RATIONALITY
AND
SOCIETY

Public preferences to trade-off gains in total health for health equality: Discrepancies between an abstract scenario versus the real-world scenario presented by COVID-19

David A Comerford [ORCID]
Economics Division, University of Stirling, UK

Angela Tufte-Hewett
Birmingham City University, UK

Emma K Bridger
School of Psychology and Vision Sciences, University of Leicester, UK

Abstract

Policymakers must ration healthcare. This necessity became salient during the COVID-19 pandemic. Some policymakers took that opportunity to reduce inequality of health outcomes at the expense of overall health gains. There is a literature that seeks to quantify the optimal trade-off between efficiency and equality in health outcomes: economists employ surveys to

Corresponding author:
David A Comerford, Economics Division, University of Stirling, Stirling FK9 4LA, UK.
Email: david.comerford@stir.ac.uk

quantify the public's preferred level of equity/efficiency trade-off. An odd result from these studies is that a non-trivial subsample of respondents choose to "level down" i.e., they choose as though an additional year of life delivers negative utility to society if it accrues to the most privileged. In an experiment of US and UK respondents (n = 495), we compare equity/ efficiency trade-offs across an abstract scenario along the lines of that presented in previous surveys versus a COVID-19 scenario, where it is made explicit that healthcare rationing is a real and current necessity occasioned by the pandemic. We find that preference for "levelling down" is reduced in the COVID-19 scenario relative to the abstract scenario. This result implies that, at least in the context of the COVID-19 pandemic, previous results have overestimated the public's willingness to sacrifice overall gains in population health in order to reduce inequality of health outcomes.

Keywords
Distributional cost-effectiveness analysis, empirical ethics, empirical social choice, health inequality, inequality aversion

JEL codes: D63, I14, C9

Introduction

The recent COVID-19 pandemic made salient a subtle but perennial truth regarding the allocation of healthcare – it necessitates making trade-offs. The current research focuses on the trade-off between efficiency and equity. One example of where this came into play during the pandemic is the case of vaccination. Vaccination against a lethal and rapidly spreading virus saves lives directly – the vaccinated themselves are at reduced risk – and indirectly – anyone who comes into contact with the vaccinated is at reduced risk of contracting the virus from them. Lives are saved therefore by en-rolling people in a vaccination program as quickly as possible. Yet, a review of the US' pandemic response shows that policymakers sought to delay vaccination. They did so to reduce inequality: "Fauci [Chief Medical Officer] and Slaoui [head of Operation Warp Speed] said they actually wanted Moderna to slow down overall enrolment in order to ensure they enrolled more minorities" (Loftus, 2022, square brackets contain clarifications added by the current authors).

There is evidence to suggest that the public endorses a policy of sacri-ficing efficiency to reduce inequality. Over the past two decades, an aca-demic literature has emerged measuring the public's preferences around the distribution of health and healthcare resources across groups. The standard

approach in this literature is to show survey respondents two allocations of healthcare. One allocation might deliver greater gains in health to a group that is already privileged and lesser gains in health to a group that is relatively deprived whereas the other allocation might deliver smaller but equal gains to each group. Respondents are asked which allocation they prefer.[1] By presenting such choices repeatedly, ratcheting up or dialling down the gains to each group in each iteration, we can infer for each respondent their point of indifference between equity and efficiency.

There is an oddity in the results delivered by these stated preference studies, however. A substantial minority of respondents choose as though a gain in health delivers *negative* utility. For instance, a majority of respondents - 558 of 973 - from a representative sample of the Spanish population, chose a gain in health of 2 years for the lower socioeconomic group and a gain of zero years for the higher socioeconomic group in preference to a gain of 2 years for each of the two groups. The authors of that study, Abasolo and Tsuchiya (2004), refer to this pattern of choice as a violation of monotonicity because it implies that, all else being equal, an additional year of life delivers *negative* utility to society if it goes to high socioeconomic groups. Later researchers describe this pattern of choice as "levelling down" (Cookson et al., 2018). Despite placing negative value on human life, levelling down is a common finding in other studies (e.g., Ali et al., 2017; Cookson et al., 2018).

Taking levelling down at face value suggests placing greater weight on equality and less weight on efficiency when making the sort of real-world decisions faced during the COVID-19 pandemic. For instance, during the pandemic health officials in the state of Utah devised a scorecard to prioritize the allocation of scarce medications (see Online appendix 1). Some specifics of the scoring protocol are that it awards points to pregnant women and to older people. The Utah scorecard did not prioritize on the basis of income level but, given that COVID-19 had particularly damaging effects on the health of lower income groups (Brandily et al., 2021; Jung et al., 2021), the results of stated preference studies suggest that the public would have welcomed doing so.

The research question that motivates the current study is to discern whether levelling down is a robust preference, i.e., is negative valuation of an additional year of life as prevalent a preference as is suggested by the previous literature? This question is important given that policy decisions around the allocation of scarce healthcare resources have life and death consequences.

The current research asks respondents to choose among healthcare allocations and experimentally manipulates whether they do so in the context of the COVID-19 pandemic or, as is typical of the literature, in an unspecified abstract context. We find that preference for levelling

down is reduced when respondents are presented with a real-world case of healthcare rationing relative to when they are choosing among identical options in an abstract scenario. Convergent evidence from our experiment indicates that levelling down is at least partially explained as a symptom of confusion among respondents who have been confronted with an unfamiliar and complicated task. As such, the current research suggests that the prior literature has overestimated how common it is for members of the public to value an additional year of life negatively.

The next section describes our methods. Then follows the results of our study. We conclude with suggestions for future research.

Motivation

Research undertaken during the 1990s demonstrated that the public evaluated medical treatments according to a broader set of criteria than was being used by the medical and policymaking communities (e.g., Nord et al., 1993). The state-of-the-art had been to value a medical treatment by aggregating its health benefits to patients e.g., the number of quality adjusted life years (QALYs) the treatment gained for society. Surveys demonstrated however that the public valued not merely the total health benefits but also how fairly those benefits were distributed (e.g., Richardson, 1994; Ubel et al., 1996). Results such as these suggested a need to incorporate equity concerns when calculating the societal value of a medical treatment (Avanceña and Prosser, 2021; Nord et al., 1999). Shaw et al. (2001) devised the first questionnaire aimed at capturing a precise measure of the public's willingness to sacrifice overall health gains for other goals (e.g., equality). Their paradigm has been widely adopted and adapted (e.g., Abasolo and Tsuchiya, 2004; Ali et al., 2017; Cropper et al., 2016; Dolan and Tsuchiya, 2011; Edlin et al., 2012; Robson et al., 2017).

Surveys used for measuring equity/efficiency trade-offs share some general characteristics: they measure public preferences by asking the respondent to choose among two allocations of health program. They measure the public's willingness to trade-off total health gains against equity by having the respondent choose between an allocation with higher total health gain and an allocation with lower total health gain but greater equality of outcomes across groups. By experimentally manipulating across choice sets the total health gains delivered and the distribution of gains across groups, it is possible to infer the public's preferred level of trade-off between efficiency and equity.

Some researchers using the Shaw et al. (2001) paradigm have given respondents a choice set that affords an opportunity to "level down," i.e., they have asked respondents to choose between allocation A versus B where

A delivers more health to both groups than B (e.g., Abasolo and Tsuchiya, 2004, 2013; Ali et al., 2017; Cookson et al., 2018). One interpretation of this choice set is that allocation A strictly dominates allocation B. This is true in the sense that neither group gains as much health from allocation B as they would have gained if they had received allocation A. This view would predict all respondents to choose allocation A. In fact, in any survey we could find where respondents have been presented with choice sets of this type a non-trivial proportion of respondents has chosen allocation B i.e., they have opted to "level down" (Abasolo and Tsuchiya, 2004, 2013; Ali et al., 2017; Cookson et al., 2018).

It is important to gain insight on the mechanism that explains levelling down because there are very different implications depending on whether levelling down is a mere survey artefact or whether levelling down represents a substantive preference. If levelling down is merely a survey artefact then it misrepresents preferences. Moreover, because levelling down is the most extreme form of equity/efficiency trade-off, these misrepresented preferences will overstate the extent to which the public are willing to sacrifice efficiency for equity. On the other hand, if levelling down manifests a considered and sincere preference then there is an argument that levelling down preferences should be honoured in policy decisions.

Cookson et al. (2018) speculate that levelling down might be a symptom of respondents' incomplete thinking when presented with an unfamiliar choice task. In this view, respondents satisfice – they answer with a response that is "good enough" for the purposes of getting the survey completed but that lacks the contemplation that the same respondent would bring to the question were there real-world consequences to their response (Krosnick, 1991).

An important feature of equity/efficiency trade-off surveys is that they present the choice sets in an abstract context (though see below for one important exception). The health condition that is being treated is not named and the intervention is unspecified e.g., it is termed "program A". This methodological choice is material because research in cognitive science demonstrates that people process information less effectively when it is presented in an abstract manner than they do when the same information is scaffolded by a familiar context. The seminal study here concerns the Wason selection task – a puzzle in which people are shown four cards, each of which has a number on one side and a letter on the other. The task is to test a logical proposition (e.g., any card with an odd number on one side also has a vowel on its reverse side) and to do so by turning over as few cards as possible. Performance on this puzzle improves dramatically when it is contextualised with a real-world scenario. If the cards are presented as individuals sitting in a bar and have on one side what that person is drinking and on its reverse side that person's age, then people perform extremely well at the logically

equivalent task of deducing in the minimum number of steps which individual is an underage drinker. The general insight is that, when charged with tasks that involve thinking through consequences, people achieve better results when the motivating context of the task is framed to align with concrete realities than when it is framed abstractly (Fiddick et al., 2000). Applying this insight to the domain of thinking through trade-offs in health, we hypothesise that framing the trade-offs in the context of allocating scarce COVID-19 medications will facilitate respondents' thinking through the consequences of their choices. To the extent that negative valuation of an additional year of life is a distortion induced by confusion we expect the concrete scenario to reduce levelling down. Our preregistered hypothesis is that the concrete scenario induces a shift in preference away from levelling down and towards a compensatory trade-off between equity and efficiency.

A result that suggests our hypothesis comes from Ali et al. (2017). Like the current research, it tested the effect of presenting a real-world medical condition and treatment – it asked respondents about the allocation of bowel cancer screening across high- and low-income groups. It found that preferences for levelling down were lower when the treatment was bowel cancer screening than when the treatment was described abstractly. Ali et al. (2017) speculate that social desirability bias might account for levelling down; in this view, those who wish to make a good impression are especially likely to endorse pro-egalitarian preferences.

An additional suggestion offered by Cookson and co-authors (2018) is that levelling down signals an endorsement of sacred values. The literature on sacred values identifies that for some people there are certain trade-offs that are simply taboo (Tetlock, 2000). It might be that levelling down responses are concentrated among such people. One mechanism through which sacred values would explain levelling down is if respondents choose to level down as a protest against a question that is – from the perspective of sacred values – morally repugnant and meaningless.

Working against the survey artefact interpretation is that real-world consequential choices manifest behaviour that resembles levelling down. In ultimatum games, where there are real stakes, we observe that participants choose to forego unambiguous gains (Thaler, 1988). In these games, a proposer presents a recipient with an ultimatum concerning the division of a sum of money. If the recipient rejects the ultimatum, neither party receives any money. It has been shown through the mapping of neurological responses that recipients in ultimatum games who are experiencing anger in response to their ultimatum are more likely to choose to reject (i.e., to forego gains, Crockett, 2009). Perhaps it is anger – or some other affective motive such as spite, envy, resentment or indignation – that leads to levelling down. For instance, Diermeier and Niehues (2022) find using the European Social

Survey that affective motivations drive some of the preference against immigration.

The current research tests the effect on levelling-down of presenting respondents with the real-world dilemma presented by COVID-19 instead of the abstract scenario that is typically presented in such surveys. The primary function served by this manipulation is that it tests whether the results of previous surveys would have served as a valid guide for real-world policymaking in the context of the COVID-19 pandemic. If they are valid then we should observe that the preferences delivered by the abstract scenarios presented in previous research match the preferences expressed when respondents are asked to make trade-offs in the specific context of COVID-19. A second function is to help with the interpretation of levelling down responses. To the extent that levelling down is driven by motives such as anger, spite etc., we would expect those motives to be more active in the emotive setting of a real-world decision than in the detached setting presented by an abstract scenario (Schwartz et al., 2018). This mechanism would predict increased levelling down in the COVID-19 scenario relative to the abstract scenario.

If levelling down is a symptom of confusion, however, then we would expect the opposite result: reduced levelling down when the choice scenario presents respondents with the real-world context of the COVID-19 pandemic as a scaffold with which to think through the implications of their choices. Additionally, we measure sacred values and test whether they can account for levelling down.

It is important to determine the explanation for levelling down because each candidate explanation has different implications. If levelling down results from the confusion induced by a complicated task then there is every reason to dismiss these confused responses. Dropping the confused responses from the data will give a more accurate signal of the public's preferences, just as dropping confused responses from surveys of inflation expectations gives a clearer signal of inflationary beliefs (Comerford, 2023). If anger or some similar affective state is driving levelling down then this raises an interesting question. Equity/efficiency trade-off surveys purport to offer guidance to policy makers (Avanceña and Prosser, 2021). Should we give the same weight in policy formulation to preferences that are driven by emotions as we give to those that derive from cool-headed deliberation? A similar quandary is presented if levelling down turns out to be driven by those who hold sacred values. In the context of equity/efficiency trade-offs, sacred values are problematic because they imply that no amount of efficiency gain will compensate for an increase in inequality. That stance is of course anathema to the entire project of cost-benefit analysis, which relies on aggregating net benefits.

There are important implications if levelling down is a considered and sincerely held preference. Those who level down in these data imply that, as long as excess deaths were even larger among the rich than among the poor, they would rather live in a world with higher COVID-19 death tolls than one with lower COVID-19 death tolls. If we interpret this pattern of choice as indicating a preference for a world with more suffering to one with strictly less suffering then it seems perverse. Perhaps what explains it is a conviction that inequality itself induces suffering: there exists evidence that, independent of their absolute conditions, people in lower rank positions experience lower life satisfaction, higher rates of mental distress and higher suicidality (Boyce et al., 2010; Wetherall et al., 2015). Still, sacrificing years of life seems an extraordinary price to pay for reducing inequality in health outcomes. Yet, we will see that over a quarter of our sample choose to level down. If this preference is rooted in a considered evaluation of costs and benefits then our findings raise questions about societal priorities, not merely in healthcare allocation but also in resource allocation more generally. We return to consider these implications in the discussion.

Methods

Respondents were presented with a sequence of screens asking them to make pairwise selections among two health programs, A and some other program (B – F, described in Table 1).

In one experimental manipulation the health programs were described either as treating COVID-19 or else the condition that is treated was left

Table 1. Sequence of pairwise choices respondents were asked to choose between.

Screen	Program(s)	Description of program
1 – Program A vs. Program B	Program A	I give 1.7 years to the rich and 1.3 years to the poor (always one of the alternatives presented)
	Program B	I give 1.3 years to the rich and 2.1 years to the poor
2 – Program A vs. Program C	Program C	I give 1.3 years to the rich and 1.7 years to the poor
3 – Program A vs. Program D	Program D	I give 1.3 years to the rich and 1.5 years to the poor
4 – Program A vs. Program E	Program E	I give 1.3 years to the rich and 1.3 years to the poor
5 – Program A vs. Program F	Program F	I give 1.3 years to the rich and 1.1 years to the poor

unspecified, as is standard in stated preference studies. All other features of programs A – F were identical except for the health condition that is treated.

In a second orthogonal manipulation, respondents were asked either "which would you choose?" or "which would you wish to see chosen?" This manipulation was included because both variants have been asked in prior surveys investigating equity/efficiency trade-offs (e.g., compare Cookson et al., 2018; McNamara et al., 2021). The "which would you choose?" wording elicits a preference. It grants agency to the respondent, making them responsible for determining which group is privileged. That agency might grant various sources of responsibility utility, e.g., warm glow or signalling value (Comerford and Lades, 2022). By contrast, the "which would you wish to see chosen" wording casts the respondent as a passive recipient of a policy outcome and hence absolves them of responsibility. It elicits a desirance, i.e., a rank ordering over exogenously determined outcomes. Comerford and Lades (2022) present theory and experimental evidence that preference orderings can differ from desirance orderings. It is insightful therefore to investigate whether these differences in wordings make a difference to trade-offs. If they do, then an important question for future research will be to investigate whether a preference or desirance is the more meaningful measure for informing policy.

In summary, the survey randomly assigned respondents to 1 of 4 conditions: Abstract No Responsibility; Abstract Responsibility; COVID No Responsibility; COVID Responsibility.

Our design iterates towards finding each respondent's point of indifference. Program A is present in each pair, and it always gives a 1.7-years gain in life expectancy to the richest fifth and a 1.3-years gain in life expectancy to the poorest fifth. In the first pairwise choice, the alternative to Program A is Program B, which gives a 1.3 years gain to the richest fifth and a 2.1 years gain to the poorest fifth. Program B thus offers a larger gain in total health and – for those who believe the rich to currently live longer than the poor – a reduction in health inequality. On each subsequent screen, the life expectancy gains to the poorest offered by the alternative to Program A reduces. So, Program C offers 1.7 years additional life expectancy. Program D offers 1.5 years, Program E offers 1.3 years and Program F offers just 1.2 years to the poorest fifth of the population. Respondents first chose between Program A and Program B, then on subsequent screens between Program A and Program C; Program A and Program D; Program A and Program E; Program A and Program F.

Respondents could express indifference by endorsing the answer: "I would toss a coin to decide - Program A and Program X are equally good". The "Toss a coin" wording is intended to make clear to respondents that a choice will be taken even if the respondent avoids making a choice.

To include a range of gains that might plausibly be delivered by COVID-19 therapies, the gains offered by our hypothetical programs range from 1.2 years to 2.1 years. We explain to respondents that the gain in lifespan is an expected gain. The relevant text informed respondents that the estimated gain of 1.7 years means that "some members of the group will gain more than 1.7 years e.g., if they are especially responsive to the program, and some will gain less than 1.7 years, e.g., if they do not suffer from the health problem that the program is designed to treat." In the COVID-19 condition of the survey, the phrase "health problem that the program is designed to treat" was replaced by "COVID-19"; Online appendix 2 presents the relevant text for each condition.

At the close of the survey, we ask respondents to estimate the life expectancy of the richest 20% of the population and the poorest 20% of the population. This allows us to assess the extent to which perceptions of current health inequality explain preferences for reducing health inequality.

The key manipulation in the study is whether the health program is described in abstract terms (i.e., "health program A") or whether it is described as a specific program to treat a specific disease (i.e., "a program for distributing COVID-19 medication"). To give the COVID-19 medication some context, respondents were presented with the following preamble:

"Given the extreme scarcity of COVID-19 treatments some jurisdictions have designed programmes to determine who gets medication.

For example, in the US state of Utah, some factors automatically qualify a person for treatment e.g., unvaccinated pregnant women. For those who do not automatically qualify, a score was developed to include other relevant factors. An example of such a factor is age - because older people are more at risk of hospitalization and death, older people are given a higher score and so are prioritized for treatment.

Another relevant factor is income because evidence from a number of countries shows that a person's income predicts COVID-19 mortality risk.

The questions that follow ask you how you would wish to see COVID-19 treatments allocated across the richest 20% and poorest 20% of the population, as measured by household income."

Those in the abstract condition did not see this preamble; instead, they were simply directed to the screen that called on respondents to make their selection (reproduced below as Figure 2). On that screen, we instruct respondents that:

- We cannot pay for both programs

- A choice must be made
- "Equally good" means you don't mind which program is chosen
- Both programs cost exactly the same and have identical consequences for everyone except the poorest 20% and the richest 20%.

Description of categories

We categorize respondents according to their choices as: pro-rich; health maximizer; weighted prioritarian; maximin and extreme egalitarian, as set out in Table 2.

Respondents who always prefer Program A are classified as *Pro-rich* because they prefer a program that increases the health of the rich even at the expense of increasing total health.

Respondents are classified as *Health Maximizer* if they always choose the program delivering more total health.

The term *Weighted Prioritarian* refers to people who give priority to the worse off but who also find positive utility in gains for the rich. We categorize respondents as Weighted Prioritarians if they choose Program C over Program A and also choose Program A over Program E. Their choice of Program A over Program E indicates that that, all else being equal, they prefer larger than smaller gains in lifespan for the rich. Their choice of Program C over Program A indicates they would prefer that an additional year of life goes to the poor than that it goes to the rich.

Whereas a weighted prioritarian prefers larger than smaller gains in lifespan for the rich, a *Maximin* respondent is indifferent between the richest group gaining an additional year of life or not, all else being equal. They are solely concerned with promoting the health outcomes of the worse off. With that in mind, they are indifferent between Programs E and A but will always choose Program A over Program F.

Table 2. Categories implied by respondents' choices.

Category	Program choice
Pro-rich	A, A, A, A, A
Health maximizer	B, A or C, A, A, A
Weighted prioritarian	B, C, A or D, A, A
Maximin	B, C, D, A̶, A
Extreme egalitarian	B, C, D, E, A̶

Note: A, A, A, A, A indicates that the respondent chose option A in each of five pairwise choices. A̶ denotes that the respondent selected a choice other than A. So, B, C, D, E, A̶ indicates a respondent who chose option F or answered that they would toss a coin in the last of the five pairwise choices.

The numbers in the table are the expected gains in lifespan for each group compared to a scenario where no program was introduced.

	Gain in lifespan for richest 20%	Gain in lifespan for poorest 20%
Program A	+1.7 years	+1.3 years
Program F	+1.3 years	+1.2 years

Which program would you wish to see chosen?

○ Program A - gives 1.7 years to the rich & 1.3 years to the poor

○ Program F - gives 1.3 years to the rich & 1.2 years to the poor

○ Toss a coin to decide - Program A and Program F are equally good

Figure 1. Screenshot of the choice scenario that identified levelling down.

The final category is those whose egalitarian preferences are so strong that they choose to level down. We follow Cookson et al. (2018) in labelling this group *Extreme Egalitarian*. This group comprises those who give priority to the poor in each pairwise choice and then, in the final pairwise choice, fail to choose Program A. (A screenshot of the choice respondents faced between Program A and Program F is depicted in Figure 1).

Relative to Program F, choice of Program A delivers gains in health for both the rich and the poor. The only logic for choosing Program F in preference to Program A is that Program F sacrifices health gains amongst the rich to a greater degree than it sacrifices health gains among the poor. Relative to choosing Program A, choosing Program F leaves both rich and poor worse off but, assuming the rich outlive the poor to begin with, the gap in lifespan between rich and poor is reduced.

Participants

Data were collected using Prolific.co. Prolific.co (formerly Prolific Academic) is an online platform on which academics post surveys for completion by a pool of participants. It has been demonstrated to produce good quality data (Peer et al., 2022). We preregistered that we would collect data from 480 respondents, split evenly between the US and the UK.

Data were collected between March 11th and March 17th, 2022. For context, this was about 2 years into the pandemic. At the time of data collection, 77% of the US population and 92% of the UK population had received at least one dose of the vaccine against COVID-19

(USAFACTS.ORG, 2023; GOV.UK, 2023a). Still, both the US and UK were recording tens of thousands of cases, e.g., the 7-days average of new cases for the US on March 14th was 49,048 (New York Times, 2023) and the 7-days average number of cases for the UK on March 14th was 81,856 (GOV.UK, 2023b). The Oxford Stringency Index captures the intensity and pervasiveness of regulatory response to COVID-19 and shows that both countries were in a state of vigilance at the time of data collection. On March 11th, 2022, the Oxford Stringency Index showed a score of 34.6 for the US (all-time high: 75.5 in November 2020; all-time low: 28.3 in December 2022 when the data series ends) and the UK scored 23.2 (all-time high: 88.0 in January 2021; all-time low: 5.6 in December 2022 when the data series ends).

Ultimately, 520 respondents consented to take part in our study but some of these were routed out for failing our attention screening question. The final sample therefore comprised 499 respondents, 252 respondents from the UK and 247 from the US. Four hundred ninety-five of these respondents gave answers to all choices presented to them. The mean age of respondents was 38 (age range: 18 – 85) and 70% identified as female.

Procedures

The survey opened with the attention screening question.[2] Then respondents were randomized to 1 of 4 conditions: Abstract no responsibility; Abstract responsibility; COVID no responsibility and COVID responsibility. Those assigned to the Abstract conditions then saw an instructional screen that explained the task. Since this screen is helpful for interpreting our results we present it as Figure 2.

Before seeing the screen depicted in Figure 2, respondents in the COVID-19 condition saw the screen containing the contextualizing preamble on healthcare rationing in the state of Utah (see the Methods section for the precise wording). An additional change is that the text presented in Figure 2 was amended to refer to COVID medication instead of "programs" (online appendix 2 presents the text from each condition).

Following the sequence of five screens in which respondents made their choices, they were asked to estimate the lifespans (in years) of the average member of the richest 20% of their population and the poorest 20% of their national population (following Kiatpongsan and Norton, 2014). Then followed the five items of the sacred values scale (Tanner et al., 2009; see online appendix 3 for items), a question on subjective socioeconomic status (Adler et al., 2000), and questions about education level, income level, gender identity, race and age. Because US and UK respondents were asked nationality-specific questions about race, we could not compare these directly across groups and so we do not include these as control variables in our analyses.

Over the screens that follow you will find a series of tables summarizing the expected effects of two candidate programs for improving population health. You will be asked for your preference.

The only difference between the programs is the gain to the poorest 20% and richest 20%. Low income people have different health conditions than the rich. This means that any program that attempts to influence population health will impact a different number of people from each income group and will have different consequences for their health.

When making a decision, it is important to remember the following:
- We cannot pay for both programs
- A choice must be made
- "Equally good" means you don't mind which one is chosen
- Both programs cost exactly the same and have identical consequences for everyone except the poorest 20% and the richest 20%.

Figure 2. Screenshot of instructional text from the control condition.

Results

Our attention screening question routed 22 respondents (4.2% of the potential sample) out of the survey as having failed to read the question.

Levelling down

Of the 495 respondents who gave an answer to all of the choice scenarios, 31.4% opted to level down i.e., chose Program F over Program A. The likelihood that a respondent chose to level down was lower in the COVID-19 condition than in the abstract condition: 26.4% chose to level down in the COVID-19 condition compared to 36.2% in the abstract condition. Model 1 of Table 3 reports a probit model that tests whether COVID-19 condition significantly reduced levelling down. It finds a statistically significant effect ($z = 2.34$, $p = .019$). Since the COVID-19 condition was randomly assigned, the effect described by the simple Model 1 is expected to be independent of

Table 3. Predictors of preference for levelling down.

	Model 1	Model 2
COVID condition	−0.098 (0.041)*	−0.103 (0.042)*
Responsibility cond		0.019 (0.042)
Age		−0.001 (0.002)
Female		−0.072 (0.042)
Income		−0.001 (0.015)
Subjective socioeconomic status		0.006 (0.015)
UK education		0.035 (0.031)
US respondent		−0.046 (0.109)
US education		0.072 (0.070)
Sacred values		−0.001 (0.005)
Predicted P at x-bar	0.313	0.301
Number of observations	496	490

*$p < .05$; **$p < .01$. Standard errors in parentheses. Marginal effects from probit models of levelling down, i.e., choosing Program F over Program A. Education is modelled separately for the US and UK because US and UK respondents were asked to report their highest level of education on response scales that reflected the different education systems in each country.

any other determinants of levelling down. This assumption is tested in Model 2 of Table 3. It adds control variables and demonstrates that the effect of the COVID-19 condition remains statistically significant ($z = 2.47, p = .013$). For our purposes, the key result of Model 2 is that the reduction in levelling down achieved by the COVID-19 condition cannot be explained by respondents' characteristics.

Incoherent responses

Of the 495 respondents who gave an answer to all of the choice scenarios, 30% made incoherent responses in the sense that later choices contradicted an earlier choice.

An additional result of interest in Model 2 is that there is no respondent characteristic that reliably predicts levelling down. Even when we run simple regressions of levelling down on each characteristic sequentially, none emerges as a significant predictor of levelling down. The closest a characteristic comes to predicting levelling down is the US education variable though it fails to attain even marginal significance ($z = 1.62$, $p = .106$).

As an example of such a response pattern, consider the choices made by respondent 498. They chose program A, sacrificing larger health gains for the poor in order to give priority to smaller health gains for the rich and then chose programs C, D and E sacrificing larger health gains for the rich in order to give priority to smaller health gains for the poor; finally they chose program A in preference to program F, thereby delivering larger health gains for both rich and poor relative to the available alternative program. There is no respondent characteristic or experimental manipulation that predicts incoherence except one: respondents who levelled down were more likely to deliver incoherent preferences (Model 5, Table 4).

Types of preferences

Recall that the survey randomly assigned respondents to 1 of 4 conditions: Abstract No Responsibility; Abstract Responsibility; COVID No Responsibility; COVID Responsibility. Figure 3 depicts for each survey condition how respondents who gave coherent preferences are categorised. Given the structure of our data, we run non-parametric tests as specified in our preregistration.

Of those respondents whose choices could be categorised ($n = 321$), Kruskal-Wallis tests of category (scored 1–5) finds no difference in preference across the randomly assigned responsibility conditions: respondents who were asked "which do you choose?" delivered preferences that did not differ significantly from those who were asked "which you wish to see chosen?" ($X^2(1) = 1.92$, $p = .161$). There was a significant difference across the abstract and the COVID-19 conditions ($X^2(1) = 4.14$, $p = .042$). Drilling down into what explains the difference in rank across the COVID-19 versus abstract scenario, Figure 3 depicts that the abstract condition was more likely than the COVID-19 condition to return extreme egalitarian preferences, i.e., the difference across conditions was entirely explained by greater levelling down in the abstract condition.

The null effect of the COVID-19 condition on pro-rich, health maximizing, weighted prioritarian and maximin preferences speaks to the following counterfactual: if respondents who level down in response to the abstract scenario were instead answering by the COVID-19 scenario, what preference would they return? We had preregistered that such respondents would shift to becoming weighted prioritarians. If that hypothesis were

Table 4. Predictors of delivering incoherent preferences.

	1	2	3	4	5
COVID condition	0.011	0.008	0.009	0.006	0.060
	(0.041)	(0.041)	(0.041)	(0.041)	(0.043)
Responsibility cond	0.015	0.014	0.009	0.013	0.003
	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)
Age	−0.002	−0.002	−0.002	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female	−0.047	−0.047	−0.048	−0.049	−0.020
	(0.040)	(0.040)	(0.040)	(0.041)	(0.042)
Income	−0.001	−0.002	0.000	−0.001	0.000
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)
Subj. SES	−0.011	−0.011	−0.013	−0.012	−0.017
	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)
Education	0.037	0.021	0.020	0.020	0.007
	(0.028)	(0.031)	(0.031)	(0.031)	(0.033)
US respondent	0.008	−0.106	−0.121	−0.124	−0.125
	(0.047)	(0.107)	(0.107)	(0.107)	(0.111)
US education		0.082	0.091	0.094	0.069
		(0.069)	(0.069)	(0.069)	(0.072)
Sacred values			0.007	0.007	0.009
			(0.005)	(0.005)	(0.005)
Response time				−0.056	−0.033
				(0.047)	(0.047)
Levelled down					0.450
					(0.046)**
Predicted P at x-bar	0.289	0.289	0.286	0.286	0.259
N. observations	490	490	489	489	489

$*p < .05$; $**p < .01$. Standard Errors in Parentheses. Marginal effects from a probit model of delivering incoherent preferences. Model 1 treats the education response scale as consistent across the US and UK, which it was not because the US and UK education systems differ. Model 2 takes accounts of that difference by modelling the US and UK education scales separately. Model 3 additionally controls for scores on the sacred values scale. Model 4 includes Box-Cox normalised response time. Model 5 adds an indicator of whether the respondent chose to level down.

supported then the COVID-19 condition would show not only lower rates of levelling down but also higher rates of weighted prioritarian preferences. But the COVID-19 condition does not show higher rates of weighted prioritarian preferences. Nor does the COVID-19 condition show higher rates of in-coherent preferences (Table 4). It only shows lower rates of levelling down. This result suggests that the marginal respondent who levels down in the abstract condition is drawn with equal probability from each of the pro-rich,
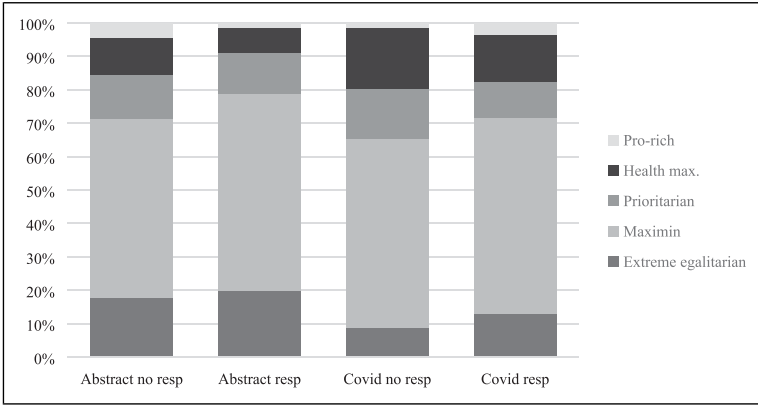
Figure 3. Categories of preference by condition (*n* = 321). *Note:* Abstract = Program treats an unspecified illness; COVID = Program treats COVID-19; No resp = Respondent is asked "which you wish to see chosen?"; Resp = Respondent is asked "which would you choose?".

health maximizing, weighted prioritarian and maximin categories i.e., had they been randomly assigned to the COVID-19 condition, they would be equally likely to switch to expressing pro-rich preferences as to switch to expressing maximin preferences. In the discussion section we elaborate on how this result implies that the previous literature has overstated preferences to sacrifice efficiency to reduce inequality of a health outcome.

The effect of covariates on preferences

Scores on the sacred values scale failed to explain levelling down or delivery of incoherent responses or any of the five categories of preference. In principle, scores on the scale could range from 7 (least endorsing of sacred values) to 35 (most endorsing). In practice, respondents' scores ranged from 11 to 35, with a mean of 22 and an interquartile range of 19 – 25. In univariate binary logistic regressions on each of the categories (pro-rich; health maximizer; weighted prioritarian; egalitarian; extreme egalitarian) respondent's sacred values scale score is consistently non-significant; all $ps >$ .140. A potential explanation for this effect is that the sacred values scale showed only modest interitem reliability (Cronbach's Alpha = .556; a desirable level would be at least .70, Nunnally and Bernstein, 1994).

A univariate binary logistic regression finds that respondent's subjective socioeconomic status negatively and significantly predicted health maximizing preferences ($z = 2.02$, $p = .044$), but no other category of preference. In a multivariate regression that controls for age, gender, income, sacred

values and whether the respondent is based in the US, this effect of subjective socioeconomic status is subsumed by income ($z = 2.44$, $p = .015$). This result suggests that respondents who are better off are more willing than worse off respondents to sacrifice efficiency in order to reduce inequality. The only other covariate that predicts this preference is that females were more likely than males to endorse pro-rich allocations (univariate result: $z = 2.71$, $p = .007$; multivariate result: $z = 2.00$, $p = .045$).

At the close of the survey, we elicited from respondents their perceived life expectancies for the richest quintile and poorest quintile of the population. US-based respondents were asked to make these estimates of the US population and UK-based respondents were asked it of the UK population.

In order to conduct this analysis, we first needed to remove from the data responses that could not be reliably interpreted (e.g., respondent no. 116 reported a lifespan of "70+" for the rich). We also dropped respondents who delivered implausibly small responses (lower than 10 years) and one respondent who delivered an implausibly large response, who answered 50,000 for both groups.

Cleaning the data in the manner described above removed 10 respondents from the dataset. We then calculated as a measure of perceived health inequality the gap in life expectancy across the richest quintile and poorest quintile (i.e., we subtracted each respondents' estimate of life expectancy for the poorest quintile from their estimate for the richest quintile). The mean gap estimated was 11.2 years for the UK ($n = 247$) and 12.2 years for the US ($n = 242$). There were just nine respondents who estimated that the poor live longer than the rich and five of those delivered preferences that were incoherent. Three of the remaining four endorsed health maximizing and the final one endorsed extreme egalitarianism.

Applying the same model as specified in Table 4, larger gaps were reported by respondents with higher household income ($t = 2.80$, $p = .005$) and lower gaps were reported by females ($t = 2.67$, $p = .005$) and, among UK respondents only, by those with a higher level of education ($t = 2.67$, $p = .008$).

Perceived health inequality, as measured by the gap variable, predicted two categories of preference: in univariate regressions, gap positively predicted preferences that categorized the respondent as an Extreme Egalitarian ($z = 3.12$, $p = .002$) and negatively predicted preferences that categorized the respondent as a Health Maximizer ($z = 3.37$, $p = .001$). These results are robust to the inclusion of the controls reported in Model 3 of Table 4 (Extreme Egalitarian: $z = 2.80$, $p = .005$; Health Maximizer: $z = 3.30$, $p = .001$). These results are substantively unchanged if we drop from the sample those respondents who estimated that the poor live longer than the rich (both ps from multivariate regressions $<.01$).

The magnitude of these effects is impressive: each additional year in gap reduced the likelihood of being a Health Maximizer by nine percentage points and increased the likelihood of being an Extreme Egalitarian by six percentage points.

One final null effect is worth noting: none of our preference measures showed a significant difference across UK and US respondents (all $ps > .10$). Nor were preferences explained by interactions between respondent nationality and the experimental treatments (all $ps > .20$). These null effects suggest that the results reported above apply more or less equally to US and UK respondents.

Discussion

This research was motivated by the question "what are the public's preferences for allocating scarce health resources across income groups when, as in the COVID-19 pandemic, such choices are real-world policy dilemmas?" The key result from this analysis is that levelling-down was reduced in the real-world scenario relative to the abstract scenario. This result is noteworthy because previous research has employed the abstract scenario. This difference was obtained despite the fact that the abstract and concrete scenarios presented here were substantively identical; in every consequential dimension, the concrete scenario was indistinguishable from the abstract scenario. The only difference across the two scenarios was that the concrete scenario was a specific, real-world instance that exemplifies the more abstract scenario. If the research question is "how do the public wish to allocate scarce health resources across income groups in the COVID-19 pandemic?", then we can see no reason to take the results of the abstract scenario as more valid than the results of our COVID-19 scenario.

Our results suggest that levelling down, which implies a negative valuation of an additional year of life, is at least partly explained by confusion. Relevant data from our study that speaks to this mechanism is that levelling down was concentrated among respondents who contradicted themselves (Table 4) – where 17.5% of those who gave a coherent preference opted to level down, 63.9% of those who gave incoherent preferences did.[3] There can be no doubt that some of the observed instances of levelling down are a symptom of respondents' failure to think through the consequences of their choice.

Our results suggest that the previous literature has overestimated health inequality aversion. Levelling down is consistent with extreme inequality aversion – after all it implies that the respondent would prefer both rich and poor to enjoy shorter lifespans if that achieves greater equality of lifespan across rich and poor. Also, because our survey used an experimental design, we can observe how respondents who levelled down in response to the abstract scenario would have answered by the COVID-19 scenario. The data suggest that, had they been randomly assigned to the COVID-19 condition, the marginal respondent who levels down in the abstract condition would be no more likely to express preferences that are consistent with inequality aversion as to express pro-rich preferences. This adds independent evidence supporting

our claim that levelling down derives from confusion on the part of survey respondents. It also leaves little doubt that the abstract scenario used in previous research has overestimated inequality aversion in the population.

Questions for further research

Despite reducing the prevalence of levelling down, the COVID-19 scenario still returned levelling down preferences from over a quarter of respondents (26.4%). This result is striking. If people are truly willing to take years of life from both rich and poor in order to reduce inequality in health outcomes, then what other sacrifices might they be willing to make to promote equality? We don't yet know because the survey paradigm that allows identification of levelling down preferences is absent from many domains. It would be instructive to learn, for instance, whether those who levelled down in this survey would also sacrifice years of education from the top and bottom of the income distribution if doing so would reduce educational inequality. Presumably they would because anyone willing to sacrifice years of life to reduce inequality in health outcomes would, assuming they recognize socioeconomic differences in education as one of the drivers of health inequality (e.g., Bridger et al., 2023), surely also be willing to sacrifice years of education to promote equality.

A related question is whether these results flag a disconnect between the public's priorities and the goals of state and supranational institutions. In each condition in our study, a majority implied that there was no gain in utility accruing from an additional year of life to the rich (see Figure 3, which shows that extreme egalitarian and maximin accounted for >60% of preferences). If we take this result at face value, then it suggests widespread indifference to many forms of efficiency gain. By contrast, increasing efficiency is a central goal of many public institutions, such as central banks and governments. It is worth interrogating this apparent mismatch as it might offer a partial explanation for the anti-elitist sentiment that pervades political discourse in Western countries at the current moment. Note however that there are other paradigms for measuring equity-efficiency trade-offs (e.g., Tepe et al., 2021) in which laypeople are very sensitive to efficiency gains.

An important question for future research therefore is to continue to clarify the mechanisms that explain levelling down preferences. These data suggest that sacred values are not the driver, although we recognize that our sacred values scale was noisy and so we welcome further work on this point. Anger, spite or some other affective response might be a driver. Previous research in psychology (Schwartz et al., 2018) suggests that, relative to an abstract scenario, our COVID-19 scenario would be more emotive and hence might potentially increase knee-jerk levelling down responses. Our results do not rule out that this mechanism was at play; they merely show that any

such effect was swamped by the reduction in levelling down that was induced by the decision-making scaffold offered by our COVID-19 scenario. There is compelling evidence in these data that some levelling down preferences are a symptom of respondents' confusion. Even among those in the COVID-19 condition, there are likely some respondents who did not think through the implications of their choice. We suspect then that our COVID-19 condition also overstates the true number of respondents in these data who believe that the world would be a better place if policymakers had implemented policy F instead of policy A during the COVID-19 pandemic.

A more general research question suggested by the current research is how best to measure equity/efficiency trade-offs. Designing stated preference surveys that are robust to bias is a difficult project. A limitation that the current study shares with all research in this paradigm is that we cannot be confident that what we have measured is anything other than noise. By noise, we do not mean pure noise (e.g., random button clicking in a websurvey). We mean something more nuanced. The literature on stated preferences that emerged from the Exxon Valdez oil spill on environmental valuation has documented many sources of noise and bias. For instance, Kahneman et al. (1999) finds that people are willing to pay more to protect turtles than they are to protect all reptiles (including turtles). Results such as this have led some to conclude that stated "preferences" are not in fact preferences in the sense that economists use that term but rather are mere attitude expressions. So, when we say that these data might be noise what we mean is that they may not predict the choices these respondents would actually make. We look forward to future developments in survey measurement of preferences.

Our finding that preferences for reallocation are explained by perceptions of health inequality warrants future research. At one extreme, one could dismiss this finding as a mere tautology. It may be that the preference measures returned by these sorts of surveys are better described as attitude expressions than as policy preferences (see the previous paragraph and Kahneman et al., 1999). If that is the case, then the correlation between responses to the life expectancy questions and the preference measures might be due to the simple fact that both are picking up the same attitudes. At the other extreme, this result might reflect considered and informed decision making on the part of respondents. It might be that those who have witnessed inequality at its most intense have been moved to recognize the value in reducing health inequality and are willing to incur sacrifices to bring about reductions in health inequality. The current research suggests that our measure of perceived health inequality – gap in perceived lifespan across rich and poor – is a useful and meaningful measure for resolving which, if either, of these mechanisms explains the association.

In this study, preferences were no different when respondents were asked "which would you choose?" versus when they were asked "which would you

wish to see chosen?" On the one hand, this result is reassuring because the literature has been using these two variants of question wording interchangeably (compare, for instance, Cookson et al., 2018; McNamara et al., 2021). On the other hand, we should be careful about extrapolating from this one data point. Theory suggests that the agentic variant will be answered with reference to a distinct utility function relative to the non-agentic variant and experimental evidence demonstrates that these variants in question wording can matter (Comerford and Lades, 2022). Future research should attend to these nuances of question wording because they may have substantive implications for optimal resource allocation.

This survey considered one dimension of health inequality – income. Income differences in health have been a focus of much research in economics (e.g., Ali et al., 2017; Cookson et al., 2018; Costa-Font and Cowell, 2019) and in health psychology (e.g., Bridger et al., 2023). Of course, there are other important dimensions of health inequality, such as race (e.g., Abedi et al., 2021) and age (e.g., Lloyd-Sherlock et al., 2022). Future research might test whether public preferences regarding equity/efficiency trade-offs across these other dimensions of inequality are also sensitive to a concrete versus abstract framing. More generally, future research might investigate preferences to level down in domains beyond health.

Conclusion

Our results suggest that previous research has overstated how many people place a negative value on gains to human life. Specifically, the abstract scenario employed in previous research seems to confuse some respondents into endorsing an allocation of health outcomes that deviates from what they would wish to see in an actual case of healthcare rationing. This result implies that previous research overestimated the public's willingness to sacrifice total gains in health for an increase in health equality during the COVID-19 pandemic.

Declaration of conflicting interests

Funding

Ethical approval

It was approved by the Ethics committee at Birmingham City University, reference: Bridger /#10273 /sub2 /R(A) /2022 /Mar /BLSS FAEC.

ORCID iD

David A Comerford ⬥ https://orcid.org/0000-0003-2611-9504

Supplemental Material

Supplemental material for this article is available online.

Notes

1. There are other approaches to estimating the public's demand for health equity that are not considered here. They are budget allocation approaches (Costa-Font et al., 2015) and veil of ignorance approaches (Costa-Font and Cowell, 2019). These approaches invite respondents to express their preferences for healthcare provision in the distant future and so are ill-suited to allocating treatments to an immediate threat, such as the COVID-19 pandemic. As such, they are beyond the scope of the current enquiry.
2. The attention screening question was taken from Comerford (2019) and reads as follows: "Will you read each question before answering? To indicate that you will, tick only the middle two options listed below. I will answer all the questions; I will read all the questions; I will only answer questions that I have read; I will not answer all the questions". Respondents were routed out of the survey if they answered anything other than by endorsing both "I will read all the questions" and "I will only answer questions that I have read".
3. It is important to note here that the direction of causation can run in only one direction. By the time they have the option to level down, respondents have already made the choices that classify them as coherent or incoherent. Levelling down is determined by their choice between programs A and F and, if earlier choices had been coherent then choice of either A or F would also be coherent.

References

Abásolo I and Tsuchiya A (2013) Is more health always better for society? Exploring public preferences that violate monotonicity. *Theory and Decision* 74(4): 539–563.

Abasolo I and Tsuchiya A (2004) Exploring social welfare functions and violation of monotonicity: an example from inequalities in health. *Journal of Health Economics* 23(2): 313–329.

Abedi V, Olulana O, Avula V, et al. (2021) Racial, economic, and health inequality and COVID-19 infection in the United States. *Journal of Racial and Ethnic Health Disparities* 8(3): 732–742.

Adler NE, Epel ES, Castellazzo G, et al. (2000) Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy, white women. *Health Psychology* 19(6): 586.

Ali S, Tsuchiya A, Asaria M, et al. (2017) How robust are value judgments of health inequality aversion? Testing for framing and cognitive effects. *Medical Decision Making* 37(6): 635–646.

Avanceña AL and Prosser LA (2021) Examining equity effects of health interventions in cost-effectiveness analysis: a systematic review. *Value in Health* 24(1): 136–143.

Boyce CJ, Brown GD and Moore SC (2010) Money and happiness: rank of income, not income, affects life satisfaction. *Psychological Science* 21(4): 471–475.

Brandily P, Brébion C, Briole S, et al. (2021) A poorly understood disease? The impact of COVID-19 on the income gradient in mortality over the course of the pandemic. *European Economic Review* 140: 103923.

Bridger EK, Tufte-Hewett A and Comerford DA (2023) Dispositional and situational attributions for why the rich live longer than the poor. *Journal of Applied Social Psychology* 53(6): 469–481.

Comerford DA (2019) Asking for frequencies rather than percentages increases the validity of subjective probability measures: evidence from subjective life expectancy. *Economics Letters* 180: 33–35.

Comerford DA (2023) Response bias in survey measures of expectations: evidence from the survey of consumer expectations' inflation module. *Journal of Money, Credit and Banking*. Available at: https://doi.org/10.1111/jmcb.13003.

Comerford DA and Lades LK (2022) Responsibility utility and the difference between preference and desirance: implications for welfare evaluation. *Social Choice and Welfare* 58(2): 201–224.

Cookson R, Ali S, Tsuchiya A, et al. (2018) E-learning and health inequality aversion: a questionnaire experiment. *Health Economics* 27(11): 1754–1771.

Costa-Font J and Cowell F (2019) Incorporating inequality aversion in health-care priority setting. *Social Justice Research* 32(2): 172–185.

Costa-Font J, Forns JR and Sato A (2015) Participatory health system priority setting: evidence from a budget experiment. *Social Science & Medicine* 146: 182–190.

Crockett MJ (2009) The neurochemistry of fairness: clarifying the link between serotonin and prosocial behavior. *Annals of the New York Academy of Sciences* 1167(1): 76–86.

Cropper M, Krupnick A and Raich W (2016) *Preferences for equality in environmental outcomes* National Bureau of Economic Research Working Paper No. w22644). Available at: https://www.nber.org/papers/w22644.

Diermeier M and Niehues J (2022) Towards a nuanced understanding of anti-immigration sentiment in the welfare state–a program specific analysis of welfare preferences. *Rationality and Society* 34(3): 302–333.

Dolan P and Tsuchiya A (2011) Determining the parameters in a social welfare function using stated preference data: an application to health. *Applied Economics* 43(18): 2241–2250.

Edlin R, Tsuchiya A and Dolan P (2012) Public preferences for responsibility versus public preferences for reducing inequalities. *Health Economics* 21(12): 1416–1426.

Fiddick L, Cosmides L and Tooby J (2000) No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task. *Cognition* 77(1): 1–79.

GOV.UK (2023a) Vaccinations in the UK. Available at: https://coronavirus.data.gov.uk/details/vaccinations?areaType=overview&areaName=United_Kingdom (accessed 4 July 2023).

GOV.UK (2023b) Cases in UK. Available at: https://coronavirus.data.gov.uk/details/cases?areaType=overview&areaName=United_Kingdom (accessed 15 April 2023).

Jung J, Manley J and Shrestha V (2021) Coronavirus infections and deaths by poverty status: the effects of social distancing. *Journal of Economic Behavior and Organization* 182: 311–330.

Kahneman D, Ritov I, Schkade D, et al. (1999) Economic preferences or attitude expressions? An analysis of dollar responses to public issues. In: Fischoff B and Manski CF (eds) *Elicitation of Preferences*. Dordrecht: Springer, 203–242.

Kiatpongsan S and Norton MI (2014) How much (more) should CEOs make? A universal desire for more equal pay. *Perspectives on Psychological Science* 9(6): 587–593.

Krosnick JA (1991) Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5(3): 213–236.

Lloyd-Sherlock P, Guntupalli A and Sempé L (2022) Age discrimination, the right to life, and COVID-19 vaccination in countries with limited resources. *Journal of Social Issues* 78(4).

Loftus P (2022) *The Messenger: Moderna, the Vaccine, and the Business Gamble that Changed the World*. Brighton, MA: Harvard Business Press.

McNamara S, Tsuchiya A and Holmes J (2021) Does the UK-public's aversion to inequalities in health differ by group-labelling and health-gain type? A choice-experiment. *Social Science and Medicine* 269: 113573.

New York Times (2023) Coronavirus in the U.S.: latest map and case count. Available at: https://www.nytimes.com/interactive/2021/us/covid-cases.html (accessed 15 April 2023).

Nord E, Richardson J and Macarounas-Kirchmann K (1993) Social evaluation of health care versus personal evaluation of health states: evidence on the validity of four health state scaling instruments using Norwegian and Australian surveys. *International Journal of Technology Assessment in Health Care* 9: 463–478.

Nord E, Pinto JL, Richardson J, et al. (1999) Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* 8(1): 25–39.

Nunnally JC and Bernstein IH (1994) *Psychometric Theory*. 3rd edition. New York, NY: McGraw-Hill.

Peer E, Rothschild D, Gordon A, et al. (2022) Data quality of platforms and panels for online behavioral research. *Behavior Research Methods* 54(4): 1643–1662.

Richardson J (1994) Cost-utility analysis: what should be measured? *Social Science and Medicine* 39: 7–21.

Robson M, Asaria M, Cookson R, et al. (2017) Eliciting the level of health inequality aversion in England. *Health Economics* 26(10): 1328–1334.

Schwartz A, Eyal T and Tamir M (2018) Emotions and the big picture: the effects of construal level on emotional preferences. *Journal of Experimental Social Psychology* 78: 55–65.

Shaw R, Dolan P, Tsuchiya A, et al. (2001) *Development of a Questionnaire to Elicit Public Preferences Regarding Health Inequalities*. Centre for Health Economics, University of York Working Paper. Available at: https://pure.york.ac.uk/portal/en/publications/development-of-a-questionnaire-to-elicit-public-preferences-regar

Tanner C, Ryf B and Hanselmann M (2009) Geschützte Werte Skala (GWS): konstruktion und validierung eines messinstrumentes [Sacred Value Measure (SVM): construction and validation of an instrument to assess sacred values]. *Diagnostica* 55(3): 174–183.

Tepe M, Paetzel F, Lorenz J, et al. (2021) Efficiency loss and support for income redistribution: evidence from a laboratory experiment. *Rationality and Society* 33(3): 313–340.

Tetlock PE (2000) Coping with trade-offs: psychological constraints and political implications. In: Lupia A, McCubbins MD and Popkin SL (eds) *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*. New York, NY: Cambridge University Press, 239–263.

Thaler RH (1988) Anomalies: the ultimatum game. *Journal of Economic Perspectives* 2(4): 195–206.

Ubel PA, Loewenstein G, Scanlon D, et al. (1996) Individual utilities are inconsistent with rationing choices. *Medical Decision Making* 16: 108–116.

USAFACTS.ORG (2023) Percent of people receiving vaccines in the US. Available from: https://usafacts.org/visualizations/covid-vaccine-tracker-states (accessed 15 April 2023).

Wetherall K, Daly M, Robb KA, et al. (2015) Explaining the income and suicidality relationship: income rank is more strongly associated with suicidal thoughts and attempts than income. *Social Psychiatry and Psychiatric Epidemiology* 50: 929–937.