ for all α in $(0,1)$. So let $\succsim$ exhibit preference for life chances, fix i and consider ℓ, $\ell' \in \triangle(\mathcal{X})$. Assume first that $\ell \sim_i \ell'$ By acceptance, $V(i, \ell) = V(i, \ell')$. Hence, by preference for life chances,

$$\begin{aligned}
& V(i, \alpha\ell + (1-\alpha)\ell') \\
\geq \;& V(\alpha[i] + (1-\alpha)[i], \ell) \quad \text{(by preference for life chances)} \\
= \;& V(i, \ell) \\
= \;& \alpha V(i, \ell) + (1-\alpha) V(i, \ell') \quad \text{(since } V(i, \ell) = V(i, \ell')\text{)},
\end{aligned}$$

as desired.

Assume henceforth that $\ell \succ_i \ell'$ (and, by acceptance, $V(i, \ell) > V(i, \ell')$). By absence of unanimity, there must exist a j such that $V(j, \ell) < V(j, \ell')$. There are three cases to consider.

(a) If $V(i, \ell') \geq V(j, \ell)$ then, by the representation in lemma 1, there exists z' (of the form $\beta[i] + (1-\beta)[j]$) such that $V(z', \ell) = V(i, \ell')$. Thus, for all α in $(0,1)$,

$$\begin{aligned}
& V(i, \alpha\ell + (1-\alpha)\ell') \\
\geq \;& V(\alpha[i] + (1-\alpha) z', \ell) \quad \text{(by preference for life chances)} \\
= \;& \alpha V(i, \ell) + (1-\alpha) V(z', \ell) \\
= \;& \alpha V(i, \ell) + (1-\alpha) V(i, \ell') \quad \text{(since } V(z', \ell) = V(i, \ell')\text{)},
\end{aligned}$$

as desired.

Assume henceforth that $V(j,\ell) > V(i,\ell')$ (which implies $V(j,\ell') > V(i,\ell')$).

(b) If $V(j,\ell') \geq V(i,\ell)$ then, by the representation in lemma 1, there exists z (of the form $\beta[i] + (1-\beta)[j]$) such that $V(z,\ell') = V(i,\ell)$. Thus, for all α in $(0,1)$,

$$\begin{aligned} & V(i, \alpha\ell' + (1-\alpha)\ell) \\ \geq\ & V(\alpha[i] + (1-\alpha)z, \ell') \quad \text{(by preference for life chances)} \\ =\ & \alpha V(i,\ell') + (1-\alpha)V(z,\ell') \\ =\ & \alpha V(i,\ell') + (1-\alpha)V(i,\ell) \quad (\text{since } V(z,\ell') = V(i,\ell)), \end{aligned}$$

as desired.

(c) Finally, let $V(i,\ell) > V(j,\ell') > V(j,\ell) > V(i,\ell')$. By the continuity of V, there exist β^0, β_0 in $(0,1)$ such that $\beta^0 > \beta_0$, and such that $V(i, \beta^0\ell + (1-\beta^0)\ell') = V(j,\ell')$ and $V(i, \beta_0\ell + (1-\beta_0)\ell') = V(j,\ell)$. Denote $\ell_0 = \beta_0\ell + (1-\beta_0)\ell'$. Then, similarly to part (a),

$$V_i(\gamma\ell + (1-\gamma)\ell_0) \geq \gamma V_i(\ell) + (1-\gamma)V_i(\ell_0)$$

for all $\gamma \in (0,1)$. Next, denote $\ell^0 = \beta^0\ell + (1-\beta^0)\ell'$. Then, similarly to part (b),

$$V_i(\gamma\ell' + (1-\gamma)\ell^0) \geq \gamma V_i(\ell') + (1-\gamma)V_i(\ell^0)$$

for all $\gamma \in (0,1)$. Therefore, restricted to the line segment $[\ell',\ell]$, the graph of V_i lies weakly above the line connecting $(\ell', V_i(\ell'))$ and $(\ell^0, V_i(\ell^0))$ (as does the point $(\ell_0, V_i(\ell_0))$) and weakly above the line connecting $(\ell_0, V_i(\ell_0))$ and $(\ell, V_i(\ell))$ (as does the point $(\ell^0, V_i(\ell^0))$). Hence, $V_i(\alpha\ell + (1-\alpha)\ell') \geq \alpha V_i(\ell) + (1-\alpha)V_i(\ell')$ for all $\alpha \in (0,1)$. ■

Proof of Theorem 4 (Utilitarianism I): In the text.

Proof of Proposition 5 (Different Risk Attitudes). First, notice that if $\mathcal{U}_{ji}$ is not empty then it is a closed interval. If $\mathcal{U}_{ji}$ has an empty interior then the proposition holds trivially true. Therefore, assume that $\mathcal{U}_{ji} = [\underline{u}_{ji}, \bar{u}_{ji}]$ where $\underline{u}_{ji} < \bar{u}_{ji}$.

To prove that $\phi_i^{-1} \circ \phi_j$ convex is sufficient, fix $\ell, \ell', \tilde{\ell}$ and $\tilde{\ell}'$ such that $V(i,\ell) = V(j,\ell')$ and $V(i,\tilde{\ell}) = V(j,\tilde{\ell}')$. We want to show that $V(i, \alpha\tilde{\ell} + (1-\alpha)\ell) \geq V(j, \alpha\tilde{\ell}' + (1-\alpha)\ell')$. By construction, both $U_j(\ell')$ and $U_j\left(\tilde{\ell}'\right)$ lie in $\mathcal{U}_{ji}$. Moreover, we have $U_i(\ell) = \phi_i^{-1} \circ \phi_j[U_j(\ell')]$ and $U_i\left(\tilde{\ell}\right) = \phi_i^{-1} \circ \phi_j\left[U_j\left(\tilde{\ell}'\right)\right]$ Applying the representation we obtain,

$$
\begin{aligned}
& V\left(i, \alpha\tilde{\ell} + (1-\alpha)\ell\right) \\
&= \phi_i\left[U_i\left(\alpha\tilde{\ell} + (1-\alpha)\ell\right)\right] && \text{(by the representation)} \\
&= \phi_i\left[\alpha U_i\left(\tilde{\ell}\right) + (1-\alpha) U_i(\ell)\right] && \text{(by affinity of } U_i\text{)} \\
&= \phi_i\left[\alpha\phi_i^{-1} \circ \phi_j\left[U_j\left(\tilde{\ell}'\right)\right] + (1-\alpha)\phi_i^{-1} \circ \phi_j[U_j(\ell')]\right] && \text{(by the representation)} \\
&\geq \phi_i\left[\phi_i^{-1} \circ \phi_j\left[\alpha U_j\left(\tilde{\ell}'\right) + (1-\alpha) U_j(\ell')\right]\right] && \text{(by convexity of } \phi_i^{-1} \circ \phi_j\text{)} \\
&= \phi_j\left[U_j\left(\alpha\tilde{\ell}' + (1-\alpha)\ell'\right)\right] && \text{(by affinity of } U_j\text{)} \\
&= V\left(j, \alpha\tilde{\ell}' + (1-\alpha)\ell'\right) && \text{(by the representation)}
\end{aligned}
$$

□

To prove that $\phi_i^{-1} \circ \phi_j$ convex is necessary, fix v,w in $\mathcal{U}_{ji}$. By the definition of $\mathcal{U}_{ji}$, there exists outcome lotteries $\ell, \ell' \in \triangle(\mathcal{X})$ such that $U_j(\ell') = v$ and $U_i(\ell) = \phi_i^{-1} \circ \phi_j(v)$; and there exists outcome lotteries $\tilde{\ell}, \tilde{\ell}' \in \triangle(\mathcal{X})$ such that $U_j\left(\tilde{\ell}'\right) = w$ and $U_i\left(\tilde{\ell}\right) = \phi_i^{-1} \circ \phi_j(w)$. By construction, we have $V(i,\ell) = V(j,\ell')$ and $V\left(i,\tilde{\ell}\right) = V\left(j,\tilde{\ell}'\right)$. Therefore, for all α in $(0,1)$

$$
\begin{aligned}
\phi_i\left[U_i\left(\alpha\tilde{\ell} + (1-\alpha)\ell\right)\right] &\geq \phi_j\left[U_j\left(\alpha\tilde{\ell}' + (1-\alpha)\ell'\right)\right] \quad \Rightarrow \\
\alpha U_i\left(\tilde{\ell}\right) + (1-\alpha) U_i(\ell) &\geq \phi_i^{-1} \circ \phi_j\left[\alpha U_j\left(\tilde{\ell}'\right) + (1-\alpha) U_j(\ell')\right] \quad \Rightarrow \\
\alpha\phi_i^{-1} \circ \phi_j(w) + (1-\alpha)\phi_i^{-1} \circ \phi_j(v) &\geq \phi_i^{-1} \circ \phi_j(\alpha w + (1-\alpha)v)
\end{aligned}
$$

Since v and w were arbitrarily, the last inequality corresponds to the convexity of $\phi_i^{-1} \circ \phi_j$ on $\mathcal{U}_{ji}$.

■

Proof of Proposition 6 (Common ϕ-function) . To show that a common ϕ function is sufficient, it is enough to show that the representation $V(\cdot,\cdot)$ (as defined in the proposition) satisfies weak independence over outcome lotteries. If $(i,\ell) \sim (j,\ell')$ then $\phi\left[\hat{U}_i(\ell)\right] = \phi\left[\hat{U}_j(\ell')\right]$, hence $\hat{U}_i(\ell) = \hat{U}_j(\ell')$. Similarly, $(i,\tilde{\ell}) \succsim (j,\tilde{\ell}')$ implies $\hat{U}_i\left(\tilde{\ell}\right) \geq \hat{U}_j\left(\tilde{\ell}'\right)$. By affinity of $\hat{U}_i$ and $\hat{U}_j$, we have

$$\begin{aligned} V(i, \alpha\tilde{\ell} + (1-\alpha)\,\ell) &= \phi\left[\alpha\hat{U}_i\left(\tilde{\ell}\right) + (1-\alpha)\,\hat{U}_i\,(\ell)\right] \text{ and} \\ V(j, \alpha\tilde{\ell}' + (1-\alpha)\,\ell') &= \phi\left[\alpha\hat{U}_j\left(\tilde{\ell}'\right) + (1-\alpha)\,\hat{U}_j\,(\ell')\right]. \end{aligned}$$

Hence $(i, \alpha\tilde{\ell} + (1-\alpha)\,\ell) \succsim (j, \alpha\tilde{\ell}' + (1-\alpha)\,\ell')$. □

To show that weak independence implies a common ϕ, first notice that weak independence over outcome lotteries and the acceptance principle imply that all individuals satisfy independence over outcome lotteries. Thus the conditions of theorem 2 are met. Let $V(z, \ell) = \sum_{i=1}^{I} z_i \phi_i [U_i(\ell)]$ be the corresponding representation.

Suppose individuals i and j are such that the interval $\mathcal{U}_{ji}$ as defined in Proposition 5 has a non-empty interior. Clearly, the corresponding interval $\mathcal{U}_{ij}$ must also have non-empty interior. We argued in the text that weak independence implies that the impartial observer is indifferent between facing similar risks as person i or person j. By Proposition 5, it follows that $\phi_i^{-1} \circ \phi_j$ is affine on $\mathcal{U}_{ji}$. Since $\mathcal{U}_{ji}$ has a non-empty interior, $\phi_i^{-1} \circ \phi_j$ has a unique extension on $\mathbb{R}$. Define a new von Neumann-Morgenstern utility function $\hat{U}_j$ for agent j by the affine transformation, $\hat{U}_j(\ell) := \phi_i^{-1} \circ \phi_j [U_j(\ell)]$ for all ℓ in $\triangle(\mathcal{X})$. Define a new transformation function $\hat{\phi}_j$ for agent j by setting $\hat{\phi}_j\left(\hat{U}_j(\ell)\right) := \phi_j(U_j(\ell))$. Thus, in particular, if $\phi_j[U_j(\ell')] = \phi_i[U_i(\ell)]$ (and hence $U_i(\ell) \in \mathcal{U}_{ij}$), then $\hat{U}_j(\ell') = U_i(\ell)$, and hence $\hat{\phi}_j(u) = \phi_i(u)$ for all u in $\mathcal{U}_{ij}$. By construction, the new social welfare function with U_j replaced by $\hat{U}_j$ and ϕ_j replaced by $\hat{\phi}_j$ still has a generalized utilitarian form and represents the same preferences. With slight abuse of notation we can write $\hat{\phi}_j = \phi_i$, even if this extends the domain of ϕ_i.

To complete the proof, it is sufficient to show that any two individuals (call them j_1 and j_N) can be 'connected' by a sequence of 'intermediary' individuals (call them j_2 through j_{N-1}) such that $\mathcal{U}_{j_n j_{n+1}}$ has non-empty interior for all $n = 1, \ldots, N-1$. This is where we use our second richness condition, redistributive scope. In fact, we never need more than two such intermediaries.

As in lemma 9, let outcome lotteries ℓ^1, ℓ_2 (not necessarily distinct) and identity lotteries z^1, z_2 (not necessarily distinct) be such that $\left(z^1, \ell^1\right) \succ (z_2, \ell_2)$ and such that $\left(z^1, \ell^1\right) \succsim (z, \ell) \succsim (z_2, \ell_2)$ for all identity-outcome lotteries (z, ℓ). That is, the product lottery $\left(z^1, \ell^1\right)$ is weakly better than all other product lotteries, and the product lottery (z_2, ℓ_2) is weakly worse than all other product lotteries. And (symmetric to lemma 9), let ℓ_1 and ℓ^2 (not necessarily distinct) be such that $\left(z^1, \ell^1\right) \succsim \left(z^1, \ell\right) \succsim \left(z^1, \ell_1\right)$ and $\left(z_2, \ell^2\right) \succsim (z_2, \ell) \succsim (z_2, \ell_2)$ for all outcome lotteries ℓ. That is, given identity lottery z^1, the outcome lottery ℓ_1 is (weakly) worse than all other outcome lotteries; and, given identity lottery z_2, the outcome lottery ℓ^2 is (weakly) better than all other outcome lotteries. The existence of these special lotteries follows from continuity of $\succsim$, non-emptyness of $\succ$, and the compactness of $\Delta(\mathcal{I}) \times \Delta(\mathcal{X})$.

By independence over identity lotteries, we can take z^1 and z_2 each to be a degenerate identity lottery. Let these be $i^1$and i_2 respectively. But then, by weak independence over outcome lotteries, we can take ℓ^1, ℓ_1, ℓ^2, and ℓ_2 each to be a degenerate identity lottery. Let these be x^1, x_1, x^2, and x_2 respectively.

The following lemma is symmetric to lemma 9. The proof is essentially the same with redistributive scope playing the role of absence of unanimity.

Lemma 10 *Assume redistributive scope holds and that the impartial observer satisfies independence over identity lotteries and weak independence over outcome lotteries. Let i^1, i_1, x^1, x_1, x^2, and x_2 be defined as above. Then (a) either $\left(i^1, x_1\right) \sim (i_2, x_2)$, or $\left(i_2, x^2\right) \sim \left(i^1, x^1\right)$, or $\left(i_2, x^2\right) \succ \left(i^1, x_1\right)$. And (b), for all product lotteries (z, ℓ), either $\left(i^1, x^1\right) \succsim (z, \ell) \succsim \left(i^1, x_1\right)$ or $\left(i_2, x^2\right) \succsim (z, \ell) \succsim (i_2, x_2)$ or both.*

Given acceptance and the fact that individual preferences are not degenerate, $\left(i^1, x^1\right) \succ \left(i^1, x_1\right)$ and $\left(i_2, x^2\right) \succ (i_2, x_2)$. Hence an immediate consequence of lemma 10 is that $\mathcal{U}_{i^1 i_1}$ has non-empty interior and that, for all individuals j in $\mathcal{I}$, either $\mathcal{U}_{ji^1}$ has non empty interior or $\mathcal{U}_{ji_2}$ has non empty interior (or both). Thus all individuals are connected as desired.

The uniqueness of $\phi \circ U_i$ up to common affine transformations follows from Lemma 1. For the uniqueness of the $\hat{U}_i$ functions, notice that $(i, \ell) \sim (j, \ell')$ implies $\hat{U}_i(\ell) = \hat{U}_j(\ell')$, and that (by the redistributive scope), for each i there exists a j such that $\widehat{\mathcal{U}}_{ji}$ has a non-empty interior. ■

Remark. If we drop redistributive scope, we can still construct the representation but we would lose the uniqueness result. For example, consider a two person society in which $(i, \ell) \succsim (j, \ell')$

for all ℓ, ℓ' in $\triangle(\mathcal{X})$. In this case, let $\hat{U}_i$ be an affine transformation of U_i such that $\hat{U}_i(\ell) > U_j(\ell')$ if $(i, \ell) \succ (j, \ell')$ and such that $\hat{U}_i(\ell) = U_j(\ell')$ if $(i, \ell) \sim (j, \ell')$. And let $\hat{\phi}_i$ be such that $\hat{\phi}_i \circ \hat{U}_i \equiv \phi_i \circ U_i$. Then simply set $\phi := \hat{\phi}_i$ on range of $\hat{U}_i$ and $\phi := \phi_j$ on the range of U_j. These ranges have at most one point in common.

Proof of Theorem 7 (Utilitarianism II). It is clear that (b) implies (a). We will show (a)$\Rightarrow$(b). Clearly, strong independence implies weak independence over outcome lotteries, hence proposition 6 applies. Let $V(z, \ell) = \sum_{i=1}^{I} z_i \phi\left[\hat{U}_i(\ell)\right]$ be as defined there. It is enough to show that the common ϕ-function is affine. Since the proof is long, we will break it into 6 steps, and we will signpost some parts.

Step 1 of the proof consists of the following lemma showing that the function $\phi^{-1} \circ V$ is affine on $\triangle(\mathcal{X})$.

Lemma 11 *Suppose that absence of unanimity and redistributive scope both apply, and that the impartial observer satisfies the acceptance principle, independence over identity lotteries and strong independence over outcome lotteries. Let V and ϕ be defined as in proposition 6; that is, $V(z, \ell) = \sum_{i=1}^{I} z_i \phi\left[\hat{U}_i(\ell)\right]$. Then for each z in $\triangle(\mathcal{I})$ the function $\phi^{-1} \circ V(z, \cdot) : \triangle(\mathcal{X}) \to \mathbb{R}$ is affine.*

Proof. Fix an identity lottery z and an individual i. Similar to the notation in proposition 5, let $\widehat{\mathcal{U}}_i \subset \mathbb{R}$ be the interval such that $u \in \widehat{\mathcal{U}}_i$ implies that there exists an ℓ such that $\hat{U}_i(\ell) = u$. We will first show that $\phi^{-1} \circ V(z, \cdot)$ is affine on the inverse image of $\widehat{\mathcal{U}}_i$; that is, on the subset of outcome lotteries $\{\ell \in \triangle(\mathcal{X}) : \phi^{-1} \circ V(z, \ell) \in \widehat{\mathcal{U}}_i\}$. If this inverse image is empty then affinity is trivial. Hence consider two outcome lotteries ℓ, ℓ' (not necessarily distinct) such that $\phi^{-1} \circ V(z, \ell) \in \widehat{\mathcal{U}}_i$ and $\phi^{-1} \circ V(z, \ell') \in \widehat{\mathcal{U}}_i$. By the definition of $\widehat{\mathcal{U}}_i$, there exists two outcome lotteries $\bar{\ell}$ and $\bar{\ell}'$ such that $\phi^{-1} \circ V(z, \ell) = \hat{U}_i(\bar{\ell})$ and $\phi^{-1} \circ V(z, \ell') = \hat{U}_i(\bar{\ell}')$; that is, $(z, \ell) \sim (i, \bar{\ell})$ and $(z, \ell') \sim (i, \bar{\ell}')$. Applying independence over outcome lotteries, yields

$$(z, \alpha\ell + (1-\alpha)\ell') \sim \left(i, \alpha\bar{\ell} + (1-\alpha)\bar{\ell}'\right)$$

for all α in $[0,1]$ (hence $\phi^{-1} \circ V(z, \alpha\ell + (1-\alpha)\ell') \in \widehat{\mathcal{U}}_i$). Applying the representation yields:

$$\begin{aligned}
\phi^{-1} \circ V(z, \alpha\ell + (1-\alpha)\ell') &= \hat{U}_i\left(\alpha\bar{\ell} + (1-\alpha)\bar{\ell}'\right) \\
&= \alpha\hat{U}_i\left(\bar{\ell}\right) + (1-\alpha)\hat{U}_i\left(\bar{\ell}'\right) \quad \text{(by affinity of } \hat{U}_i\text{)} \\
&= \alpha\phi^{-1} \circ V(z,\ell) + (1-\alpha)\phi^{-1} \circ V(z,\ell').
\end{aligned}$$

where the third line is by definition of $\bar{\ell}$ and $\bar{\ell}'$. This argument holds for all i.

An immediate consequence of the lemma 10 is that there exist two individuals i^1 and i_2 such that range $\left[\phi^{-1} \circ V(z,\cdot)\right] \subseteq \widehat{\mathcal{U}}_{i^1} \cup \widehat{\mathcal{U}}_{i_2}$ and the inverse image of $\widehat{\mathcal{U}}_{i^1} \cup \widehat{\mathcal{U}}_{i_2}$ is $\Delta(\mathcal{X})$. We know that $\phi^{-1} \circ V(z,\cdot)$ is affine on the inverse image of $\widehat{\mathcal{U}}_{i^1}$ and $\widehat{\mathcal{U}}_{i_2}$. Moreover, by lemma 10, the interior of $\widehat{\mathcal{U}}_{i^1 i_2}$ (= interior $\widehat{\mathcal{U}}_{i^1} \cap \widehat{\mathcal{U}}_{i_2}$) is not empty. Hence $\phi^{-1} \circ V(z,\cdot)$ is affine on $\Delta(\mathcal{X})$. This argument holds for all z. □

Let the individuals i^1, i_1, i^2 and i_2 (not necessarily distinct) and outcome lotteries ℓ^1 and ℓ_2 (not necessarily distinct) be defined as in lemma 9. By strong independence over outcome lotteries, we can take ℓ^1 and ℓ_2 each to be a degenerate identity lottery. Let these be x^1 and x_2 respectively. Recall that, given our representation with a common ϕ-function, $(i,\ell) \sim (j,\ell')$ implies $\hat{U}_i(\ell) = \hat{U}_j(\ell')$. Hence, by lemma 9(a) either $\hat{U}_{i_1}(x^1) = \hat{U}_{i_2}(x_2)$, or $\hat{U}_{i^2}(x_2) = \hat{U}_{i^1}(x^1)$, or $\hat{U}_{i^2}(x_2) > \hat{U}_{i_1}(x^1)$. And, by lemma 9(b), for all product lotteries (z,ℓ), we have $\phi^{-1}\left(\sum_i z_i \phi\left[\hat{U}_i(\ell)\right]\right) \in \left[\hat{U}_{i_1}(x^1), \hat{U}_{i^1}(x^1)\right] \cup \left[\hat{U}_{i_2}(x_2), \hat{U}_{i^2}(x_2)\right]$. We will first concentrate on the interval $\left[\hat{U}_{i_1}(x^1), \hat{U}_{i^1}(x^1)\right]$, but we will return to the interval $\left[\hat{U}_{i_2}(x_2), \hat{U}_{i^2}(x_2)\right]$ in step 5. If $\hat{U}_{i_1}(x^1) = \hat{U}_{i^1}(x^1)$ then affinity is trivial.[28] Hence assume $\hat{U}_{i_1}(x^1) < \hat{U}_{i^1}(x^1)$.

Defining $\hat{x}$ and $\bar{u}$. Since $(i^1, x^1) \succ (i_1, x^1)$, by redistributive scope, there exists an outcome $\hat{x}$ such that $(i_1, \hat{x}) \succ (i^1, \hat{x})$. Consider the outcome lotteries $\ell_{[\lambda]}$ defined by $\ell_{[\lambda]} := \lambda[\hat{x}] + (1-\lambda)\left[x^1\right]$. By continuity of both $\hat{U}_{i_1}$ and $\hat{U}_{i^1}$, there must exist an outcome lottery $\bar{\ell}$ ($:= \ell_{[\bar{\lambda}]}$) such that $\left(i^1, \bar{\ell}\right) \sim \left(i_1, \bar{\ell}\right)$. Let $\bar{u}$ be given by

$$\bar{u} := \phi^{-1}\left[V\left(i^1, \bar{\ell}\right)\right] = \phi^{-1}\left[V\left(i_1, \bar{\ell}\right)\right] \tag{1}$$

The level of utility $\bar{u}$ is going to be important in the argument below. By the definition of x^1, if

[28] In this case, lemma 9 implies $\hat{U}_{i^2}(x_2) = \hat{U}_{i^1}(x^1)$ and hence showing ϕ is affine on $[\hat{U}_{i_2}(x_2), \hat{U}_{i^2}(x_2)]$ would be enough.

$\bar{u}$ does not lie in the interval $\left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$ then $\bar{u} < \hat{U}_{i_1}\left(x^1\right)$.

Step 2 of the proof is to show that, for all u' and $u'' \in \left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$ and all α and β in $[0,1]$,

$$\begin{aligned} & \phi^{-1}\left[\beta\phi\left[\alpha u' + (1-\alpha)\bar{u}\right] + (1-\beta)\phi\left[\alpha u'' + (1-\alpha)\bar{u}\right]\right] \\ = \; & \alpha\phi^{-1}\left[\beta\phi\left(u'\right) + (1-\beta)\phi\left(u''\right)\right] + (1-\alpha)\bar{u}. \end{aligned} \tag{2}$$

To show this, fix u' and $u'' \in \left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$ such that $u' < u''$. Denote by $z' = \beta'\left[i^1\right] + \left(1-\beta'\right)\left[i_1\right]$ and $z'' = \beta''\left[i^1\right] + \left(1-\beta''\right)\left[i_1\right]$, the identity lotteries with support just on i^1 and i_1 for which $\phi^{-1}\left[V\left(z', x^1\right)\right] = u'$, and $\phi^{-1}\left[V\left(z'', x^1\right)\right] = u''$. Also fix α and β in $[0,1]$, and define u^β by:

$$u^\beta := \phi^{-1}\left[V\left(\beta z' + (1-\beta)z'', x^1\right)\right] \tag{3}$$

By the definition of $\bar{u}$ and the fact that V is affine in identity lotteries, we have

$$\bar{u} = \phi^{-1}\left[V\left(\beta z' + (1-\beta)z'', \bar{\ell}\right)\right] \tag{4}$$

By lemma 11, the function $\phi^{-1} \circ V\left(\beta z' + (1-\beta)z'', \cdot\right)$ is affine on $\Delta(\mathcal{X})$, hence combining expressions (3) and (4), we get

$$\phi^{-1}\left[V\left(\beta z' + (1-\beta)z'', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right)\right] = \alpha u^\beta + (1-\alpha)\bar{u} \tag{5}$$

Our two affinity properties allows us to expand the left side of expression . First, by the affinity of $V\left(\cdot, \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right)$ on $\Delta(\mathcal{I})$, we get

$$\begin{aligned} & V\left(\beta z' + (1-\beta)z'', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right) \\ = \; & \beta V\left(z', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right) + (1-\beta)V\left(z'', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right) \\ = \; & \beta\phi\left[\phi^{-1} \circ V\left(z', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right)\right] + (1-\beta)\phi\left[\phi^{-1} \circ V\left(z'', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right)\right] \end{aligned} \tag{6}$$

Second, by the affinity of $\phi^{-1} \circ V\left(z', \cdot\right)$ and $\phi^{-1} \circ V\left(z'', \cdot\right)$ on $\Delta(\mathcal{X})$, we have

$$\left[\phi^{-1} \circ V\left(z', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right)\right] = \left[\alpha\phi^{-1} \circ V\left(z', x^1\right) + (1-\alpha)\phi^{-1} \circ V\left(z', \bar{\ell}\right)\right] \quad \text{and} \tag{7}$$

$$\left[\phi^{-1} \circ V\left(z'', \alpha\left[x^1\right] + (1-\alpha)\bar{\ell}\right)\right] = \left[\alpha\phi^{-1} \circ V\left(z'', x^1\right) + (1-\alpha)\phi^{-1} \circ V\left(z'', \bar{\ell}\right)\right]. \tag{8}$$

Substituting $u' = \phi^{-1} \circ V\left(z', x^1\right)$, and $u'' = \phi^{-1} \circ V\left(z'', x^1\right)$, and $\bar{u} = \phi^{-1} \circ V\left(z', \bar{\ell}\right) = \phi^{-1} \circ V\left(z'', \bar{\ell}\right)$, expressions (7) and (8) become

$$[\alpha u' + (1-\alpha)\bar{u}] \quad \text{and} \quad [\alpha u'' + (1-\alpha)\bar{u}]$$

respectively. Substituting these back into expression (6) and then substituting back into the left side of expression (5), yields

$$\phi^{-1}\left[\beta\phi\left[\alpha u' + (1-\alpha)\bar{u}\right] + (1-\beta)\phi\left[\alpha u'' + (1-\alpha)\bar{u}\right]\right] = \alpha u^{\beta} + (1-\alpha)\bar{u}. \tag{9}$$

Using the definition of u^{β} in expression (3) and the affinity of $V\left(\cdot, x^1\right)$ on $\Delta(\mathcal{I})$, we have

$$\begin{aligned} u^{\beta} &= \phi^{-1}\left[V\left(\beta z' + (1-\beta) z'', x^1\right)\right] \\ &= \phi^{-1}\left[\beta V\left(z', x^1\right) + (1-\beta) V\left(z'', x^1\right)\right] \\ &= \phi^{-1}\left[\beta\phi\left(\phi^{-1}\left[V\left(z', x^1\right)\right]\right) + (1-\beta)\phi\left(\phi^{-1}\left[V\left(z'', x^1\right)\right]\right)\right] \\ &= \phi^{-1}\left[\beta\phi(u') + (1-\beta)\phi(u'')\right] \end{aligned} \tag{10}$$

where the last line follows from the definition of u' and u''. Substituting expression (10) back into expression (9) yields expression (2) as desired. Our choice of u', u'', α and β was arbitrary, so this completes step 2. $\square$

Re-normalization. Recall that functions $\left[\hat{U}_i\right]_{i \in \mathcal{I}}$ are unique only up to a common affine transformation and that the composite functions $\left[\phi \circ \hat{U}_i\right]_{i \in \mathcal{I}}$ are also unique only up to a common affine transformation. Hence we can re-normalize such that the utility level $\bar{u} = 0$. With slight abuse of notation, we will continue to use ϕ and $\left[\hat{U}_i\right]_{i \in \mathcal{I}}$ to denote these re-normalized functions. With this re-normalization, expression (2) becomes

$$\phi^{-1}\left[\beta\phi(\alpha u') + (1-\beta)\phi(\alpha u'')\right] = \alpha\phi^{-1}\left[\beta\phi(u') + (1-\beta)\phi(u'')\right] \tag{11}$$

Since $u' < u''$ were arbitrary, expression (11) holds (for all α and β in $[0,1]$) for all utility pairs in the (re-normalized) interval $\left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$. Recall that 0 need not lie in this interval.

Step 3 is to show that expression (11) also holds (for all α and β in $[0,1]$) for all utility pairs $u' < u''$ in $\left[0, \hat{U}_{i^1}\left(x^1\right)\right]$ even if $0 < \hat{U}_{i_1}\left(x^1\right)$; that is, even if $\bar{u}$ does not lie in $\left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$.

To show this, we will establish that expression (11) holds in each interval in a sequence of intervals $I_0, I_1, \ldots$, with (i) $I_0 := \left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$; (ii) $I_n \cap I_{n+1}$ an interval with positive length (and so having a non-empty interior), for all $n = 0, 1, \ldots$; and (iii) $\bigcup_{n=0}^{\infty} I_n = \left(0, \hat{U}_{i^1}\left(x^1\right)\right]$.

Fix an $\tilde{\alpha} \in (0,1)$, for which $\tilde{\alpha}\hat{U}_{i^1}\left(x^1\right) > \hat{U}_{i_1}\left(x^1\right)$ $(> \tilde{\alpha}\hat{U}_{i_1}\left(x^1\right))$. Set $I_n := \left[\tilde{\alpha}^n \hat{U}_{i_1}\left(x^1\right), \tilde{\alpha}^n \hat{U}_{i^1}\left(x^1\right)\right]$. By construction $I_n \cap I_{n+1}$ is an interval with positive length and $\bigcup_{n=0}^{\infty} I_n = \left(0, \hat{U}_{i^1}\left(x^1\right)\right]$. To see that I_n satisfies (11), consider a pair of utilities u' and u'' in I_n and fix α, β in $[0,1]$. By construction both $\hat{u}' := u'/\tilde{\alpha}^n$ and $\hat{u}'' := u''/\tilde{\alpha}^n$ are in $\left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$. Since $\tilde{\alpha}^n$ and $\alpha\tilde{\alpha}^n$ are in $[0,1]$, expression (11) implies

$$\phi^{-1}\left[\beta\phi\left(\tilde{\alpha}^n\hat{u}'\right) + (1-\beta)\phi\left(\tilde{\alpha}^n\hat{u}''\right)\right] = \tilde{\alpha}^n\phi^{-1}\left[\beta\phi\left(\hat{u}'\right) + (1-\beta)\phi\left(\hat{u}''\right)\right] \tag{12}$$

and

$$\phi^{-1}\left[\beta\phi\left(\alpha\tilde{\alpha}^n\hat{u}'\right) + (1-\beta)\phi\left(\alpha\tilde{\alpha}^n\hat{u}''\right)\right] = \alpha\tilde{\alpha}^n\phi^{-1}\left[\beta\phi\left(\hat{u}'\right) + (1-\beta)\phi\left(\hat{u}''\right)\right]. \tag{13}$$

Substituting u' for $\tilde{\alpha}^n\hat{u}'$ and u'' for $\tilde{\alpha}^n\hat{u}''$ and then combining expressions (12) and (13), we obtain

$$\phi^{-1}\left[\beta\phi\left(\alpha u'\right) + (1-\beta)\phi\left(\alpha u''\right)\right] = \alpha\phi^{-1}\left[\beta\phi\left(u'\right) + (1-\beta)\phi\left(u''\right)\right],$$

as required. □

Step 4 consists of the following lemma showing that ϕ must be an affine transformation of a homogenous function.

Lemma 12 *Suppose $\phi(.)$ satisfies equation* (11) *for all u', u'' in $\left[\min\left\{0, \hat{U}_{i_1}\left(x^1\right)\right\}, \hat{U}_{i^1}\left(x^1\right)\right]$ and all α and β in* $[0,1]$*, then*

$$\phi(u) = \begin{cases} Cu^k + D & u \geq 0 \\ -C(-u)^k + D & u < 0 \end{cases}$$

for some C, k in $\mathbb{R}_{++}$ and some D in $\mathbb{R}$.

Case 1.[29] $\hat{U}_{i_1}\left(x^1\right) \geq 0$. We shall show that

$$\phi\left(\alpha u''\right) - \phi\left(\alpha u'\right) = \gamma(\alpha)\left[\phi\left(u''\right) - \phi\left(u'\right)\right]. \tag{14}$$

[29] This draws on Moulin's [1988 p45] proof of Robert's [1980] theorem that a social welfare ordering that is additively separable and independent of common utility scale admits a generalized utilitarian representation with a power function.

for all $u', u'' \in \left[0, \hat{U}_{i^1}\left(x^1\right)\right]$ and all $\alpha \in (0,1)$.

Consider four positive numbers u_1, $\hat{u}_1$, u_2, $\hat{u}_2$ in $\left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$, such that $u_1 < \hat{u}_1$, $u_2 > \hat{u}_2$ and suppose (contra-hypothesis) that for some $\alpha \in (0,1)$,

$$\frac{\phi\left(\alpha u_1\right) - \phi\left(\alpha \hat{u}_1\right)}{\phi\left(\alpha u_2\right) - \phi\left(\alpha \hat{u}_2\right)} \neq \frac{\phi\left(u_1\right) - \phi\left(\hat{u}_1\right)}{\phi\left(u_2\right) - \phi\left(\hat{u}_2\right)} =: -r. \tag{15}$$

Then we have,

$$\phi^{-1}\left[\frac{1}{1+r}\phi\left(u_1\right) + \frac{r}{1+r}\phi\left(u_2\right)\right] = \phi^{-1}\left[\frac{1}{1+r}\phi\left(\hat{u}_1\right) + \frac{r}{1+r}\phi\left(\hat{u}_2\right)\right]$$

And (15) implies that

$$\phi^{-1}\left[\frac{1}{1+r}\phi\left(\alpha u_1\right) + \frac{r}{1+r}\phi\left(\alpha u_2\right)\right] \neq \phi^{-1}\left[\frac{1}{1+r}\phi\left(\alpha \hat{u}_1\right) + \frac{r}{1+r}\phi\left(\alpha \hat{u}_2\right)\right].$$

But (11) implies (setting $\beta = 1/(1+r)$)

$$\begin{aligned}
\phi^{-1}\left[\frac{1}{1+r}\phi\left(\alpha \hat{u}_1\right) + \frac{r}{1+r}\phi\left(\alpha \hat{u}_2\right)\right] &= \alpha\phi^{-1}\left[\frac{1}{1+r}\phi\left(\hat{u}_1\right) + \frac{r}{1+r}\phi\left(\hat{u}_2\right)\right], \\
\text{and } \phi^{-1}\left[\frac{1}{1+r}\phi\left(\alpha u_1\right) + \frac{r}{1+r}\phi\left(\alpha u_2\right)\right] &= \alpha\phi^{-1}\left[\frac{1}{1+r}\phi\left(u_1\right) + \frac{r}{1+r}\phi\left(u_2\right)\right],
\end{aligned}$$

leading to a contradiction. Hence, (14) obtains.

The continuous solutions of (14) are known (Aczel [1966]) to be

$$\phi(u) = C^+ u^{k^+} + D^+$$

for some C^+, k^+ in $\mathbb{R}_{++}$ and some D^+ in $\mathbb{R}$.

Case 2. $0 \in \left(\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right)$. By an analogous argument to the one employed in case 1, for u in the sub-interval $\left[0, \hat{U}_{i^1}\left(x^1\right)\right]$, we obtain $\phi(u) = C^+ u^{k^+} + D^+$; and for u in the sub-interval $\left[\hat{U}_{i_1}\left(x^1\right), 0\right) \subset \mathbb{R}_-$, we obtain $\phi(u) = -C^-(-u)^{k^-} + D^-$, for some C^-, k^- in $\mathbb{R}_{++}$ and some D^- in $\mathbb{R}$. Continuity of ϕ implies $D^+ = D^- =: D$. Thus we obtain:

$$\phi(u) = \begin{cases} C^+ u^{k^+} + D & u \geq 0 \\ -C^-(-u)^{k^-} + D & u < 0 \end{cases}. \tag{16}$$

It remains to show $k^+ = k^-$ and $C^+ = C^-$.

To show $k^+ = k^-$, we again exploit expression (11). Consider $u', u'' \in \left(\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right)$ such that $u' < 0$, $u'' > 0$. Then, for any α, β in $(0,1)$

$$\begin{aligned}
\beta\phi\left(\alpha u'\right) + (1-\beta)\,\phi\left(\alpha u''\right) &= -\beta C^-\left(-\alpha u'\right)^{k^-} + (1-\beta)\,C^+\left(\alpha u''\right)^{k^+} + D \\
\text{and } \beta\phi\left(u'\right) + (1-\beta)\,\phi\left(u''\right) &= -\beta C^-\left(-u'\right)^{k^-} + (1-\beta)\,C^+\left(u''\right)^{k^+} + D.
\end{aligned}$$

Choose α, β in $(0,1)$, such that

$$-\beta C^-\left(-\alpha u'\right)^{k^-} + (1-\beta)\,C^+\left(\alpha u''\right)^{k^+} > 0 \text{ and } -\beta C^-\left(-u'\right)^{k^-} + (1-\beta)\,C^+\left(u''\right)^{k^+} > 0.$$

Therefore, on the left side of (11) we have

$$\phi^{-1}\left(\beta\phi\left(\alpha u'\right) + (1-\beta)\,\phi\left(\alpha u''\right)\right) = \left(\frac{-\beta C^-\alpha^{k^-}\left(-u'\right)^{k^-} + (1-\beta)\,C^+\left(\alpha u''\right)^{k^+}}{C^+}\right)^{1/k^+}$$

and on the right side of (11) we have

$$\begin{aligned}
\alpha\phi^{-1}\left(\beta\phi\left(u'\right) + (1-\beta)\,\phi\left(u''\right)\right) &= \alpha\left(\frac{-\beta C^-\left(-u'\right)^{k^-} + (1-\beta)\,C^+\left(u''\right)^{k^+}}{C^+}\right)^{1/k^+} \\
&= \left(\frac{-\beta C^-\alpha^{k^+}\left(-u'\right)^{k^-} + (1-\beta)\,C^+\left(\alpha u''\right)^{k^+}}{C^+}\right)^{1/k^+}.
\end{aligned}$$

This is possible only if $k^+ = k^- =: k$.

It only remains to show that $C^+ = C^-$. Recall that, by redistributive scope and strong independence over outcome lotteries, there exists an outcome $\hat{x}$ such that $\hat{U}_{i_1}(\hat{x}) > \hat{U}_{i^1}(\hat{x})$. We used this fact to construct $\bar{u}$. Case 2 (i.e., $\bar{u} = 0 \in \left(\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right)$) corresponds to the situation in which $\hat{U}_{i_1}(\hat{x}) > 0 > \hat{U}_{i_1}\left(x^1\right)$.

Recalling the notation we used to define $\bar{u}$, let $\ell_{[\lambda]} := \lambda\left[\hat{x}\right] + (1-\lambda)\left[x^1\right]$. By strong independence over outcome lotteries and our construction, $\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)$ is linear and decreasing in λ, and is positive at $\lambda = 0$ and negative at $\lambda = 1$; and $\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)$ is linear and increasing in λ, and is negative at $\lambda = 0$ and positive at $\lambda = 1$. Let $\bar{\lambda}$ correspond to $\bar{u}$; that is, $\ell_{\left[\bar{\lambda}\right]} = \bar{\ell}$ and $\hat{U}_{i_1}\left(\ell_{\left[\bar{\lambda}\right]}\right) = \hat{U}_{i^1}\left(\ell_{\left[\bar{\lambda}\right]}\right) = 0$. By outcome independence, $\bar{\lambda}$ is implicitly given by

$$\bar{\lambda}\hat{U}_{i_1}(\hat{x}) + \left(1-\bar{\lambda}\right)\hat{U}_{i_1}\left(x^1\right) = 0 = \bar{\lambda}\hat{U}_{i^1}(\hat{x}) + \left(1-\bar{\lambda}\right)\hat{U}_{i^1}\left(x^1\right). \tag{17}$$

Using $\bar{\lambda}$ we can write for $\lambda > \bar{\lambda}$, $\hat{U}_{i_1}\left(\ell_{[\lambda]}\right) = \hat{U}_{i_1}\left(\hat{x}\right)\left(\lambda - \bar{\lambda}\right) / \left(1 - \bar{\lambda}\right)$ and $\hat{U}_{i^1}\left(\ell_{[\lambda]}\right) = \hat{U}_{i^1}\left(\hat{x}\right)\left(\lambda - \bar{\lambda}\right) / \left(1 - \bar{\lambda}\right)$, and so,

$$\frac{\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)}{\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)} = \frac{\hat{U}_{i_1}\left(\hat{x}\right)}{\hat{U}_{i^1}\left(\hat{x}\right)}.$$

Similarly, for $\lambda < \bar{\lambda}$, $\hat{U}_{i_1}\left(\ell_{[\lambda]}\right) = \hat{U}_{i_1}\left(x^1\right)\left(\bar{\lambda} - \lambda\right)/\bar{\lambda}$ and $\hat{U}_{i^1}\left(\ell_{[\lambda]}\right) = \hat{U}_{i^1}\left(x^1\right)\left(\bar{\lambda} - \lambda\right)/\bar{\lambda}$, and so,

$$\frac{\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)}{\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)} = \frac{\hat{U}_{i_1}\left(x^1\right)}{\hat{U}_{i^1}\left(x^1\right)}.$$

Furthermore, again from equation (17), since $-\hat{U}_{i_1}\left(x^1\right)/\hat{U}_{i_1}\left(\hat{x}\right) = \bar{\lambda}/\left(1 - \bar{\lambda}\right) = -\hat{U}_{i^1}\left(x^1\right)/\hat{U}_{i^1}\left(\hat{x}\right)$, we have $\hat{U}_{i_1}\left(\hat{x}\right)/\left[-\hat{U}_{i^1}\left(\hat{x}\right)\right] = \left[-\hat{U}_{i_1}\left(x^1\right)\right]/\hat{U}_{i^1}\left(x^1\right)$. Hence

$$\frac{\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)}{\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)} = \frac{\hat{U}_{i_1}\left(x^1\right)}{\hat{U}_{i^1}\left(x^1\right)} \tag{18}$$

for all $\lambda \neq \bar{\lambda}$.

Also to simplify notation, let $z_{[\gamma]} := \gamma\left[i_1\right] + (1 - \gamma)\left[i^1\right]$. Using this notation, case 2 implies $V\left(z_{[0]}, \ell_{[0]}\right) \geq V\left(z_{[1]}, \ell_{[1]}\right) > V\left(z_{[1]}, \ell_{[0]}\right)$ and $V\left(z_{[1]}, \ell_{[1]}\right) > V\left(z_{[0]}, \ell_{[1]}\right)$. Let $\bar{v} := V\left(z_{[1]}, \ell_{\left[\bar{\lambda}\right]}\right) = \phi(0)$. By independence over identity lotteries, $V\left(z_{[\gamma]}, \ell_{\left[\bar{\lambda}\right]}\right) = \bar{v}$ for all γ.

By construction, for all $\lambda < \bar{\lambda}$, $V\left(z_{[0]}, \ell_{[\lambda]}\right) > V\left(z_{[0]}, \ell_{\left[\bar{\lambda}\right]}\right) > V\left(z_{[1]}, \ell_{[\lambda]}\right)$; and for all $\lambda > \bar{\lambda}$, $V\left(z_{[0]}, \ell_{[\lambda]}\right) < V\left(z_{[0]}, \ell_{\left[\bar{\lambda}\right]}\right) < V\left(z_{[1]}, \ell_{[\lambda]}\right)$. Thus, by the affinity of $V(\cdot, \ell)$ on $\Delta(\mathcal{I})$, for all λ, $V\left(z_{[\gamma]}, \ell_{[\lambda]}\right)$ is affine in γ. Thus there exists a unique $\bar{\gamma} \in (0, 1)$ such that $V\left(z_{[\bar{\gamma}]}, \ell_{[0]}\right) = \bar{v}$. That is, $\left(z_{[\bar{\gamma}]}, \ell_{[0]}\right) \sim \left(z_{[\bar{\gamma}]}, \ell_{\left[\bar{\lambda}\right]}\right)$. An immediate implication of independence over outcome lotteries, is that $V\left(z_{[\bar{\gamma}]}, \ell_{[\lambda]}\right) = \bar{v}$ for all $\lambda \leq \bar{\lambda}$. We claim that $V\left(z_{[\bar{\gamma}]}, \ell_{[\lambda]}\right) = \bar{v}$ for all λ. Suppose not: that is, without loss of generality, there exists a $\lambda > \bar{\lambda}$ such that $V\left(z_{[\bar{\gamma}]}, \ell_{[\lambda]}\right) > \bar{v}$. Then, by independence over outcome lotteries, by mixing with $\left(z_{[\bar{\gamma}]}, \ell_{[0]}\right)$, we would have $V\left(z_{[\bar{\gamma}]}, \ell_{[\lambda]}\right) > \bar{v}$ for all $\lambda > 0$, a contradiction. Thus $V\left(z_{[\bar{\gamma}]}, \ell_{[\lambda]}\right) = \bar{v}$ for all λ.

We can solve for $\bar{\gamma}$ using the definition of V and the fact that

$$\bar{\gamma}\phi\left(\hat{U}_{i_1}\left(x^1\right)\right) + (1 - \bar{\gamma})\,\phi\left(\hat{U}_{i^1}\left(x^1\right)\right) = \bar{\gamma}\phi\left(\hat{U}_{i_1}\left(\hat{x}\right)\right) + (1 - \bar{\gamma})\,\phi\left(\hat{U}_{i^1}\left(\hat{x}\right)\right)$$

Hence

$$\bar{\gamma} = \frac{\phi\left(\hat{U}_{i^1}\left(x^1\right)\right) - \phi\left(\hat{U}_{i^1}\left(\hat{x}\right)\right)}{\left(\phi\left(\hat{U}_{i_1}\left(\hat{x}\right)\right) - \phi\left(\hat{U}_{i_1}\left(x^1\right)\right)\right) + \left(\phi\left(\hat{U}_{i^1}\left(x^1\right)\right) - \phi\left(\hat{U}_{i^1}\left(\hat{x}\right)\right)\right)}$$

By the definition of $\bar{\gamma}$, we have

$$\bar{\gamma}\phi\left[\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)\right]+(1-\bar{\gamma})\,\phi\left[\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right]=\bar{v} \tag{19}$$

for all λ. By the affinity of $\hat{U}$, we have $\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)=\lambda\hat{U}_{i_1}(\hat{x})+(1-\lambda)\,\hat{U}_{i_1}\left(x^1\right)$ and $\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)=\lambda\hat{U}_{i^1}(\hat{x})+(1-\lambda)\,\hat{U}_{i^1}\left(x^1\right)$. Since this holds for all λ, and since ϕ is differentiable almost everywhere, we have

$$\bar{\gamma}\phi'\left[\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)\right]\left(\hat{U}_{i_1}(\hat{x})-\hat{U}_{i_1}\left(x^1\right)\right)+(1-\bar{\gamma})\,\phi'\left[\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right]\left(\hat{U}_{i^1}(\hat{x})-\hat{U}_{i^1}\left(x^1\right)\right)=0 \tag{20}$$

at almost all λ.

Indeed, for all $\lambda\neq\bar{\lambda}$ such that $\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)$ and $\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)$ lie in $\left(\hat{U}_{i_1}\left(x^1\right),\hat{U}_{i^1}\left(x^1\right)\right)$, we can use our homogenous expression for ϕ and obtain:

$$\phi'\left[\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)\right]=K\phi'\left[\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right], \tag{21}$$

where $K:=(1-\bar{\gamma})\left(\hat{U}_{i^1}\left(x^1\right)-\hat{U}_{i^1}(\hat{x})\right)/\left[\bar{\gamma}\left(\hat{U}_{i_1}(\hat{x})-\hat{U}_{i_1}\left(x^1\right)\right)\right]$ is a constant (that is, does not depend on λ).

Since ϕ is a power function we have by plugging in to expression (21), for $\lambda>\bar{\lambda}$,

$$kC^{+}\left(\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)\right)^{k-1}=KkC^{-}\left(-\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right)^{k-1}.$$

This reduces to

$$\left[\frac{\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)}{-\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)}\right]^{k-1}=K\frac{C^{-}}{C^{+}}.$$

Similarly for for $\lambda<\bar{\lambda}$,

$$kC^{-}\left(-\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)\right)^{k-1}=KkC^{+}\left(\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right)^{k-1}.$$

This reduces to

$$\left[\frac{-\hat{U}_{i_1}\left(\ell_{[\lambda]}\right)}{\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)}\right]^{k-1}=K\frac{C^{+}}{C^{-}}.$$

But, by expression (18), the ratio in the left side of both these expressions is equal to $-\hat{U}_{i_1}\left(x^1\right)/\hat{U}_{i^1}\left(x^1\right)$ for all $\lambda\neq\bar{\lambda}$. Thus we have shown that $C^{+}=C^{-}$. $\square$

Step 5 extends the argument to cover the interval $\left[\hat{U}_{i_2}(x_2),\hat{U}_{i^2}(x_2)\right]$. So far we have shown that ϕ must have the form given in Lemma 12 – that is, an affine transformation of a homogenous function – on the interval $\left[\hat{U}_{i_1}\left(x^1\right), U_{i^1}\left(x^1\right)\right]$. We next show that the same function extends over $\left[\hat{U}_{i_2}\left(x_2\right), \hat{U}_{i^2}\left(x_2\right)\right]$.

We can repeat step 2 through step 4 above focussing on the interval $\left[\hat{U}_{i_2}\left(x_2\right), \hat{U}_{i^2}\left(x_2\right)\right]$. The argument is the same except that we need to be careful about the normalization that set $\bar{u} = 0$ prior to step 3. Since we are re-normalizing a second time, we have to keep track of how this second re-normalization is related to the first.

In particular, to be consistent with our notational convention above, let $\left[\hat{U}_i\right]_{i\in\mathcal{I}}$ and ϕ be the individual levels and ϕ-function given the normalization that set $\bar{u} = 0$ above. In these utility units, let the utility level that is analogous to $\bar{u}$ (see expression (1) for the definition) for our analysis of the interval $\left[\hat{U}_{i_2}\left(x_2\right), \hat{U}_{i^2}\left(x_2\right)\right]$ be $\bar{u}_2$. Denote our re-normalized utility function for each individual i by $\tilde{U}_i(\ell) := \hat{U}_i(\ell) - \bar{u}_2$ (so that the utility level $\bar{u}_2$ is re-normalized to zero as before). For each utility level u, let $\tilde{u}$ denote the corresponding re-normalized individual utility level and let $\tilde{\phi}$ denote the correspondingly re-normalized ϕ-function. Then we can re-normalize $\tilde{\phi}$ such that for all u in $\mathbb{R}$, $\tilde{\phi}[\tilde{u}] = \tilde{\phi}[(u - \bar{u}_2)] = \phi[u]$.

By repeating steps 2 to 4, we know that $\tilde{\phi}[\tilde{u}]$ must have a form analogous to that in Lemma 12 on the interval $\left[\hat{U}_{i_2}\left(x_2\right), \hat{U}_{i^2}\left(x_2\right)\right]$. With slight abuse of notation, we can keep track of the re-normalization by writing

$$\tilde{\phi}(u) = \begin{cases} \tilde{C}\left[u - \bar{u}_2\right]^{\tilde{k}} + \tilde{D} & u - \bar{u}_2 \geq 0 \\ -\tilde{C}\left(-\left(u - \bar{u}_2\right)\right)^{\tilde{k}} + \tilde{D} & u - \bar{u}_2 < 0 \end{cases}.$$

By lemma 9, we know that $\left[\hat{U}_{i_2}\left(x_2\right), \hat{U}_{i^2}\left(x_2\right)\right] \cap \left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$ has a non-empty interior. Thus, the $\tilde{\phi}(u)$ and $\phi(u)$ must coincide on this interval. Clearly if either function were affine then both functions must be affine and we would be done. Suppose then that $k \neq 1$ and $\tilde{k} \neq 1$. We will show that this implies $\bar{u}_2 = 0$; that is, the two normalizations must be the same.

Suppose first that the overlap $\left[\hat{U}_{i_2}\left(x_2\right), \hat{U}_{i^2}\left(x_2\right)\right] \cap \left[\hat{U}_{i_1}\left(x^1\right), \hat{U}_{i^1}\left(x^1\right)\right]$ contains a subinterval in which both $u > 0$ and $u - \bar{u}_2 > 0$. Then we know that

$$\tilde{C}\left[u - \bar{u}_2\right]^{\tilde{k}} + \tilde{D} = Cu^k + D \tag{22}$$

for all u in that subinterval. Differentiating yields

$$\tilde{k}\tilde{C}\left[u-\bar{u}_2\right]^{\tilde{k}-1}=kCu^{k-1}.$$

Notice that, if $k=\tilde{k}=2$, then we would have $\left[u-\bar{u}_2\right]/u=C/\tilde{C}$ and, since the right side is constant, this implies $\bar{u}_2=0$. Therefore assume $k\neq 2$ or $\tilde{k}\neq 2$. Differentiating again, dividing the second derivative by the first, and rearranging yields

$$\frac{\left[u-\bar{u}_2\right]}{u}=\frac{\tilde{k}-1}{k-1}.$$

But again the right side is constant, implying $\bar{u}_2=0$. The argument on subintervals where either $u<0$ or $u-\bar{u}_2<0$ is similar.

Since $\bar{u}_2=0$ (if $k\neq 1$), the first derivative reduces to $u^{k-\tilde{k}}=\tilde{k}\tilde{C}/kC$ but again the right side is a constant hence $k=\grave{k}$ so that $\tilde{k}\tilde{C}/kC=u^{k-\tilde{k}}=1$, and hence $C=\tilde{C}$. Finally, using expression (22), we obtain $D=\tilde{D}$. In other words, the two functions ϕ and $\tilde{\phi}$ must be the same. □

Step 6 completes the proof by showing that $k=1$. To do this, we invoke our third richness condition, three-player richness.

Recall from the proof of Lemma 12, for the outcome lottery $\ell_{[\bar{\lambda}]}:=\bar{\lambda}\left[\hat{x}\right]+\left(1-\bar{\lambda}\right)\left[x^1\right]$ we had $\hat{U}_{i_1}\left(\ell_{[\bar{\lambda}]}\right)=\hat{U}_{i^1}\left(\ell_{[\bar{\lambda}]}\right)$. That is, $\left(i^1,\ell_{[\bar{\lambda}]}\right)\sim\left(i_1,\ell_{[\bar{\lambda}]}\right)$. Hence by three-player richness, there exists another individual $\hat{\imath}$, such that $\left(\hat{\imath},\ell_{[\bar{\lambda}]}\right)\nsim\left(i_1,\ell_{[\bar{\lambda}]}\right)$. That is, $\hat{U}_{\hat{\imath}}\left(\ell_{[\bar{\lambda}]}\right)\neq 0$. Consider the graphs of $\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)$ and $\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right)$ as functions of $\lambda\in[0,1]$ to $\mathbb{R}$. Both are lines. The first passes through the point $\left(\bar{\lambda},0\right)$, while the second does not. And, by the definition of x^1, $\hat{x}$ and i^1, the line $\hat{U}_{i^1}\left(\ell_{[\cdot]}\right)$ is strictly decreasing. Suppose $\hat{U}_{\hat{\imath}}\left(\ell_{[\bar{\lambda}]}\right)>0$ – the argument for the case $\hat{U}_{\hat{\imath}}\left(\ell_{[\bar{\lambda}]}\right)<0$ is similar. Then we can find λ and λ' such that $0<\lambda<\lambda'<\bar{\lambda}$ and such that the vectors $\left(\hat{U}_{i^1}\left(\ell_{[\lambda]}\right),\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right)\right)\gg 0$ and $\left(\hat{U}_{i^1}\left(\ell_{[\lambda']}\right),\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)\gg 0$. Moreover, since $\hat{U}_{\hat{\imath}}\left(\ell_{[\bar{\lambda}]}\right)\neq 0$, these vectors are not colinear.

By the affinity of $\phi^{-1}\circ V$ (lemma 11), for all z,ℓ,ℓ' and all α,

$$\begin{aligned}&\phi^{-1}\left[\sum_i z_i\phi\left[\alpha\hat{U}_i\left(\ell\right)+\left(1-\alpha\right)\hat{U}_i\left(\ell'\right)\right]\right]\\ =\;&\alpha\phi^{-1}\left[\sum_i z_i\phi\left[\hat{U}_i\left(\ell\right)\right]\right]+\left(1-\alpha\right)\phi^{-1}\left[\sum_i z_i\phi\left[\hat{U}_i\left(\ell'\right)\right]\right]\end{aligned}\tag{23}$$

In particular, this must hold for $z = (1/2)\left[i^1\right] + (1/2)[\hat{\imath}]$, $\ell_{[\lambda]}$ and $\ell_{[\lambda']}$. Substituting in these values along with our homogenous functional forms $\phi(u) = Cu^k + D$ and $\phi^{-1}(v) = [(v - D)/C]^{1/k}$, The left side of expression (23) becomes:

$$\begin{aligned}
& \phi^{-1}\left[\frac{1}{2}\phi\left(\alpha\hat{U}_{i^1}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{i^1}\left(\ell_{[\lambda']}\right)\right) + \frac{1}{2}\phi\left(\alpha\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)\right] \\
= \; & \phi^{-1}\left[\frac{1}{2}C\left[\left(\alpha\hat{U}_{i^1}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{i^1}\left(\ell_{[\lambda']}\right)\right)^k + \left(\alpha\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)^k\right] + D\right] \\
= \; & \frac{1}{2^{1/k}}\left[\left(\alpha\hat{U}_{i^1}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{i^1}\left(\ell_{[\lambda']}\right)\right)^k + \left(\alpha\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)^k\right]^{1/k}
\end{aligned}$$

And the right side of expression (23) becomes:

$$\frac{1}{2^{1/k}}\left(\alpha\left[\left(\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right)^k + \left(\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right)\right)^k\right]^{1/k} + (1-\alpha)\left[\left(\hat{U}_{i^1}\left(\ell_{[\lambda']}\right)\right)^k + \left(\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)^k\right]^{1/k}\right)$$

Combining these yields:

$$\begin{aligned}
& \left[\left(\alpha\hat{U}_{i^1}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{i^1}\left(\ell_{[\lambda']}\right)\right)^k + \left(\alpha\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right) + (1-\alpha)\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)^k\right]^{1/k} \\
= \; & \alpha\left[\left(\hat{U}_{i^1}\left(\ell_{[\lambda]}\right)\right)^k + \left(\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda]}\right)\right)^k\right]^{1/k} + (1-\alpha)\left[\left(\hat{U}_{i^1}\left(\ell_{[\lambda']}\right)\right)^k + \left(\hat{U}_{\hat{\imath}}\left(\ell_{[\lambda']}\right)\right)^k\right]^{1/k} \qquad (24)
\end{aligned}$$

Notice that if "$=$" were replaced by "$\leq$" then expression (24) becomes the Minkowski inequality. Recall that if the Minkowski inequality holds with equality (and the vectors involved are not colinear) then $k = 1$. Since the vectors we chose were not colinear, we have $k = 1$, completing the proof. ■

Remark. Notice that our third richness condition, three-player richness, was only used in the last step (step 6) of the proof. Specifically, it allowed us to construct vectors that were not colinear, and hence to apply the Minkowski inequality.[30]

The previous step (step 5) illustrates how our counterexample (example 1) relies on their being two outcomes and two agents. Recall the construction of $\bar{u}$. Starting from the interval $\left[\hat{U}_{i_1}\left(x^1\right), U_{i^1}\left(x^1\right)\right]$, redistributive scope ensured there existed an outcome $\hat{x}$ such that $(i_1, \hat{x}) \succ \left(i^1, \hat{x}\right)$. Continuity then ensures there exits an outcome lottery $\ell_{[\bar{\lambda}]}$ between x^1 and $\hat{x}$ such that

[30] Even here, we only need this condition if $\hat{U}_{i_1}(\hat{x}) \leq 0$. If $\hat{U}_{i_1}(\hat{x}) > 0$, then we have $\left(i_1, x^1\right) \succ (i_1, \hat{x})$ and $\left(i^1, x^1\right) \succ \left(i^1, \hat{x}\right)$. In this case, our first richness condition (absence of unanimity) already implies there exists an $\hat{\imath}$ such that $(\hat{\imath}, \hat{x}) \succ \left(\hat{\imath}, x^1\right) \succsim \left(i_1, x^1\right)$, hence $\hat{U}_{\hat{\imath}}\left(\ell_{[\bar{\lambda}]}\right) > 0$.

$\left(i_1, \ell_{[\bar{\lambda}]}\right) \sim \left(i^1, \ell_{[\bar{\lambda}]}\right)$, and $\bar{u}$ corresponded to the utility level at that lottery. Similarly, starting from the interval $\left[\hat{U}_{i_2}(x_2), \hat{U}_{i^2}(x_2)\right]$, redistributive scope ensured there exists an outcome $\hat{x}_2$ such that $(i_1, \hat{x}_2) \succ \left(i^1, \hat{x}_2\right)$ and continuity ensures there exits an outcome lottery $\ell_{[\bar{\lambda}_2]}$ between x_2 and $\hat{x}_2$ such that $\left(i_2, \ell_{[\bar{\lambda}_2]}\right) \sim \left(i^2, \ell_{[\bar{\lambda}_2]}\right)$, and $\bar{u}_2$ corresponded to the utility level at that lottery. An implication of step 5 is that if $\bar{u} \neq \bar{u}_2$ then ϕ is affine. In the example, there are only two outcomes and two agents, hence $\ell_{[\bar{\lambda}]}$ and $\ell_{[\bar{\lambda}_2]}$ must be the same lottery, and therefore $\bar{u}$ and $\bar{u}_2$ are trivially equal. But in a world with three agents or three outcomes, such a coincidence is knife edge.

Proof of Theorem 8. Remark: in the proof of lemma 1, for each product lottery (z, ℓ) except the best and the worst, we found some lottery ℓ' and two individuals i and j such that $(i, \ell') \succ (z, \ell) \succ (j, \ell')$. Absence of unanimity ensured us that such a lottery and pair of individuals existed. We then constructed a 'local' representation $V(z, \ell)$ that solves

$$(V(z, \ell)[i] + (1 - V(z, \ell))[j], \ell') \sim (z, \ell).$$

When we attempt to generalize this idea without absence of unanimity, there might exist "problem" product lotteries (z, ℓ), other than just the best or the worst, such that no lottery ℓ' and individuals i and j exist with the property above. If (z, ℓ) is a such a "problem" product lottery then all product lotteries in its indifference set have the same problem.

More formally, let $\tilde{V}$ be a continuous utility function representing $\succsim$ that we can use as a benchmark to label indifference sets. Without loss of generality, assume that $\tilde{V}$'s image is equal to $[0, 1]$. Let us define the set of "problem" indifference levels as follows

$$\vartheta = \left\{v \in (0, 1) : \nexists i, j \in \mathcal{I} \text{ and } \ell \in \triangle(\mathcal{X}) \text{ s.t. } \tilde{V}(i, \ell) > v > \tilde{V}(j, \ell)\right\}.$$

We claim that the set ϑ is closed (relative to $(0,1)$). Suppose not. That is, let $v_n \to v$ be a sequence in ϑ (where $v \in (0,1)$) and assume there exists $i, j \in \mathcal{I}$ and $\ell \in \triangle(\mathcal{X})$ such that $\tilde{V}(i, \ell) > v > \tilde{V}(j, \ell)$. Then for sufficiently large n, $\tilde{V}(i, \ell) > v_n > \tilde{V}(j, \ell)$: a contradiction.

For all v in ϑ, continuity of $\tilde{V}$ implies $\tilde{V}(z, \ell) = v$ for some product lottery (z, ℓ). By the definition of ϑ, if $\tilde{V}(z, \ell) = v \in \vartheta$ then either $\min_i \tilde{V}(i, \ell) \geq v$ or $\max_i \tilde{V}(i, \ell) \leq v$. By independence over outcome lotteries, $\max_i \tilde{V}(i, \ell) \geq \tilde{V}(z, \ell) \geq \min_i \tilde{V}(i, \ell)$. Hence, $\tilde{V}(z, \ell) = v$ implies there exists at least one individual j such that $\tilde{V}(j, \ell) = v$. And, using independence over identity lotteries again, $\tilde{V}(z,\ell) = v$ implies $V(i,\ell) = v$ for all individuals i in the support of the identity lottery z.

Moreover, we claim that, for all v in the interior of ϑ, if $\tilde{V}(j,\ell) = v$ for some individual j then $\tilde{V}(i,\ell) = v$ for all individuals i. Suppose not. That is, without loss of generality, let $\tilde{V}(i,\ell) < v$. Then for all v' such that $\tilde{V}(i,\ell) < v' < v$, we have $\tilde{V}(i,\ell) < v' < \tilde{V}(j,\ell)$. But this contradicts v being interior.

The proof proceeds in three cases.

Case 1: the set ϑ is empty. Lemma 1 is a special case of this case, and this case yields the same uniqueness conditions as lemma 1. Fix $\varepsilon > 0$ and denote

$$BL^{\varepsilon} = \left\{(z,\ell) \in \triangle(\mathcal{I}) \times \triangle(\mathbf{X}) : 1 - \varepsilon \geq \tilde{V}(z,\ell) \geq \varepsilon\right\}$$

Since ϑ is empty, for each t in $(0,1)$ we can find an outcome lottery ℓ^t for which there exist individuals i and j such that $\tilde{V}(i,\ell^t) > t > \tilde{V}(j,\ell^t)$. Let

$$BL^{t} = \left\{(z,\ell) \in \triangle(\mathcal{I}) \times \triangle(\mathbf{X}) : \tilde{V}\left(i,\ell^t\right) > \tilde{V}(z,\ell) > \tilde{V}\left(j,\ell^t\right)\right\}.$$

As in the proof of lemma 1, construct a function V^t that represents $\succsim$ on the closure of BL^t and which is affine in identity lotteries. Since $\{BL^t\}_{\varepsilon \leq t \leq 1-\varepsilon}$ is an open cover of the compact set BL^{ε}, we can find a finite cover $\{BL^{t_1}, \ldots, BL^{t_K}\}$. The intersection of any two adjacent sets is non-empty. Therefore, we can therefore re-normalize these 'local' representations to find an affine function V^{ε} that represents $\succsim$ on BL^{ε}. For ε' in $(0,\varepsilon)$, the affine function $V^{\varepsilon'}$ can be chosen to agree with V^{ε} on BL^{ε}. By continuity, the limit function $V = \lim_{\varepsilon \to 0} V^{\varepsilon}$ is well defined. By affinity, as in the proof of lemma 1, we can write $V(z,\ell) = \Sigma_i z_i V_i(\ell)$ where, for each i in $\mathcal{I}$, by acceptance, the function $V_i(\ell) := V(i,\ell)$ represents $\succsim_i$ on $\triangle(\mathbf{X})$.

Case 2: the set ϑ is finite. Then we can write $\vartheta = \{v_1, \ldots, v_{K-1}\}$ where $k' > k$ implies $v_{k'} > v_k$. Let $v_0 := 0$ and $v_K := 1$. Fix an interval of the form $[v_{k-1}, v_k]$, $k = 1, \ldots, K$. By independence over identity lotteries, if $\tilde{V}(z,\ell) \in (v_{k-1}, v_k)$ then $\min_i \tilde{V}(i,\ell) < v_k$. Hence, by the definition of ϑ, $\max_i \tilde{V}(i,\ell) \leq v_k$. Similarly, $\max_i \tilde{V}(i,\ell) > v_{k-1}$ and hence $\min_i \tilde{V}(i,\ell) \geq v_{k-1}$. That is, if $\tilde{V}(z,\ell) \in (v_{k-1}, v_k)$ then $\tilde{V}(i,\ell) \in [v_{k-1}, v_k]$ for all i. Moreover, if $\tilde{V}(z,\ell) = v_{k-1}$ (resp.

v_k) then $\tilde{V}(i,\ell) = v_{k-1}$ (resp. v_k) for all i in the support of z. Therefore, following the method of case 1, we can construct a function $V^k(z,\ell) = \Sigma_i z_i V_i^k(\ell)$ that represents $\succsim$ on $\{(z,\ell) \in \triangle(\mathcal{I}) \times \triangle(\mathbf{X}) : \tilde{V}(z,\ell) \in [v_{k-1}, v_k]\}$. To complete the representation, we simply re-normalize these K functions such that they agree on the product lotteries (z,ℓ) such that $\tilde{V}(z,\ell) = v_k$ for $k = 1, \ldots, K$. For example, we can re-normalize such that the range of V^k is $[k-1,k]$. Notice that, as this construction suggests, we do not have uniqueness in this case.

Case 3: the set ϑ is infinite. Choose $\tau \in (0,1) \setminus \vartheta$ Define

$$\begin{aligned} \tau^+ &= \min\{1, \min\{v \in \vartheta : v > \tau\}\} \\ \tau^- &= \max\{0, \max\{v \in \vartheta : v < \tau\}\} \end{aligned}$$

Clearly $\tau^- < \tau < \tau^+$ (recall ϑ is a closed set). As in cases 1 and 2, define functions V and V_i on $[\tau^-, \tau^+]$. The set $(0,1) \setminus \vartheta$ is covered by a countable number of disjoint intervals of the form (τ^-, τ^+), and hence the functions V and V_i can be constructed (inductively and continuously) on their closed union. In this way the functions are also defined for ϑ^0, the set of ϑ's boundary points.

Let v be an interior point of ϑ and let

$$\begin{aligned} v^+ &= \min\left\{1, \min\left\{v' \in \vartheta^0 : v' > v\right\}\right\} \\ v^- &= \max\left\{0, \max\left\{v' \in \vartheta^0 : v' < v\right\}\right\} \end{aligned}$$

Clearly, $v^- < v < v^+$ and all $\succsim_i$ agree on $\{\ell \in \triangle(\mathbf{X}) : \tilde{V}(z,\ell) \in (v^-, v^+)$ for some $z\}$. Choose V that (with continuity) agrees with V of the former step at the indifference sets that are associated with v^- and v^+, such that it represents $\succsim$ on this set. To conclude, define

$$V_i : \{\ell \in \triangle(\mathbf{X}) : \tilde{V}(z,\ell) \in \left(v^-, v^+\right) \text{ for some } z\} \to \mathbb{R}$$

by $V_i(\ell) = V(z,\ell)$ and note that $V(z,\ell) = \Sigma_i z_i V_i(\ell)$ is trivially satisfied.

Finally, as the number of (non trivially) open components of ϑ is also countable, the construction of the desired functions can be carried out easily. ■

References

Aczél, J (1966): *Lectures of Functional Equations and Their Applications*, New York, Dover Press.

Broome, John (1984): "Uncertainty and Fairness," *Economic Journal*, 94, 624-632.

Broome, John (1991): *Weighing Goods*, Oxford, Blackwell.

Diamond, Peter A. (1967): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment." *Journal of Political Economy* 75, 765-66.

Elster, Jon (1989): *Solomonic Judgements: Studies in the Limitation of Rationality*, Cambridge: Cambridge University Press

Epstein, Larry G. and Uzi Segal (1992): "Quadratic Social Welfare," *Journal of Political Economy* 100, 691-712.

Peter C. Fishburn (1982): *Foundations of Expected Utility.* Dordrecht: D. Reidel.

Grant, Simon, Atsushi Kajii and Ben Polak (1998): "Intrinsic Preference for Information," *Journal of Economic Theory*, 83, 233-259.

Hammond, Peter (1981): " Ex-Ante and Ex-Post Welfare Optimality under Uncertainty ," *Economica*, 48, 235-250.

Hammond, Peter (1981): "Utilitarianism, Uncertainty and Information," in A.K. Sen and B. Williams (eds) *Utilitarianism and Beyond*, Cambridge University Press, ch 4, 85 - 102.

Hardy Godfrey H., Littlewood John E. and George Polya (1934): *Inequalities*, Cambridge, Cambridge University Press.

Harsanyi, John C. (1953): "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking." *Journal of Political Economy* 61, 434-5.

Harsanyi, John C. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment." *Journal of Political Economy* 63, 309-21.

Harsanyi, John C. (1975): "Nonlinear Social Welfare Functions: Do Welfare Economists Have a Special Exemption from Bayesian Rationality?" *Theory and Decision* 6, 311-32.

Karni, Edi (1996): "Social welfare functions and fairness." *Social Choice and Welfare* 13, 487-96.

Karni, Edi and Zvi Safra (2000): "An extension of a theorem of von Neumann and Morgenstern with an application to social choice." *Journal of Mathematical Economics* 34, 315-27.

Karni, Edi and Zvi Safra (2002): "Individual sense of justice: a utility representation" *Econometrica* 70(1), 263-284.

Kreps, David M. and Evan L. Poreus (1979): "Temporal von Neumann-Morgenstern and Induced Preferences", *Journal of Economic Theory*, 20, 81-109.

Meyer, Margaret (1991): "A Social Welfare Function Approach to the Measurement of Ex Post Inequality under Uncertainty," working paper Nuffield College.

Mongin, Philippe (2001): "The Impartial Observer Theorem of Social Ethics," *Economics & Philosophy*, 17, 147-179.

Moulin, Herve (1988): *Axioms of Cooperative Decision Making.* Cambridg: Cambridge University Press.

Myerson, Roger (1981): "Utilitarianism, egalitarianism and the timing effect in social choice problems," *Econometrica* 49, 883-97.

Pattanaik, Prasanta K. (1968): "Risk, Impersonality, and the Social Welfare Function." *Journal of Political Economy* 76, 1152-69.

Roberts, Kevin (1979): "Possibility Theorems with Interpersonally Comparable Welfare Levels," *Review of Economic Studies*, 47, 409-420.

Safra, Zvi and Einat Weissengrin (2002): "Harsanyi's impartial observer theorem with a restricted domain." *Social Choice and Welfare*, 20(2), 95-111.

Sen, Amartya, K. (1970): *Collective Choice and Social Welfare.* San Francisco: Holden-Day.

Sen, Amartya, K. (1977): "On weights and measures: informational constraints in social welfare analysis." *Econometrica* 45, 1539-72.

Sen, Amartya, K. (1992): *Inequality Reexamined*, Harvard University Press.

Weymark, John (1991): "A reconsideration of the Harsanyi-Sen debate on utilitarianism" in *Interpersonal Comparisons of Well-being* edited by Jon Elster and John Roemer. Cambridge: CUP.