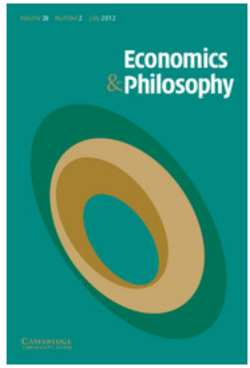
Luc Bovens and Marc Fleurbaey

Evaluating life or death prospects



EVALUATING LIFE OR DEATH PROSPECTS

Luc Bovens and Marc Fleurbaey

Economics and Philosophy / Volume 28 / Special Issue 02 / July 2012, pp 217 - 249
DOI: 10.1017/S0266267112000235, Published online: 29 August 2012

Economics and Philosophy, **28** (2012) 217–249
doi:10.1017/S0266267112000235

EVALUATING LIFE OR DEATH PROSPECTS

LUC BOVENS
London School of Economics and Political Science, UK
l.bovens@lse.ac.uk

MARC FLEURBAEY
Princeton University, USA
mfleurba@princeton.edu

We consider a special set of risky prospects in which the outcomes are either life or death (or, more generally, binary utilities). There are various alternatives to the *utilitarian* objective of minimizing the expected loss of lives in such prospects. We start off with the two-person case with independent risks and construct taxonomies of *ex ante* and *ex post* evaluations for such prospects. We examine the relationship between the *ex ante* and the *ex post* in this restrictive framework: There are more possibilities to respect *ex ante* and *ex post* objectives simultaneously than in the general framework, i.e. without the restriction to binary utilities (*cf.* Harsanyi's aggregation theorem). We extend our results to n persons and to dependent risks. We study optimal strategies for allocating risk reductions given different objectives. We place our results against the backdrop of various *pro-poorly off* (or *prioritarian*) value functions (Diamond 1967; Rabinowicz 2002; Fleurbaey 2010) for the evaluation of risky prospects.

We are grateful to comments from Wlodek Rabinowicz, Martin van Hees and Alex Voorhoeve. Luc Bovens is grateful for the support of the Swedish Collegium for Advanced Studies (SCAS) in Uppsala, Sweden, the Grantham Research Institute on Climate and the Environment and the Centre for Climate Change Economics and Policy, which is funded by the Economic and Social Research Council, and the Dutch Organisation for Scientific Research (NWO – project nr. 236-20-005).

1. INTRODUCTION

We investigate a special set of prospects, viz. prospects in which utilities are two-valued. One natural interpretation of such prospects is that they represent social actions in which people may either live or die.[1] One may think of public works, strategies on the battle field, public health policies, medical triage etc. A policy analyst is charged with ordering sets of such prospects. She may have various objectives and orderings will differ depending on these objectives.

What kind of objectives might an analyst have? Clearly, she might aim to minimize the expected loss of human lives. But this is not all that an analyst typically cares about. Consider the following two alternative objectives. First, when evaluating two public works, she may want to avoid a public works in which the risks are highly focused on a few people and opt instead for a public works in which they are more evenly distributed. Second, in evaluating military strategies, she may want to maximize the *chance* that all of the forces or some strategically important portion of the forces will survive. As we will see, the former objective is one type of *ex ante* objective and the latter is one type of *ex post* objective.

We will construct a taxonomy which captures such objectives in a systematic way and spells out the interrelations between them. Some of these objectives will lead to exactly the same orderings. Some will lead to orderings that are only marginally different. Others will lead to radically different orderings. We do not defend one particular normative stance here, but it is important to understand the synergies and tensions that are present in the space of normative possibilities.

The analyst's objectives affect how she will invest available resources in reducing the risks given the available technology. Should she focus all her resources on reducing the risk to one person or should she spread them and reduce the risks to multiple people? Should she aim to bring about greater risk reductions or should she focus her resources foremost on reducing the risk to people who are at higher risk (even if she could reduce the risk of people at lower risk to a greater extent)? We show how different objectives determine different risk reduction strategies.

We start from a simple two-person model with independent risks (section 2), model *ex ante* objectives (section 3) and *ex post* objectives (section 5), and investigate investment decisions in risk reduction (section 4). We then study the relationship between *ex ante* and *ex post* objectives (section 6). We extend the model to correlated and anti-correlated risks (section 7) and to the n-person case (section 8). We show how our work relates to Keeney's seminal paper on equity and risk reduction (1980)

[1] One might also think of other interpretations, e.g. admission policies with uncertain binary outcomes such as success or failure in the programme of study.

(section 6), to Harsanyi's aggregation theorem (sections 6.6 and 9.1), and to conflicting models of prioritarianism in Diamond (1967), Rabinowicz (2002) and Fleurbaey (2010) (section 9).

2. A TWO-PERSON MODEL, INDEPENDENT RISKS AND DIFFERENT OBJECTIVES

Let p_i be the chance of a fatality for person i and assume that fatality chances are independent. Then we can represent a life or death prospect as a vector of fatality chances $<p_1, p_2>$. The challenge of evaluating such prospects is to construct an ordering over the set $\{<p_1, p_2>|<p_1, p_2> \in [0, 1] \times [0, 1]\}$.

We will distinguish between two types of analysts. An *ex ante* analyst considers the value of the prospect to be determined by the expected utilities of the people affected. We set the utility for person i at $u_i = 1$ if i survives and at $u_i = 0$ if i incurs a fatality. Then the expected utility $E[u_i]$ of the prospect for person i is $(1 - p_i) = 0p_i + 1(1 - p_i)$, i.e. her survival chance. An *ex post* analyst considers the value of the prospect to be determined by the expectation of the social utility of the state of the world that is actualized. The social utility of the state is 0 if 0 people survive, 1 if both people survive and β if exactly 1 person survives, with $0 \leq \beta \leq 1$.

Suppose that an *ex ante* analyst can improve the expected utility (i.e. the survival chance) of a person with a small amount ε. Then her objectives may align with the following priorities:

(i) Improving survival chances (for a person) at a lower survival-chance level matters infinitely more to her than doing so (for a person) at a higher survival-chance level (*maximin*);
(ii) Improving survival chances at a lower survival-chance level matters somewhat more to her than doing so at a higher survival-chance level (*prioritarian*);
(iii) Improving survival chances at a lower survival-chance level matters equally to her as doing so at a higher survival-chance level (*utilitarian*);
(iv) Improving survival chances at a higher survival-chance level matters somewhat more to her than doing so at a lower survival-chance level (*anti-prioritarian*).
(v) Improving survival chances at a higher survival-chance level matters infinitely more to her than doing so at a lower survival-chance level (*maximax*).

An *ex post* analyst may also have different objectives:

(i) All that matters to her is that both survive (*omnitarian*);
(ii) The survival of one (or the other) matters little to her, whereas the survival of both matters much to her (*majoritarian*);
(iii) The survival of each additional person matters equally to her (*paritarian*);
(iv) The survival of one (or the other) person matters very much to her and the survival of an additional person is of little concern to her (*minoritarian*);
(v) All that matters to her is that at least one survives (*unitarian*).

What orderings over $\{<p_1, p_2> | <p_1, p_2> \in [0,1] \times [0,1]\}$ do these objectives generate? Are some of these objectives functionally equivalent in that they generate the same orderings? And how do they affect investment decisions?

3. *EX ANTE* EVALUATIONS

For an *ex ante* analyst, the value v of a prospect is a function of $E[u_1]$ and $E[u_2]$. As we saw before,

$$E[u_i] = (1 - p_i) \text{ for } i = 1, 2 \tag{3.1}$$

i.e. i's survival chance. To differentiate between the various *ex ante* objectives, we construct a transform φ_α of the expected utility $E[u_i]$. Let us first define the φ_α function:

$$\begin{aligned}
\varphi_\alpha(E[u_i]) = \; & {}^1\!/_{(1-\alpha)}\big((E[u_i])^{(1-\alpha)} - 1\big) && \text{if } \alpha \geq 0, \alpha \neq 1; \\
& -{}^1\!/_{(1+\alpha)}\big((1 - E[u_i])^{(1+\alpha)} - 1\big) && \text{if } \alpha < 0, \alpha \neq -1; \\
& \lim_{\alpha \to 1} \big[{}^1\!/_{(1-\alpha)}\big((E[u_i])^{(1-\alpha)} - 1\big)\big] = \ln(E[u_i]) && \text{if } \alpha = 1; \\
& \lim_{\alpha \to -1} \Big[-{}^1\!/_{(1+\alpha)}\big((1 - E[u_i])^{(1+\alpha)} - 1\big)\Big] = -\ln(1 - E[u_i]) && \\
& && \text{if } \alpha = -1.
\end{aligned} \tag{3.2}$$

There are two questions. What does the function in (3.2) do for us in modelling *ex ante* objectives? And why did we choose this function rather than some other function? We will take up the first question here and address the second question towards the end of this section.

These transforms indicate the importance that the analyst assigns to an ε change (i.e. an infinitesimal increment) in the expected utility for a person at lower fatality chances relative to the importance of such a

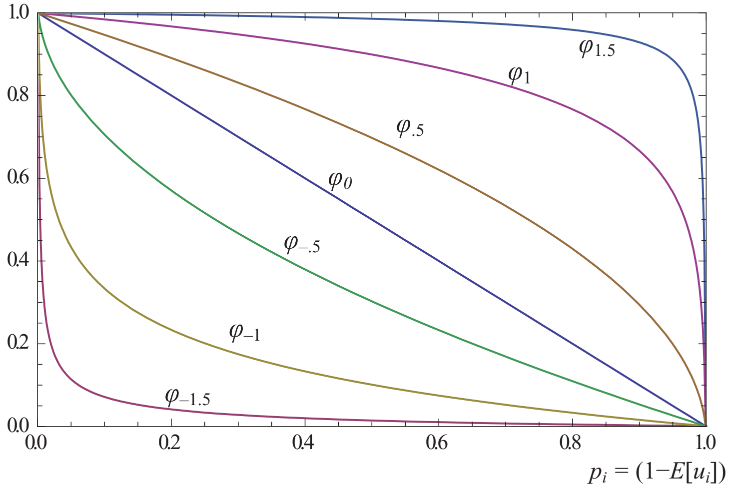


FIGURE 1. Linear transforms of φ_α as a function of the fatality chance p. The parameters for the linear transforms are chosen so that $\varphi_\alpha(p_i = 0) = 0$ and $\varphi_\alpha(p_i = 1) = 1$.

change at higher fatality chances. To see this, we have plotted φ_α as a function of p_i for various values of α in Figure 1.[2]

For $\alpha \in (0, \infty)$, note that an ε change at high fatality chance matters more than at low fatality chance, i.e. $\varphi_\alpha(p_i + \varepsilon) - \varphi_\alpha(p_i)$ is larger for higher values of p_i than for lower values of p_i. This is the *prioritarian* objective, i.e. the *ex ante* objective (ii) in section 2. Similarly, $\alpha = 0$ maps onto the *utilitarian* objective; $\alpha \in (-\infty, 0)$ maps onto the *anti-prioritarian* objective. As α goes to ∞, we have the *maximin* objective and, as α goes to $-\infty$, we have the *maximax* objective.

The value of a prospect is the sum of φ_α transforms of people's expected utilities:

$$v_\alpha(<p_1, p_2>) = \sum_{i=1,2} \varphi_\alpha(E[u_i]) \tag{3.3}$$

We construct an ordering over $\{<p_1, p_2> | <p_1, p_2> \in [0, 1] \times [0, 1]\}$ for a given α:

$$<p_1^*, p_2^*> \succcurlyeq_\alpha <p_1^\#, p_2^\#> \;\Leftrightarrow v_\alpha(<p_1^*, p_2^*>) \geq v_\alpha(<p_1^\#, p_2^\#>) \tag{3.4}$$

[2] Or more accurately, we have plotted positive linear transformations of φ_α – this makes no difference, since positive linear transformations produce the same orderings.

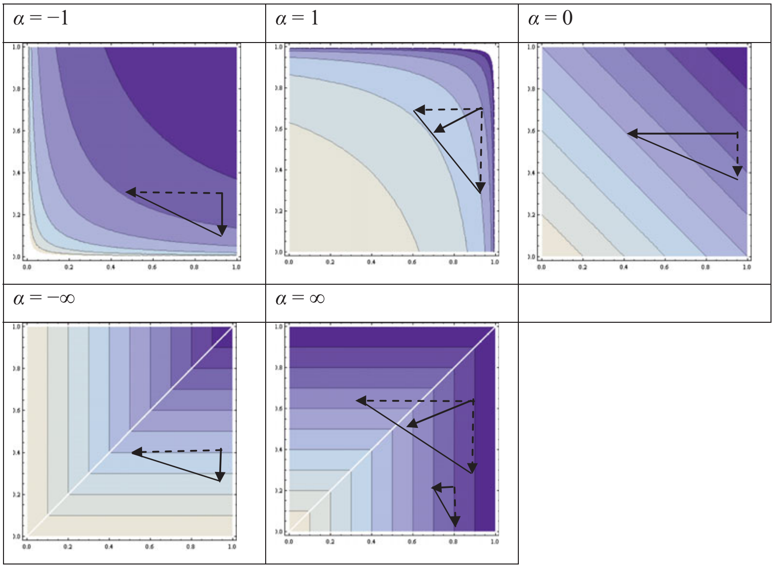


FIGURE 2. Technology with risk reductions imposed on five *ex ante* evaluations. The analyst will implement the risk reduction indicated by the full-line arrow (rather than the dashed arrows).

and represent the orderings for different values of α by means of contour plots in Figure 2. (Ignore the arrows in Figure 2 for now.) Note how for $\alpha = -1$ (an instance of *anti-prioritarianism*), the contour lines are convex to the origin; For $\alpha = 1$ (an instance of *prioritarianism*), the contour lines are concave to the origin; For $\alpha = 0$ (*utilitarianism*), the contour lines are diagonals; For α approaching $-\infty$ (*maximax*), the contour lines become angular towards the origin; For α approaching ∞ (*maximin*), the contour lines become angular away from the origin.

These contour plots represent the orderings of analysts whose objectives are characterized by particular αs. For example, a *prioritarian* with $\alpha = 1$ is more willing to make trade-offs that favour higher fatality chances rather than lower fatality chances. To see this, look at the points S^* and $S^\#$ in Figure 3. Both points are on the same contour line, i.e. the analyst is indifferent between them. Note how the analyst is willing to incur a huge increase δ in fatality chance for person 2 (who is at low fatality chance) to secure a small decrease ε in fatality chance for person 1 (who is at high fatality chance). It is in this respect that *prioritarianism* favours people at high fatality chance.

Similarly, the contour plots show that *anti-prioritarianism* (e.g. $\alpha = -1$) favours people at low fatality chance; that *maximax* ($\alpha \to -\infty$) only cares

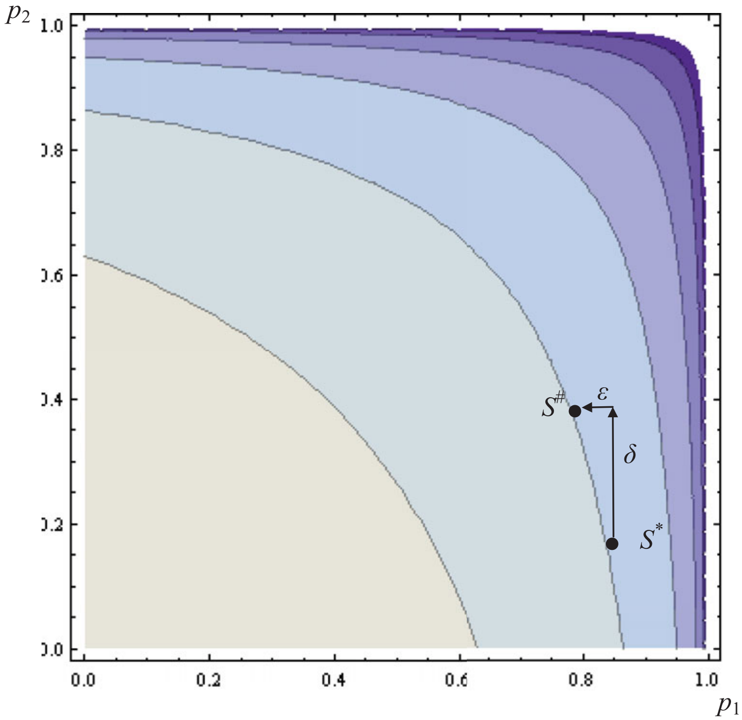


FIGURE 3. Contour plots for $v_\alpha(<p_1, p_2>) = \sum_{i=1,2} \varphi_\alpha(E[u_i])$ for $\alpha = 1$.

about the person who has the higher survival chance; that *maximin* ($\alpha \to \infty$) only cares about the person who has the lower survival chance; and that *utilitarianism* ($\alpha = 0$) does not care who is at lower or higher survival chance – a *utilitarian* analyst is indifferent between two prospects if an increase in one person's survival chance is offset by a decrease *of the same size* in the other person's survival chance.

Let us now turn to the second question: Why did we choose the function in (3.2)? For $\alpha \geq 0$, a variant of this function is commonly used in a particular social welfare function. We find a presentation in Atkinson (1970: 251). Let y_i be the income level of each person in a society.

$$\text{(3.5)} \qquad \varphi_\alpha(y_i) = \begin{cases} {}^{1}\!/_{(1-\alpha)} y_i^{(1-\alpha)} & \text{if } \alpha \geq 0, \alpha \neq 1; \\ \ln(y_i) & \text{if } \alpha = 1 \end{cases}$$

The social welfare function is the sum of the $\varphi_\alpha(y_i)$s for all persons i. α is a measure of inequality aversion. One income distribution is weakly preferred to another if the social welfare function for the former yields a value that weakly exceeds the value of the social welfare function for the latter.

This social welfare function has a number of attractive properties. One such property is that it generates orderings over income distributions that satisfy the Pigou–Dalton condition, i.e. if we take a fixed sum away from a higher-income person and transfer it to a lower-income person without thereby changing their relative rank, then this transfer should yield a distribution that is weakly preferred to the original distribution.[3]

For $\alpha \geq 0$, our function in (3.2) is a linear transformation of Atkinson's function in (3.5) substituting expected utilities (or survival chances) for income levels. Also here, α is a measure of inequality aversion and the Pigou–Dalton condition in our framework says that if we reduce survival chances with ε for a person at a higher survival chance level at the benefit of increasing survival chances with ε for a person at a lower survival chance level without thereby changing their relative rank, then the distribution that we arrive at is weakly preferred to the original one.

Admittedly, there is a whole class of functions that satisfy the Pigou–Dalton condition. So what is the attraction of choosing this one? By choosing this function, (i) we no longer need to *stipulate* that φ_α is the logarithm for $\alpha = 1$, but rather it can be shown that the value of φ_α equals the logarithm in the limit with α approaching 1, and, as we will see, (ii) we can identify a particular *ex ante* evaluation (viz. with $\alpha = 1$) that yields precisely the same ordering as the (*ex post*) *omnitarian* evaluation.

For $\alpha < 0$, the function satisfies an 'anti-Pigou–Dalton' condition, viz. if we increase survival chances with ε for a person at a higher survival chance level at the cost of decreasing survival chances with ε for a person at a lower survival chance level, then the distribution that we arrive at is preferred to the original one. Again, there is a whole class of functions that satisfies this anti-Pigou–Dalton condition. But we have chosen the function in (3.2) because (i) it yields mirror image contour plots for α and $-\alpha$, (ii) it can be shown to equal a logarithmic function in the limit with α approaching -1, and, as we will see, (iii) there is a particular *ex ante* evaluation (viz. with $\alpha = -1$) that yields precisely the same orderings as the (*ex post*) *unitarian* evaluation.[4]

[3] An extensive discussion and defence of this social welfare function can be found in Adler (2011: 307–404), with references to Boadway and Bruce (1984: 159–160), Bossert and Weymark (2004: 1159–1164) and Roberts (1980).

[4] In the context of inequality measures for income distributions, Atkinson's function in (3.5) has the additional property that it is scale invariant. This is arguably a welcome property, since, one might say, it should not affect an inequality ordering whether incomes are measured in, say, francs or in euros. But scale invariance is not a *motivation* for our function in (3.2). For one thing, for $\alpha \geq 0$, there is scale invariance in survival chances, whereas for $\alpha < 0$, there is scale invariance in fatality chances. Now rather than taking scale invariance as a motivating property, it is interesting that one *ex ante–ex post* equivalence entails scale

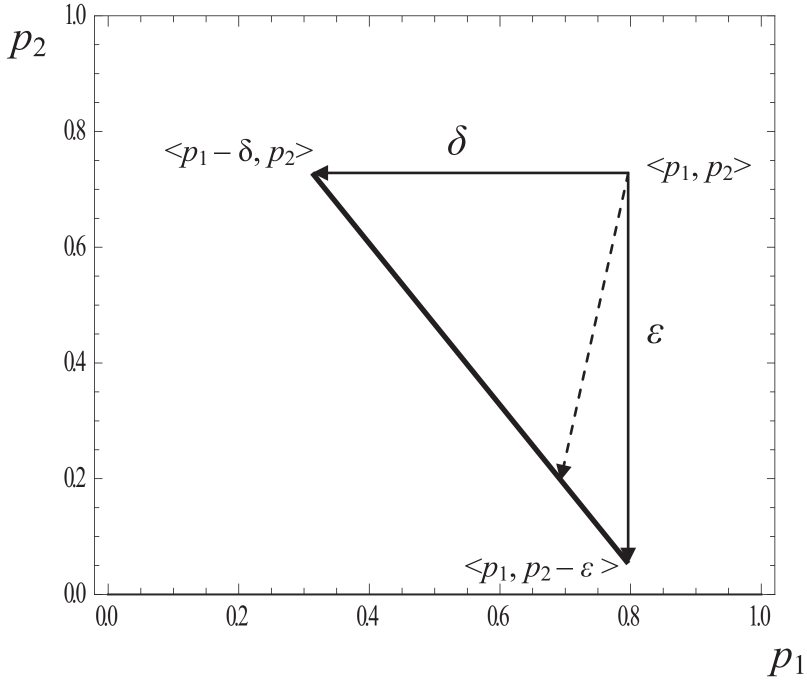


FIGURE 4. Technology determining the possibility set for risk reduction

4. TECHNOLOGY

We bring in the technology in Figure 4. Suppose that we start from a prospect $<p_1, p_2>$, represented by the tails of the arrows. We have a budget for risk reductions. If we focus all of our resources on person 1 we can reduce the risk for her with δ. If we focus all of our resources on person 2 we can reduce the risk for him with ε. Or we can choose any linear combination of these risk reductions. The possibility set is the triangle defined by the points $<p_1, p_2>$, $<p_1-\delta, p_2>$ and $<p_1, p_2-\varepsilon>$. The thick line is the Pareto frontier of the possibility set. The full-line arrows represent *corner solutions*, i.e. risk reductions that are fully focused on one person. The dashed arrow points to an *interior solution*, i.e. a risk reduction in which the resources are divided between the two people.

Return to Figure 2. We want to move to the point on the Pareto frontier at which v_α is maximal. This means that we are moving from the northeast (in which fatality chances are higher) to the southwest (in which fatality

invariance in survival chances, whereas another *ex ante–ex post* equivalence entails scale invariance in fatality chances.

chances are lower) and choose an arrow so that its head is in an area in which v_α is maximal – i.e. the shading in the contour plot is lightest. We track the risk reductions for five *ex ante* evaluations in Figure 2. The full arrows represent the risk reduction that is chosen, the dashed arrows represent risk reductions that are not chosen.

For an *anti-prioritarian* ($\alpha = -1$), the solution is a corner solution. Furthermore, the analyst is more inclined to favour people at low risk, i.e. she tends to favour a smaller risk reduction for a low-risk person over a greater risk reduction for a high-risk person. (Note how she favours person 2 over person 1 in Figure 2 with $\alpha = -1$.)

For a *prioritarian* ($\alpha = 1$), the solution may be a corner solution or an interior solution. Furthermore, the analyst favours people at high risk, i.e. she prefers to shift the bulk of the risk reduction towards the person at high risk. (Note how she favours person 1 over person 2 in Figure 2 with $\alpha = 1$.)

For a *utilitarian* ($\alpha = 0$), if $\delta = \varepsilon$, any solution on the Pareto frontier is open; if $\delta \neq \varepsilon$, the solution is a corner solution and the greatest risk reduction will be implemented, i.e. $<p_1 - \delta, p_2>$ if $\delta > \varepsilon$ and $<p_1, p_2 - \varepsilon>$ if $\varepsilon > \delta$. (Note how the analyst chooses the greater risk reduction in Figure 2 with $\alpha = 0$.)

For a *maximaxer* ($\alpha \to -\infty$), the solution is the corner solution that contains maximal higher survival chance (i.e. minimal lower fatality chance) of the points $<p_1 - \delta, p_2>$ and $<p_1, p_2 - \varepsilon>$. (Note how the head of the vertical arrow contains minimal lower fatality chance in Figure 2 with $\alpha \to -\infty$.)

For a *maximiner* ($\alpha \to \infty$), if the Pareto frontier crosses the diagonal (with $p_1 = p_2$) then the solution is an interior solution, viz. the point at which $p_1 = p_2$; If it does not cross the diagonal, then the solution is the corner solution that affords maximal lower survival chance (i.e. minimal higher fatality chance). (Note the equal fatality chances solution and the solution at the end of the horizontal arrow with the minimal higher fatality chance for person 1 in Figure 2 with $\alpha \to \infty$.)

In summary, as we move from the *maximax* over *anti-prioritarianism* to *utilitarianism*, solutions tend to be corner solutions and we tend to move away from favouring higher survival chances towards aiming for maximal risk reductions. As we move from *utilitarianism* over *prioritarianism* to the *maximin*, possibilities of interior solutions open up and we move away from aiming for maximal risk reduction towards favouring lower survival chances, which yields greater risk equity.

5. *EX POST* EVALUATIONS

The *ex post* analyst attaches utilities to *states*. There are four possible states, viz. (i) all die (i.e. S_0), (ii) person 1 survives and person 2 dies, (iii) person 1

$U(S_i)$		Omnitarian	Majoritarian	Paritarian	Minoritarian	Unitarian
S_2	All survive	1	1	1	1	1
S_1	Exactly 1 survives	0	(0, .5)	.5	(.5, 1)	1
S_0	All die	0	0	0	0	0

TABLE 1. *Ex post* utility functions for *omnitarians, majoritarians, paritarians, minoritarians* and *unitarians*

dies and person 2 survives and (iv) all survive (i.e. S_2). Let us assume anonymity, i.e. the analyst does not favour one person over the other. Hence the states (ii) and (iii) are of equal value to her and we can represent their disjunction as S_1. Then we can capture all the different types of *ex post* analysts in Table 1.

Remember that the utility of exactly one person surviving $U(S_1)$ equals β. So for the *ex post* analyst, the value v_β of the public works is

$$v_\beta(<p_1, p_2>) = E[U(S_i)] = \sum_{i=0,1,2} P(S_i)U(S_i) \quad (5.1)$$
$$= 0P(S_0) + \beta P(S_1) + 1P(S_2)$$
$$= \beta P(S_1) + P(S_2)$$

Assuming independence between fatality events:

$$P(S_1) = p_1(1-p_2) + (1-p_1)p_2 \quad (5.2)$$

$$P(S_2) = (1-p_1)(1-p_2) \quad (5.3)$$

We construct an ordering over $\{<p_1, p_2> | <p_1, p_2> \in [0, 1] \times [0, 1]\}$ for a given β:

$$<p_1^*, p_2^*> \succcurlyeq_\beta <p_1^\#, p_2^\#> \Leftrightarrow v_\beta(<p_1^*, p_2^*>) \geq v_\beta(<p_1^\#, p_2^\#>) \quad (5.4)$$

How this ordering behaves for values of $\beta \in [0, 1]$ will become clear when we compare the orderings for $\succcurlyeq_a$ and $\succcurlyeq_\beta$ in the next section.

6. COMPARING *EX ANTE* AND *EX POST* EVALUATIONS

6.1 Three shared rankings

The following three equivalences hold.[5] The (*ex ante*) *utilitarian* and (*ex post*) *paritarian* analyst share the same ranking:

(6.1.1) $<p_1^*, p_2^*> \succcurlyeq_{\alpha=0} <p_1^{\#}, p_2^{\#}> \Leftrightarrow <p_1^*, p_2^*> \succcurlyeq_{\beta=.5} <p_1^{\#}, p_2^{\#}>$

Furthermore, there is a particular (*ex ante*) *prioritarian* who shares the same ranking with the (*ex post*) *omnitarian*:

(6.1.2) $<p_1^*, p_2^*> \succcurlyeq_{\alpha=1} <p_1^{\#}, p_2^{\#}> \Leftrightarrow <p_1^*, p_2^*> \succcurlyeq_{\beta=0} <p_1^{\#}, p_2^{\#}>$

And finally, there is a particular (*ex ante*) *anti-prioritarian* who shares the same ranking with the (*ex post*) *unitarian*:

(6.1.3) $<p_1^*, p_2^*> \succcurlyeq_{\alpha=-1} <p_1^{\#}, p_2^{\#}> \Leftrightarrow <p_1^*, p_2^*> \succcurlyeq_{\beta=1} <p_1^{\#}, p_2^{\#}>$

How can we interpret these equivalences? (6.1.1) states that an *ex ante* analyst who is indifferent between making ε reductions at high or low fatality chances will construct the same ranking as an *ex post* analyst who values the survival of each additional person equally. So the contour plots for the *utilitarian* and the *paritarian* analysts are the same. According to (6.1.2), the ranking of an *ex ante* analyst who favours ε reductions at high fatality chances to a particular degree coincides with that of an *ex post* analyst who only cares about the survival of all. The last equivalence, (6.1.3), expresses that an *ex ante* analyst who favours ε reductions at low fatality chances to a particular degree will have the same ranking as an *ex post* analyst who only cares about the survival of at least one.

In Table 2, we have juxtaposed *ex ante* and *ex post* positions and have introduced some additional terms. We have constructed a continuum of *weak*, *moderate* and *extreme* (*anti*)-*prioritarianism*. *Moderate prioritarianism* maps onto *omnitarianism* (6.1.2) and *anti-prioritarianism* maps onto *unitarianism* (6.1.3). *Utilitarianism* maps onto *paritarianism* (6.1.1). We have also introduced *ex post maximin* and *maximax* labels, which we discuss in section 6.4. This table invites a number of observations which we will discuss in the remainder of this section.

[5] Note that the v_α and v_β functions that determine the rankings in (6.1.1) are order-preserving (and similarly in (6.1.2) and in (6.1.3)): for (6.1.1), $v_{\beta=.5}(<p_1, p_2>) = .5(p_1(1-p_2) + (1-p_1)p_2) + (1-p_1)(1-p_2) = 1-.5p_1-.5p_2$ and $v_{\alpha=0}(<p_1, p_2>) = (1-p_1) + (1-p_2) = 2(1-.5p_1 - .5p_2)$; for (6.1.2), $v_{\beta=0}(<p_1, p_2>) = (1-p_1)(1-p_2)$ and $v_{\alpha=1}(<p_1, p_2>) = \ln(1-p_1) + \ln(1-p_2) = \ln((1-p_1)(1-p_2))$; and for (6.1.3), $v_{\beta=1}(<p_1, p_2>) = p_1(1-p_2) + (1-p_1)p_2 + (1-p_1)(1-p_2) = 1-p_1p_2$ and $v_{\alpha=-1}(<p_1, p_2>) = -\ln(p_1)-\ln(p_2) = -\ln(p_1p_2)$.

Ex Ante			*Ex Post*	
	α		β	
Maximax *Extreme Anti-Prior^{ism}*	$-\infty$			
Strong Anti-Prior^{ism}	$(-\infty, -1)$			
Moderate Anti-Prior^{ism}	-1	$=$	1	*Unitarian* *Maximax*
Weak Anti-Prior^{ism}	$(-1, 0)$		$(.5, 1)$	*Minoritarian*
Utilitarian	0	$=$	.5	*Paritarian*
Weak Prior^{ism}	$(0, 1)$		$(0, .5)$	*Majoritarian*
Moderate Prior^{ism}	1	$=$	0	*Omnitarian* *Maximin*
Strong Prior^{ism}	$(1, \infty)$			
Maximin *Extreme Prior^{ism}*	∞			

TABLE 2. *Ex ante* and *ex post* positions

6.2 Keeney on risk equity and catastrophe avoidance

There is a connection between our work and Keeney's work (1980). Keeney shows that there is a conflict between the goal of risk equity, i.e. the goal of distributing risk equally between individuals, and the goal of catastrophe avoidance, i.e. minimizing the risk that a large number of individuals will be hit. Indeed, the best way to avoid a catastrophe is to concentrate the risk on a few individuals, but this sacrifice of a small group is anti-egalitarian.

The same point comes out in our analysis. In Table 2, focus on the row indicating the identity of the value functions of *unitarians* and *moderate anti-prioritarians*. If all we care about is to avoid the catastrophe of both dying, then we are *ex post* analysts with $\beta = 1$, i.e. *unitarians*. *Unitarians* order prospects in the same way as the *ex ante* analyst with $\alpha = -1$, i.e. the *moderate anti-prioritarian*. The *moderate anti-prioritarian* favours risk reductions for people with higher rather than lower survival chances, which will lead to more unequal distributions of risk. Hence, Keeney's theorem to the effect that there is a tension between catastrophe avoidance and risk equity is illustrated in our analysis.

6.3 Strong and extreme *ex ante* positions have no *ex post* correlate

Note in Table 2 that there are no *ex post* orderings that map onto *ex ante maximax* (i.e. *extreme anti-prioritarianism*) nor onto *strong anti-prioritarianism*; nor are there *ex post* orderings that map onto *ex ante* *maximin* (i.e. *extreme prioritarianism*) nor onto *strong prioritarianism*. What does this mean? Let us focus on the *prioritarianism* side.

Think of the *ex post omnitarians*, who aim to save everyone. This requires a certain *ex ante* concern for the plight of the weak – people at higher fatality chances matter more than people at lower fatality chances. But how much more?

Clearly not so much more that the analyst's *only* concern is the plight of the person with the lower survival chance. To see this, suppose that we have two prospects S_1 and S_2, each containing a person with high fatality chance and a person with low fatality chance. Let us suppose that S_1 contains a person with a higher fatality chance of .9 and S_2 contains a person with a higher fatality chance of .88.

The *ex ante maximiner* is so concerned with the plight of the worse off that she does not even need to know about the person with lower fatality chance. Her mind is already made up—she ranks S_2 over S_1.

However, the *omnitarian* is not so extreme in her favouring of the worst off. She ranks prospects in the same way as the *moderate prioritarian*. For $\beta = 0$ in (5.1) – (5.3) and $\alpha = 1$ in (3.1) – (3.3), the analyst is indifferent when the *products* of the survival chances are equal. She is willing to accept a decrease in the higher fatality chance, say, from .9 to .88, as long as this is not at the cost of an excessive increase in the lower fatality chance. For example, if the lower fatality chance is originally at .4 then it should not increase above .5.[6]

Clearly, the *ex ante* range of positions can go much further than the *ex post* range of positions. $\alpha = 1$ is just a mid-point in the *ex ante* range of positions. But $\beta = 0$ is the extreme point in the *ex post* range of positions. The same reasoning holds *mutatis mutandis* for the *anti-prioritarian* at $\alpha = -1$ and the *unitarian*.

6.4 The *ex ante minimax* and *maximax* versus the *ex post minimax* and *maximax*

Here is another way of making the same point. Note that for the *omnitarian*, $U(S_i) = \min\{u_1(S_i), u_2(S_i)\}$ for $i = 0, 1$ and 2: The social utility of zero people surviving $U(S_0)$ is the minimum of $\{u_1(S_0) = 0, u_2(S_0) = 0\}$, i.e. 0; the social utility of both people surviving $U(S_2)$ equals $\min\{u_1(S_2) = 1, u_2(S_2) = 1\} = 1$; and the social utility of one person surviving $U(S_1)$ equals $\min\{u_1(S_1) = 0, u_2(S_1) = 1\} = \min\{u_1(S_1) = 1, u_2(S_1) = 0\} = 0$. So we can

[6] Note that in terms of survival chances $(1-p_i)$, $(.1 \times .6) \leq (.12\times x)$ for $x \geq .5$.

rewrite the *omnitarian* value function as:

$$(6.4.1) \qquad v_{\beta=0}(<p_1, p_2>) = E[U(S_i)] = \sum_{i=0,1,2} P(S_i)U(S_i)$$
$$= \sum_{i=0,1,2} P(S_i) \min\{u_1(S_i), u_2(S_i)\}$$

Similarly for the *unitarian* $U(S_i) = \max\{u_1(S_i), u_2(S_i)\}$ for $i = 0, 1$ and 2. And so we can rewrite the *unitarian* value function as:

$$(6.4.2) \qquad v_{\beta=0}(<p_1, p_2>) = E[U(Si)] = \sum_{i=0,1,2} P(S_i)\,U(S_i)$$
$$= \sum_{i=0,1,2} P(S_i) \max\{u_1(S_i), u_2(S_i)\}$$

Hence the *omnitarian* position is an *ex post maximin* position and the *unitarian* position is an *ex post maximax* position. But the *ex post maximin* or *omnitarian* position maps onto *moderate prioritarianism* and not onto the *ex ante maximin*, i.e. *extreme prioritarianism*. The *ex post maximax* or *unitarian* position maps onto *moderate anti-prioritarianism* and not onto the *ex ante maximax*, i.e. *extreme anti-prioritarianism*.

6.5 *Weak prioritarians* versus *majoritarians* and *weak anti-prioritarians* versus *minoritarians*

What about the positions of the *ex ante weak prioritarian* and the *ex post majoritarian*? How do their orderings compare? To compare their respective orderings, it is instructive to focus first on the identities in our table. Note that the following claim holds true, since the contour plots for $\alpha = -1$, 0 and 1 are the same as the contour plots for $\beta = 1$, .5 and 0, respectively:

(6.5.1) For all values of α in the set {–1, 0, 1}, there exists a value β, such that α and β will yield the same ordering over $\{<p_1, p_2>|<p_1, p_2> \in [0, 1] \times [0, 1]\}$.

Does a similar claim hold for *weak prioritarian* values of α and *majoritarian* values of β? This would be a similar claim:

(6.5.2) For all *weak prioritarian* values α in the open interval (0, 1), there exists a *majoritarian* value β, such that α and β will yield the same ordering over $\{<p_1, p_2>|<p_1, p_2> \in [0, 1] \times [0, 1]\}$.

But this claim is false. There does not exist a *weak prioritarian* of the form (3.4) who holds the same ordering over $\{<p_1, p_2> | <p_1, p_2> \in [0, 1] \times [0, 1]\}$ as a *majoritarian*. Both the *weak prioritarian* and the *majoritarian* are characterized by contour plots that are concave to the origin, but the contour lines do not overlap. And similarly, there does not exist a *weak anti-prioritarian* – with values for α in the open interval $(-1, 0)$ – who holds the same ordering over $\{<p_1, p_2> | <p_1, p_2> \in [0, 1] \times [0, 1]\}$ as a *minoritarian*. Both the *weak anti-prioritarian* and the *minoritarian* are characterized by contour plots that are convex to the origin, but the contour lines do not overlap.

When it comes to implementations of the technology through risk reductions, clearly *ex post* analysts that are characterized by the same contour plots as *ex ante* analysts will make the same decisions; *ex post* analysts who are characterized by similar contour plots (in the sense that they are convex to the origin or concave to the origin) will make similar decisions. So our observations regarding the technology in section 4 for the various *ex ante* objectives can be repeated here for the matching *ex post* objectives.

6.6 Representing the *ex post* evaluation as a non-additively separable function of the expected utilities

One could ask if the negative result of the previous subsection – i.e. the non-overlapping contour plots – is due to the fact that we have adopted a special functional form in φ_α that is used in the *ex ante* objectives in (3.2). Could we find another functional form for the *ex ante* objectives, that would establish a correspondence with the (*ex post*) *majoritarian* and *minoritarian* objectives? In this subsection we show that the answer is negative. There is a connection here to Harsanyi's aggregation theorem (1955). We will see that in our framework of binary utilities the conditions in Harsanyi's theorem are much less constraining than in the general framework.

It is possible to write the *ex post* evaluation as a function of expected utilities. By the probability calculus, the chance of exactly one person surviving is the chance of person 1 surviving plus the chance of person 2 surviving minus twice the chance of both people surviving, and so $P(S_1) = E[u_1] + E[u_2] - 2\,E[u_1]E[u_2]$. Hence, the *ex post* analyst maximizes

$$(6.6.1) \qquad v_\beta(<p_1, p_2>) = \beta P(S_1) + P(S_2) = \beta(E[u_1] + E[u_2] - 2\,E[u_1]E[u_2]) + E[u_1]E[u_2]$$

So the *ex post* value function can be written as a function of expected utilities and hence one could conceive of it as an *ex ante* objective. But this function is not additively separable in $E[u_1]$, $E[u_2]$, as required in (3.3), except when $\beta = 0$, .5 or 1.

Furthermore, this non-additively separable function can be written as a weighted average of additive and multiplicative functions of the expected utilities. The *majoritarian* objective ($0 < \beta < .5$) can be expressed as a weighted average of additive and multiplicative functions of $E[u_i]$, since (6.6.1) can be written as:

$$\beta(E[u_1] + E[u_2]) + (1 - 2\beta)\, E[u_1]E[u_2] \tag{6.6.2}$$

The *minoritarian* objective ($.5 < \beta < 1$) can be expressed as a weighted average of additive and multiplicative functions of $(1-E[u_1])$, since (6.6.1) can also be written as:

$$1-(1 - \beta)(1 - E[u_1]+1 - E[u_2])-(2\beta - 1)(1 - E[u_1])(1 - E[u_2]) \tag{6.6.3}$$

However, we may reasonably expect that an *ex ante* analyst will insist on additive separability. When we laid out the *ex ante* position in section 2 we said that how much an ε improvement in, say, person 1's survival chance matters to the value of the prospect is determined by person 1's level of survival chance. If we stipulate in addition that how much an ε improvement in person 1's survival chance matters to the value of the prospect is determined *only* by person 1's level of survival chance – and *not* by person 2's level of survival chance – then we build in additive separability into the *ex ante* objectives. It is indeed quite natural to read this as a conversational implicature in our presentation of the *ex ante* objectives.

So, assuming additive separability, one can make the following claim about the intrapersonal motivations of a single analyst. If she wishes to construct an ordering that is sensitive to both *ex ante* and *ex post* objectives and that fully respects each of these objectives, then she must be either (i) both a *utilitarian* and a *paritarian*, or (ii) both a *moderate prioritarian* and an *omnitarian*, or (iii) both a *moderate anti-prioritarian* and a *unitarian*. For any other *ex ante* and *ex post* positions, one cannot construct an ordering over $\{<p_1, p_2>|<p_1, p_2>\in [0, 1] \times [0, 1]\}$ that simultaneously respects *ex ante* and *ex post* objectives. Of course an analyst might value both objectives and give weights to each objective to strike a balance. But this balance will not respect both objectives taken in isolation.

Similarly, one can make the following claim about interpersonal agreement between analysts. An *ex ante* and an *ex post* analyst can agree on an ordering just in case (i) the former is a *utilitarian* and the latter is a *paritarian*, (ii) the former is a *moderate prioritarian* and the latter is an *omnitarian*, or (iii) the former is a *moderate anti-prioritarian* and the latter is a *unitarian*. They have different motivations to back up their positions, but they agree on an ordering. But if they do not occupy these particular positions and do not pair up as laid out in (i), (ii) and (iii), then they will favour different orderings and they will need to compromise if they wish to strike an agreement.

Some pairs of *ex ante* and *ex post* objectives (or analysts) may not fall into (i), (ii) or (iii), but their preference orderings may nonetheless be quite close. Say, if one analyst is a mid-range *weak prioritarian* and the other a mid-range *majoritarian*, then their respective preference orderings are both characterized by contour plots that are slightly convex to the origin. These preference orderings may be sufficiently close so that (intrapersonal or interpersonal) compromise may be easily reached, but the initial orderings do not display exact overlap.

How do our results compare to Harsanyi's theorem (1955; see also Broome 1991: 160)? According to this theorem, an evaluation that is (i) a function of individual expected utilities (i.e. an *ex ante* objective) and (ii) can also be represented as the expected value of social utility (i.e. as an *ex post* objective) must be linear in individual utilities, i.e. it must be the *utilitarian* evaluation (or a weighted variant of it in case a violation of anonymity is accepted).

What we see here is that in the special case of two-valued utilities, the two conditions (i) and (ii) in Harsanyi's theorem put no constraint on the evaluations because every *ex post* objective can be written as a function of individual expected utilities (viz. survival chances). Moreover, even under the additional constraint that (iii) the *ex ante* evaluation be an additively separable function of individual expected utilities, the *moderate prioritarian* and *moderate anti-prioritarian* evaluations remain possible, in addition to the *utilitarian* evaluation.

7. CORRELATIONS

So far we have assumed that fatality chances are independent. But this is not always the case. It may be the case that both people are subject to positively correlated risks – if one incurs a fatality, then the other one is more likely to do so as well. Or alternatively, it may be the case that one person shields the other one from risks – if one person incurs a fatality, then the other one is less likely to do so. In this case risks are negatively correlated. How does this affect our analysis?

Correlations or anti-correlations do not affect the *ex ante* analysis since v_α is a function of the expectations of each person. As long as these expectations are held fixed, it does not matter whether risks are correlated or anti-correlated. But it does affect the *ex post* analysis.

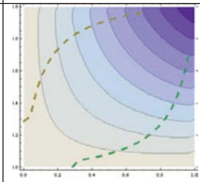
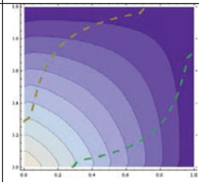
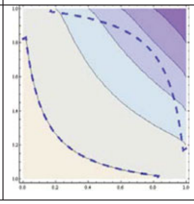
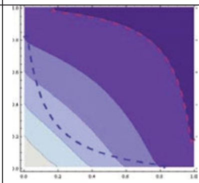
	unitarian	*omnitarian*
$r = .3$		
$r = -.3$		

FIGURE 5. *Unitarian* and the *omnitarian* contour plots with positive correlation ($r = .3$) and negative correlation ($r = -.3$). There exists a probability model for the values of $<p_1, p_2>$ inside the dashed lines.

The standard way to measure correlations between two variables X and Y is by means of the Pearson product-moment correlation coefficient:

$$r = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{7.1}$$

To calculate joint probabilities as a function of this correlation coefficient we follow Lucas (1995: 79):

$$P(S_0) = p_1 p_2 + r\sqrt{p_1(1 - p_1)p_2(1 - p_2)} \tag{7.2}$$

$$P(S_1) = p_1(1 - p_2) + (1 - p_1)p_2 - 2r\sqrt{p_1(1 - p_1)p_2(1 - p_2)} \tag{7.3}$$

$$P(S_2) = (1 - p_1)(1 - p_2) + r\sqrt{p_1(1 - p_1)p_2(1 - p_2)} \tag{7.4}$$

In Figure 5 we have plotted the contour plots for the *unitarian* and the *omnitarian* with positive ($r = .3$) and negative ($r = -.3$) correlations. Only the values of $<p_1, p_2>$ inside the dotted lines are meaningful. (The values of $<p_1, p_2>$ outside these lines do not have a probability model due to the constraints imposed by r – the image of these values is outside the interval $[0, 1]$.)

We presented the contour plots for the *moderate anti-prioritarian* with independent fatality chances in Figure 2 for $\alpha = -1$. This contour plot is the same as the contour plot for the *unitarian* with independent fatality chances, as we showed in the previous section. We now introduce correlations in the two left figures in Figure 5. For the *unitarian*, the contour lines are more convex to the origin when there are positive correlations and less convex to the origin when there are negative correlations than in the contour plot in Figure 2 for $\alpha = -1$, in which fatality chances are independent. Similarly, for the *omnitarian*, the contour lines are more concave to the origin when there are positive correlations and less concave to the origin when there are negative correlations in comparison to independence between fatality chances than in the contour plot in Figure 2 for $\alpha = 1$.

Can we get an intuitive feel for these results? We will restrict ourselves here to the case of *omnitarians* and positive correlations – the other cases can be explained in a similar vein. *Omnitarians*, like *moderate prioritarians*, favour people with lower survival chances. Now if we introduce positive correlations, then it becomes more likely that there are joint survivals – which is all *omnitarians* care about. If it becomes easier to secure joint survivals, then they will be even more intent on favouring people with lower survival chances. And this is precisely what increased concavity to the origin means.

8. AN *n*-PERSON MODEL

The results that we have obtained for two people also hold for n people, *mutatis mutandis*. The *moderate prioritarian* and the *omnitarian* (who cares only that all n survive) stand by the same ordering; similarly for the *moderate anti-prioritarian* and the *unitarian* (who cares only that at least one of the n survive); and for the *utilitarian* and the *paritarian* (with $U(S_0) = 0$; $U(S_1) = 1/n$; $U(S_2) = 2/n$; ...; $U(S_3) = 1$).

But extending the problem to n people does open up an additional entry in the *ex post* continuum, viz. what if our objective is not that at least one or all survive, but rather that at least j people survive for $1 < j < n$? Let us call this the (*ex post*) *strategist* position. We chose this term because a military strategist might act on the knowledge that she can win the battle if at least j of the n troops would still be standing.

Let us analyse this case for $n = 3$. For $j = 2$, we have utility assignments $U(S_0) = U(S_1) = 0$ and $U(S_2) = U(S_3) = 1$ with S_i being a state of the world with i people surviving. So we calculate the value of the public works for the *ex post strategist* as before:

$$(8.1) \qquad v(<p_1, p_2, p_3>) = E[U(S_i)] = \sum_{i=0,\ldots,3} P(S_i)U(S_i)$$
$$= P(S_2) + P(S_3)$$

Assuming independence between fatality events:

$$P(S_2) = p_1(1-p_2)(1-p_3) + (1-p_1)p_2(1-p_3) + (1-p_1)(1-p_2)p_3 \tag{8.2}$$

$$P(S_3) = (1-p_1)(1-p_2)(1-p_3) \tag{8.3}$$

We construct a contour plot for the *ex post strategist* in 3D space (Figure 6). Both the *strategist* and the *paritarian* occupy an intermediate position between the *unitarian* (with contour planes concave to the origin) and the *omnitarian* (with contour planes convex to the origin). For the *paritarian* the contour planes are flat. (The graph is omitted since they are simply 3D versions of the 2D version represented in Figure 2 with $\alpha = 0$.) But the *strategist* occupies a different intermediate position. Her position is characterized by contour planes that are convex to the origin at low fatality chances, relatively flat at mid-level fatality chances, and concave to the origin at high fatality chances. Note also that there is no *ex ante* objective that yields the same ordering as the *strategist*.

This has interesting consequences for decisions about implementing the technology. The recommendations for *unitarians, paritarians* and *omnitarians* are the same as in the two person case. However, the case of the *strategist* is novel. At low-level fatality chances, contour planes are convex to the origin and hence the analyst will make technology choices in the same way as the *unitarian*, i.e. choose corner solutions favouring people with lower fatality chances. At mid-range fatality chances, contour planes are relatively flat and hence the analyst will make technology choices in the same way as the *paritarian*, i.e. choose the corner solution that maximally reduces fatality chances. At high-level fatality chances, contour planes are concave to the origin and hence the analyst will make technology choices in the same way as the *omnitarian*, i.e. she will typically choose interior solutions favouring people with higher fatality chances.

For the *strategist*, the mathematical analysis yields results that exceed our intuitive grasp. What can be said is that the *strategist*, just like the *paritarian*, is a step on the *ex post* continuum from the *unitarian* to the *omnitarian*. The *paritarian* with flat contour planes occupies one intermediate position between *unitarian* convexity and *omnitarian* concavity to the origin. The *strategist* occupies another intermediate position with convexity to the origin at low-level fatality chances, relative flatness at mid-range fatality chances and concavity to the origin at high-level fatality chances.

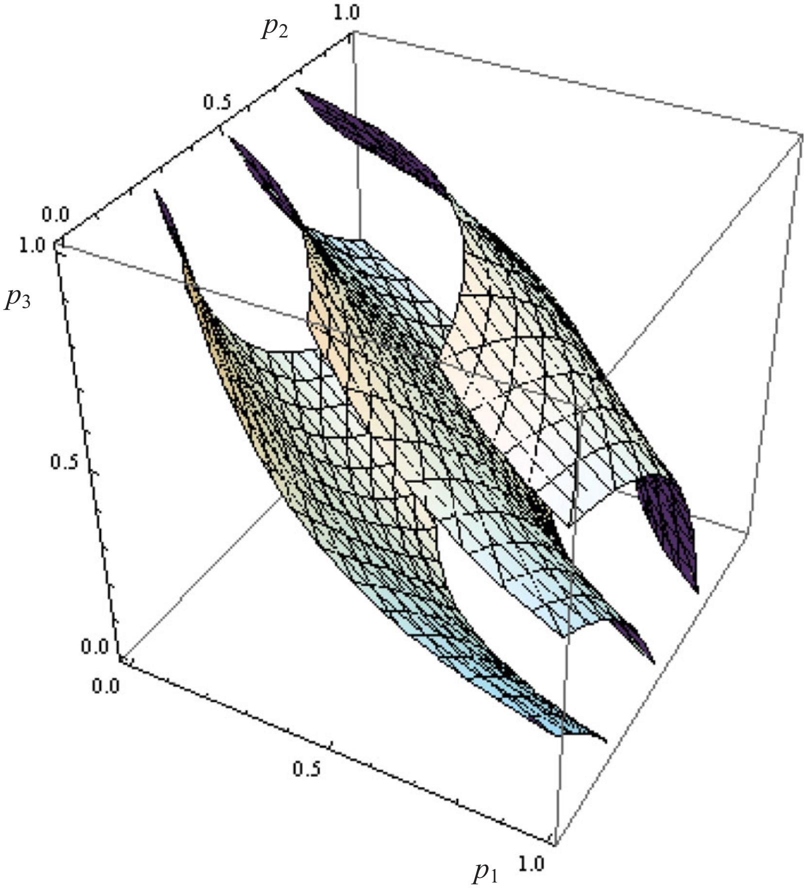


FIGURE 6. Contour plot for the *ex post strategist*.

9. *PRO-POORLY OFF* EVALUATIONS

9.1 Concerns for the poorly off

Our problem is a special case of the problem of the evaluation of prospects in general which has been extensively discussed. It is special in that we restricted ourselves to a framework of two-valued utilities. Let us look at the general framework and lift the restriction of two-valued utilities. Consider the prospects in Table 3 and Table 4. Table 3 presents a prospect that is uncertain and contains inequalities. Table 4 presents a prospect that is both certain and equal.

A variant of Harsanyi's theorem (1955) (see subsection 6.6 above) provides an axiomatic justification for the evaluation of prospects by

$L^{\mathbf{u\&u}}$	State 1 $p = .3$	State 2 $(1-p) = .7$
Person 1	2	.1
Person 2	.2	.4

TABLE 3. An uncertain and unequal prospect

$L^{\mathbf{c\&e}}$	$p = 1$
Person 1	.5
Person 2	.5

TABLE 4. A certain and equal prospect

means of a value function that takes the sum of people's expected utilities. So the value of a prospect L is

(9.1.1) $$v_U(L) = \sum_{i=1}^{n}\sum_{j=i}^{m} p_j u_{ij} \text{ for persons } i = 1, \ldots, n \text{ and}$$
states $j = 1, \ldots, m$ and u_{ij} is person i's utility if state j actualizes.

This is the *utilitarian* value function. We order prospects as follows:

(9.1.2) $$L^* \succcurlyeq L^\# \Leftrightarrow v_U(L^*) \geq v_U(L^\#)$$

On this utilitarian ordering, $L^{\mathbf{u\&u}} \succ L^{\mathbf{c\&e}}$, since $v_U(L^{\mathbf{u\&u}}) = ((.3\times 2) + (.7\times .1)) + ((.3\times .2) + (.7\times .4)) = 1.01 > 1 = .5 + .5 = v_U(L^{\mathbf{c\&e}})$.

But now suppose that some analyst would voice the following objection: 'Granted, I see the value of prospect $L^{\mathbf{u\&u}}$: person 1 has a relatively small chance of ending up at a high level of utility. But person 2 is bound to do worse than she would in prospect $L^{\mathbf{c\&e}}$ and person 1 has a relatively large chance of ending up substantially worse than she would in prospect $L^{\mathbf{c\&e}}$. On balance, I prefer $L^{\mathbf{c\&e}}$ to $L^{\mathbf{u\&u}}$.'

This does not seem an unreasonable position. The challenge is to spell out a calculus that would support this alternative ordering. We will present three value functions that are prominent in the literature and that support this alternative ordering. All these value functions could be interpreted as attempts to capture concerns for the poorly off in the evaluation of prospects.

Our *ex ante* model with $\alpha > 0$ is a special case of one of these value functions as applied to our special framework with binary utilities. Our *ex post* model with $\beta < .5$ is a special case of another one of these value functions as applied to our special framework with binary utilities. And the third value function is problematic because it fails to capture any concern for the poorly off when we restrict ourselves to the framework of binary utilities.

We will call the orderings based on these value-functions '*pro-poorly off* evaluations'. These evaluations are sometimes presented as contesting interpretations of what moral philosophers call 'prioritarianism' following Parfit's 'Priority View' (1991). Now we could follow this terminology, but then we would need to distinguish between this sense of 'prioritarianism' and the more narrow sense of 'prioritarianism' as one side of the continuum of *ex ante* evaluations, as defined above. To avoid confusion, we will talk about *pro-poorly off* evaluations instead.

9.2 Diamond's, Rabinowicz's and Fleurbaey's *pro-poorly off* value functions

Pro-poorly off analysts may calculate the expected utility for each person, construct a strictly concave φ transform of this expectation and sum these φ transforms. This value function is commonly ascribed to Diamond (1967)[7]:

(9.2.1) $$v_D(L) = \sum_{i=1}^{n} \varphi\left(\sum_{j=i}^{m} p_j u_{ij}\right) \text{ for persons } i = 1, \ldots, n \text{ and states } j = 1, \ldots, m \text{ and } \varphi \text{ is a strictly concave function.}$$

According to Rabinowicz, *pro-poorly off* analysts should construct a strictly concave φ transform of each utility value u_{ij}, sum these transforms within each state to determine the social utility of the state and then construct the expectation of the social utility of a state. So, for Rabinowicz,

(9.2.2) $$v_R(L) = \sum_{j=1}^{m} p_j \sum_{i=i}^{n} \varphi(u_{ij}) \text{ for persons } i = 1, \ldots, n \text{ and states } j = 1, \ldots, m \text{ and } \varphi \text{ is a strictly concave function.}$$

According to Fleurbaey, *pro-poorly off* analysts should construct a strictly concave transform φ of each utility value u_{ij}, calculate the average value

[7] Diamond (1967) actually does not lay out this value function. He offers two prospects that are structurally analogous to the prospects in Table 6 and objects that Harsanyi's social welfare function assigns equal value to them. The value function $v_D(L)$ is commonly attributed to Diamond (1967), because it does indeed assign greater value to one of the prospects in Table 6 than to the other, as we will see below. A statement of the social welfare function in (9.2.1) can be found in McCarthy (2006: 339) and Adler and Sanchirico (2006: 306).

	$L^{\mathbf{u\&u}}$	$L^{\mathbf{c\&e}}$	
v_D	$\sqrt{}((.3(2)+.7(.1))) + \sqrt{}((.3(.2)+.7(.4)) = 1.40$	$\sqrt{.5}+\sqrt{.5} = 1.41$	$L^{\mathbf{c\&e}} > L^{\mathbf{u\&u}}$
v_R	$.3(\sqrt{2}+\sqrt{.2}) + .7(\sqrt{.1}+\sqrt{.4}) = 1.22$	$\sqrt{.5}+\sqrt{.5} = 1.41$	$L^{\mathbf{c\&e}} > L^{\mathbf{u\&u}}$
v_F	$.3((\sqrt{2}+\sqrt{.2})/2)^2 + .7((\sqrt{.1}+\sqrt{.4})/2)^2 = .42$	$((\sqrt{.5}+\sqrt{.5})/2)^2 = .5$	$L^{\mathbf{c\&e}} > L^{\mathbf{u\&u}}$

TABLE 5. Comparing $L^{\mathbf{u\&u}}$ and $L^{\mathbf{c\&e}}$ on Diamond's, Rabinowicz's and Fleurbaey's value functions

of these transforms within each state, and construct the inverse transform φ^{-1} of these averages. This is the *equally distributed equivalent* (following Kolm 1968; Atkinson 1970: 250) of the state which is a measure of the value of the state. They then construct the expectation of the *equally distributed equivalent*. So, for Fleurbaey,

$$\text{(9.2.3)} \quad v_F(L) = \sum_{j=1}^{m} p_j \varphi^{-1}\left(\left(\sum_{i=i}^{n} \varphi(u_{ij})\right)/n\right) \text{ for persons } i = 1, \ldots, n$$
$$\text{and states } j = 1, \ldots, m \text{ and } \varphi \text{ is a strictly concave function.}$$

Let us explain the role of the *equally distributed equivalent* of a state. Take a state with distribution of utility 25 for person 1 and utility 9 for person 2. Then the *equally distributed equivalent* is the utility value which is such that the analyst would be indifferent between a state with person 1 at 25 and person 2 at 9 and a state in which every person has the utility of the *equally distributed equivalent*. So suppose that the φ transform is the square root function (which is a strictly concave transform) and hence φ^{-1} is the square function. Then the *equally distributed equivalent* equals $((\sqrt{25} + \sqrt{9})/2)^2 = 16$. So the analyst is indifferent between a state in which one person is at 25 and the other is at 9 on the one hand and a state in which both are at 16. (A utilitarian, on the other hand, is indifferent between the former state and a state in which both people are at average utility, i.e. at 17.)

Notice that all of these proposals with the square root as the φ transform provide us with the desired result, i.e. they overturn the utilitarian ordering $L^{\mathbf{u\&u}} > L^{\mathbf{c\&e}}$ – as Table 5 shows.

But of course v_D, v_R and v_F do not provide us with the same orderings over all sets of prospects. Much can be said about the desirability of these contesting conceptions, but we will not enter this debate here. A discussion can be found in Adler and Sanchirico (2006) and Fleurbaey (2010: 649–652). We are interested in how these different *pro-poorly off*

<table>
<tr><td>suicide mission
$L^{\mathbf{sm}}$</td><td>State 1
p</td><td>State 2
$(1-p)$</td><td></td><td>drawing straws
$L^{\mathbf{ds}}$</td><td>State 1
p</td><td>State 2
$(1-p)$</td></tr>
<tr><td>Person 1</td><td>0</td><td>0</td><td></td><td>Person 1</td><td>0</td><td>1</td></tr>
<tr><td>Person 2</td><td>1</td><td>1</td><td></td><td>Person 2</td><td>1</td><td>0</td></tr>
<tr><td colspan="3">$v_D(L^{\mathbf{sm}}) = \sqrt{0} = \sqrt{1} = 1$</td><td>$<$</td><td colspan="3">$v_D(L^{\mathbf{ds}}) = \sqrt{(1-p)} + \sqrt{p}$</td></tr>
<tr><td colspan="3">$v_R(L^{\mathbf{sm}}) = p\sqrt{1} + (1-p)\sqrt{1} = 1$</td><td>$=$</td><td colspan="3">$v_R(L^{\mathbf{ds}}) = p\sqrt{1} + (1-p)\sqrt{1} = 1$</td></tr>
<tr><td colspan="3">$v_F(L^{\mathbf{sm}}) =$
$p((\sqrt{1})/2)^2 + (1-p)((\sqrt{1})/2)^2 = 1/4$</td><td>$=$</td><td colspan="3">$v_F(L^{\mathbf{ds}}) =$
$p((\sqrt{1})/2)^2 + (1-p)((\sqrt{1})/2)^2 = 1/4$</td></tr>
</table>

TABLE 6. **Suicide mission** and **drawing straws** with $0 < p < 1$

evaluations play out within our restricted framework of two-valued utilities.

9.3 Two concerns for the poorly off

There are two concerns for the poorly off in our framework and we wish to investigate how these concerns fare on v_D, v_R and v_F.

The first concern relates to the observation that an analyst is typically more tolerant towards distributed risk than towards focused risk. She may take some risks and be tolerant of an expected fatality count greater than 1 but she does not like to send a single person into certain death. One could interpret this as a concern for the poorly off – the analyst abhors the focused risk of sending someone into certain death. This kind of concern is captured by the prospects in Table 6. Even if someone is bound to die, one might think that it is better if the analyst determines that straws be drawn rather than that she designates who will go on a suicide mission. By doing so the risk is distributed rather than focused.

The second concern is the concern that misery loves company. An analyst may not like it that states actualize in which some people flourish and others do not. Consider the following three prospects. Person 1 and person 2 may both have an equal chance of dying kept constant across all prospects. But on prospect 1, the risks are fully correlated; on prospect 2, they are independent; and on prospect 3, the risks are fully anti-correlated. In Table 7, we present these prospects with the risk to each person set at .5. On a 'misery loves company' understanding of our concern for the poorly off, it is better to be subjected to correlated risk

L^{cor}	$P(S_1)$ $=.5$	$P(S_2)$ $=.5$
P1	1	0
P2	1	0

L^{ind}	$P(S_1)$ $=.25$	$P(S_2)$ $=.25$	$P(S_3)$ $=.25$	$P(S_4)$ $=.25$
P1	0	1	0	1
P2	0	0	1	1

L^{ac}	$P(S_1)$ $=.5$	$P(S_2)$ $=.5$
P1	0	1
P2	1	0

$v_D(L^{\text{cor}}) = \sqrt{5} + \sqrt{.5}$	$=$	$v_D(L^{\text{ind}}) = \sqrt{.5} + \sqrt{.5}$	$=$	$v_D(L^{\text{ac}}) = \sqrt{.5} + \sqrt{.5}$
$v_R(L^{\text{cor}}) =$ $.5(2\sqrt{1}) + .5(2\sqrt{0}) = 1$	$=$	$v_R(L^{\text{ind}}) =$ $.25(2\sqrt{0}) + .5(\sqrt{1}) + .25(2\sqrt{1}) = 1$	$=$	$v_R(L^{\text{ac}}) =$ $.5(\sqrt{1}) + .5(\sqrt{1}) = 1$
$v_F(L^{\text{cor}}) =$ $.5(\sqrt{1})^2 + .5(\sqrt{0})^2 = 1/2$	$>$	$v_F(L^{\text{ind}}) =$ $.25(\sqrt{0})^2 = .5((\sqrt{1})/2)^2 + .25(\sqrt{1})^2 = 3/8$	$>$	$v_F(L^{\text{ac}}) =$ $.5((\sqrt{1})/2)^2 + .5((\sqrt{1})/2)^2 = 1/4$

TABLE 7. **Correlated risk**, **independent risk** and **anti-correlated risk**

than to independent risk and it is better to be subjected to independent risk than to anti-correlated risk.

Notice the equalities and inequalities in Tables 6 and 7. v_D captures the first type of concern for the poorly off, v_F captures the second type of concern, but v_R captures neither. Hence, in a framework with binary variables, Rabinowicz's approach cannot capture a concern for the poorly off neither in the sense of steering clear from focused risk, nor in the sense of steering clear from solitary misery.

Rabinowicz's response (in personal communication) is that he simply does not wish to capture such concerns. He only wishes to capture what he takes to be Parfit's notion of prioritarianism, viz. that ε-changes in utilities should count for more at a lower rather than at a higher level. And since expressing such comparisons involves more than two utility values, it is no wonder that his value function does not capture any pro-poorly off concerns in a framework of binary utilities.

Our concern is broader. We wish to capture concerns for the poorly off of any kind that are not captured by the utilitarian value function. There clearly are such concerns in a framework of binary utilities.

9.4 Our *ex ante* and *ex post* evaluations compared with Diamond's and Fleurbaey's value functions

Our *ex ante* evaluation for $\alpha > 0$ in (3.2)–(3.4) is a particular instantiation of v_D. We impose a particular family of φ transforms on the expected utility of each person. Our choice of this family is motivated by the fact that by manipulating a single parameter we can slide from *utilitarianism* via *prioritarianism* to the *ex ante maximin* and on the route we can capture the *moderate prioritarian*, whose ordering coincides with the *omnitarian*.

Our *ex post* evaluation for $\beta < .5$ can be interpreted as a particular instantiation of v_F. Define the following strictly concave φ transform with parameter γ:

(9.4.1) $\varphi_\gamma(x) = x^{1/(\gamma+1)}$ for $0 < \gamma \leq \infty$ and undefined otherwise.

And so $\varphi_\gamma{}^{-1}(x) = x^{(\gamma+1)}$. The concavity of the function increases with greater values of γ expressing greater concern for the poorly off.

As in section 5, we distinguish between three types of states, viz. states in which nobody survives (S_0), states in which one person survives (S_1) and states in which both survive (S_2). The *equally distributed equivalent* in S_0 is clearly 0 and in S_2 is clearly 1. In S_1 the *equally distributed equivalent* is

(9.4.2) $$\varphi_\gamma^{-1}((\varphi_\gamma(0) + \varphi_\gamma(1))/2) = \varphi_\gamma^{-1}(.5)$$

We can now define v_F within our two-person framework:

(9.4.3) $$\begin{aligned} v_F(L) &= 0P(S_0) + \varphi_\gamma^{-1}(.5)P(S_1) + 1P(S_2) \\ &= \varphi_\gamma^{-1}(.5)P(S_1) + P(S_2) \end{aligned}$$

Notice that $\varphi_\gamma{}^{-1}(.5)$ is a monotonically decreasing function of γ from domain $[0, \infty]$ to range $[0, .5]$. Now compare (9.4.3) with (5.1): $v_F(L)$ equals $v_\beta(<p_1, p_2>)$ if we equate $\varphi_\gamma{}^{-1}(.5)$ in the former with β in the latter. In other words, the γ-value which expresses the concern for the poorly off within Fleurbaey's *pro-poorly off* value function, maps onto our β value, which expresses the *majoritarian* concern within our special framework.

This formal result has an intuitive interpretation. The *majoritarian* concern can indeed be interpreted in the same style as Fleurbaey's concern for the poorly off: To set β low in $v_\beta(<p_1, p_2>)$ is to be strongly concerned that not just one but both people live; To set γ high in $v_F(L)$ is to be strongly concerned about the unsaved life in the presence of the saved life. Both concerns are tantamount.

9.5 Suicide mission and drawing straws

How do our *ex ante* value function v_α and our *ex post* value function v_β fare with **suicide mission** and **drawing straws**?

Our *ex ante* approach is an extension of the v_D from the *prioritarian* orderings into the *anti-prioritarian* orderings for binary utility values. For **suicide mission**,

(9.5.1) $$v_\alpha(<p_1 = 1, p_2 = 0>) = \sum_{i=1,2} \varphi_\alpha(E[u_i]) = \varphi_\alpha(0) + \varphi_\alpha(1).$$

For **drawing straws**,

$$(9.5.2) \qquad v_\alpha(<p_1 = p,\, p_2 = (1-p)>) = \varphi_\alpha(1-p) + \varphi_\alpha(p).$$

It is easy to verify that

$$(9.5.3) \quad v_\alpha(<p_1 = p,\, p_2 = (1-p)>) \geq v_\alpha(<p_1 = 1,\, p_2 = 0>) \Leftrightarrow \alpha \geq 0.$$

Hence *prioritarians* prefer **drawing straws**, *anti-prioritarians* prefer **suicide mission**, and *utilitarians* are indifferent.

Our *ex post* approach cannot distinguish between **suicide mission** and **drawing straws**. In either case, one person will be left standing. So the v_β will simply equal β, i.e. the value of one person surviving. And this coincides with Fleurbaey's approach on which $v_F(L^{\mathbf{sm}}) = v_F(L^{\mathbf{ds}})$, i.e. the *equally distributed equivalent* of one person left standing.

This result maps onto our presentation in section 7. Notice that we have perfect anti-correlation between fatality chances in both **suicide mission** and **drawing straws** and so $r = -1$. In the lower graphs in Figure 5, the area between the dashed lines will, for $r = -1$, shrink to a single contour line, viz. the northwest-southeast diagonal containing all values of $<p_1 = p, p_2 = (1-p)>$ for all probability values p. And this is the case in the *omnitarian* (left), the *paritarian* (not represented), and *unitarian* (right) contour plots. Hence $<p_1 = 1, p_2 = 0>$ (**suicide mission**) and $<p_1 = p, p_2 = (1-p)>$ (**drawing straws**) are on the same contour line in any *ex post* evaluation.

9.6 Correlated risk, independent risk and anti-correlated risk

How does our *ex ante* approach and *ex post* approach fare with **correlated risk** $L^{\mathbf{cor}}$, **independent risk** $L^{\mathbf{ind}}$ and **anti-correlated risk** $L^{\mathbf{ar}}$?

Like the *pro-poorly off* value function v_D, our *ex ante* approach does not make any distinctions for *prioritarian*, *utilitarian* or *anti-prioritarian* values of α. In each case, the *ex ante* evaluation is $v_\alpha(<p_1 = .5, p_2 = .5>)$ since each person's expectation $\mathrm{E}[u_i]$ is precisely the same.

On our *ex post* approach, $v_\beta = .5$ for $L^{\mathbf{cor}}$ (since there is a .5 chance of two people standing), $v_\beta = .25 + .5\beta$ for $L^{\mathbf{ind}}$ (since there is .25 chance of two people standing and a .5 chance of one person standing) and $v_\beta = \beta$ for $L^{\mathbf{ac}}$ (since there is a certainty of one person standing). Hence like the *pro-poorly-off* value function v_F, our *ex ante* approach ranks $L^{\mathbf{cor}} \succ L^{\mathbf{ind}} \succ L^{\mathbf{ac}}$ for *omnitarian* and *majoritarian* values of $\beta < .5$. Furthermore, for *paritarians* (i.e. for $\beta = .5$), $L^{\mathbf{cor}} \sim L^{\mathbf{ind}} \sim L^{\mathbf{ac}}$; and for *minoritarians* and *unitarians* (i.e. for $\beta > .5$), $L^{\mathbf{ac}} \succ L^{\mathbf{ind}} \succ L^{\mathbf{cor}}$.

9.7 *Ex post* evaluations and the principle of personal good.

It is well-known that *ex post* evaluations do not respect the **principle of personal good**. The **principle of personal good** says that if every person's expectation is at least as good in one prospect as in another and at least one person's expectation is better, then the former prospect is better than the latter. Here is a case in which an *ex post* evaluation violates the **principle of personal good**. Take **drawing straws** in Table 7 and set p at .5. Then each person's expected utility is .5. Now suppose that we give each person $.5-\varepsilon$ for sure for small ε. (Note that we bring in a third utility value.) Then the expectation of both people is greater in **drawing straws** than in the sure prospect and so, by the **principle of personal good, drawing straws** is better than the sure prospect. But an analyst using the *ex post* value function v_F will be averse to the distribution in **drawing straws** and rank the sure prospect over **drawing straws**.

Can we have the same phenomenon in life or death prospects (i.e. within the framework of binary utilities) with independent risks? The answer is no: Within our framework with binary utilities and on the assumption of independent risks, *ex post* evaluations always respect the **principle of personal good**. To see this in the two-person case, consider the *ex post* value function in (5.1) with independent chances set as in (5.2) and (5.3). Now take the derivatives of this value function toward the expectations, i.e. toward the survival chances $(1-p_i)$ for $i = 1, 2$. These derivatives equal $(1-\beta)(1-p_j) + \beta p_j$ for $j = 1, 2$ and $i \neq j$. Note that these derivatives are always non-negative for the admissible values of the parameters. Hence the *ex post* value function is a non-decreasing function of the individual expectations *ceteris paribus*. Hence, it cannot be the case that one person's expectation is raised *ceteris paribus* in prospect $L^{\#}$ relative to L^* and yet the *ex post* value function would come to rank L^* over $L^{\#}$. (The same holds when the number of people $n > 2$.)

10. SUMMARY

In ordering life or death prospects, there are two grand distinctions. First, do we base our judgement on the expected survival chances of the participants (*ex ante*) or on the chance that at least a particular fraction of the population will survive (*ex post*)? Second, are we soft-hearted or hard-hearted about the matter? Soft-hearted *ex ante* analysts have a soft spot for participants at high fatality risk. Soft-hearted *ex post* analysts are willing to take chances so that all (or most) participants may survive. Hard-hearted *ex ante* analysts favour participants at low fatality risk. Hard-hearted *ex post* analysts are willing to sacrifice people so as to avoid catastrophes, i.e. situations in which nobody (or only few) would survive.

We have laid out a range of *ex ante* and *ex post* evaluations that range from soft-heartedness to hard-heartedness and the aim to minimize the expected loss of human lives is a mid-point on both of these continua. To conclude, we list the following 12 findings:

1. There is a range of *ex post* evaluations and a range of *ex ante* evaluations with additively separable value functions that yield *roughly* parallel orderings. Contour plots display concavity towards the origin for soft-heartedness and convexity towards the origin for hard-heartedness.
2. However, *ex ante* evaluations can go to extremes of soft-heartedness (*maximin*) and hard-heartedness (*maximax*) whereas *ex post* evaluations cannot.
3. *Ex post* evaluations and *ex ante* evaluations with additively separable functions agree on *precisely* the same orderings if and only if (i) the *ex post* aim is to save at least one (*unitarian*), in which case it coincides with the *moderate anti-prioritarian* value function; (ii) the *ex post* aim is to save all (*omnitarian*), in which case it coincides with the *moderate prioritarian* value function; or (iii) we attribute equal value to each life saved (*paritarian*) in which case it coincides with the *utilitarian* value function.
4. Within the more restrictive framework of binary utilities, the conditions in Harsanyi's theorem put no constraints on the set of permissible evaluations since every *ex post* value function can be written in terms of an *ex ante* value function of individual expectations. If we add the constraint that *ex ante* value functions should be additively separable, it still leaves room for *moderate prioritarian* and *moderate anti-prioritarian* evaluations, aside from the *utilitarian* evaluation.
5. There are multiple continua from soft-heartedness to hard-heartedness. We have spelled out three – one over the *ex ante* route and two over the *ex post* route. Continua over the *ex post* route pass via the *paritarian* objective or via the *strategist* objective – i.e. aiming to save at least m out of the n people for $1 < m < n$. These continua all yield different orderings.
6. The *strategist* objective for $n = 3$ resembles the *unitarian* ordering at low-level fatality chances, the *paritarian* ordering on mid-level fatality chances, and the *omnitarian* ordering at high-level fatality chances.
7. As to technology choices, the *utilitarian* objective to minimize expected fatalities focuses risk reductions on the person who can receive the greatest reduction in her fatality chance. Unless possible risk reductions are equal for both people, this will lead to a corner solution (i.e. an allocation benefitting a single person).

Hard-heartedness yields corner solutions favouring people at low risk. With soft-heartedness the possibility of interior solutions (i.e. allocations benefiting multiple people) opens up – these allocations are slanted towards favouring people at high risk.

8. Positive correlations between fatality chances affect *ex post* orderings. They yield orderings that map onto orderings for more extreme positions – i.e. the *omnitarian* ordering tends toward the *ex ante maximin* and the *unitarian* ordering tends towards the *ex ante maximax* ordering; negative correlations between fatality chances yield *ex post* orderings that tend towards the *utilitarian* ordering.
9. Within our framework with binary utilities, Rabinowicz's *pro-poorly off* value function is not sensitive to a concern for the poorly off, neither as a concern for focused (i.e. undistributed) risk nor as a concern for solitary misery. It simply reduces to a utilitarian evaluation.
10. Our *ex ante prioritarian* evaluation is an instance of Diamond's *pro-poorly off* value function. As such it is sensitive to the poorly off in that it is concerned about focused risk (but not about solitary misery).
11. Our *ex post majoritarian* and *omnitarian* evaluations are instances of Fleurbaey's *pro-poorly off* value function. As such they are sensitive to the poorly off in that these evaluations are concerned about solitary misery (but not about focused risk).
12. Within our framework of binary utilities and assuming independent risk, the *ex post* evaluation respects the **principle of personal good**, unlike in the general framework.

REFERENCES

Adler, M. 2011. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford: Oxford University Press.

Adler, M.D. and C.W. Sanchirico 2006. Inequality and uncertainty: theory and legal applications. *University of Pennsylvania Law Review* 155: 279–377.

Atkinson, A.B. 1970. On the measurement of inequality. *Journal of Economic Theory* 2: 244–263.

Boadway, R. and N. Bruce 1984. *Welfare Economics*. Oxford: Basil Blackwell.

Bossert, W. and J.A. Weymark 2004. Utility in social choice. In *Handbook of Utility Theory*, vol. 2 (*Extensions*), eds. S. Barberà, P.J. Hammond and C. Seidl, 1099–1177. Boston: Kluwer Academic.

Broome, J. 1991. *Weighing Goods*. Oxford: Blackwell.

Diamond, P. 1967. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: comment. *Journal of Political Economy* 75: 765–766.

Fleurbaey, M. 2010. Assessing risky social situations. *Journal of Political Economy* 118: 649–680.

Harsanyi, J. 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy* 3: 309–321.

Keeney, R.L. 1980. Equity and public risk. *Operations Research* 28: 527–534.

Kolm, S.C. 1968. The optimal production of social justice. In *Public Economics*, eds. H. Guitton and J. Margolis, 145–200. London: Macmillan.

Lucas, D.J. 1995. Default correlations and credit analysis. *Journal of Fixed Income* 4: 76–87.

McCarthy, D. 2006. Utilitarianism and prioritarianism I. *Economics and Philosophy* 22: 335–63.

Parfit, D. 1991. Equality or priority? *The Lindley Lecture*. Lawrence: University of Kansas. Reprinted in *The Ideal of Equality*, ed. M. Clayton and A. Williams, 2002, 81–125. Basingstoke: Palgrave.

Rabinowicz, W. 2002. Prioritarianism for prospects. *Utilitas* 14: 2–21.

Roberts, K.W.S. 1980. Interpersonal comparability and social choice theory. *Review of Economic Studies* 47: 421–439.