The Efficiency-Equity Trade-off, Self-Interest, and Moral Principles in Health and Safety Valuation

Danae Arroyos-Calvera[a†]
Judith Covey[b]
Graham Loomes[c]
Rebecca McDonald[d2]

[a]Warwick Business School, University of Warwick
Coventry, CV4 7AL, UK

[b]Department of Psychology, University of Durham
j.a.covey@durham.ac.uk
Upper Mount Joy, South Rd,
Durham, DH1 3LE, UK

[c]Warwick Business School, University of Warwick
graham.loomes@wbs.ac.uk
Coventry, CV4 7AL, UK

[d]Department of Economics, University of Birmingham
r.l.mcdonald@bham.ac.uk
Edgbaston, B15 2SB, UK

[†]Corresponding author;
d.arroyoscalvera@bham.ac.uk; +44(0)1214146656

Acknowledgements:
We are grateful to the Leverhulme Trust for their financial support (Grant Number RP2012-V-022). Graham Loomes also received support from the Economic and Social Research Council's funding of the Network for |Integrated Behavioural Science (Grant Numbers ES/K002201/1 and ES/P008976/1.
We should like to thank Michael Jones-Lee, Jose Luis Pinto-Prades, Andrea Isoni, Matteo Galizzi, Tigran Melkonyan, Richard Cookson, and the participants of the 2018 Experimental Workshop on Social Preferences at Birmingham, and the Summer 2018 Health Economists Study Group (particularly Nils Gutacker) for valuable comments.

Abstract

Policy makers try to take account of public preferences when making trade-offs between policy options. Yet most estimates of the value of health and safety reflect only individuals' self-interested preferences, neglecting their preferences over the *distribution* of public resources. We conduct an experiment in which participants choose between policy options that differ in their efficiency (expected number of fatalities or cases of ill health they would prevent) and their equity (defined in terms of the balance of risk reductions for different sections of the population). The policy options were framed as interventions to improve a hypothetical city's water supply that would reduce the risk of death or ill health for people in different areas of the city to varying degrees. In order to examine whether self-interest would affect the trade-offs, we asked half of the sample about scenarios where they would personally benefit from some options. Our results suggest that efficiency is the most important single factor determining preferences between policy options, but decisions were influenced almost as much by equity as by efficiency. The effect of self-interest was smaller than that of the general concern for efficiency. We also elicited participants' stated moral principles regarding trade-offs between equity, efficiency and self-interest, and found that their expressed principles were well-aligned with their choices. Our findings contribute to the growing evidence that distributional concerns matter when evaluating health interventions. [228]

Keywords: efficiency, equity, self-interest, moral principles

1. Introduction

When deciding how to allocate resources to maximise societal welfare, difficult trade-offs are inevitable. Should resources be spent on education or healthcare? Are some health improvements more valuable than others? Such trade-offs involve weighing benefits and costs of different magnitudes, at different times, and to different recipients.

Policy makers try to take account of public preferences when making trade-offs between policy options. This often involves using estimates of the monetary value of policy outcomes (e.g., fatalities prevented, health improvements, or environmental damage prevented or rectified), obtained using methods including contingent valuation (e.g., Carthy et al., 1999). However, most of these estimates reflect only the preferences that individuals express for their personal benefits and costs, inferred from their stated willingness to pay for reductions in their own risks or improvements in their own health or benefits to their own enjoyment of environmental amenities. To some extent, this self-interested approach is legitimate, since it avoids double counting of the benefits of policies (for a useful discussion in relation to the Value of Statistical Life, see Jones-Lee, 1991). However, there is good reason to suppose that people have preferences which extend beyond concern for themselves. These preferences include concern for efficiency (maximising the expected number of fatalities or cases of illness prevented) and equity (balancing risk reductions for different sections of the population). By 'equity' we mean 'gains egalitarianism' rather than 'outcome egalitarianism' (see Tsuchiya and Dolan (2009)). Focusing exclusively on individuals' values for their own outcomes may neglect their preferences over efficiency and equity considerations and thereby fail to adequately represent important sources of societal welfare.

To ensure that resource allocations maximise social welfare, policies should ideally reflect the trade-offs that individuals would make between self-interest, efficiency and other distributional concerns. However, we do not yet have an adequate account of these trade-offs in applied contexts. This paper

presents evidence that begins to address this issue: we elicit preferences in the context of health and safety policy choices using a structured experimental design that helps us to disentangle the various factors entering into preferences for different policy options.

Section 2 discusses the relevant literature. Section 3 presents our experimental design. In section 4 we outline our results, and in Section 5 we discuss their implications.

2. Theoretical Background and Research Questions

Much of standard economic theory supposes that people are *self-interested, utility-maximising* agents. However, extensive evidence suggests that this assumption fails to describe behaviour adequately. This evidence includes experiments and surveys suggesting that individuals are also concerned about the equity and efficiency of different gains (e.g. Tsuchiya and Dolan, 2009). Yet rather little attention has been paid to ways in which self-interest interacts with the equity-efficiency trade-off.

Experimental economists have provided evidence that people are sometimes willing to reduce their own payoffs in order to improve the payoffs of others, contrary to the assumption of perfectly self-interested agents (see a 2011 meta-analysis by Engel and a 2014 review by Güth and Kocher). This laboratory evidence is accompanied by a wealth of theoretical models that account for preferences for others' welfare (often modelled as interdependent preferences, such as in Bergstrom, 1999) and for distributional concerns (e.g., fairness models including Rabin, 1993; Fehr and Schmidt, 2000; and Bolton and Ockenfels, 2000).

However, evidence is mixed regarding how closely behaviour in laboratory experiments corresponds to behaviour in the context of richer scenarios and in the field (Laury & Taylor, 2008; Voors et al., 2012; Galizzi & Navarro-Martínez, 2018). It is not clear that we can make confident inferences from the laboratory evidence when considering the equity-efficiency trade-off in policy relevant contexts.

There is a long and rich literature on the notion of fairness in public policymaking in the context of safety. Writing in *Science,* Arrow et al. (1996, p. 222) state:

> *"Although benefit-cost analysis should focus primarily on the overall relation between benefits and costs, a good analysis will also identify important distributional consequences."*

Despite the recognition that distributional consequences matter, it is not straightforward to define exactly what is involved. For example, when evaluating measures to reduce risks of premature death, should we value each prevented fatality equally irrespective of the ages of the potential beneficiaries or should we value each year of increased life expectancy equally even though this tends to favour younger people over the elderly? (For more on this issue, including a discussion of the 'senior discount' controversy, see Viscusi 2014; also Jones-Lee et al. 2015).

More progress has been made in the health economics literature. Bobinac et al. (2012) provide a useful review of the literature about trade-offs between efficiency and equity when cost-effectiveness is measured in quality adjusted life years (QALYs). They advocate consideration of both efficiency and equity concerns, where equity can be determined by patient characteristics such as age. Light (1992) discusses the efficiency-equity trade-offs embedded in the healthcare system, beginning with the premise that equity and efficiency are both desirable. Wagstaff (1991) proposes a Social Welfare Function approach to incorporate these efficiency and equity concerns into health resource allocation decisions. However, neither Light nor Wagstaff provide empirical evidence about public preferences over the trade-off between efficiency and equity. Lindholm, Rosen and Emmelin (1996) examine the requirements for meaningful empirical estimates of the trade-off, and in a pilot study, over two thirds of the 68 Swedish politicians responsible for healthcare that took part in their study stated that they would be willing to give up efficiency to achieve more equity.

Patrick et al., (1973) and Nord (1992) use a Person Trade-Off approach, a framework for eliciting trade-offs. These trade-offs are embedded in choices between helping different people to achieve different

levels of health. Dolan (1998) extends this method and provides an experimental test demonstrating a general preference for fairness. More recently, Bleichrodt et al. (2005) find that people are averse to inequalities in the domain of health, and propose that policy values for health effects should be adjusted using equity weights.

Ubel et al. (1996a) conducted a survey to elicit equity and efficiency trade-offs from a sample including the general public and specialists including medical ethicists. Between 41% and 56% of their participants recommended:

> "offering [a] less effective screening test to everyone, even though 100 more lives would have been saved by offering [a] more expensive test to only a portion of the population." (p. 1174)

In follow up studies, Ubel et al. (2000) provide evidence that the trade-off between efficiency and equity is not continuous, but rather the preference for equity is 'all or none'. When offered a choice between a more efficient versus a more equitable policy, unless the equitable option helped 100% of the population, respondents tended to choose the efficient one. In constrast, Johannesson and Gerdtham (1996) found that respondents were "willing to give up 1 QALY in the group with more QALYs to gain 0.45 QALYs in the group with fewer QALYs" (pp. 365-366). They used a veil of ignorance approach, and a sample of 80 students.

However, all of the health studies mentioned above elicit social preferences from impartial observers, setting aside self-interest. This is problematic, since this approach cannot account for individuals' preferences for the equity and efficiency of the policies that affect *them*. Ubel et al. (1996b) present an experiment that directly compared self-interested utility scores with impartial person trade-off responses to ask whether "people place the same values on health care conditions when thinking of their own health as when thinking about health care policy" (p. 109). They find that social and self-interested perspectives generate different implied distributions of health. The utilities of the worst conditions implied by impartial choices were much lower than those elicited when participants took a

self-interested perspective. Similarly, Nord et al. (1999) consider different distributional preferences including aversion to inequalities in health and discuss how these preferences can be combined with a self-interested measure of health state utilities. However, in both studies the social perspective and self-interested perspective are elicited, or proposed to be elicited, using completely different methods. Specifically, when engaged in the self-interested tasks, there was no scope to express preferences for equity.

We present a study in which participants trade off different concerns within the same design. In some cases, participants were asked to consider their own self-interest alongside any efficiency and distributional concerns, while in other cases they were not themselves part of the population at risk and could express efficiency-equity trade-offs in the absence of self-interest.

We also explore whether the 'magnitude' of the outcome at stake can influence individuals' willingness to trade efficiency for equity. We manipulate whether the policies would prevent illnesses (low stakes) or fatalities (high stakes). Camerer and Hogarth (1999) provide a useful review of studies presenting evidence that changing monetary stakes can influence social preferences. However, to our knowledge, no comparable research has been conducted in the context of health and physical risk.

Self-interest, efficiency and equity trade-offs are liable to reflect a set of underlying moral principles. For example, an individual who chooses a policy that gives everyone in the population a risk reduction, despite offering a lower expected risk reduction overall, would appear to subscribe to some equity principle. An underexplored question is whether it is feasible to elicit individuals' degrees of (dis)agreement with statements of principle; and whether there is a reliable relationship between people's support for different moral principles and their expressed choices between policy outcomes. If so, knowledge about the nature, strength and distribution of people's moral principles may be a useful additional input into policy making.

To summarise, this study aims to address three questions. First, how do people trade-off equity, efficiency and self-interest when choosing between different policies in the context of health and safety? Second, do these trade-offs differ when the stakes are changed? Third, are the preferences revealed through participants' choices between policy options aligned with the principles which they endorse?

3. Methods

The methods were developed on the basis of the findings of a pilot study (n=107) conducted with student participants at the University of Durham. Results are available on request.

3.1. Scenario

Participants were asked to consider scenarios involving a city with an East and a West zone, each with 100,000 inhabitants. These inhabitants were at risk of adverse health effects from different bacteria present in their water supply. Due to the different prevalence of bacteria across the zones, the baseline risk in the East zone is 18/100,000 whereas that in the West is 28/100,000. The full description is shown in Figure A1 in the Appendix.

We manipulated self-interest between subjects by telling half of the participants that they lived in the East zone. The other half were not told they lived in the city and could therefore take an impartial view, analogous to that of a social planner. We use the choices of the impartial participants to investigate trade-offs between efficiency and equity, and compare these with the choices of those in the self-interest condition.

Between subjects, we manipulated the harm caused by the bacteria. Half of the subjects saw a scenario in which the bacteria would cause fatality. For the other half, the bacteria would cause gastroenteritis (see Figure A2 in Appendix).

We defined four policy options, outlined in Table 1. The policies differed in terms of which bacteria they targeted (and hence which zone they helped), and in the magnitude of the risk reductions. We label each policy to reflect the nature of the risk reduction. Policy EO ('East-only') reduces risk in the East zone by 10 cases. Policy WO ('West-only') reduces risk in the West Zone by 20 cases. Policy BC ('both constant') gives *both* zones the same risk reduction: it reduces risk by 8 cases in each zone. Lastly, policy BR ('both relative') also offered a risk reduction in both areas: by 4 cases in the zone with the higher baseline (the West), and by 3 cases in the East.

Participants saw shorter policy names (E instead of EO, W instead of WO, B instead of BC, and X instead of BR). Find the instructions and policy descriptions in the Appendix.

Table 1. Policy Options

Policy Option	EAST ZONE (Baseline Risk = 18/100,000)		WEST ZONE (Baseline Risk = 28/100,000)	
	Risk Reduction	**Final Risk**	**Risk Reduction**	**Final Risk**
EO	10/100,000	8/100,000	0	28/100,000
WO	0	18/100,000	20/100,000	8/100,000
BC	8/100,000	10/100,000	8/100,000	20/100,000
BR	3/100,000	15/100,000	4/100,000	24/100,000

To summarise, we employed a 2x2x2 factorial between-subjects design, manipulating self-interest, stake size, and task order. Further details are provided below.

3.2. Tasks

Participants completed four tasks: a rating task and a comprehension task, where participants learned about the policies; a principles task, in which participants rated their agreement with various statements of moral principles; and a preference task, in which participants made choices between policy options. By randomising the order of the latter two tasks, we tested for the possibility that any alignment between principles and choices was simply a result of participants' desire to appear consistent. Responses to the first task that respondents encountered cannot be influenced by this desire to appear consistent, so comparing responses to a task when it comes first with responses to

the same task when it comes second enables us to observe any such effects and, if necessary, make allowance for them.

Rating and comprehension tasks

These tasks were designed to familiarise participants with the policy options and to ensure that they had carefully considered the policy options before beginning the principles and preferences tasks. For the rating task, each policy option was presented on the screen in turn and participants were asked to rate them on a scale from *very poor* to *very good* (an example is shown in Figure A3 in the Appendix). Participants then answered either four or five comprehension questions (depending on the condition) about the important properties of the policies. In each comparison, the information about the policies was displayed on screen, and the respondent was asked to identify the correct answer. Participants immediately received the correct answer with an explanation. The questions were presented in random order, and the order of the options on the screen was also randomised. The comprehension questions and percentage of participants who provided correct answers are shown in Table A1 in the Appendix.

Principles task

In the principles task, participants saw six pairs of statements (see Table 2). Each statement embeds a trade-off between two relevant principles, representing opposite ways of thinking about how policies should be prioritised. Each statement favoured one principle over another, and participants were asked to explicitly express a relative preference between competing moral principles. Participants indicated which statement in the pair best described their opinion on a 7-point scale ("strongly favour A over B", "moderately favour A over B", "slightly favour A over B", "equally favour A and B", "slightly favour B over A", "moderately favour B over A", "strongly favour B over A").

We asked participants to trade off the following principles: efficiency (to maximise the expected number of cases prevented); equity (to evenly spread the risk reduction across the city); self-interest

(to prioritise own risk reduction); helping those most at risk (to reduce the risk of those with the highest baseline); and inequity (offering a higher risk reduction to only a few citizens).

From this task, we constructed an efficiency preference score that indicates the preference for efficiency against other concerns. The score is the number the times the participant favoured efficiency over the competing concern. The trade-offs that did not involve efficiency were included to induce participants to consider principles in the wider sense, and not only as a trade-off against efficiency.

Table 2. Principles Task Questions

	Trade-Off	Principle A The chosen option should…	Principle B The chosen option should…
1	**Equity vs. Efficiency**	…make the water safer for everyone who lives in the city even if that means that fewer lives are saved.	…save the most lives even if that means that the water is made safer for only some of the people who live in the city.
2	**Helping those most at risk vs. Efficiency**	…make the water safer for the people who are most at risk from the bacteria even if that means that fewer lives are saved.	…save the most lives even if that means that the water is not made safer for the people who are most at risk from the bacteria.
3	**Equity vs. Inequity**	…make the water a little safer for everyone rather than a lot safer for only some of the people in the city.	…make the water a lot safer for only some of the people in the city rather than a little safer for everyone.
4	**Equity vs. Helping those most at risk**	…make the water a little safer for everyone rather than a lot safer for the people who are most at risk from the bacteria.	…make the water a lot safer for the people who are most at risk from the bacteria rather than a little safer for everyone.
5	**Self-interest vs. Helping those most at risk***	…make the water safer for me even if that means that the water is not made safer for the people who are most at risk from the bacteria.	…make the water safer for the people who are most at risk from the bacteria even if that means that the water is not made safer for me.
6	**Self-interest vs. Efficiency***	…make the water safer for me even if that means fewer lives are saved.	…save the most lives even if that means the water is not made safer for me.

Note. *self-interest condition only

Preference task

In this task respondents made choices between policies (and indicated their strength of preference for the chosen one). We informed participants that due to the cost of treating the bacteria in the reservoirs, the government would not be able to implement all the policies and that "The following questions will ask you to tell us your opinion about which of the options you would prefer the government to choose. There are no right or wrong answers. We are simply interested in your honest opinion." With four policies, there are six pairwise comparisons (BC vs WO; EO vs WO; BR vs WO; EO vs BC; BR vs BC; BR vs EO). Every respondent saw every comparison, with the presentation order randomised between participants. We also randomised which options were presented on the left and right of the display.

Figure 1. Preference Elicitation Task Example (self-interest condition)

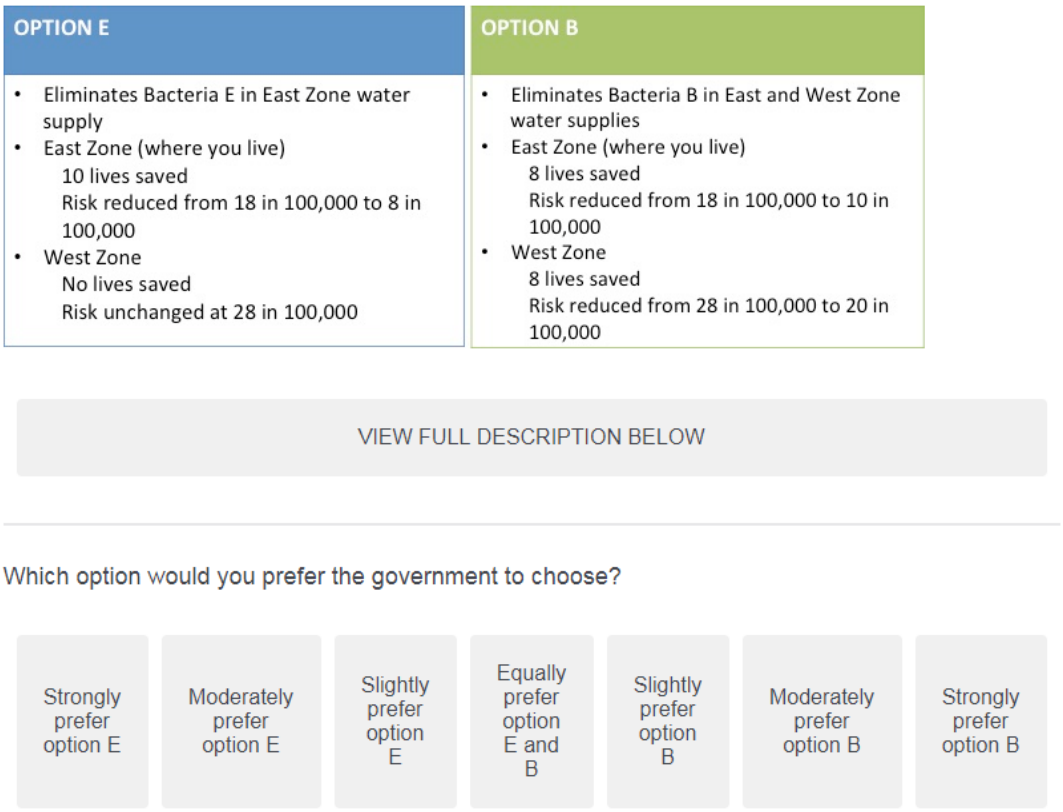


Figure 1 shows a typical preference elicitation screen. Participants saw the longer summaries of the policies (as shown in Table A2 in the Appendix) if they selected "view full description below".

3.3. Estimation

We assume a latent continuous variable P, capturing preference for the option that maximises efficiency at the expense of other considerations. Table 3 outlines the observable characteristics, both of the task and of the respondents, whose relationship with preference is explored.

Table 3. Core Model and Statement of Principle Explanatory Variables (label, description)

Main effects of treatment	
StatF	Dummy for principle statements task completed before preference task
Ftl	Dummy for fatality scenario (otherwise gastroenteritis)
Self	Dummy for 'living in the East zone' treatment
Main effects of choice scenario	
Eff	Difference in total efficiency between Y and Z: number of cases prevented by Policy Y (the most efficient option in any given comparison) minus the number of cases prevented by Policy Z (the alternative to Y).
Eq	Difference in equity between Y and Z: difference between cases prevented in the West and in the East by policy Y, minus the difference between the cases prevented in the West and in the East by policy Z. Higher *Eq* indicates that the efficient option is also the most equitable.
BenE	Difference in benefit to the East zone between Y and Z: cases prevented in the East zone by policy Y, minus cases prevented in the East zone by policy Z.
BenW	Difference in benefit to the West zone between Y and Z: cases prevented in the West zone by policy Y, minus cases prevented in the West zone by policy Z.
Interactions	
X	Vector of the 9 pairwise interactions between the main effects of choice scenario and treatment, and the interaction between *BenE*, *Self* and *Ftl*.
Demographics	
Vector of the demographic characteristics:	
Age	Respondents' age in years
Fem	Dummy for female respondent

HH	Number of close relatives living in the household (e.g., children, partner, parents)
Principles	
FavEff	Efficiency preference score based on the principles task, capturing the extent to which participants favoured efficiency over the equity, self-interest and helping those most at risk.
Eff x FavEff	Interaction term
Eq x FavEff	Interaction term

Notes. Variables capturing the effect of choice scenario characteristics are expressed in a common unit: cases prevented. *BenW* is highly collinear with *BenE* (ρ = -0.904, p-value = 0.0134) and hence will not be included in our models.

We propose the following relationship between preference and our core explanatory variables:

$$P = \beta_0 + \beta_1 StatF + \beta_2 Ftl + \beta_3 Self + \beta_4 Eff + \beta_5 Eq + \beta_6 BenE + \boldsymbol{X\beta} + \boldsymbol{Demog\ \beta} \qquad (1)$$

To explore the relationship between preference and the endorsement of various moral principles, we add the efficiency preference score based on the principles task (*FavEff*) to our core model (which captures the preference for efficiency against other concerns such as equity, self-interest and helping those most at risk) along with its interactions with *Eff* and *Eq.*

Preference is not directly observable, so we infer it from participants' choices. In estimating the relationship proposed in (1), we use participants' choices in the preference task as our dependent variable. Given the nature of this variable (7-point scale of strength of preference between policies), an ordered logistic model might appear to be suitable. However, meaningful interpretation of the coefficients in such a model requires that the proportional odds assumption holds. This assumption implies that the effect of the predictors should be the same for all levels of the dependent variable. A test using STATA command *omodel* revealed that this assumption fails to hold for our data ($p < 0.001$), making the ordered logistic model inappropriate. We therefore transformed our variable into a binary variable taking the value 1 if the more efficient policy in the pair was chosen, and 0 otherwise. Sacrificing granularity in the data allows us to perform logistic regression.

3.4. Participants

The study was completed online by a non-student sample of UK residents between January and February 2017. We obtained 322 completed responses. Our sample is 48% female, ranging in age from 20 to 71 (M=38.6, SD=10.83), with an average household size of 4 (M=3.9, SD=1.57). The randomisation to treatments was successful (with around 40 participants per cell), with no significant differences in observables, nor in the percentage of comprehension questions correctly answered (87% overall), between conditions. The study was implemented in Qualtrics and distributed through Prolific Academic, an online labour market. The median completion time was 13 minutes. Participants received a fixed payment of £2.50.

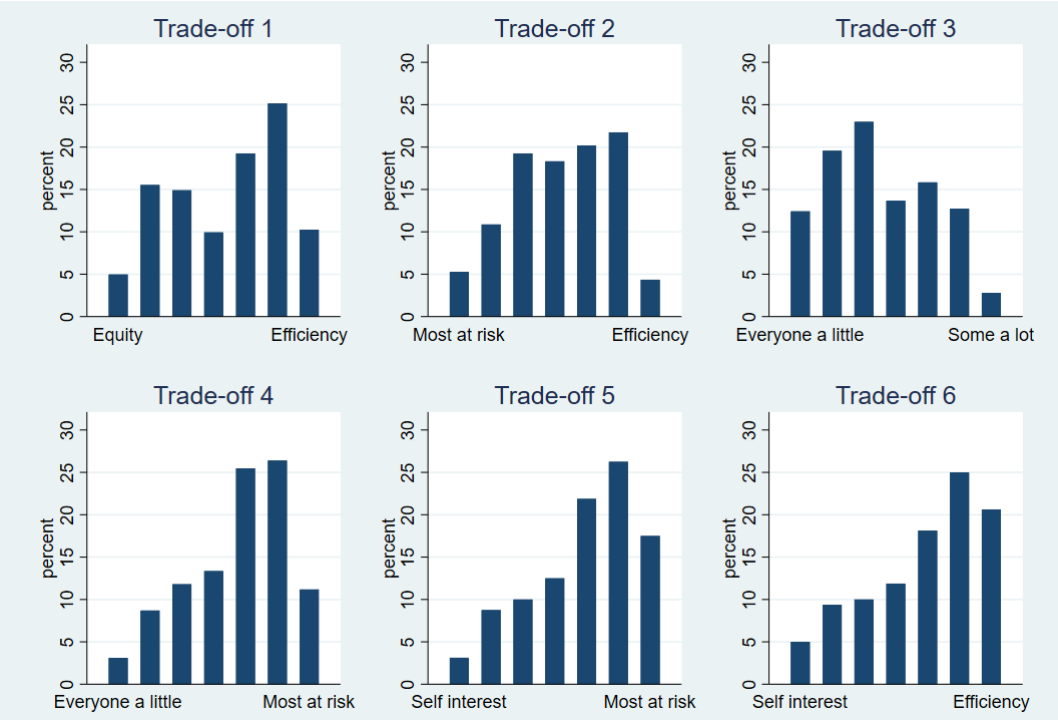
3. Results

3.1. Descriptive Results

Stated Principles

Participants stated their strength of preference (strong, moderate, or slight) in favour of one principle over another in each trade-off, or else they indicated that they equally favoured the two. The distributions of the responses are given in Figure 2. Each plot presents a trade-off between the two principles on the x-axis. For example, in Trade-off 1 the first bar represents the percentage of participants who strongly favoured equity over efficiency.

Figure 2. (Dis)agreement with Statements of Principle



Trade-offs involving self-interest (trade-offs 5 and 6) gave rise to the greatest degree of skewness. When self-interest was traded off against "helping those most at risk", 43.8% of respondents moderately or strongly favoured helping those most at risk over self-interest; and when self-interest was traded off against efficiency, this percentage increased to 45.6%. This pattern might reflect social norm adherence, by not prioritising oneself. It is less clear what the social norm would be when trading off the other principles, and this is reflected in more finely balanced response patterns: the most evenly balanced trade-off was between efficiency and helping those most at risk: respondents were almost equally distributed between equally favouring the statements, slightly favouring efficiency and slightly favouring helping those most at risk, with an overall slight preference for efficiency (Wilcoxon signed rank test $p<0.05$).

We quantify how strongly participants favoured one principle over its comparator in each trade-off. Scores range from -3 for "strongly prefer Principle 1" to +3 for "strongly prefer Principle 2" where Principle 1 is the first one of the pair under consideration and Principle 2 is the second one of the pair.

For example, in the “Equity vs Efficiency” trade-off, a score of +3 indicates that the efficiency principle was strongly favoured over the equity principle.

The data analysis is presented in Table 4. We found no significant difference between the self-interest and the impartial conditions, and no difference depending on the order in which participants completed the principle statements and the preference task. However, preferences for efficiency over equity, and for efficiency over helping those most at risk, were stronger in the fatality risk condition, compared to gastroenteritis. Preferences revealed over the principles are transitive for the sample as a whole, with the revealed preference ordering: efficiency $\succ$ helping those most at risk $\succ$ equity $\succ$ self-interest. These results also suggest that equity and efficiency are both important, and that self-interest is not sufficient to override them.

Table 4. (Dis)agreement with Statements of Principle, by Condition

Trade-off between Principle Statements			Wilcoxon signed-rank test p-value H_a: Mean ≠ 0		Wilcoxon rank-sum test p-value								
					H_a: Self-Interested ≠ Impartial			H_a: Death ≠ Gastroenteritis			H_a: Before ≠ After		
	Principle 1 (-3)	Principle 2 (+3)	Mean† (SE)	p-value	Mean Self-Int. (SE)	Mean Impartial (SE)	p-value	Mean Death (SE)	Mean Gastro. (SE)	p-value	Mean Before (SE)	Mean After (SE)	p-value
1	**Equity vs. Efficiency**		0.39*** (0.10)	0.0001	0.36 (0.14)	0.43 (0.14)	0.6050	0.64 (0.13)	0.15* (0.15)	0.0257	0.46 (0.14)	0.33 (0.14)	0.5013
2	**Helping those most at risk vs. Efficiency**		0.20* (0.09)	0.0176	0.27 (0.12)	0.13 (0.13)	0.4755	0.48 (0.12)	-0.09*** (0.13)	0.0010	0.23 (0.13)	0.17 (0.12)	0.6766
3	**Equity vs. Inequity**		-0.50*** (0.09)	0.0000	-0.53 (0.14)	-0.46 (0.13)	0.6101	-0.41 (0.13)	-0.59 (0.13)	0.3165	-0.44 (0.13)	-0.55 (0.13)	0.5739
4	**Equity vs. Helping those most at risk**		0.73*** (0.09)	0.0000	0.74 (0.12)	0.73 (0.13)	0.8834	0.72 (0.13)	0.75 (0.12)	0.9659	0.67 (0.13)	0.80 (0.12)	0.5499
5	**Self-interest vs. Helping those most at risk**		0.90*** (0.13)	0.0000	0.90 (0.13)	n/a	n/a	0.79 (0.19)	1.01 (0.18)	0.3970	0.72 (0.19)	1.09 (0.18)	0.2402
6	**Self-interest vs. Efficiency**		0.86*** (0.14)	0.0000	0.86 (0.14)	n/a	n/a	0.69 (0.21)	1.04 (0.19)	0.2505	0.94 (0.21)	0.78 (0.20)	0.4563

Notes. † -3 = strongly favour the first concern over the second; 3 = strongly favour the second concern over the first. E.g., in trade-off 1, "-3" = strongly favouring equity over efficiency and "3" = strongly favouring efficiency over equity. *** p< .001, ** p< .01, * p< .05.

Preferences

Having established that efficiency and equity both matter in principle, we next turn to preferences expressed through pairwise choices between policy options that embed these trade-offs implicitly. We define an option as 'chosen' if it was slightly, moderately, or strongly preferred over its comparator. Option BC was the most frequently preferred (chosen in 38% of the pairwise comparisons in which it appeared); followed by WO (chosen in 33% of the comparisons), with options EO and BR being chosen in a similar proportion of choices (12% and 10% respectively).

For each comparison, we quantify participants' choices according to the expressed strength of preference. Scores range from +3 for "strongly prefer Y" to -3 for "strongly prefer Z" where Y and Z are the policy options under consideration, and option Y is the more *efficient* option of the pair (offering a greater overall risk reduction). For example, in the WO vs. BC comparison, a score of +3 indicates that the more efficient WO was strongly preferred to the less efficient BC.

Table 5 presents the mean choices, alongside separate means for sub-groups by treatment. Differences by subgroup occurred between self-interested and impartial participants in four cases, in the anticipated direction. No significant differences were found between the gastroenteritis and fatality conditions, and only one significant difference was found in relation to task order: participants who completed the principles task before the preferences task made choices consistent with being more sensitive to baseline risks in the comparison between BR and BC.

Table 5. Policy Options Preferences, by Condition

Policy Choice			Wilcoxon signed-rank test p-value H_a: Mean ≠ 0		Wilcoxon rank-sum test p-value								
					H_a: Self-Interested ≠ Impartial			H_a: Death ≠ Gastroenteritis			H_a: Before ≠ After		
	Option Z (-3)	Option Y (+3)	Mean† (SE)	p-value	Mean Self-Int. (SE)	Mean Impartial (SE)	p-value	Mean Death (SE)	Mean Gastro. (SE)	p-value	Mean Before (SE)	Mean After (SE)	p-value
1	BC^E	WO	-0.07 (0.10)	0.4930	-0.51 (0.15)	0.36*** (0.13)	0.0000	-0.10 (0.14)	-0.04 (0.15)	0.7546	0.06 (0.15)	-0.21 (0.15)	0.1843
2	EO^E	WO	1.40*** (0.09)	0.0000	0.94 (0.13)	1.85*** (0.09)	0.0000	1.41 (0.12)	1.39 (0.12)	0.7913	1.36 (0.11)	1.45 (0.13)	0.2606
3	BR^E	WO	1.14*** (0.10)	0.0000	0.78 (0.16)	1.51*** (0.13)	0.0011	1.17 (0.14)	1.12 (0.15)	0.7895	1.17 (0.14)	1.12 (0.15)	0.9976
4	EO^E	BC	1.96*** (0.07)	0.0000	1.76 (0.12)	2.17* (0.08)	0.0494	2.01 (0.10)	1.91 (0.11)	0.3403	2.01 (0.10)	1.92 (0.11)	0.8177
5	BR	BC^E	2.21*** (0.06)	0.0000	2.24 (0.09)	2.19 (0.09)	0.4978	2.25 (0.08)	2.18 (0.10)	0.7877	2.10 (0.10)	2.34* (0.08)	0.0341
6	BR	EO^E	0.44*** (0.10)	0.0000	0.48 (0.14)	0.40 (0.13)	0.5349	0.46 (0.13)	0.41 (0.14)	0.9038	0.35 (0.14)	0.53 (0.14)	0.3524

Notes. † -3 = strongly prefer option Z over option Y; +3 = strongly prefer option Y over option Z. Superscript E denotes the option that was better for the East zone population, i.e., the self-interested participants. *** p< .001, ** p< .01, * p< .05.

3.2. Summary regressions

So far, our analysis has suggested that efficiency, equity and self-interest all play a role in determining the preferences between different policy options. However, these analyses did not involve quantifying the relative importance of different motivations, stated principles, and experimental treatments. To address all of these issues, we next present a series of regression analyses.

Table 6 presents five logistic regressions where the dependent variable takes value 1 if the most efficient option of the pair was 'chosen' (i.e., if it was slightly, moderately, or strongly preferred over its comparator). Because coefficients are not directly interpretable in logistic regressions, we report odds ratios instead. We consider the latter more appropriate than marginal effects. Marginal effects require evaluation at specific values of the covariates, typically the variable means. In our setup, the variables that capture differences in efficiency and equity across policy options (non-treatment main effects) take arbitrary values (-3 to +3). Hence, means (or any other point in their distribution) would not provide a meaningful benchmark to evaluate the effects. Odds ratios are interpretable as relative effects, making them more appropriate for establishing the relative importance of competing concerns.

Odds ratios reveal the relative likelihood that the efficient option is chosen, given the level of the independent variable. The six choices per participant were pooled and standard errors were clustered at the participant level. Model (1) includes main effects only. Model (2) adds the demographics, which do not significantly modify the odds of choosing the most efficient policy option. Model (3) adds the interactions. Lastly, model (4) captures the relationship between preferences for efficiency versus the other concerns stated in the both the preference and the principles tasks.

No significance was found for the main effects or interactions for the other treatment variables. That is, there were no differences between the fatality versus gastroenteritis treatments (*Ftl*), nor between the order treatments (*StatF*). This suggests that, within the domain of health and physical risk,

preferences may not be highly stake-sensitive, and are not influenced by prior deliberation about principles.

In contrast, we found statistically significant differences relating to self-interest (*Self*). Specifically, 'self-interested' participants were significantly less likely to choose the most efficient option than their impartial counterparts, indicated by the *Self* odds ratio being less than 1. This implies that self-interested participants were more likely to sacrifice efficiency for other concerns compared to their impartial counterparts.

This conclusion is supported by the effect of the variable which represented the benefit in the East zone *(BenE)*. There is no significant main effect of this predictor at the 5% level, but as expected, the interaction between the *Self* and *BenE* is positive and significant: participants prefer policies that benefit the East zone, but only if they are in the self-interested condition. More precisely, whilst impartial participants are marginally less likely to select the efficient option if it benefits the East zone, self-interested participants were around 9% *more* likely to opt for it. However, a Wald test indicated that the effect of the *Self x BenE* interaction was significantly smaller than that of overall efficiency; this suggests that self-interest does not override affected participants' concern for efficiency.

While self-interest may draw people toward choosing the policy that helps those in the East zone the most, the concern for helping those most at risk would have the opposite effect. A variable measuring the difference in the benefit offered to the West zone (*BenW*) would capture this effect. However, this variable is almost perfectly correlated with *BenE* (ρ=-0.904). Therefore, including *BenE* in our models allows us both to capture the effect of self-interest (through the interaction the Affected condition dummy variable) and to control for any preference for helping those most at risk thanks to the high correlation between them. In fact, at the 10% level, participants are around 4% less likely to choose the efficient option if it saves an additional life in the East zone. We attribute this to a concern for wanting to help those most at risk (i.e., those living in the West zone). Altogether, this allows us to

interpret the other coefficients in the model as free from the effect of any concern for helping those most at risk.

The two other main effects had a significant impact on choices. *Eff* is a positive and significant predictor, and this effect is robust across model specifications. For each additional incidence of harm prevented by Policy Y compared to Policy Z, respondents are between 11.5% and 18.5% more likely to choose the efficient option. *Eq* is another robust, positive indicator for choice: when the efficient policy is also more equitable than the other policy, it is statistically significantly more likely to be chosen. Specifically, for a one-person difference in *Eq* between the policies, the efficient option is between 8% and 11% more likely to be chosen. A Wald test indicated that the coefficients of the variables capturing efficiency and equity concerns were significantly different (p-value < 0.05), with *Eff* having a greater influence on choice than *Eq*.

Finally, we turn to the relationship between choices and principles. Model (4) suggests that for every additional principle trade-off in which efficiency was favoured over the competing principle, the odds of choosing policy Y over Z increased by 58%. This indicates that endorsements of principles are strongly aligned with choices.

The odds ratios of the interactions between the *FavEff* variable and the main effects *Eff* and *Eq* can be regarded as multiplier effects. The *FavEff x Eff* interaction is not significant (p>0.05) and the *FavEff x Eq* is in the anticipated direction, but its magnitude and significance are weak (0.01<p<0.05).

Table 6. Odds Ratios (robust standard errors, in 322 clusters at the participant level)

Choice of the efficient option	Model (1)	Model (2)	Model (3)	Model (4)
Main effects of treatment				
StatF	1.272	1.254	1.259	
	(0.211)	(0.211)	(0.297)	
Ftl	1.085	1.113	1.076	
	(0.180)	(0.186)	(0.252)	
Self	0.511***	0.502***	0.569*	
	(0.085)	(0.085)	(0.146)	
Main effects of choice scenario				
Eff	1.151***	1.151***	1.185***	1.115***
	(0.013)	(0.013)	(0.029)	(0.030)
Eq	1.083***	1.084***	1.109***	1.114***
	(0.008)	(0.008)	(0.019)	(0.014)
BenE	1.004	1.004	0.961	
	(0.010)	(0.0102)	(0.023)	
Interactions (*X*)				
*Eff*Self*			0.957	
			(0.023)	
*Eff*Ftl*			1.001	
			(0.024)	
*Eff*StatF*			0.990	
			(0.024)	
*Eq*Self*			0.969	
			(0.016)	
*Eq*Ftl*			0.996	
			(0.015)	
*Eq*StatF*			0.999	
			(0.015)	
*BenE*Self*			1.094**	
			(0.030)	
*BenE*Ftl*			1.009	
			(0.026)	
*BenE*StatF*			0.973	
			(0.019)	
*BenE*Self *Ftl*			0.998	
			(0.033)	
Principles				
FavEff				1.580***
				(0.192)
*Eff*FavEff*				1.029
				(0.016)
*Eq*FavEff*				0.987*
				(0.007)
Demographics				
Age		1.014	1.014	
		(0.009)	(0.009)	
Fem		0.859	0.857	
		(0.144)	(0.146)	
HH		0.966	0.965	
		(0.054)	(0.055)	
Constant	2.541***	1.807	1.799	
	(0.461)	(0.771)	(0.856)	
Observations	1,920	1,920	1,920	

Note. *** < .001, ** < .01, * < .05.

4. Discussion

In this paper, we have provided evidence that metrics which only capture private interest may not provide an adequate representation of people's preferences: considerations of efficiency and equity also matter. To explore the relative impacts of, and interactions between, these factors we developed an experimental design that provides an empirical framework for generating new and more appropriate metrics.

Our results suggest that efficiency is the more important of those two factors. However, equity was also an important concern and participants in our sample were willing to make trade-offs between them and sacrifice some overall benefit if the policy benefited more people.

Participants in the self-interest condition favoured policies that benefited them personally, and as a result their preference for the most efficient option was weaker than for the impartial respondents. Nonetheless, self-interest was not sufficient to nullify the influence of overall efficiency. Our results imply that neither taking a purely self-interested nor a purely impartial perspective will provide an accurate basis for policymaking. Álvarez and Rodríguez-Míguez (2001) identified self-interest as a driver of disparities between patients' and general population preferences. By quantifying the difference between 'self-interested' and 'impartial' respondents' preferences, we contribute to the debate on whose preferences should be taken into account to inform policy (see Dolan (1999)).

We also explored the relationship between the endorsement of moral principles and respondents' policy choices. Respondents revealed a transitive ordering of principles: efficiency $\succ$ equity $\succ$ self-interest. The relationship between principles and choices was independent of the order in which the tasks were presented. Such transitivity and stability encourages further exploration of the possibility that data about popular endorsement of moral principles could be an important complement to stated preferences elicited through choice.

The policy decisions in our experiment were hypothetical, and so the responses may have embodied a *socially desirability* concern. However, if their stated preferences express how they believe societal resources ought to be distributed, these preferences may be as valid a basis for public policy as preferences elicited under incentive-compatibility.

To conclude, in order for policy decision-making to respect the preferences of the affected population, it is important to take account of the extent to which individuals would be willing to trade efficiency for equity, and to understand how such considerations may be balanced against the self-interested preferences of different sections of the population. This paper offers an empirical framework that can be used to address these issues and reports some indicative results from a study implementing that framework.

References

Álvarez, B., & Rodríguez-Míguez, E. (2011). Patients' self-interested preferences: Empirical evidence from a priority setting experiment. *Social Science & Medicine*, *72*(8), 1317-1324.

Arrow, K. J., Cropper, M. L., Eads, G. C., Hahn, R. W., Lave, L. B., Noll, R. G., Portney, P. R., Russell, M., Schmalensee, R., Smith, V. K., & Stavins, R. N. (1996). Is there a role for benefit-cost analysis in environmental, health, and safety regulation?. *Science*, 272(5259), 221-222.

Bergstrom, T. C. (1999). Systems o Bergstrom benevolent utility functions. *Journal of Public Economic Theory*, 1(1), 71-100.

Bleichrodt, H., Doctor, J., & Stolk, E. (2005). A nonparametric elicitation of the equity-efficiency trade-off in cost-utility analysis. *Journal of Health Economics*, 24(4), 655-678.

Bobinac, A., van Exel, N. J. A., Rutten, F. F., & Brouwer, W. B. (2012). Inquiry into the relationship between equity weights and the value of the QALY. *Value in Health*, 15(8), 1119-1126.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 166-193.

Camerer, C., & Thaler, R. H. (1995). Anomalies: Ultimatums, dictators and manners. *The Journal of Economic Perspectives*, 9(2), 209-219.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty*, 19(1-3), 7-42.

Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M. W., Loomes, G., Pidgeon, N., Robinson, A., & Spencer, A. (1999). On the contingent valuation of safety and the safety of contingent valuation: Part 2-The CV/SG" chained" approach. Journal of Risk and Uncertainty, 17(3), 187-214.

Dolan, P. (1998). The measurement of individual utility and social welfare. *Journal of health economics*, 17(1), 39-52.

Dolan, P. (1999). Whose preferences count? *Medical Decision Making*, 19(4), 482-486.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14(4), 583-610.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817-868.

Fried, C. (1969). The value of life. *Harvard Law Review*, 82(7), 1415-1437.

Galizzi, M. M., & Navarro-Martínez, D. (2018). On the external validity of social preference games: a systematic lab-field study. *Management Science*.

Güth, W., & Kocher, M. G. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization*, 108, 396-409.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367-388.

Johannesson, M., & Gerdtham, U. G. (1996). A note on the estimation of the equity-efficiency trade-off for QALYs. *Journal of Health Economics*, 15(3), 359-368.

Jones-Lee, M.W. (1991). Altruism and the value of other people's safety. *Journal of Risk and Uncertainty,* 4(2), 213-219.

Jones-Lee, M., Chilton, S., Metcalf, H., & Nielsen, J.S. (2015). Valuing gains in life expectancy: clarifying some ambiguities. *Journal of Risk and Uncertainty,* 51(1), 1-21.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 728-741.

Laury, S. K., & Taylor, L. O. (2008). Altruism spillovers: Are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good? *Journal of Economic Behavior & Organization*, *65*(1), 9-29.

Light, D. W. (1992). Equity and efficiency in health care. *Social Science & Medicine*, 35(4), 465-469.

Lindholm, L., Rosen, M., & Emmelin, M. (1996). An epidemiological approach towards measuring the trade-off between equity and efficiency in health policy. *Health Policy*, 35(3), 205-216.

Nord, E. (1992). Methods for quality adjustment of life years. *Social Science & Medicine*, 34(5), 559-569.

Nord, E., Pinto, J. L., Richardson, J., Menzel, P., & Ubel, P. (1999). Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics*, 8(1), 25-39.

Patrick, D. L., Bush, J. W., & Chen, M. M. (1973). Methods for measuring levels of well-being for a health status index. *Health Services Research*, 8(3), 228.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, 1281-1302.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. *The American Economic Review*, 1068-1095.

Tsuchiya, A., & Dolan, P. (2009). Equality of what in health? Distinguishing between outcome egalitarianism and gain egalitarianism. *Health Economics*, 18(2), 147-159.

Ubel, P. A., DeKay, M. L., Baron, J., & Asch, D. A. (1996a). Cost-effectiveness analysis in a setting of budget constraints—is it equitable? *New England Journal of Medicine*, 334(18), 1174-1177.

Ubel, P. A., Loewenstein, G., Scanlon, D., & Kamlet, M. (1996b). Individual utilities are inconsistent with rationing choices: A partial explanation of why Oregon's cost-effectiveness list failed. *Medical Decision Making*, 16(2), 108-116.

Ubel, P. A., Baron, J., Nash, B., & Asch, D. A. (2000). Are preferences for equity over efficiency in health care allocation" all or nothing"?. *Medical Care*, 38(4), 366-373.

Viscusi, W. K. (2014). The value of individual and societal risks to life and health. Chapter 7 in Handbook of the Economics of Risk and Uncertainty, Volume 1, eds Machina, M. J. and Viscusi, W.K., North Holland: Amsterdam.

Voors, M., Turley, T., Kontoleon, A., Bulte, E., & List, J. A. (2012). Exploring whether behavior in context-free experiments is predictive of behavior in the field: Evidence from lab and field experiments in rural Sierra Leone. *Economics Letters*, *114*(3), 308-311.

Wagstaff, A. (1991). QALYs and the equity-efficiency trade-off. *Journal of Health Economics*, 10(1), 21-41.

Appendix

Figure A1. Bacteria Introduction & Policy Options Rating (no self-interest, fatality condition)

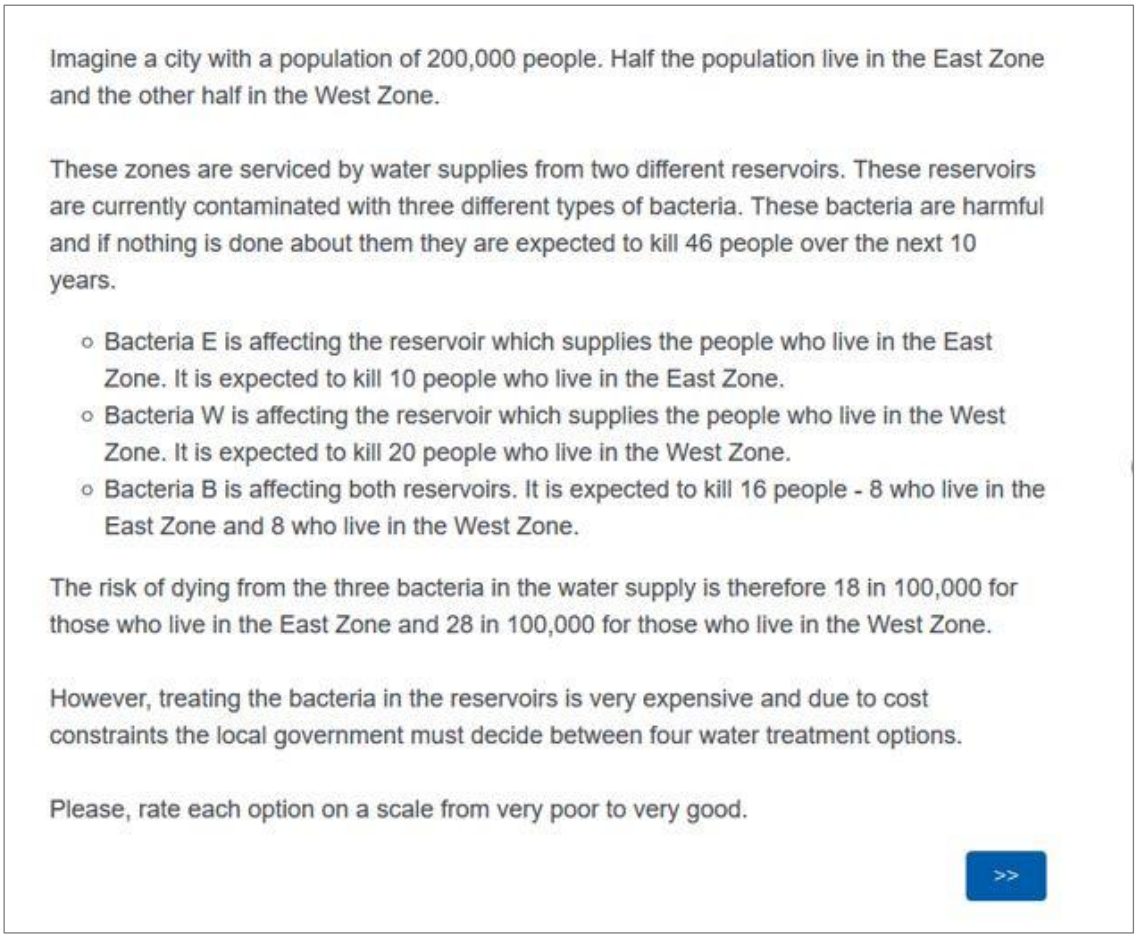
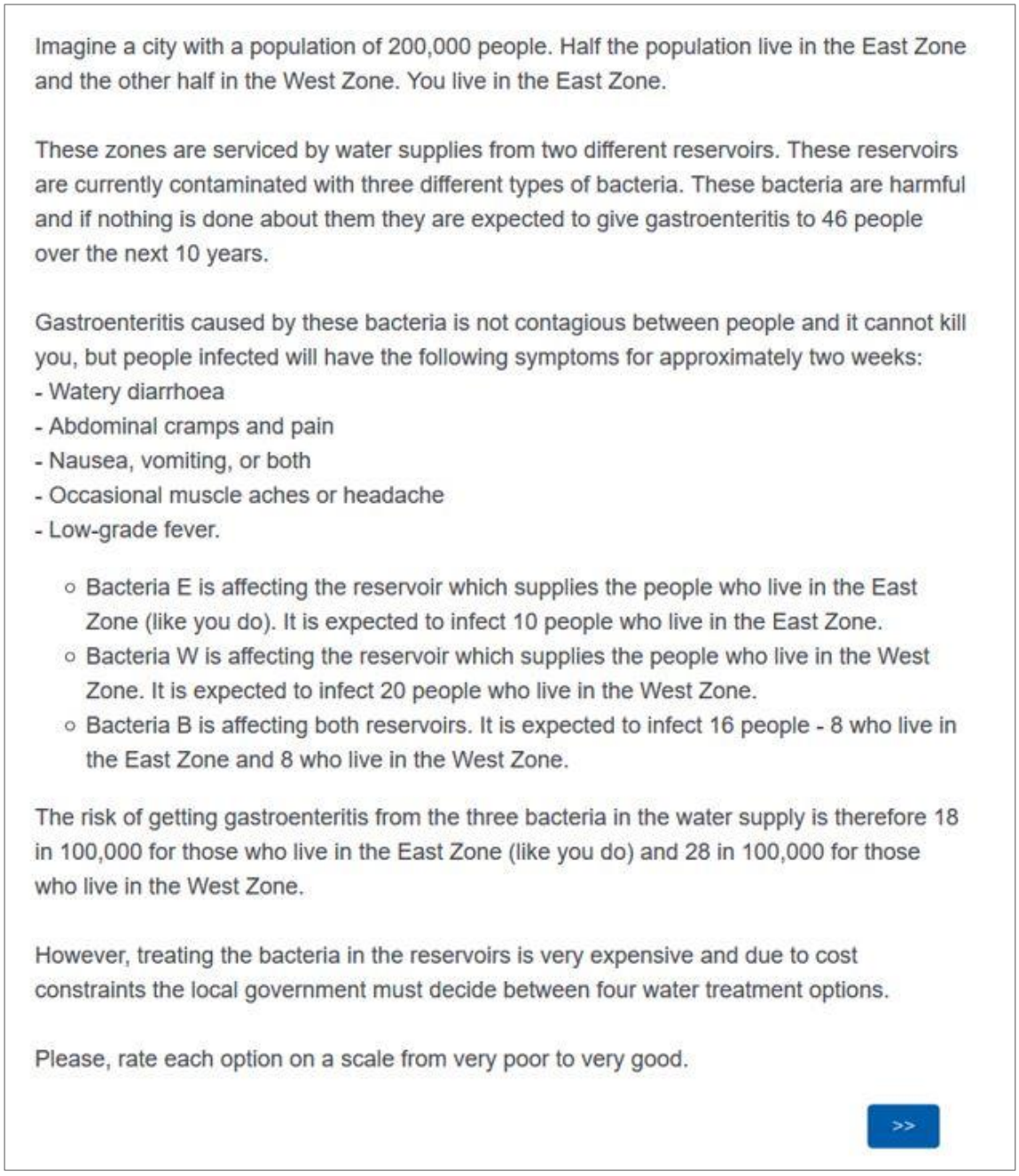


Figure A2. Bacteria Introduction & Policy Options Rating (self-interest, gastroenteritis condition)



Imagine a city with a population of 200,000 people. Half the population live in the East Zone and the other half in the West Zone. You live in the East Zone.

These zones are serviced by water supplies from two different reservoirs. These reservoirs are currently contaminated with three different types of bacteria. These bacteria are harmful and if nothing is done about them they are expected to give gastroenteritis to 46 people over the next 10 years.

Gastroenteritis caused by these bacteria is not contagious between people and it cannot kill you, but people infected will have the following symptoms for approximately two weeks:
- Watery diarrhoea
- Abdominal cramps and pain
- Nausea, vomiting, or both
- Occasional muscle aches or headache
- Low-grade fever.

- Bacteria E is affecting the reservoir which supplies the people who live in the East Zone (like you do). It is expected to infect 10 people who live in the East Zone.
- Bacteria W is affecting the reservoir which supplies the people who live in the West Zone. It is expected to infect 20 people who live in the West Zone.
- Bacteria B is affecting both reservoirs. It is expected to infect 16 people - 8 who live in the East Zone and 8 who live in the West Zone.

The risk of getting gastroenteritis from the three bacteria in the water supply is therefore 18 in 100,000 for those who live in the East Zone (like you do) and 28 in 100,000 for those who live in the West Zone.

However, treating the bacteria in the reservoirs is very expensive and due to cost constraints the local government must decide between four water treatment options.

Please, rate each option on a scale from very poor to very good.

>>

Figure A3. Example Rating Question

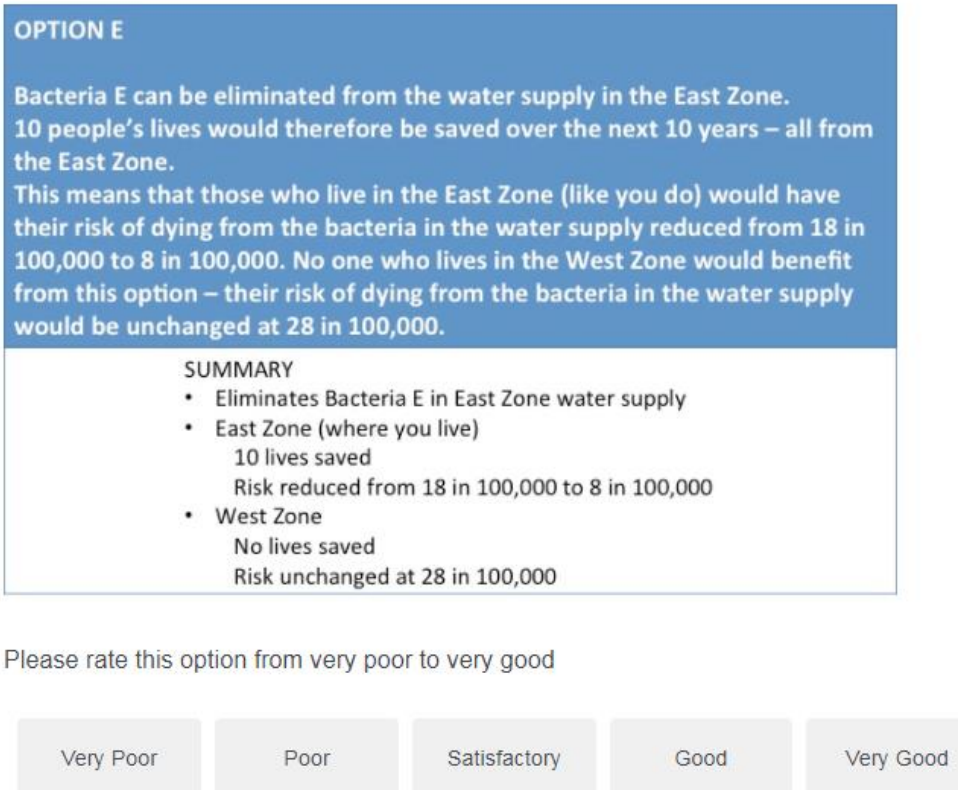


Table A1. Comprehension Questions

	Policy Comparison	Question	Correct answer (%)
1	**EO vs. WO**	Which option would make the water safer for the people who are most at risk from the bacteria?	WO (81%)
2	**WO vs. BR**	Which option would make the water a lot safer for some of the people who live in the city rather than a little safer for everyone who lives in the city?	WO (84%)
3	**BC vs. EO**	Which option saves lives (avoids gastroenteritis cases) in both zones of the city rather than in just one zone of the city?	BC (95%)
4	**WO vs. BC**	Which option would save the most lives (avoid the most gastroenteritis cases)?	WO (86%)
5	**BC vs. BR***	Which option would make the water safest for you and your household?	BC (94%)

Note. *self-interest condition only.

Table A2. Policy Options Descriptions for Participants

Policy B (BC)	Bacteria B can be eliminated from the water supply in the East Zone and the West Zone. 16 people's lives would therefore be saved / 16 fewer people will have gastroenteritis over the next 10 years – 8 lives / people from each zone. This means that those who live in the East Zone (like you do) would have their risk of dying / getting gastroenteritis from the bacteria in the water supply reduced from 18 in 100,000 to 10 in 100,000. Those who live in the West Zone would have their risk of dying / getting gastroenteritis from the bacteria in the water supply reduced from 28 in 100,000 to 20 in 100,000.
Policy E (EO)	Bacteria E can be eliminated from the water supply in the East Zone. 10 people's lives would therefore be saved / 10 fewer people will have gastroenteritis over the next 10 years – all from the East Zone. This means that those who live in the East Zone (like you do) would have their risk of dying / getting gastroenteritis from the bacteria in the water supply reduced from 18 in 100,000 to 8 in 100,000. No one who lives in the West Zone would benefit from this option – their risk of dying / getting gastroenteritis from the bacteria in the water supply would be unchanged at 28 in 100,000.
Policy W (WO)	Bacteria W can be eliminated from the water supply in the West Zone. 20 people's lives would therefore be saved / 20 fewer people will have gastroenteritis over the next 10 years – all from the West Zone. This means that those who live in the West Zone would have their risk of dying from the bacteria in the water supply reduced from 28 in 100,000 to 8 in 100,000. No one who lives in the East Zone (like you do) would benefit from this option – your/ their risk of dying from the bacteria in the water supply would be unchanged at 18 in 100,000.
Policy X (BR)	All three bacteria which affect the water supplies in the East Zone and West Zone can be treated. Although the bacteria would not be completely eliminated by this treatment, 7 people's lives will be saved / fewer people will have gastroenteritis over the next 10 years – 3 lives / people from the East Zone and 4 lives / people from the West Zone. This means that those who live in the East Zone (like you do) would have their risk of dying / getting gastroenteritis from the bacteria in the water supply reduced from 18 in 100,000 to 15 in 100,000. Those who live in the West Zone would have their risk of dying / getting gastroenteritis from the bacteria in the water supply reduced from 28 in 100,000 to 24 in 100,000.

Note. The text in blue was only shown to those participants in the self-interest condition. The text in red corresponds to the alternative versions for the death and the gastroenteritis conditions.