TI 2024-045/V
Tinbergen Institute Discussion Paper

Responsibility-Sensitive Welfare Weights for Health

Matthew Robson[1]
Owen O'Donnell[2]
Tom Van Ourti[3]

1 Erasmus University Rotterdam & Tinbergen Institute
2 Erasmus University Rotterdam & Tinbergen Institute
3 Erasmus University Rotterdam & Tinbergen Institute

Responsibility-Sensitive Welfare Weights for Health

Matthew Robson † Owen O'Donnell ‡ Tom Van Ourti §¶

July 5, 2024

Abstract

We estimate responsibility-sensitive welfare weights for health that facilitate inequality- and inequity-sensitive policy evaluation. In a UK general population sample, 569 online experiment participants distribute constrained resources to determine the health of hypothetical individuals distinguished by randomly generated resource productivity as well as sex, income and smoking (41,460 observations). We elicit beliefs about responsibility for income and smoking, and use their associations with the allocations to estimate responsibility-sensitive weights that reflect inequality aversion and health prioritisation by the non-health characteristics. There is slight, moderate and substantial prioritisation of females, the poor and non-smokers, respectively. Inequality aversion lowers weights on females and non-smokers, who are health-advantaged, and raises the weight on the poor, who are health-disadvantaged. As beliefs about responsibility for income and smoking strengthen, weights on the poor decrease and weights on non-smokers significantly increase.

Keywords: Experiment, Social Preferences, Inequality Aversion, Equity

JEL: C90, D30, D63, I14, I38.

---

† Corresponding Author. Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute, the Netherlands. Email: robson@ese.eur.nl. Website: www.mrobson92.com.

‡ Erasmus School of Economics, Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Tinbergen Institute, the Netherlands. Email: odonnell@ese.eur.nl.

§ Erasmus School of Health Policy and Management, Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute, the Netherlands. Email: vanourti@ese.eur.nl.

¶ This study was funded by the Erasmus Initiative Smarter Choices for Better Health and the Wellcome Trust (205427/Z/16/Z). Ethics approval given by the Internal Review Boards of the Erasmus Research Institute of Management and Erasmus School of Economics. We would like to thank Richard Cookson, Tim Doran, Raf Van Gestel, Niek Mouter, Erik Schokkaert, Aki Tsuchiya, Sandy Tubeuf and participants at various seminars and conferences for helpful comments.

1 Introduction

Resolution of competing claims on public resources involves trade-offs between efficiency, equality and equity. To inform those charged with making these trade-offs, we use a general public sample to elicit willingness to prioritise health by equity-relevant characteristics for which there may be personal responsibility, while simultaneously identifying aversion to health inequality. We estimate responsibility-sensitive welfare weights for health that facilitate inequality- and inequity-sensitive policy evaluation. The weights are *generalised* (Saez and Stantcheva, 2016) — they potentially capture non-welfarist concerns for equity that would be missed by consequentialist evaluation of the health distribution.

Responsibility-sensitive egalitarianism contends that personal responsibility for characteristics over which outcomes vary moderates aversion to inequality (Dworkin, 1981; Arneson, 1989; Cohen, 1989; Roemer, 1998; Fleurbaey, 2008; Fleurbaey and Maniquet, 2011).[1] One strand of this theory of justice maintains that people be held responsible for what they control and compensated for disadvantage arising from what they do not control (Arneson, 1989; Cohen, 1989; Roemer, 1998). The normative appeal of this attempt to build a philosophical bridge between inequality aversion and respect for preferences is open to debate (Fleurbaey, 1995; Anderson, 1999; Scheffler, 2003). From a positive perspective on distributive justice (Yaari and Bar-Hillel, 1984; Konow, 2001; Scott et al., 2001; Konow, 2003; Gaertner and Schokkaert, 2011; Weinzierl, 2014), gauging public support for the theory can inform democratically minded social decision makers.

In our online experiment, participants allocate constrained resources between hypothetical individuals who turn resources into health with different productivity. This forces an equality-efficiency trade-off that identifies aversion to inequality in the univariate distribution of health. Initially, the individuals are anonymous. In subsequent treatments, we randomly

[1] Responsibility-sensitive egalitarianism is also labelled *luck egalitarianism* and the *accountability principle* (Konow, 1996). Those subscribing to it may respect the Pareto Principle, unlike pure egalitarians.

label them by different categories of sex, income or smoking, which enables identification of health prioritisation by each of those characteristics. To estimate responsibility-sensitive prioritisation, we elicit beliefs about responsibility for income and smoking (assuming there is no responsibility for sex) and use the association of beliefs with participants' allocations of resources to individuals distinguished by the respective characteristic.

Using data from a broadly representative United Kingdom (UK) sample of 569 participants with 41,460 observations, we estimate a random behavioral model of resource allocations and find substantial aversion to health inequality, slight prioritisation of females' health, stronger prioritisation of the health of poorer individuals and even stronger prioritisation of the health of non-smokers over that of smokers. Willingness to prioritise health by income and smoking behavior vary with beliefs about responsibility. Participants who perceive greater personal responsibility for income prioritise the health of poorer individuals to a lesser degree, although this relationship is not statistically significant in parametric analysis. Those who ascribe greater personal responsibility for smoking prioritise the health of non-smokers more.

Using our estimates of responsibility-sensitive health prioritisation and health inequality aversion, together with external estimates of health differences between population groups, we estimate welfare weights that indicate, for example, that the social value of a marginal health gain to a non-smoker is around 60% greater than the value of an equivalent health gain to a smoker. This is the net result of direct prioritisation of non-smokers attenuated by aversion to health inequality that is to their advantage. The estimated social marginal welfare weight on the health of a poor individual (bottom 20%) is about 45% larger than the weight for a rich individual (top 20%). This reflects both prioritisation of the poor and aversion to health inequality that is to their disadvantage.

Previous attempts to estimate responsibility-sensitive social preferences for health estimated the extent of health prioritisation and univariate health inequality aversion separately, conditional on an assumption (Dolan and Tsuchiya, 2009) or information (Edlin et al., 2012) about responsibility for a particular non-health characteristic (smoking) and its causal effect on health. We estimate all parameters of responsibility-sensitive social preferences simultaneously without making any assumptions about responsibility (except for sex) or causality. This is possible because we elicit beliefs about responsibility.

Participants of qualitative studies often ascribe substantial personal responsibility for health and claim this weakens their aversion to health inequality (Cookson and Dolan, 1999; Lundell et al., 2013; Asada et al., 2022). Systematic reviews of stated preference studies document widespread support for giving lower priority to patients deemed responsible for their ill-health (Whitty et al., 2014; Gu et al., 2015).[2] The design of these studies does not allow identification of the strength of support for the responsibility principle relative to other equity and efficiency criteria for allocating health resources (Whitty et al., 2014; Asada et al., 2022).

We do not ask about responsibility for health differences. Rather, we elicit beliefs about responsibility for characteristics (income and smoking) that can have health consequences, and we use these beliefs, together with allocations between individuals distinguished by each characteristic, to infer the extent to which there is prioritisation on the basis of perceived responsibility. By separating the assignment of responsibility for a characteristic from prioritisation by that characteristic, we allow for (and experimentally confirm) a less punitive form of responsibility-sensitive egalitarianism that does not hold people fully accountable for the consequences of their choices (Temkin, 2003; Cappelen and Norheim, 2005; Schmidt, 2009).[3]

[2]Outside of the health domain, aversion to income inequality and support for redistribution are weaker when it is known, or believed, that people have greater responsibility for their incomes (Konow, 1996; Konow, 2003; Schokkaert and Devooght, 2003; Cappelen et al., 2007; Saez and Stantcheva, 2016; Almås et al., 2020).

[3]Someone may view smoking as largely a personal responsibility and yet support public subsidies for medical treatment of smoking-related disease due to a) a compassionate response to basic needs (Segall, 2009), b) a conviction that social responsibility to the worse off trumps personal responsibility for that predicament (Wikler, 2004), or c) a belief that liability for engaging in risky behavior is limited to bearing its expected cost (possibly levied through the tax system) (Le Grand, 1991; Cappelen and Norheim, 2005).

We do not only gauge support for responsibility-sensitive egalitarianism. Willingness to prioritise groups is unlikely to depend entirely on beliefs about responsibility. It can arise from aversion to inequality in overall wellbeing (Hausman, 2007; Deaton, 2013; Hausman, 2013). The health of females or the poor may be prioritised in order to lessen cumulative disadvantage experienced by these groups, even without any perceived causal impact of sex or income on health.

Our primary contribution is to estimate responsibility-sensitive welfare weights for health that capture not only willingness to sacrifice maximisation of health for its equalisation but also responsibility-contingent prioritisation by non-health characteristics. Our method, which consists of an experiment with a resource allocation task, elicitation of responsibility beliefs and a random behavioral model, offers opportunities to estimate welfare weights for health in other contexts and for other outcomes.

2 Experiment design

2.1 General setup

Participants of an online experiment on the Prolific platform allocate constrained resources to three hypothetical individuals to determine their health. Resource productivity can vary between individuals, forcing a trade-off between equalisation and maximisation of health. Treatments differ in the information participants are given about the individuals. In one treatment, which was previously reported (Robson et al., 2024), income information is given. Here, we add two treatments that provide information on sex and smoking behavior, the elicitation of beliefs about responsibility for income and smoking, and another UK sample.

In each round of each treatment, a participant is given a randomly generated resource *budget*, $m \in \{180, 240, 360\}$, and told that the *health* of each *individual* (i) is the product of the *resources* (y_i) allocated to the individual and a productivity factor, aka *multiplier* (p_i). Participants are told that health is the number of years an individual lives adjusted for illness or disability. To clarify the concept, they are given an example derivation of quality-adjusted life years (QALYs).

Participants use sliders to allocate resources to the three individuals. Interactive graphics display the resources allocated to each individual and the resulting health (Appendix A.1 Figure A1). The interface also displays measures of resource and health gaps between individuals, the aggregate health generated and the remaining budget, which must be exhausted.

2.2 Treatments

In treatment A, participants are given no information about the three individuals other than their multipliers, which change across 10 rounds (Appendix A.2 Table A1). Between individual differences in the multipliers force equality-efficiency trade-offs that we use to identify aversion to inequality in the univariate distribution of health.

In treatments B_1, B_2 and B_3, participants are also given information about the individuals' sexes, incomes and smoking behavior, respectively. In each round of B_1, each individual is labelled with a sex, $x_{1i} \in \{Male, Female\}$, that is selected randomly conditional on there always being at least one individual of each sex. In each round of B_2, each individual is labelled with a randomly selected (without replacement) income from $x_{2i} \in \{£5,000, £10,000, £25,000, £50,000, £100,000\}$. In each round of B_3, one individual is labelled as a *Non-Smoker* ($= x_{3i}$) and the other two are described by a randomly selected (without replacement) smoking behavior from $x_{3i} \in \{$*Light-Smoker*, *Moderate-Smoker*, *Heavy-Smoker*$\}$. We tell participants that the average daily number of cigarettes smoked is 0, 5, 10 and 30 for Non-Smokers, Light-Smokers, Moderate-Smokers and Heavy-Smokers, respectively.

In each variant of treatment B, the distribution of the respective non-health characteristic (sex, income or smoking) varies over 10 rounds in a way that is orthogonal to the multipliers that are distributed as in treatment A. This allows identification of both prioritisation by each non-health characteristic and aversion to univariate health inequality.

In all treatments, the order of the 10 rounds is randomised across participants, the individuals are labelled by random initials, in addition to some non-health characteristic in B, and their screen positions (from low to high or high to low for B_2 (income) and B_3 (smoking)) are orthogonal to their multipliers.

All participants complete treatment A before B, which reveals more information. A subset is given treatment B_2 and a separate subset is given both B_1 and B_3. We randomise the order of the latter two treatments.

2.3 Responsibility beliefs

To estimate responsibility-sensitive prioritisation of health by income and smoking behavior, we ask: "To what extent are people responsible for how much income they earn/their smoking behavior?". Participants can report on a 0-10 scale, where 0 is "Not Responsible at All" and 10 is "Entirely Responsible". This approach treats responsibility as fluid (Roemer, 1993; Roemer, 1998) rather than binary (Fleurbaey and Schokkaert, 2009). Reported beliefs can reflect assessment of a) the relative importance of income and smoking determinants that individuals can control to varying degrees and b) latitude for the exercise of free will more generally (Roemer and Trannoy, 2015). The strength of these beliefs would be expected to influence responsibility-sensitive egalitarians when allocating health resources across individuals with different incomes and smoking behavior. We assume that all participants believe that no individual is responsible for their sex, which the participants are told is biological sex at birth.

2.4 Timing

After two pilots, the experiment was carried out in two phases (Appendix A.3 Figure A2). Between December 14, 2021 and January 5, 2022 we selected Sample 1 and conducted treatments A and B_2. Data from this sample were used to address a different research question in Robson et al. (2024). Between April 7, 2022 and April 30, 2022 we selected Sample 2 and conducted treatments A, B_1 and B_3. We use data from both samples and all treatments. Within each phase, the experiment was conducted over two sessions. The first session included instructions, an interactive tutorial, comprehension questions, treatment A (10 rounds) and a questionnaire (Appendix A.4). The second session included a reduced tutorial, treatment B (10 rounds for each variant) and a final questionnaire. The median completion time was 55.5 minutes: 28.3 minutes for the first session and 26.4 minutes for the second. Participants were paid a flat £3.50 for the first session and £5 for the second.[4]

2.5 Data

Over the two samples, 723 participants completed the first session (402 and 321 from Samples 1 and 2, respectively). Of these, 68 (9.4%) failed to either start or complete the second session. We drop an additional 66 participants (9.1%) who gave incorrect answers to ≥ 3 out of 5 comprehension questions. We drop a further 20 (2.8%) participants with incomplete response for responsibility beliefs or demographic characteristics, which are used to construct sample weights. This leaves an analysis sample of 569 participants (325 and 244 from Samples 1 and 2, respectively). We check robustness to sample selection.

Participants allocate resources to three individuals in each of 20 (Sample 1) or 30 (Sample 2) rounds. This gives 41,460 observations in total. From treatments A, B_1, B_2 and B_3, there

[4]There is no consensus on the impact of using real incentives Moffatt et al. (2009), but the prevailing view is that for the type of tasks in this experiment, which involve distributing health resources to others, hypothetical choices are not necessarily less informative than financially motivated choices.

are 17,070, 7,320, 9,750 and 7,320 observations, respectively. Each sample was intended to be representative of the UK adult population with respect to sex, age and ethnicity. We apply sample weights (Appendix B) to improve balance between Sample 1 and Sample 2.

3 Specification and estimation of welfare weights

3.1 Without responsibility sensitivity

We assume that each participant allocates resources between individuals in each round as if maximising (subject to random errors) a weighted utilitarian social welfare function (SWF) (Fleming, 1952; Harsanyi, 1955),

$$W = \sum_{i=1}^{N} \omega_i U(h_i), \qquad 0 \leq \omega_i \leq 1 \, \forall_i, \quad \sum_{i=1}^{N} \omega_i = 1, \tag{1}$$

where $U(.)$ represents a participant's social utility derived from the health (h_i) of an individual irrespective of that individual's non-health characteristics, ω_i is a Pareto weight on that utility (Piketty and Saez, 2013) and N is the population size (= 3 in each round).[5] Utility is a function only of health because that is the only outcome participants can determine when allocating resources. This restriction is consistent with our objective of eliciting social preferences over the distribution of health in order to evaluate policies that impact that distribution. The weights may depend on non-health characteristics and, as explained in the next sub-section, possibly on responsibility for those characteristics. That leaves scope

[5] Although participants are encouraged to think of health as QALYs, which is a utility metric, many are likely to interpret it as raw health rather than wellbeing derived from health. Even if it were interpreted as health utility, a concave transformation by $U()$ and then aggregation over this measure may still capture social preferences — such as those emanating from a prioritarian ethical concern for the wellbeing of the worst off (Parfit, 2000) — better than a linear aggregation (Bleichrodt et al., 2005; Piketty and Saez, 2013; Lakdawalla and Phelps, 2020).

for social preferences that are non-welfarist, reflecting concern about inequity (and not only inequality) in the distribution of health.

We specify social utility as iso-elastic (Atkinson, 1970),[6]

$$U(h_i) = \begin{cases} (h_i^{1-\varepsilon} - 1)/(1-\varepsilon) & \text{if } \varepsilon \geq 0 \;\&\; \varepsilon \neq 1, \\ \ln h_i & \text{if } \varepsilon = 1. \end{cases} \tag{2}$$

Increasing ε represents increasing aversion to univariate health inequality: willingness to sacrifice maximisation of health for less variation in health between individuals irrespective of their non-health characteristics. As $\varepsilon \to \infty$, preferences approach *maximin* (Rawls, 1971): maximisation of health of the least healthy.

When estimating with data from treatment A, in which participants get no information on non-health characteristics, the Pareto weights are constant, $\omega_i = 1/N \; \forall_i$. When using data from treatment B, in which individuals are labelled by sex, income or smoking, we allow the weights, and so prioritisation, to vary with these potentially equity-relevant characteristics. Social welfare can then depend on both equity in the distribution of health by non-health characteristics (captured by ω_i) and the trade off between efficiency and equality in the univariate distribution of health (captured by ε). Both concerns are reflected in relative (generalised) social marginal welfare weights (SMWW) (Piketty and Saez, 2013; Saez and Stantcheva, 2016) — the social decision maker's marginal rate of substitution between the health of individuals,[7]

$$\text{Relative}\, SMWW_{ij} = -\left.\frac{dh_j}{dh_i}\right|_{dW=0} = \frac{\omega_i}{\omega_j}\left(\frac{h_j}{h_i}\right)^{\varepsilon}. \tag{3}$$

[6]Robson et al. (2024) find that this specification, which allows for aversion to relative inequality, fits the Sample 1 data slightly better than another that allows for aversion to absolute inequality.

[7]Given the linear health production technology, $h_i = p_i \times y_i$, the relative SMWW on *resources* is equal to (3) multiplied by the ratio of the respective productivities, p_i/p_j. The optimal allocation of resources is determined by these productivities in addition to ω_i and ε.

Health policies can be evaluated by applying these welfare weights to policy-induced health changes across individuals distinguished by non-health characteristics potentially associated with initial health.[8]

We assume the ω_i (equity weights, hereafter) are a function of ranks of individuals ordered by the respective characteristic, such that $x_{ki} \leq x_{ki+1} \forall_i$, $k \in \{1, 2, 3\}$ for sex, income and smoking, respectively. We order sex from females to males, incomes from low to high and smoking from non-smoker to heavy smoker. We specify the weights as

$$\upsilon_{ki} = \frac{[P(X \geq x_{ki})]^{\beta_k} - [P(X > x_{ki})]^{\beta_k}}{N \times P(X = x_{ki})}, \tag{4}$$

$$\omega_{ki} = \begin{cases} \upsilon_{ki} & \text{if } \varepsilon \leq 1, \\ {\upsilon_{ki}}^{\varepsilon} / \sum_{j=1}^{N} {\upsilon_{kj}}^{\varepsilon} & \text{if } \varepsilon > 1. \end{cases} \tag{5}$$

where $P()$ is a probability. Equation (4) is a standard specification of rank-dependent weights (Donaldson and Weymark, 1980; Donaldson and Weymark, 1983; Yitzhaki, 1983; Robson et al., 2024). The denominator ensures that $\sum_i \upsilon_{ki} = 1$ even when there are ties, which happens for sex. The $\beta_k > 0$ parameter captures the direction and intensity of preferences for prioritisation by rank of the respective non-health characteristic. With $\beta_k > 1$, υ_{ki} decreases monotonically with increasing rank, giving greater weight to the health of individuals with relatively less of the non-health characteristic, i.e. females, those with lower incomes and those who smoke less. With $1 < \beta_k < 2$, $\beta_k = 2$ and $\beta_k > 2$, these weights decline concavely, linearly and convexly, respectively, with the ranks. As $\beta_k \to \infty$, they approach zero for all but the lowest ranked individual. With $0 < \beta_k < 1$, these weights increase monotonically with increasing rank and the direction of prioritisation is reversed.

[8] We provide code in an online repository (https://data.mendeley.com/datasets/6wdpxx2hfp) which allows users to calculate SMWWs using the preference parameters we estimate. Users can input data on health (QALEs) and, if desired, sex, income and/or smoking status of individuals or groups in a population.

With $\beta_k = 1$, the weights are constant and there is no prioritisation by that non-health characteristic.

The transformation of the equity weights in (5) for $\varepsilon > 1$ deals with an issue arising from the iso-elastic utility function (Thöni, 2015; Inukai et al., 2022). As $\varepsilon \to \infty$, that parameter becomes dominant: the model with untransformed weights predicts equalisation of health irrespective of the values of those weights.[9] In this case of extreme aversion to univariate health inequality, there is no scope for the weights, if they are not transformed, to capture a co-existing preference for health prioritisation by non-health characteristics. A given set of allocations can be consistent with different values of the weights. Hence, without the transformation, as ε increases the weights become more difficult to identify and to interpret as the importance attached to the health of types of individuals (Inukai et al., 2022). The transformation of the weights in (5) for $\varepsilon > 1$ deals with both problems (Senhadji, 1997; Inukai et al., 2022).[10] It ensures that the weights always matter, and so they can be interpreted as the importance attached to health on the basis of non-health, equity-relevant characteristics irrespective of the degree of aversion to univariate health inequality. This helps not only with the interpretation of the β_k parameter but also with its identification.

[9] Maximisation of (1) given (2), subject to the linear health production function, $h_i = p_i \times y_i$, and the resource constraint, $m = \sum_i y_i$, with untransformed weights, $\omega_i = \upsilon_i$, gives the optimal allocation, $y_i^* = \frac{m}{1+\sum_{j \neq i} \frac{p_i}{p_j} \left(\frac{\upsilon_j \, p_j}{\upsilon_i \, p_i} \right)^{\frac{1}{\varepsilon}}} \quad \forall_i$. As $\varepsilon \to \infty$, this approaches $y_i^* = m / \left(p_i \sum_j 1/p_j \right)$. So, the relative allocations to two individuals i and g are $y_i^*/y_g^* = p_g/p_i \ \forall \upsilon_i \ \& \, \upsilon_g$, which implies equal optimal health given the linear technology. With iso-elastic utility and untransformed equity weights, as $\varepsilon \to \infty$ the resulting (2-person) L-shaped social indifference curves over health must also be symmetric. With this specification, it is not possible to capture a preference for holding the relative health of two individuals (distinguished by some non-health characteristic) constant at any ratio other than 1 (Thöni, 2015).

[10] With the transformation, the optimal allocation is $y_i^* = \frac{m}{1+\sum_{j \neq i} \frac{p_i}{p_j} \frac{\upsilon_j}{\upsilon_i} \left(\frac{p_j}{p_i} \right)^{\frac{1}{\varepsilon}}} \quad \forall_i$. As $\varepsilon \to \infty$, this approaches $y_i^* = \frac{m}{\frac{p_i}{\upsilon_i} \sum_j \frac{\upsilon_j}{p_j}}$. In this case, the relative allocations to two individuals i and g are $\frac{y_i^*}{y_g^*} = \frac{p_g}{p_i} \frac{\upsilon_i}{\upsilon_g}$, which implies $h_i^*/h_g^* = \upsilon_i/\upsilon_g$. The transformation of the weights in the case of $\varepsilon > 1$ makes it possible to capture preference for holding relative health constant at a ratio determined by the weights and not necessarily equal to 1. Indifference curves can be L-shaped and asymmetric (Thöni, 2015).

3.2 With responsibility sensitivity

The specification of preferences above allows for prioritisation of health by non-health characteristics without motivating such prioritisation from a theory of justice. To obtain generalised welfare weights (Saez and Stantcheva, 2016) that are consistent with responsibility-sensitive egalitarianism (Fleurbaey, 2008), we allow the prioritisation of health by a non-health characteristic to depend on beliefs about responsibility for that characteristic. We specify,

$$\beta_k = \gamma_k {\phi_k}^{\theta_k}, \qquad \gamma_k > 0,\ \phi_k > 0,\ 0 \leq \theta_k \leq 1, \tag{6}$$

where θ_k represents beliefs about the extent to which people are responsible for their ranks by characteristic x_k. If they are believed to have no responsibility, then $\theta_k = 0$ and the parameter γ_k determines the equity weight to an individual with x_{ki}. That is the case when a participant believes that circumstances beyond an individual's control fully determine their position in the distribution of the characteristic. According to the compensation principle of responsibility-sensitive egalitarianism (Fleurbaey, 1994; Fleurbaey, 2008) and equality of opportunity theory (Roemer, 1993; Roemer, 2002), disadvantageous circumstances should be compensated and advantageous circumstances taxed. That would imply $\gamma_k > 1$ (and $\beta_k > 1$ given $\theta_k = 0$), such that the weights, ω_{ki}, decline with increasing rank of x_{ki} when the characteristic is considered advantageous. For example, the weights may decrease with increasing relative income if richer individuals are believed to be advantaged entirely due to luck of being born into wealthier families. When the characteristic is believed to be disadvantageous and determined solely by circumstances, then $\gamma_k < 1$ (and $\beta_k < 1$) and the weights increase with the rank of x_{ki}. For example, if smoking is believed to be entirely determined by addiction attributable to the interaction of social environment up to adolescence and genes, then smokers may be prioritised.

A circumstance-characteristic need not generate health advantage or disadvantage to motivate health prioritisation by that characteristic. Priority may be given to the health of individuals exposed to circumstances that constrain overall wellbeing without interfering with the production of health. For example, greater priority may be given to the health of females or poorer individuals to compensate for any general disadvantage these groups are perceived to experience. In the experiment, the multipliers that determine health resource productivity are orthogonal to each non-health characteristic and no information about general wellbeing is provided. Nonetheless, participants may believe that a non-health characteristic affects the production of health or that it is associated with wellbeing. Such beliefs may influence allocations of health resources — the only available instrument of redistribution — across individuals distinguished by that characteristic.[11]

If participants believe that people are at least partly responsible for characteristic x_k, then $\theta_k > 0$ and preferences for the prioritisation of health by that characteristic are also reflected in the parameter ϕ_k. Such beliefs presumably stem from a perception that people can, to some extent, exert effort to change their incomes or smoking behavior. According to the reward principle of responsibility-sensitive egalitarianism (Fleurbaey, 2008), that effort should be rewarded, possibly by giving greater weight to those who exert effort to change their position in the distribution of the respective characteristic (Roemer, 2002).

If the ranks of x_{ki} were believed to be entirely determined by effort, which corresponds to $\theta_k = 1$, then the weight parameter $\beta_k = \gamma_k \phi_k$ would either be deflated ($\phi_k < 1$) or inflated ($\phi_k > 1$) compared with the case in which the characteristic is entirely attributed to circumstance ($\theta_k = 0$, $\beta_k = \gamma_k$). If those with higher ranks would get lower weights in that comparison case ($\gamma_k > 1$), then $\phi_k < 1$ would reflect reward for effort made to

[11] Robson et al. (2024) find that, at least on average, resource allocations are not (statistically significantly) responsive to manipulated beliefs about income having a causal effect on health. However, there is heterogeneity in this response and allocations are associated with beliefs about income causation elicited prior to manipulation.

increase the rank. For example, the weights may decrease less (or even increase) with relative income if it is believed that people get richer principally through hard work and, therefore, deserve relatively greater priority. On the other hand, if those with more of a disadvantageous characteristic would get larger weights if the characteristic were believed to be entirely determined by circumstances ($\gamma_k < 1$), then, under different beliefs, $\phi_k > 1$ would reflect a propensity to penalise lack of effort to reduce the characteristic. This combination of parameter values would, for example, capture the preferences of both those who believe that smokers should be compensated for their addiction (higher weights) and those who believe that smoking is a free choice that should be penalised (lower weights).

For interior beliefs that attribute a characteristic neither entirely to circumstance nor entirely to effort, ($0 < \theta_k < 1$), the weight parameter (β_k) decreases with increasing strength of belief that effort predominates (θ_k) if $\phi_k < 1$. And β_k increases with θ_k when $\phi_k > 1$.[12] For example, $\gamma_2 > 1$ and $\phi_2 < 1$ would capture prioritisation of the health of poorer individuals that weakens with strengthening beliefs that people are responsible for their lower incomes (θ_2 increasing within $[0, 1]$). Similarly, $\gamma_3 < 1$ and $\phi_3 > 1$ would capture a propensity to prioritise the health of addicted smokers that weakens or switches to increasing prioritisation of non-smokers with stronger beliefs that people are responsible for their smoking behavior.

3.3 Estimation

We assume that, subject to random error, participants allocate to maximise welfare given by (1) and (2) subject to a resource constraint, $\sum_{i=1}^{N} y_i = m$, and the linear health production

[12] $\partial\beta_k/\partial\theta_k = \gamma_k ln\left(\phi_k\right) {\phi_k}^{\theta_k} \begin{cases} < 0 & \text{if} \quad \phi_k < 1, \\ > 0 & \text{if} \quad \phi_k > 1. \end{cases}$

function, $h_i = p_i \times y_i$. The optimal resource allocations are

$$y_i^* = \frac{m}{1 + \sum_{j \neq i}^{N} \frac{p_i}{p_j} \left(\frac{\omega_j \, p_j}{\omega_i \, p_i} \right)^{\frac{1}{\varepsilon}}} \quad \forall_i, \tag{7}$$

where ω_i is $1/N$ for observations from treatment A and it is ω_{ki}, as specified in (4) and (5), for observations from each variant of treatment B.

To identify responsibility-sensitive weights, we add (6) to the specification of the weights and use the elicited beliefs about responsibility for income and smoking (rescaled to 0-1) to fix θ_2 and θ_3, respectively, for each participant. We use these data plus the resource allocations made as the respective non-health characteristic (income or smoking) varies from round to round independently of variation in the multipliers to estimate γ_k, ϕ_k and, consequently, β_k that determines the weights. With treatment B_1 data, (4) and (5) completely specify the weights, given the assumption that all participants believe that sex is exogenous ($\theta_1 = 0$).

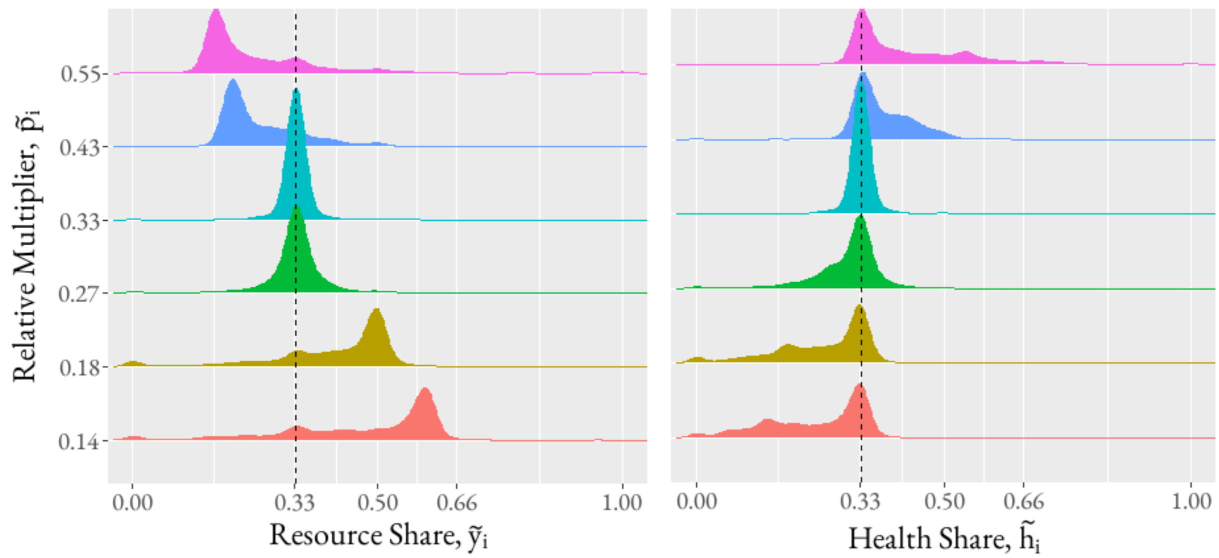
We allow actual allocations to deviate from the optimal ones by Dirichlet distributed random errors (Dirichlet, 1839; Robson, 2021). We pool data over rounds, treatments and participants, and we use maximum likelihood to estimate the preferences of a representative social decision maker (SDM) (Appendix C). We use data from treatment A, as well as B, in order to estimate ε more precisely. We apply sample weights in all analyses and use bootstrap inference (percentile method, 1000 repetitions).

4 Results

4.1 Non-parametric analysis

Health inequality aversion.—Using treatment A data, Figure 1 plots distributions of resource shares, $\tilde{y}_i = y_i / \sum_i^N y_i$, and health shares, $\tilde{h}_i = h_i / \sum_i^N h_i$, stratified by the productivity of resources in generating health, which is indicated by relative multipliers, $\tilde{p}_i = p_i / \sum_i^N p_i$.

Figure 1: Distributions of Resource and Health Shares by Productivity



Note: Data are 17,070 observations pooled over 569 participants and all rounds of treatment A. $\tilde{y}_i = y_i / \sum_i^N y_i$, $\tilde{h}_i = h_i / \sum_i^N h_i$, $\tilde{p}_i = p_i / \sum_i^N p_i$.

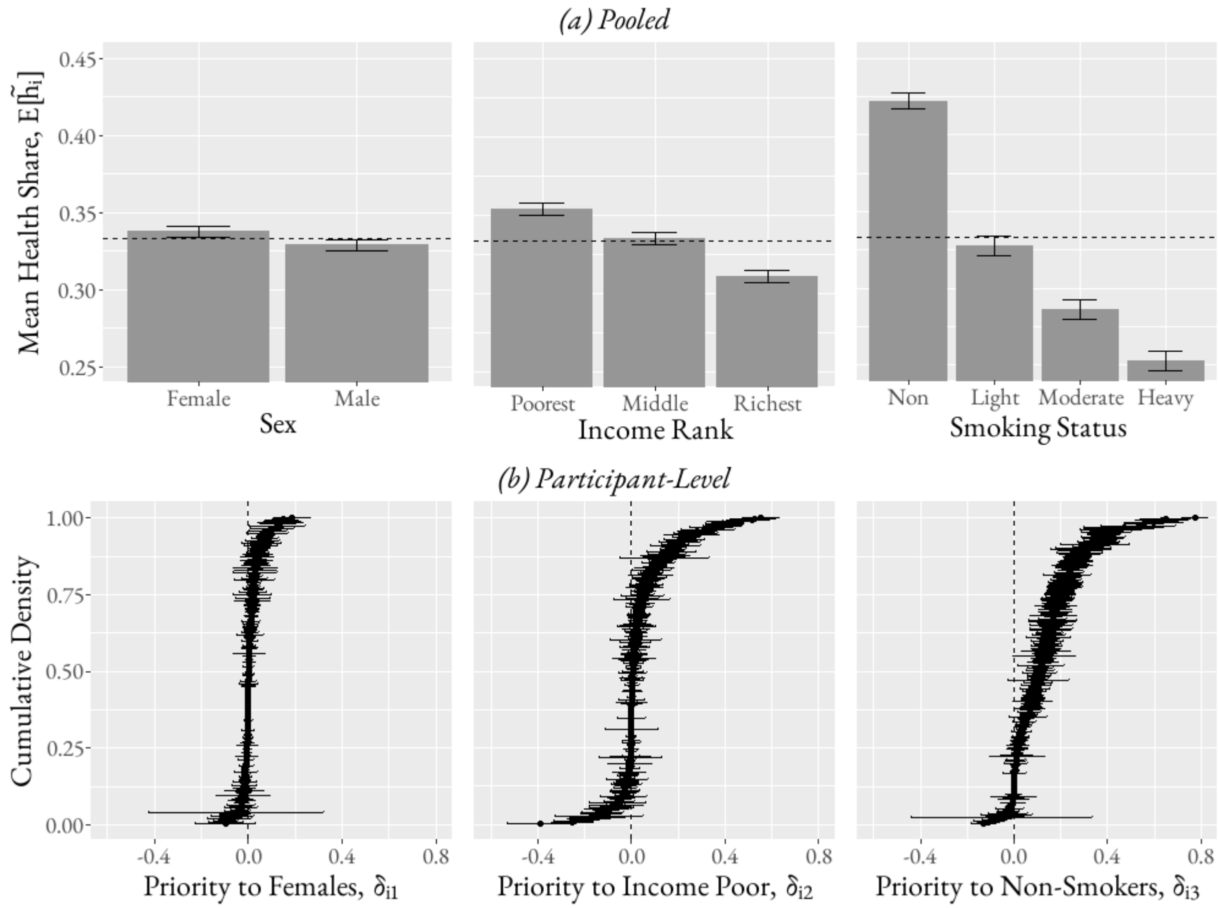
At $\tilde{p}_i = 1/3$, each individual within a round has the same productivity ($p_i = 1, \forall_i$), and so there is no trade-off between the maximisation and the equalisation of health. In this case, resources, and therefore health, are distributed equally in a vast majority of allocations. As the relative multiplier of an individual increases, a health maximising SDM would give that more productive individual a greater share of resources. A majority does the opposite. However, the reduction in resources is usually insufficient to entirely offset the productivity advantage, and so the relatively more productive individual still tends to end up with a larger share of health. There is heterogeneity. Some participants reveal maximin preferences by allocating less resources to the more productive individual to an extent sufficient to equalise the distribution of health: $\tilde{h}_i = 1/3$. Others are efficiency seeking in the sense that they give a greater share of resources, and so health, to the more productive individual. Appendix D.1 provides further evidence of equality-efficiency trade-offs.

Prioritisation.—The top panel of Figure 2 shows health shares by sex, income and smoking status averaged over all participants and all rounds of treatments B_1, B_2 and B_3, respectively. Since assignment to categories of each characteristic is orthogonal to the multipliers, any differences in mean health shares across categories implies prioritisation by that characteristic. Females are allocated slightly more health than males. Poorer individuals are given more health than richer individuals, on average. There is an even stronger gradient in mean health shares favoring those who smoke less or not at all. Regressions that allow for participant-level random effects confirm all three gradients (Appendix D.2 Table D3).[13]

The bottom panel of Figure 2 uses treatment B data to plot the distribution across participants of the difference between the health shares to a) females versus males, b) the poorest versus all other individuals, and c) non-smokers versus all other individuals. A vast majority of participants (79.5%) give equal health shares irrespective of sex. Around 14.8% give a significantly larger health share to females. A smaller percentage (5.7%) prioritise male health. A majority (60.9%) displays no significant prioritisation of the poorest. About 30.2% give the poorest a significantly larger health share, while much less (8.9%) give the poorest a smaller share. Around 70.1% give a significantly larger share to non-smokers, while most of the remainder (27.5%) do not significantly prioritise the health of non-smokers over smokers.

[13] Allocations by sex, income and smoking status are associated with participants' stated willingness to prioritise by each of these characteristics (Appendix D.2 Table D4) but not with their own characteristics (Table D5).

Figure 2: Health Shares by Non-Health Characteristics



Note: The top panel shows the health share, $\tilde{h}_{irs}$, to individuals (i) distinguished by sex, income and smoking status averaged over all participants (s) and rounds (r) of treatment B_1, B_2 and B_3, respectively. For each of these treatments, n=17,070 (= 569 × 10 × 3). Interval lines show 90% confidence intervals (CIs). The bottom panel shows empirical cumulative distributions of $\hat{\delta}_{ks} = E[\tilde{h}_{kjrs}] - E[\tilde{h}_{k!jrs}]$, where $\tilde{h}_{kjrs}$ is the health share participant s gives in round r to individual j defined as a) female if $k = 1$ (Treatment B_1), b) the poorest individual if $k = 2$ (Treatment B_2), and c) the non-smoker if $k = 3$ (Treatment B_3) and the second expectation is taken over the respective health shares $\forall_{i \neq j}$. We estimate these differences in expectations from participant-level linear regressions (n=30): $\tilde{h}_{kirs} = \delta_{0ks} + \delta_{ks} \mathbb{1}(i = j) + \upsilon_{kirs}$. Interval lines show 90% CIs. There are 569 participants.

Responsibility beliefs.—The median (rescaled) reported belief about the extent to which people are responsible for the income they earn is 0.6, indicating that a majority is closer to believing that people are entirely responsible (1) than it is to believing they are not responsible at all (0). There are much stronger beliefs that people are responsible for their smoking behavior; the median is 1. There is a great deal of heterogeneity in both beliefs

(Figure 4, top panel). Participants who believe that people are largely responsible for their incomes/smoking give smaller health shares to individuals identified as poorer/smokers compared with participants who believe people are less responsible for their incomes/smoking (Appendix D.3 Table D6). Higher-income participants tend to report that people are more responsible for the income they earn, while there is no association between smoking behavior and beliefs about responsibility for smoking (Appendix D.3 Table D7).

4.2 Parametric analysis

Without responsibility sensitivity.—The top panel of Table 1 shows estimates of the inequality aversion parameter, ε, and the equity weight parameters, β_k, obtained without allowing preferences for prioritisation by income and smoking to depend on beliefs about responsibility for those characteristics. An estimate of $\hat{\varepsilon} = 1.45$, which is significantly greater than zero, indicates aversion to health inequality across individuals irrespective of their non-health characteristics. While the point estimate of the weight parameter for females is above 1, which would imply prioritisation of female health, the null of no such prioritisation is not rejected. The weight parameter estimates are significantly greater than 1 for income and smoking status, indicating prioritisation in favor of poorer individuals and those who smoke less or not at all. The larger magnitude of the respective parameter estimate for smoking implies stronger prioritisation by smoking behavior than by income, which is evident from the non-parametric analysis.

All these findings are robust to a) estimation with data from treatments A, B_1, B_2 and B_3 separately (Appendix E.1), b) not restricting the sample and not applying sample weights (Appendix E.2), and c) order effects (Appendix E.3). Participant-specific estimation gives a median estimate of inequality aversion (ε) somewhat larger than the pooled estimate (2.91

vs 1.45) and reveals that most participants give larger weights to non-smokers, while there is more heterogeneity in the sex and income weights (Appendix E.6).

Table 1: Inequality Aversion and Equity Weight Parameter Estimates

(a) Without responsibility sensitivity

Inequality	Equity Weight			AIC	Obs.
Aversion	Sex	Income	Smoking		
ε	β_1	β_2	β_3		
1.450 [1.41, 1.50]	1.021 [0.99, 1.05]	1.078 [1.06, 1.10]	1.383 [1.34, 1.42]	-29,504.7	41,460

(b) With responsibility sensitivity

Inequality	Equity Weight						AIC	Obs.
Aversion	Sex		Income		Smoking			
ε	γ_1	ϕ_1	γ_2	ϕ_2	γ_3	ϕ_3		
1.457 [1.38, 1.50]	1.005 [0.97, 1.06]	1	1.156 [0.97, 1.38]	0.876 [0.69, 1.24]	0.773 [0.67, 1.16]	1.924 [1.11, 2.19]	-29,615.3	41,460

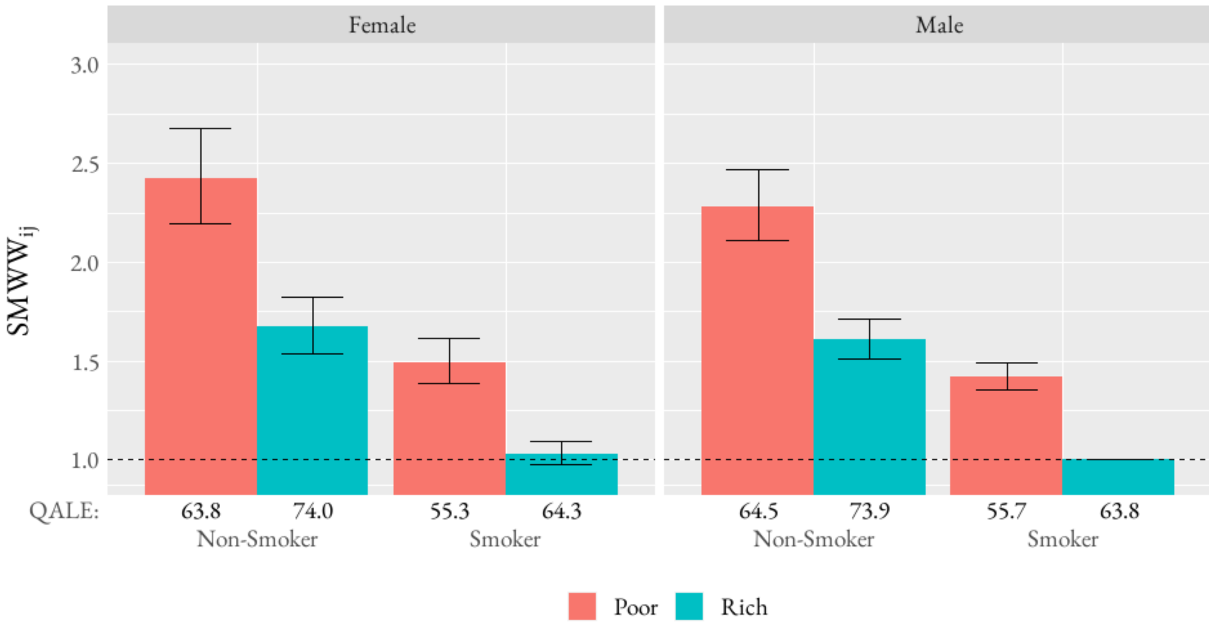
Note: Maximum likelihood estimates from data pooled over all treatments, rounds and participants (Appendix C). In the top panel, the equity weight parameters are specified in (4) and (5). In the bottom panel, the equity weight parameters are specified in (4), (5) and (6). For $\theta_k, k \in \{2, 3\}$, we use each participant's reported belief about the extent to which people are responsible for their incomes/smoking. We set $\theta_1 = 0$ and so, effectively, fix $\phi_1 = 1$. In brackets are 90% bootstrap confidence intervals (percentile method, 1000 replications). AIC is Akaike Information Criterion. Obs. is number of observations. Number of participants is 569. See Appendix Table E4 for additional estimates and statistics.

Based on the parameter estimates in the top panel of Table 1 and external estimates of health levels measured by quality adjusted life expectancy (QALE) (Appendix F), Figure 3 shows relative social marginal welfare weights defined by (3) for individuals who differ in multiple non-health characteristics.[14]

Relative to the reference category of male, rich (top 20%) smokers, the health of female, rich smokers gets a weight very close to (and not significantly different from) 1. There is approximately equal weighting of male and female health within all income-smoking categories. This is due to a slight (not significant) prioritisation of female health ($\hat{\beta}_1 > 1$) that

[14]Appendix E.4 gives relative welfare weights for individuals differing in each characteristic separately.

Figure 3: Relative Social Marginal Welfare Weights



Note: Estimates of (3) with a reference category (j) of male, rich smokers. To obtain welfare weights for multidimensional (sex, income and smoking) individuals, we calculate equity weights as the product of estimated unidimensional weights, $\upsilon_i = \prod_{k=1}^{K} \upsilon_{ki}$, where υ_{ki} is given by (4), and normalise (to sum to 1) and transform for $\varepsilon > 1$: $\omega_i = \frac{\upsilon_i}{\sum_{j=1}^{N} \upsilon_j}$ if $\varepsilon \leq 1$, and $\omega_i = \frac{\upsilon_i{}^{\varepsilon}}{\sum_{j=1}^{N} \upsilon_j{}^{\varepsilon}}$ if $\varepsilon > 1$. Parameter estimates are given in the top panel of Table 1. Health (h_i) is Quality Adjusted Life Expectancy (QALE), which is derived from external estimates (Appendix F) and is shown for each group below the x-axis. Poor and Rich are defined as the bottom and top quintile group, respectively, of an index of multiple deprivation (Appendix F). Interval lines show 90% bootstrap confidence intervals (percentile method, 1000 replications).

is offset by the combined effect of a female health (QALE) advantage of about 0.2 QALYs and aversion to health inequality ($\hat{\varepsilon} = 1.45 > 0$).

The representative SDM values a marginal health gain to male, poor (bottom 20%) smokers about 1.4 times more than a marginal improvement in the health of male, rich smokers. This reflects both direct prioritisation of the health of poorer individuals and indirect prioritisation of their lower health (QALE = 55.7 vs QALE = 63.8) due to health inequality aversion. The marginal social value of improving the health of a male, rich non-smoker is approximately 1.6 greater than the value of a health gain to a male, rich smoker. This reflects strong direct prioritisation of health of non-smokers that is constrained, to some extent, by aversion to health inequality that disadvantages smokers. For smoking, the equity weights and inequality aversion work against each other in determining the welfare weights, while for income they push in the same direction of favoring poorer individuals.

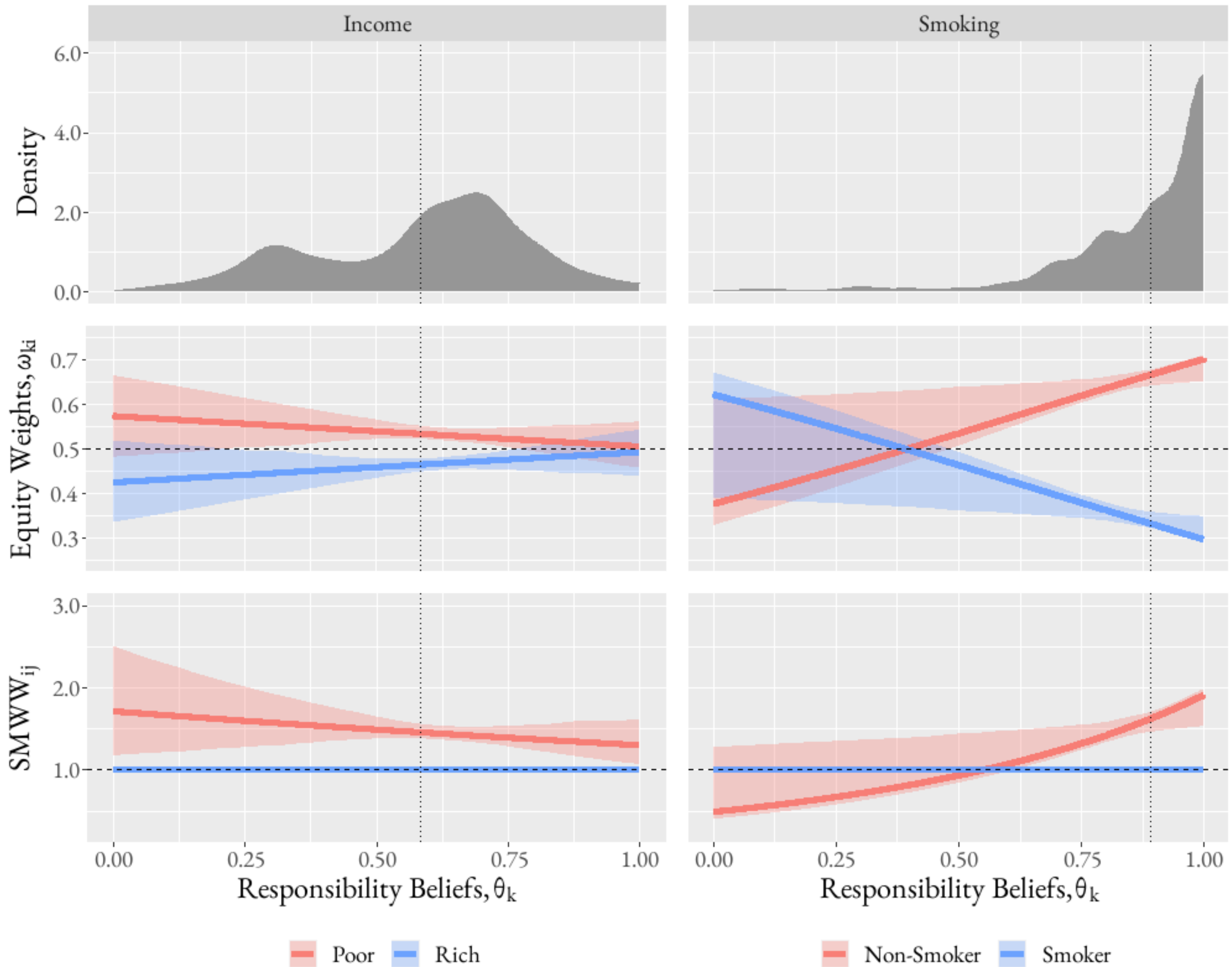
With responsibility sensitivity.—The bottom panel of Table 1 shows parameter estimates with prioritisation of health by income and smoking allowed to depend on responsibility for those characteristics. The Akaike Information Criteria show that this model fits the data better than the model without responsibility sensitivity, which, in turn, gives better fit than not allowing for equity weights (Appendix Table E4).

A point estimate of γ_2 greater than 1 suggests that the representative SDM would give greater weight to the health of poorer individuals if it were believed that they have no responsibility for their incomes. However, the 90% confidence interval estimate of this parameter includes 1.

A point estimate of ϕ_2 less than 1 indicates that sample participants with stronger beliefs that people are responsible for their incomes show less favor to the health of poorer individuals. The 90% confidence interval again includes 1, indicating non-rejection of the null that the representative SDM's prioritisation of health by income is independent of responsibility for income. However, the average may obscure heterogeneity. The middle-left panel of Figure 4 shows that, in a society with two income groups, the equity weights change from pro-poor to neutral as beliefs shift toward regarding income to be entirely a personal responsibility ($\theta_2 \rightarrow 1$). At the mean responsibility belief of $\theta_2 = 0.58$, the estimated equity weights imply prioritisation of the health of poorer individuals even though they are believed to be partly responsible for their lower incomes.

The bottom-left panel of the figure reveals that, irrespective of responsibility beliefs, a marginal increase in the health of a poor individual is always of greater social value than an equal increase in the health of a rich individual: Relative $SMWW > 1 \ \forall \ \theta_2$. The relative

Figure 4: Equity and Welfare Weights over Distributions of Responsibility Beliefs



Note: The top panel shows the distribution (over 569 participants) of elicited beliefs (θ_k) about responsibility for income/smoking, with 1 corresponding to a belief that people are entirely responsible and 0 indicating no responsibility. Vertical dotted lines show mean responsibility beliefs. The other two panels represent societies each with two equal-sized groups of homogeneous individuals: Poor vs Rich (left) and Non-smokers vs Smokers (right). We normalise the size of each group to 1. The middle panel shows equity weights, ω_i, obtained from (4), (5) and (6). Since there are only two groups, these weights sum to 1 at every value of θ_k and $\omega_i = 0.5$ corresponds to equal weights. The bottom panel shows relative welfare weights defined in (3), with the rich and smokers being the respective reference groups (j). Group-specific health (h_i) is QALE derived from external estimates (Appendix F): 62.2 and 73.3 for poor (bottom 20%) and rich (top 20%), respectively, and 69.6 and 60.2 for non-smokers and smokers, respectively. All weights are estimated using the parameter estimates in the bottom panel of Table 1. Shading around lines indicates 90% bootstrap confidence intervals (percentile method, 1000 replications).

welfare weight on the poor decreases as beliefs shift toward viewing people as more responsible for their incomes. However, it never falls to 1 because aversion to health inequality would preserve the higher social value of health gains to the poor, who are less healthy, even if full attribution of personal responsibility for income were to eliminate explicit prioritisation of the poor.

For smoking, the point estimate of γ_3 is less than 1, which is consistent with prioritisation of the health of (heavier) smokers if they were considered to have no responsibility for their smoking, although the 90% confidence interval includes 1. An estimate of ϕ_3 significantly above 1 indicates that prioritisation of the health of smokers recedes as they are perceived to have greater responsibility for their smoking. The middle-right panel of Figure 4 shows that, in a society with smokers and non-smokers, the equity weights would switch from favoring smokers to favoring non-smokers when belief (θ_3) that smoking is a personal responsibility reaches about 0.4, which is well below the mean of almost 0.9.[15] As with the estimates obtained without allowing for responsibility-sensitive preferences, less priority is given to the health of (heavier) smokers, on average. And their health is given even lower weight as beliefs that they are responsible for their smoking strengthen.

The bottom-right panel of the figure shows that, if it were believed that smokers were not at all responsible for their habit ($\theta_3 = 0$), then a marginal gain in the health of non-smokers would generate less social value than an equal improvement in the health of smokers because the first change would increase health inequality while the second would reduce it, given that non-smokers are in better health. However, in addition to the caveat that the 90% confidence interval of the relative welfare weight includes 1 at $\theta_3 = 0$, the point estimate below 1 is contingent on the functional form used for extrapolation, given that very few participants believe that there is no personal responsibility for smoking. A stronger result is that as beliefs shift toward attribution of greater personal responsibility for smoking, the relative

[15] Given $\beta_3 = \gamma_3 {\phi_3}^{\theta_3}$, with $\hat{\gamma_3} = 0.773$ and $\hat{\phi_3} = 1.924$, $\theta_3 > 0.394 \rightarrow \beta_3 > 1$.

welfare weight on the health of non-smokers increases and reaches 1 when $\theta_3 \approx 0.6$. This is above the threshold belief at which the equity weights switch to prioritise non-smokers because inequality aversion combined with the lower health of smokers maintains, to some extent, the higher social value of health gains to smokers even when they are not explicitly prioritised. The mean ($\theta_3 \approx 0.9$) and median ($\theta_3 \approx 1$) beliefs are well above the switch point, and so a majority places a higher marginal value on the health of non-smokers even taking into account that they enjoy a health advantage.

5 Conclusion

Inconsistent with the prevailing evaluation practice of weighting health gains from interventions equally, we find that preferences elicited from a UK general public sample imply greater social value from health gains to those who a) are in worse health, b) have lower incomes, and c) smoke less or not at all. Larger welfare weights on those in worse health reflect aversion to health inequality. The weights are consistent with a responsibility-sensitive egalitarian conception of distributive justice since stronger beliefs that people are responsible for their incomes are associated with weaker prioritisation of the health of the poor (although this difference is not significant in parametric analysis) and stronger beliefs that smoking is a personal responsibility are associated with greater prioritisation of the health of non-smokers.

These patterns — observable at least in the sample — are consistent with the view that, to an extent, people deserve their just health deserts (Miller, 1999; Scott et al., 2001). That is, if they choose to smoke, then it is not entirely unfair to weight any resulting ill-health less than a same-sized health loss that arises entirely by chance. And if they choose to work hard and earn more, then it would be unfair to deny them any resulting opportunities for better health. However, we do not find support for strict adherence to this desert (or reward) principle of justice (Fleurbaey, 2008; Feldman and Skow, 2020). First, in our sample, most believe that people are only moderately responsible for their incomes and many do not hold people entirely responsible for their smoking. Second, there is support for a compensation principle (Fleurbaey, 2008) of prioritising health of those who are poor or smoke partly due to circumstances beyond their control. Third, health inequality aversion moderates the sensitivity of welfare weights to responsibility. We find that even if people were held entirely responsible for their incomes, the marginal social value of health would still be greater for poorer individuals because they are in worse health. And with complete personal responsibility for smoking, the effective prioritisation of the health of non-smokers (reflected in welfare weights) would be less than the explicit prioritisation (reflected in equity weights) because non-smokers are in better health.

One limitation is that we estimate responsibility-sensitive welfare weights from between-participant association of resource allocations with elicited responsibility beliefs. Without experimental manipulation of those beliefs, we do not have causal evidence of responsibility-sensitive egalitarian preferences. Addressing this limitation would require a careful design to induce exogenous variation in beliefs — never easy — while preserving cognitive feasibility of the task for a general population sample.[16]

Another limitation is that while the data are consistent with responsibility-sensitive egalitarianism, we cannot rule out that participants, even on average, have other motivations. For example, they may believe that smoking is a personal responsibility and give smokers low priority because they disapprove of smoking, irrespective of its health consequences. There is questionnaire survey evidence of such moralistic motivation of health resource allocation (Ubel and Loewenstein, 1996). Another potential mechanism, for which there is experimental evidence (Mollerstrom et al., 2015), is the view that it is fair to give lower priority to

[16]Robson et al. (2024) find that, on average, participants' resource allocations do not respond to manipulated beliefs about the causal effect of income on health productivity.

someone's health if they choose to risk it, e.g. by smoking, even when they avoid the bad outcome, e.g. lung cancer. However, these are limitations on the normative interpretation of our evidence rather than its validity.

Notwithstanding these limitations, this study provides insight into public opinion on the equitable distribution of health and delivers responsibility-sensitive welfare weights that can be used to motivate, design and evaluate health policies. The general approach offers the potential to estimate welfare weights in other contexts and for domains other than health.

Data Availability

A data repository including estimated preference parameters and code to calculate social marginal welfare weights is available at: https://data.mendeley.com/datasets/6wdpxx2hfp. Our code allows users to calculate SMWWs using data on health (QALEs) and, if desired, sex, income and/or smoking status, and/or beliefs on responsibility for income and smoking status.

References

Almås, I., Cappelen, A. W., & Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: Are americans more meritocratic and efficiency-seeking than scandinavians? *Journal of Political Economy*, *128*(5), 1753–1788. https://doi.org/10.1086/705551

Anderson, E. S. (1999). What is the point of equality? *Ethics*, *109*(2), 287–337. https://doi.org/10.1086/233897

Arneson, R. J. (1989). Equality and equal opportunity for welfare. *Philosophical Studies*, *56*(1), 77–93. https://doi.org/10.1007/BF00646210

Asada, Y., Brown, M., McNally, M., Murphy, A., Urquhart, R., & Warner, G. (2022). Personal Responsibility for Health: Exploring Together with Lay Persons. *Public Health Ethics*, *15*(2), 160–174. https://doi.org/10.1093/phe/phac009

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263. https://doi.org/10.1016/0022-0531(70)90039-6

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. https://doi.org/10.1016/j.jhealeco.2004.10.001

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, *97*(3), 818–827. https://doi.org/10.1257/aer.97.3.818

Cappelen, A. W., & Norheim, O. F. (2005). Responsibility in health care: A liberal egalitarian approach. *Journal of Medical Ethics*, *31*(8), 476–480. https://doi.org/10.1136/jme.2004.010421

Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics*, *99*(4), 906–944. https://doi.org/10.1086/293126

Cookson, R., & Dolan, P. (1999). Public views on health care rationing: A group discussion study. *Health Policy*, *49*(1), 63–74. https://doi.org/10.1016/S0168-8510(99)00043-3

Deaton, A. (2013). Inequalities in health: Concepts, measurement and ethics. In N. Eyal, S. A. Hurst, O. F. Norheim, & D. Wikler (Eds.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199931392.003.0018

Dirichlet, P. G. L. (1839). Sur une nouvelle méthode pour la détermination des intégrales multiples. *Journal de Mathématiques Pures et Appliquées*, *4*, 164–168.

Dolan, P., & Tsuchiya, A. (2009). The social welfare function and individual responsibility: Some theoretical issues and empirical evidence. *Journal of Health Economics*, *28*(1), 210–220. https://doi.org/10.1016/j.jhealeco.2008.10.003

Donaldson, D., & Weymark, J. A. (1980). A single-parameter generalization of the gini indices of inequality. *Journal of Economic Theory*, *22*(1), 67–86. https://doi.org/10.1016/0022-0531(80)90065-4

Donaldson, D., & Weymark, J. A. (1983). Ethically flexible gini indices for income distributions in the continuum. *Journal of Economic Theory*, *29*(2), 353–358. https://doi.org/10.1016/0022-0531(83)90053-4

Dworkin, R. (1981). What is equality? part 2: Equality of resources. *Philosophy & Public Affairs*, *10*(4), 283–345.

Edlin, R., Tsuchiya, A., & Dolan, P. (2012). Public preferences for responsibility versus public preferences for reducing inequalities. *Health Economics*, *21*(12), 1416–1426. https://doi.org/10.1002/hec.1799

Feldman, F., & Skow, B. (2020). Desert. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020). Metaphysics Research Lab, Stanford University.

Fleming, M. (1952). A cardinal concept of welfare. *The Quarterly Journal of Economics*, *66*(3), 366–384. https://doi.org/10.2307/1885309

Fleurbaey, M. (1994). On fair compensation. *Theory and Decision*, *36*, 277–307. https://doi.org/10.1007/BF01079932

Fleurbaey, M. (1995). Equal opportunity or equal social outcome. *Economics and Philosophy*, *11*, 25–55. https://doi.org//10.1017/S0266267100003217

Fleurbaey, M. (2008). *Fairness, responsibility, and welfare*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199215911.001.0001

Fleurbaey, M., & Maniquet, F. (2011). *A theory of fairness and social welfare*. Cambridge University Press. https://doi.org/10.1017/CBO9780511851971

Fleurbaey, M., & Schokkaert, E. (2009). Unfair inequalities in health and health care. *Journal of Health Economics*, *28*(1), 73–90. https://doi.org/10.1016/j.jhealeco.2008.07.016

Gaertner, W., & Schokkaert, E. (2011). *Empirical social choice: Questionnaire-experimental studies on distributive justice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139012867

Gu, Y., Lancsar, E., Ghijben, P., Butler, J. R., & Donaldson, C. (2015). Attributes and weights in health care priority setting: A systematic review of what counts and to what extent. *Social Science & Medicine*, *146*, 41–52. https://doi.org/10.1016/j.socscimed.2015.10.005

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, *63*(4), 309–321. https://doi.org/10.1086/257678

Hausman, D. M. (2007). What's wrong with health inequalities? *Journal of Political Philosophy*, *15*(1), 46–66. https://doi.org/10.1111/j.1467-9760.2007.00270.x

Hausman, D. M. (2013). Injustice and inequality in health and health care. In Y. Denier, C. Gastmans, & A. Vandevelde (Eds.), *Justice, luck & responsibility in health care: Philosophical background and ethical implications for end-of-life care* (pp. 29–42). Springer Netherlands. https://doi.org/10.1007/978-94-007-5335-8_2

Inukai, K., Shimodaira, Y., & Shiozawa, K. (2022). *Revisiting CES utility functions for distributional preferences: Do people face the equality–efficiency trade-off?* (ISER Discussion Paper No. 1195). Institute of Social and Economic Research, Osaka University.

ISER, I. (2023). Understanding society: Waves 1-12, 2009-2021 and harmonised bhps: Waves 1-18, 1991-2009. *UK Data Service*, 102276. https://doi.org/10.5255/UKDA-SN-6614-18

Konow, J. (1996). A positive theory of economic fairness. *Journal of Economic Behavior & Organization*, *31*(1), 13–35. https://doi.org/10.1016/S0167-2681(96)00862-1

Konow, J. (2001). Fair and square: The four sides of distributive justice. *Journal of Economic Behavior & Organization*, *46*(2), 137–164. https://doi.org/10.1016/S0167-2681(01)00194-9

Konow, J. (2003). Which is the fairest one of all? a positive analysis of justice theories. *Journal of Economic Literature*, *41*(4), 1188–1239. https://doi.org/10.1257/002205103771800013

Lakdawalla, D. N., & Phelps, C. E. (2020). Health technology assessment with risk aversion in health. *Journal of Health Economics*, *72*, 102346. https://doi.org/10.1016/j.jhealeco.2020.102346

Le Grand, J. (1991). *Equity and choice. an essay in economics and applied philosophy*. Harper Collins Academic. https://doi.org/10.4324/9780203400098

Love-Koh, J., Schneider, P., McNamara, S., Doran, T., & Gutacker, N. (2023). Decomposition of quality-adjusted life expectancy inequalities by mortality and health-related quality of life dimensions. *PharmacoEconomics*, *41*(7), 831–841. https://doi.org/10.1007/s40273-023-01264-9

Lundell, H., Niederdeppe, J., & Clarke, C. (2013). Public views about health causation, attributions of responsibility, and inequality [PMID: 23679219]. *Journal of Health Communication*, *18*(9), 1116–1130. https://doi.org/10.1080/10810730.2013.768724

Miller, D. (1999). *Principles of social justice*. Harvard University Press. https://doi.org/10.2307/j.ctv1pdrq04

Moffatt, P., Starmer, C., Sugden, R., Bardsley, N., Cubitt, R., & Loomes, G. (2009). *Experimental economics: Rethinking the rules*. Princeton University Press.

Mollerstrom, J., Reme, B.-A., & Sørensen, E. Ø. (2015). Luck, choice and responsibility — an experimental study of fairness views. *Journal of Public Economics*, *131*, 33–40. https://doi.org/10.1016/j.jpubeco.2015.08.010

Office of National Statistics. (2023). *Deprivation and the impact on smoking prevalence, England and Wales*. Retrieved June 13, 2024, from https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/drugusealcoholandsmoking/datasets/deprivationandtheimpactonsmokingprevalenceenglandandwales

Parfit, D. (2000). The ideal of equality. In M. Clayton & A. Williams (Eds.). Palgrave MacMillian.

Piketty, T., & Saez, E. (2013). Chapter 7 - optimal labor income taxation. In A. J. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of public economics, vol. 5* (pp. 391–474). Elsevier. https://doi.org/10.1016/B978-0-444-53759-1.00007-8

Rawls, J. (1971). *A theory of justice* [Eleventh printing, 1981]. The Belknap press of Harvard University Press. https://doi.org/10.2307/j.ctvjf9z6v

Robson, M. (2021). Inequality aversion, self-interest and social connectedness. *Journal of Economic Behavior and Organization*, *183*, 744–772. https://doi.org/10.1016/j.jebo.2020.12.029

Robson, M., O'Donnell, O., & Van Ourti, T. (2024). Aversion to health inequality — pure, income-related and income-caused. *Journal of Health Economics*, *94*, 102856. https://doi.org/10.1016/j.jhealeco.2024.102856

Roemer, J. E. (1993). A pragmatic theory of responsibility for the egalitarian planner. *Philosophy & Public Affairs*, *22*(2), 146–166.

Roemer, J. E. (1998). *Equality of opportunity*. Harvard University Press.

Roemer, J. E. (2002). Equality of opportunity: A progress report. *Social Choice and Welfare*, *19*, 455–471. https://doi.org/10.1007/s003550100123

Roemer, J. E., & Trannoy, A. (2015). Chapter 4 - equality of opportunity. In A. B. Atkinson & F. Bourguignon (Eds.), *Handbook of income distribution* (pp. 217–300). Elsevier. https://doi.org/10.1016/B978-0-444-59428-0.00005-9

Saez, E., & Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, *106*(1), 24–45. https://doi.org/10.1257/aer.20141362

Scheffler, S. (2003). What is egalitarianism? *Philosophy & Public Affairs*, *31*(1), 5–39. https://doi.org/10.1111/j.1088-4963.2003.00005.x

Schmidt, H. (2009). Just health responsibility. *Journal of Medical Ethics*, *35*(1), 21–26. https://doi.org/10.1136/jme.2008.024315

Schokkaert, E., & Devooght, K. (2003). Responsibility-sensitive fair compensation in different cultures. *Social Choice and Welfare*, *21*, 207–242. https://doi.org/10.1007/s00355-003-0257-3

Scott, J. T., Matland, R. E., Michelbach, P. A., & Bornstein, B. H. (2001). Just deserts: An experimental study of distributive justice norms. *American Journal of Political Science*, *45*(4), 749–767.

Segall, S. (2009). *Health, Luck and Justice*. Princeton University Press. https://doi.org/10.1515/9781400831715

Senhadji, A. S. (1997). Two common problems related to the use of the armington aggregator in computable general equilibrium models. *Applied Economics Letters*, *4*(1), 23–25. https://doi.org/10.1080/758521826

Temkin, L. S. (2003). Egalitarianism defended. *Ethics*, *113*(4), 764–782. https://doi.org/10.1086/373955

Thöni, C. (2015). A note on ces functions. *Journal of Behavioral and Experimental Economics*, *59*, 85–87. https://doi.org/10.1016/j.socec.2015.10.001

Ubel, P. A., & Loewenstein, G. (1996). Distributing scarce livers: The moral reasoning of the general public. *Social Science & Medicine*, *42*(7), 1049–1055. https://doi.org/10.1016/0277-9536(95)00216-2

Weinzierl, M. (2014). The promise of positive optimal taxation: Normative diversity and a role for equal sacrifice. *Journal of Public Economics*, *118*, 128–142. https://doi.org/10.1016/j.jpubeco.2014.06.012

Whitty, J., Lancsar, E., Rixon, K., Golenko, X., & Ratcliffe, J. (2014). A systematic review of stated preference studies reporting public preferences for healthcare priority setting. *The patient*, *7*, 365–386. https://doi.org/10.1007/s40271-014-0063-2

Wikler, D. (2004). Personal and social responsibility for health. In S. Anand, F. Peter, & A. Sen (Eds.), *Public Health, Ethics, and Equity* (pp. 109–134). Oxford University Press. https://doi.org/10.1111/j.1747-7093.2002.tb00396.x

Xu, X., Fiacco, L., Rostron, B., Homsi, G., Salazar, E., Levine, B., Ren, C., & Nonnemaker, J. (2021). Assessing quality-adjusted years of life lost associated with exclusive cigarette smoking and smokeless tobacco use. *Preventive Medicine*, *150*, 106707. https://doi.org/10.1016/j.ypmed.2021.106707

Yaari, M., & Bar-Hillel, M. (1984). On dividing justly. *Social Choice and Welfare*, *1*, 1–24. https://doi.org/10.1007/BF00297056

Yitzhaki, S. (1983). On an extension of the gini inequality index. *International Economic Review*, *24*(3), 617–628. https://doi.org/10.2307/2526814

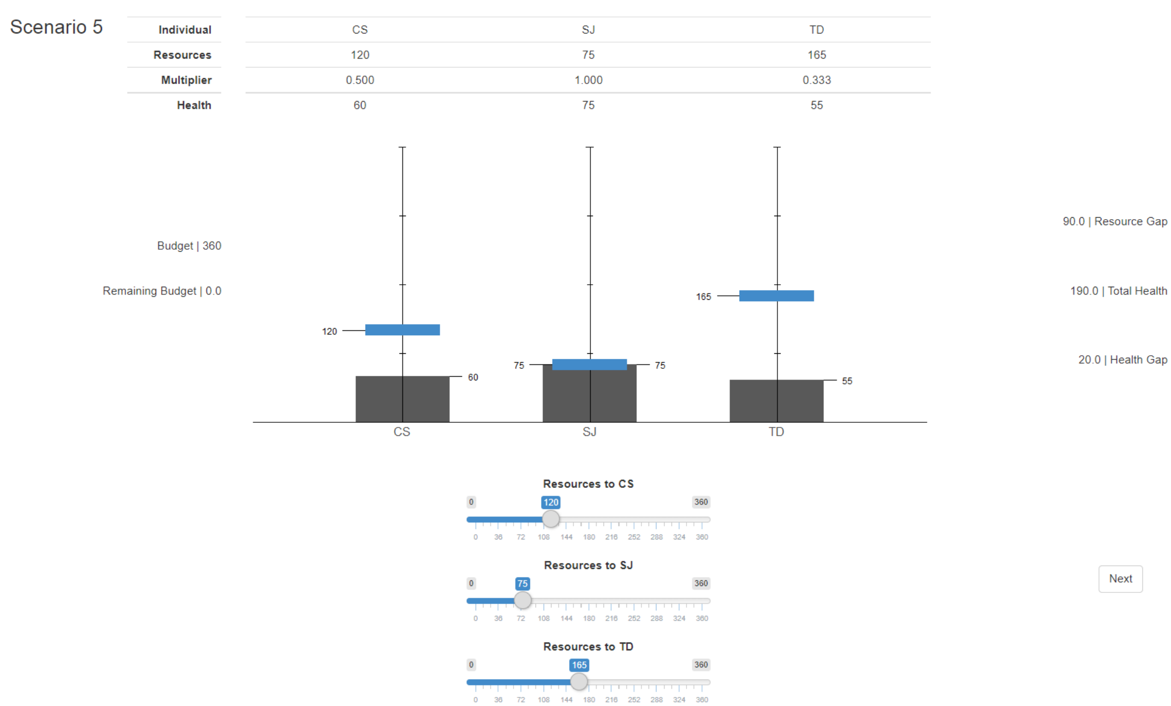
SUPPLEMENTARY MATERIAL

Appendix A Experiment Details

A.1 Interactive Interface

The experimental task is completed using an interactive online interface shown in Figure A1. The participant is asked to allocate *resources* to the three *individuals* using sliders at the bottom of the screen. Participants could also press and hold arrow keys to refine their allocations. The resources and the resulting *health* outcomes are shown by the blue and black bars, respectively. They are also shown numerically in the table at the top of the screen. The (remaining) *budget* is on the left, and summary measures are on the right. The *Resource Gap* is the largest absolute difference between resources allocated to two individuals. The *Health Gap* is the equivalent for health. *Total Health* is the sum of the health outcomes. To encourage deliberation, minimum timers were placed on each round.

Figure A1: Experiment Interface



A.2 Multipliers

Table A1 shows the multipliers and relative multipliers used across all treatments. There are 10 rounds within each treatment, the order is randomised between participants. The multipliers are orthogonal to the screen position and information about non-health characteristics of the individuals.

Table A1: Multipliers

	Absolute, p_i			Relative, $p_i / \sum p_i$		
Round	Left	Middle	Right	Left	Middle	Right
1	1	1	1	0.33	0.33	0.33
2	1	1	0.33	0.43	0.43	0.14
3	1	0.33	1	0.43	0.14	0.43
4	0.33	1	1	0.14	0.43	0.43
5	1	0.33	0.5	0.55	0.18	0.27
6	0.5	1	0.33	0.27	0.55	0.18
7	0.33	0.5	1	0.18	0.27	0.55
8	1	0.5	0.33	0.55	0.27	0.18
9	0.33	1	0.5	0.18	0.55	0.27
10	0.5	0.33	1	0.27	0.18	0.55
Mean	0.7	0.7	0.7	0.33	0.33	0.33

Note: Columns give *Multipliers* assigned to *Individuals* labelled with initials.

A.3 Experiment Overview

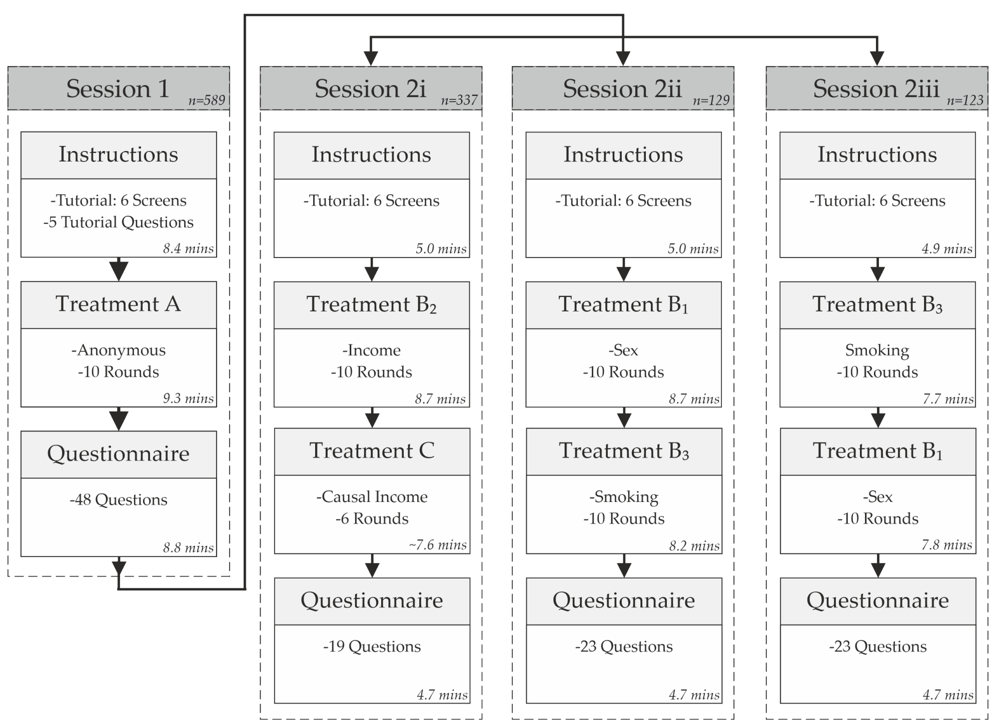
Figure A2 gives an overview of the full experiment and all treatments conducted in the first sample period between December 2021 and January 2022 (Sample 1) and the second sample period in April 2022 (Sample 2).

In each period, the experiment was run over two sessions that covered instructions, treatments, belief elicitation and two questionnaires. Participants were paid £3.50 for the first session and £5 for the second. The order participants went through the experiment is indicated by the arrows. The median times the participants completed each section are shown in minutes, in the bottom right corners.

All participants from both samples completed Session 1 (Treatment A). A between-subject design is then used for Session 2. Sample 1 participants completed Treatment B_2 in Session 2i. This sample also completed a Treatment C in Session 2i, which is reported in Robson et al. (2024). We do not use data from that treatment in this paper. A random sub-sample of Sample 2 completed Treatment B_1 and then Treatment B_3 in Session 2ii. Another random sub-sample completed Treatment B_3 and then Treatment B_1 in Session 2iii.

The numbers of participants used in our main analysis sample are shown in the bottom-right of the session headers. Median times (minutes) for completion of each section are shown in the bottom right corners.

Figure A2: Experiment Overview



A.4 Instructions and Tutorial Script

The text for the instructions, tutorial and tutorial questions below are shown to all participants in Session 1, on screen within the experiment. The instructions give an overview of the experiment to come. The six stages of a tutorial explain how to use the on-screen interface; each of the scripts are followed by an interactive on-screen tutorial. Finally, five tutorial questions are presented to check and reinforce understanding.

A.4.1 Instructions

Welcome. Thank you for taking part today.

Please Read These Instructions Carefully.

You will be asked to make decisions which determine the health of hypothetical individuals in society.

You will be given a "Budget" that you must divide between these individuals. The Budget is the total amount of "Resources" available to spend.

Resources determine "Health". Health is the number of years a person lives, adjusted for illness or disability. For example, consider someone who reached the age of 70 without any illness or disability, who then lived for a further 10 years with an illness which reduced their quality of life to half of what it was before. That person might be said to have lived for the equivalent of 75 years in full health. For shorthand, we refer to this as "Health".

Giving more Resources to an individual increases their Health. The impact of Resources on Health is determined by a number referred to as the "Multiplier". The higher the Multiplier, the higher the level of Health achieved from a given number of Resources.

On the screen, you will distribute Resources between three Individuals. You will do this a number of times. Each screen will show a different scenario. The choices you make on one screen will not affect the scenarios that follow.

There are no right or wrong answers. We are interested in the choices you make, whatever they are.

You will now go through a tutorial, which will explain how to use the computer interface and the exact nature of the experiment.

Please click Next to continue.

Tutorial 1

This tutorial will show you how to use the on-screen interface.

You will first get practice in giving Resources to only one individual, who is identified by initials (e.g. CS). Drag the horizontal slider at the bottom of the next screen to the right to give more Resources to the individual.

The amount of Resources you give is shown by the height of the blue bar in the chart above the slider. Resources are also shown by the number to the left of the blue bar and by the number in the table at the top of the page.

Once you have dragged the slider, you can use the left and right arrow keys to make precise changes to the amount. Press the arrow key for a change of 0.1 and hold the arrow key for changes of 1.

The Resources you give to an individual are taken from the Budget, which is shown on the left of the screen. As you increase the Resources, the Remaining Budget will decrease.

You must always use all of the Budget, so that the Remaining Budget is zero. When there is only one individual, this means dragging the slider all the way to the right. Later you will have to distribute the Budget between individuals.

Press Next to try out the slider. When you are done, allocate all of the Budget (100) and press Next.

Tutorial 2

The Resources you give to an individual determines their "Health". Health is the number of years a person lives, adjusted for illness or disability. Health is equal to the Resources multiplied by a number we call the "Multiplier".

The Multiplier is shown in the table at the top of the screen. When you give Resources by moving the slider, the resulting Health is shown by the height of the grey bar. The number to the right of this bar is the amount of Health, and this is also shown in the table at the top of the screen.

For this first individual the Multiplier is 1. So, if you give all of the Budget of 100 to the individual, their Health will be 100. They will live 100 years in full health.

Press next and see how Health changes as you adjust the Resources given to the individual. When you are done, allocate all of the Budget and press Next.

Tutorial 3

The Multipliers can vary from individual to individual. In the previous scenario, the Multiplier was 1. In the next scenario, it is 0.5.

Press Next and see how Health changes as you give more Resources to this individual. Notice that there is now a gap between the Resources given (blue bar) and the Health achieved (grey bar). If you give all the Budget of 100 to this individual, their Health will be 50. They will live 50 years in full health.

When you are done, allocate all of the Budget and press Next.

Tutorial 4

In each round of the experiment, there are three individuals. Individuals are identified by their initials (e.g. CS, SJ and TD) and change between rounds.

On the next screen, there are three sliders at the bottom of the screen that you can use to give Resources to each individual and so determine their Health.

The Resources and Health of each individual are shown by the blue and grey bars. The table at the top also shows the Resources, Multiplier and Health for each individual.

Now you must allocate the Budget across the three individuals. In doing so, you determine the Health of each one. In the example on the next screen, the Multipliers are the same for all three Individuals.

You must use all of the Budget, so that the Remaining Budget (on the left) equals zero.

Press Next and then give Resources to the three individuals. Remember: you can use the left and right arrow keys to make small changes to the amount of Resources. When you have used all of the Budget on the next screen, press Next.

Tutorial 5

The three individuals change from round to round.

On the previous screen, all three individuals had a Multiplier of 1. But the Multipliers can differ between individuals, as on the next screen.

Move the sliders to give Resources to the three individuals and notice how the Health achieved depends on the Multiplier of each individual. If you give the three individuals the same Resources, their Health will differ.

Take note of the size of the Budget, which can change from screen to screen.

If you are having difficulty seeing both the table and the graph on your screen, zoom out on your web browser by holding "Ctrl" and pressing "-". Hold "Ctrl" and press "+" to zoom in.

Press Next and then give Resources to the three individuals. When the Remaining Budget is zero, press Next.

Tutorial 6

The right of the screen shows further information.

"Resource Gap" is the gap between the largest and smallest amounts of Resources you give to the individuals.

"Total Health" is the total amount Health of the three individuals (e.g. Health to CS + Health to SJ + Health to TD).

"Health Gap" is the gap between the largest and smallest amounts of Health achieved by the individuals.

If you have used the whole Budget, you will not be able to move any slider to the right. If you want to give more Resources to one individual, you will need to give less to another individual first.

Press Next and then distribute Resources across the three individuals. When you have allocated all of the Budget, press Next.

Tutorial Questions

Following the tutorial, participants answered five questions to reinforce and check understanding. The questions are shown below, with the correct response in bold. After submitting answers participants were given feedback about the correct response for each question.

1. On each screen, you will give Resources to how many individuals? - Options: 2; **3**; 4; Not Sure.
2. You can make and adjust the Resources you give by (tick all that apply): **Clicking and Dragging the Allocation Sliders**; **Using the Arrow Keys**; Moving the Vertical Bar; Not Sure.
3. If you give 100 Resources to an individual with a Multiplier of 1, then the Health of that individual will be? - Options: 25; 50; **100**; Not Sure.
4. If you give 100 Resources to an Individual with a Multiplier of 0.5, then the Health of that individual will be? - Options: 25; **50**; 100; Not Sure.
5. Once you have finished giving the Resources, you proceed to the next screen by: - Options: Clicking Next; **Ensuring the Remaining Budget = 0, then Clicking Next**; Waiting; Not Sure.

In Session 2, a modified version of the above instructions and tutorial are shown. First, to remind participants of the experiment, and second, to highlight the additional information on the respective non-health characteristics of each individual. They are told that sex is an individual's "biological sex at birth", that income is an individual's "annual personal income (before tax) in pounds", and that smoking status refers to the number of cigarettes an individual smokes, on average, each day. It is pointed out that sex, income or smoking status, in the respective variant of Treatment B, is shown in the label for each individual.

Appendix B Experimental Data and Descriptives

B.1 Experimental Data

Table B1 summarises the experimental data from our analytical sample. Data includes experimental parameters and allocation decisions of 569 participants, from Sample 1 and 2, across all rounds of treatments A, B_1, B_2 and B_3.

Table B1: Summary of Experimental Data

Variable	**Notation**	**Definition**	**Mean**	**Range**	**Obs.**
Experimental					
Budget	m	m	260.86	[180-360]	41,460
Resources	y_i	y_i	86.95	[0-360]	41,460
Resource Share	$\tilde{y}_i$	y_i/m	0.333	[0-1]	41,460
Multiplier	p_i	p_i	0.7	[0.333-1]	41,460
Relative Multiplier	$\tilde{p}_i$	$p_i/\sum p_i$	0.333	[0.14-0.54]	41,460
Health	h_i	$y_i p_i$	55.21	[0-360]	41,460
Health Share	$\tilde{h}_i$	$h_i/\sum h_i$	0.333	[0-1]	41,460
Female	x_i^1	x_i^1	0.5	[0-1]	7,320
Income	x_i^2	x_i^2	£38.2k	[£5k-£100k]	9,750
Cigarettes	x_i^3	x_i^3	10.04	[0-30]	7,320

Note: Number of participants is 569.

B.2 Descriptives

Through the Prolific online platform, we aimed to recruit a sample that is representative of the UK adult population with respect to sex, age and ethnicity. We recruited Sample 1 and conducted treatments A and B_2 with this sample in December 2021 - January 2022. We recruited Sample 2 and conducted treatments A, B_1 and B_3 with this sample in April 2022.

To improve balance between Sample 1 and Sample 2 we construct and apply sample weights. We pool data from the two samples, estimate a logit model of a participant's (s) sample membership ($D_s = \mathbb{1}(s \in \text{Sample 2})$) and used it to predict the probability of a participant belonging to Sample 2: $\hat{\Lambda}(\mathbf{z}_s) = P(D_s = 1|\mathbf{Z} = \mathbf{z}_s)$, where $\hat{\Lambda}(\mathbf{z}_s)$ is the estimated cumulative density function and $\mathbf{z}_s$ includes *sex, age, ethnicity, employment status, post-graduate education* and *self-assessed health*. We then construct inverse-probability-weights, $\mathrm{w}_s = D_s + (1 - D_s)\frac{\hat{\Lambda}(\mathbf{z}_s)}{1-\hat{\Lambda}(\mathbf{z}_s)}$, and apply them in all analyses.

Table B2 shows unweighted and weighted means of characteristics for Sample 1 and Sample 2, alongside standardised differences. Before weighting, the most notable difference between the samples is that participants in Sample 2 are older. After weighting standardised differences are small across all variables.[17]

Table B3 shows unweighted and weighted means of characteristics of the 569 participants from Sample 1 and Sample 2 combined that are used in the main analysis. The right-hand column gives (weighted) means of some of these characteristics estimated from wave 10 (2018/19) of the UK Household Longitudinal Study (ISER, 2023), which is representative of the UK adult population. The variables included have comparable definitions in the two studies. After application of the weights, our sample remains representative of the UK adult population with respect to sex, but is younger and has a lower proportion of white individuals. Our sample also has fewer individuals who are married or born in the UK. The average household size and country of residence are representative, but participants generally have higher incomes, are more likely to be employed and less likely to be retired.

[17]Figure E3 additionally shows little difference between distributions of participant-level inequality aversion parameters, estimated from treatment A, further demonstrating good balance between the two samples.

Table B2: Characteristics of Sample 1 and 2: Unweighted and Weighted

	Unweighted			**Weighted**		
	Mean (S1)	Mean (S2)	Std. Diff.	Mean (S1)	Mean (S2)	Std. Diff.
Demographics						
Female	0.50	0.51	-0.02	0.50	0.51	-0.02
Age	40.16	44.71	-0.31	44.61	44.71	-0.01
Married	0.52	0.53	-0.03	0.57	0.53	0.07
Born in UK	0.78	0.84	-0.15	0.82	0.84	-0.06
White	0.80	0.81	-0.03	0.81	0.81	0.01
Household Size	2.72	2.65	0.05	2.64	2.65	-0.01
Country						
- England	0.83	0.88	-0.14	0.84	0.88	-0.11
- Wales	0.05	0.05	0.02	0.05	0.05	0.00
- Scotland	0.09	0.07	0.09	0.09	0.07	0.10
- N. Ireland	0.03	0.01	0.16	0.02	0.01	0.10
Income						
Income (£k)	29.35	27.67	0.07	30.68	27.67	0.11
Labour Market Status						
- Employed	0.62	0.68	-0.12	0.66	0.68	-0.03
- Unemployed	0.13	0.08	0.15	0.10	0.08	0.07
- Retired	0.08	0.11	-0.09	0.11	0.11	0.00
- Student	0.10	0.08	0.08	0.06	0.08	-0.07
- Other	0.06	0.05	0.05	0.06	0.05	0.06
Occupation						
- Manager/Professional	0.38	0.42	-0.08	0.38	0.42	-0.09
- Intermediate	0.31	0.34	-0.07	0.35	0.34	0.03
- Drivers/Labourers	0.02	0.02	-0.01	0.02	0.02	-0.00
- Never Worked	0.12	0.08	0.11	0.09	0.08	0.01
- Other	0.19	0.16	0.10	0.19	0.16	0.08
Highest Education						
- Postgraduate	0.24	0.19	0.11	0.20	0.19	0.02
- Undergraduate	0.36	0.40	-0.08	0.37	0.40	-0.05
- A-Level	0.25	0.25	0.01	0.25	0.25	0.01
- Secondary/Primary	0.15	0.16	-0.03	0.18	0.16	0.04
Health						
Self-Assessed Health	2.18	2.20	-0.03	2.21	2.20	0.01
Health: Likert 0-100	69.21	71.23	-0.10	69.27	71.23	-0.09
Cigarettes Smoked Per Day	1.14	0.87	0.07	1.34	0.87	0.12
Views						
Left-Right	3.93	3.91	0.01	4.11	3.91	0.09

Table B3: Characteristics of Analytical Sample and UK Representative Sample

				Mean		
	Obs	Min	Max	Unweighted	Weighted	UKHLS
Demographics						
Female	569	0	1	0.50	0.50	0.52
Age	569	18	76	42.11	44.65	51.08
Married	568	0	1	0.52	0.55	0.76
Born in UK	567	0	1	0.81	0.83	0.90
White	569	0	1	0.80	0.81	0.93
Household Size	567	0	10	2.69	2.64	2.74
Country						
- England	562	0	1	0.85	0.86	0.84
- Wales	562	0	1	0.05	0.05	0.05
- Scotland	562	0	1	0.08	0.08	0.08
- N. Ireland	562	0	1	0.02	0.02	0.03
Income						
Income (£k)	520	2	175	28.63	29.40	23.17
Labour Market Status						
- Employed	569	0	1	0.65	0.67	0.56
- Unemployed	569	0	1	0.11	0.09	0.04
- Retired	569	0	1	0.09	0.11	0.28
- Student	569	0	1	0.09	0.07	0.04
- Other	569	0	1	0.06	0.06	0.09
Occupation						
- Manager/Professional	558	0	1	0.40	0.40	.
- Intermediate	558	0	1	0.32	0.34	.
- Drivers/Labourers	558	0	1	0.02	0.02	.
- Never Worked	558	0	1	0.10	0.08	.
- Other	558	0	1	0.18	0.18	.
Highest Education						
- Postgraduate	569	0	1	0.22	0.20	.
- Undergraduate	569	0	1	0.37	0.38	.
- A-Level	569	0	1	0.25	0.25	.
- Secondary/Primary	569	0	1	0.16	0.17	.
Health						
Self-Assessed Health	569	1	5	2.19	2.20	.
Health: Likert 0-100	567	6	100	70.08	70.12	.
Cigarettes Smoked Per Day	564	0	25	1.02	1.14	.
Views						
Left-Right	552	0	10	3.92	4.02	.

Appendix C Estimation

We maximise the likelihood of the participants' resource allocations under the assumptions that each participant allocates optimally (given their preferences) on average and that deviations from the optima follow a particular distribution. Equation (7) gives the optimal allocation, y_i^*, to individual $i \in \{1, 2, 3\}$ in each round of the experiment, where the ω_i are $1/N$ for observations from Treatment A and are given by (4) and (5) for observations from each variant of Treatment B. When we allow for responsibility-sensitive preferences, the Treatment B equity weights are also defined by (6).

We assume that the observed resource shares in each round, $\tilde{y}_i = y_i/m$, are draws from the distribution of a random variable, $\tilde{Y}_i$, with an expected value equal to the optimal resource share, $E[\tilde{Y}_i] = \tilde{y}_i^*$, where $\tilde{y}_i^* = y_i^*/m$. We assume that the vector of observed resource shares ($\tilde{\boldsymbol{y}}$) is Dirichlet (1839) distributed, $f\left(\tilde{\boldsymbol{y}}; \boldsymbol{\alpha}\right) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{3} \tilde{y}_i^{\alpha_i - 1}$, where $B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^{3} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{3} \alpha_i\right)}$ and $\Gamma()$ is the gamma function (Robson, 2021).

Defining $\alpha_0 = \sum_{i=1}^{3} \alpha_i$ and using properties of the Dirichlet distribution,

$$E[\tilde{Y}_i] = \frac{\alpha_i}{\alpha_0} = \tilde{y}_i^* \,. \tag{C1}$$

We assume that the variance (Var) of the distribution from which the observed allocation shares are drawn is equal to the variance of the optimal allocation shares deflated by a precision parameter, $\sigma > 0$, that increases as noise in the observed allocations decreases, $Var(\tilde{Y}_i) = \frac{\tilde{y}_i^*(1-\tilde{y}_i^*)}{\sigma}$. Then, the distributional assumption gives,

$$Var(\tilde{Y}_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)} = \frac{\tilde{y}_i^*(1 - \tilde{y}_i^*)}{\sigma} \,, \tag{C2}$$

and, consequently,

$$\tilde{y}_i^*(\sigma - 1) = \alpha_i \qquad \forall_i \,. \tag{C3}$$

The log-likelihood contribution of participant s over all rounds r of treatment $t \in T = \{A, B_1, B_2, B_3\}$ is

$$LL_{ts} = \sum_{r=1}^{10} \log \left(\frac{\Gamma\left(\sum_{i=1}^{3} \alpha_{irts}\right)}{\prod_{i=1}^{3} \Gamma(\alpha_{irts})} \prod_{i=1}^{3} \tilde{y}_{irts}^{\alpha_{irts}-1} \right), \tag{C4}$$

where $\tilde{y}_{irts}$ is the resource share participant s allocates to individual i in round r of treatment t and α_{irts} is a function of the preference parameters that determine the respective optimal share and the precision parameter.

We estimate population averaged parameters by pooling the data over a) all rounds within each treatment, b) all treatments and c) all participants. The log-likelihood function is

$$LL = \sum_{s=1}^{S} \mathrm{w}_s \sum_{t \in T} \sum_{r=1}^{10} \log \left(\frac{\Gamma\left(\sum_{i=1}^{3} \alpha_{irts}\right)}{\prod_{i=1}^{3} \Gamma(\alpha_{irts})} \prod_{i=1}^{3} \tilde{y}_{irts}^{\alpha_{irts}-1} \right), \tag{C5}$$

where w_s is the sample weight of participant s (see Appendix B.2).

We maximise this likelihood with respect to the inequality aversion parameter ε, the weight parameters β_k with $k \in \{1, 2, 3\}$ specified in (4) and (5), as well as γ_k and ϕ_k with $k \in \{2, 3\}$ specified in (6) when allowing for responsibility-sensitivity, and the precision parameter σ. Pooled estimation of a common set of parameters provides estimates of the preferences of a representative social decision maker.

Appendix D Additional Non-Parametric Results

D.1 Health Maximisation vs Equalisation Trade-Off

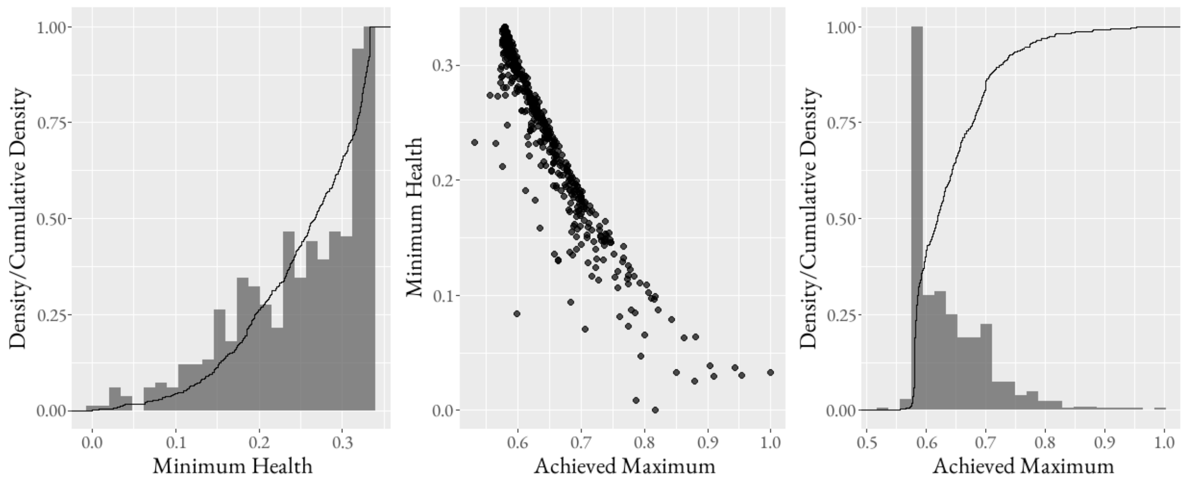
Figure D1 shows the distributions across participants of two non-parametric measures of the trade-off between maximisation of aggregate health and equalisation of health. Both measures are derived from Treatment A data.

The variable *minimum health* is the smallest health share a participant (s) gives to any individual (i) in each round (r) averaged over all 10 rounds: $\tilde{h}_s^{min} = \frac{1}{10}\sum_{r=1}^{10} \tilde{h}_{rs}^{min}$, where $\tilde{h}_{rs}^{min} = min\{\tilde{h}_{1rs}, \tilde{h}_{2rs}, \tilde{h}_{3rs}\}$. If $\tilde{h}_s^{min} = 1/3$, a participant always maximises the health of the worst-off. If $\tilde{h}_s^{min} = 0$, then they allocate resources such that at least one individual gets zero health in each round.

For each participant, the variable *achieved maximum* is aggregate health in a round as a proportion of the maximum aggregate that could have been achieved if all resources were allocated to the most productive individual(s), averaged over all rounds: $h_s^{max} = \frac{1}{10}\sum_{r=1}^{10} \frac{\sum_i^N h_{irs}}{p_{rs}^{max} \times m_{rs}}$, where $p_{rs}^{max} = max\{p_{1rs}, p_{2rs}, p_{3rs}\}$ and m_{rs} is the resource budget in round r for participant s. If $h_s^{max} = 1$, a participant always maximises aggregate health. A lower h_s^{max} indicates greater willingness to sacrifice maximisation of aggregate health for less inequality.

The left panel of Figure D1 shows the distribution of minimum health. There is a modal spike at a value of 1/3, where participants allocate health equally. A majority of participants has a value lower than 1/3, indicating that the worst-off individual is left with less than an equal share. Very few are prepared to give zero resources, and health, to the worst-off. Similarly, the right panel shows that very few participants maximise aggregate health ($h_s^{max} = 1$). There is a model spike of the achieved maximum at 0.58, which is the value consistent with the equal distribution of health.

Figure D1: Participant-Level Trade-Offs between Health Maximisation and Equalisation



Note: *Minimum health*, $\tilde{h}_s^{min} = \frac{1}{10}\sum_{r=1}^{10} \tilde{h}_{rs}^{min}$, where $\tilde{h}_{rs}^{min} = min\{\tilde{h}_{1rs}, \tilde{h}_{2rs}, \tilde{h}_{3rs}\}$ in round r. *Achieved maximum*, $h_s^{max} = \frac{1}{10}\sum_{r=1}^{10} \frac{\sum_i^N h_{irs}}{p_{rs}^{max} \times m_{rs}}$, where $p_{rs}^{max} = max\{p_{1rs}, p_{2rs}, p_{3rs}\}$ and m_{rs} is the resource budget in round r for participant s. Both variables are derived from the allocations made by each participant in Treatment A. Number of participants is 569, and number of observations 17,070.

The scatter plot in the middle panel shows a strong negative relationship between the two measures. This indicates that participants are primarily trading off between the maximisation and equalisation of health in Treatment A that gives no information on the individuals, other than their respective productivities.

The trade-offs that participants make between health maximisation and equalisation are, on average, consistent with their self-reported views about the importance of reducing health inequality versus maximising aggregate health. We asked each participant to report a number on a scale of 0 to 10, where 0 represents the view that "Reducing inequalities in health is more important than improving total population health" and 10 is the view that "Improving total population health is more important than reducing inequalities in health". Table D1 shows estimates from regressing the *minimum health* and the *achieved health* that result from the participants' allocations on their self-reports about their support for maximising total health versus reducing inequalities. Participants who express relatively stronger support for the importance of reducing health inequalities have larger values of the minimum health —

they do indeed allocate to reduce health inequality more — and lower values of the achieved maximum — their allocations leave them further from maximising aggregate health. What participants do in the experiment is consistent with what they say they believe.

Table D1: Revealed and stated preferences for health maximisation versus equalisation

	(1) Minimum health	(2) Achieved maximum
Support maximisation vs equalisation	-0.0740***	0.0595***
	(0.012)	(0.010)
Constant	0.2903***	0.6062***
	(0.007)	(0.005)
Participants	567	567
Observations	17,010	17,010

Note: Estimates from bivariate regressions of *Minimum health* and *Achieved maximum* on reported support for health maximisation versus equalisation. *Minimum health*, $\tilde{h}_s^{min} = \frac{1}{10}\sum_{r=1}^{10}\tilde{h}_{rs}^{min}$, where $\tilde{h}_{rs}^{min} = min\{\tilde{h}_{1rs}, \tilde{h}_{2rs}, \tilde{h}_{3rs}\}$. *Achieved maximum*, $h_s^{max} = \frac{1}{10}\sum_{r=1}^{10}\frac{\sum_i^N h_{irs}}{p_{rs}^{max} \times m_{rs}}$, where $p_{rs}^{max} = max\{p_{1rs}, p_{2rs}, p_{3rs}\}$ and m_{rs} is the resource budget in round r for participant s. Both variables are derived from the allocations made by each participant in Treatment A. *Support maximisation vs equalisation* is a number on a scale of 0 to 10, where 0 represents the view that "Reducing inequalities in health is more important than improving total population health" and 10 is the view that "Improving total population health is more important than reducing inequalities in health". Two participants did not self-report their support of maximisation versus equalisation. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D2 column (1) shows estimates from the regression of the resource share participant s gives to individual i in round r of Treatment A, $\tilde{y}_{irs} = y_{irs}/\sum_i y_{irs}$, on the relative multiplier, $\tilde{p}_{irs} = p_{irs}/\sum_i p_{irs}$, with allowance for participant level random effects. The negative sign on the multiplier indicates that, on average, participants give fewer resources to more productive individuals.

Table D2 column (2) shows estimates from the respective regression with health shares, $\tilde{h}_{rit} = h_{rit}/\sum_i h_{rit}$, as the dependent variable. The positive sign on the relative multiplier indicates that, on average, participants do not reduce the resources allocated to more productive individuals to an extent sufficient to offset their productivity advantage. These individuals still get a larger share of health.

Table D2: Regressions of Resource and Health Shares on Productivity

	(1) Resource Share	(2) Health Share
Relative Multiplier	-0.5126***	0.4587***
	(0.025)	(0.022)
Constant	0.3333***	0.3333***
	(0.000)	(0.000)
Participants	569	569
Observations	17070	17070

Note: Regressions of *Resource share*, $\tilde{y}_{irs}$, and *Health share*, $\tilde{h}_{irs}$ on productivity measured by the *Relative multiplier*, $\tilde{p}_{irs}$, where i is the individual in round r of Treatment A and s is the participant. The relative multiplier is re-centered around 1/3. Participant random effects model. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.2 Prioritisation by Non-Health Characteristics

Table D3 shows estimates from regressions of health shares, $\tilde{h}_{irs} = h_{irs}/\sum_i h_{irs}$, on indicators of sex, income and smoking categories from Treatment B_1, B_2 and B_3, respectively. All three regressions allow for participant-level random effects.

On average, the health share to a female is almost 1 percentage point (pp) larger than that to a male. The health share decreases monotonically with increasing income and smoking. On average, the richest individual gets a 4.3 pp smaller health share than the poorest individual. A heavy smoker gets a health share that is 16.9 pp smaller than a non-smoker, on average.

Participants' perceptions of how they allocated resources in relation to non-health characteristics were consistent with their actual allocations. We asked each participant: "In general, did you try to distribute more Health to X or Y, where X and Y were "males" and "females", "richer individuals" and "poorer individuals", and "smokers" and "non-smokers" in Treatments B_1, B_2 and B_3, respectively. Answer options were: "More to X", "More to Y", "No Difference", or "Don't Know". Using data from Treatment B_1, we regress health shares participants give to females in each round on their responses to the above question.

Table D3: Regressions of Health Shares on Non-Health Characteristics

	Health share		
	Treament B_1	Treatment B_2	Treatment B_3
	(1)	(2)	(3)
Sex (ref. Male)			
- Female	0.0085***		
	(0.002)		
Income (ref. Poorest)			
- Middle		-0.0190***	
		(0.004)	
- Richest		-0.0430***	
		(0.007)	
Smoking (ref. Non-smoker)			
- Light			-0.0945***
			(0.008)
- Moderate			-0.1357***
			(0.009)
- Heavy			-0.1692***
			(0.011)
Constant	0.3290***	0.3540***	0.4222***
	(0.001)	(0.004)	(0.006)
Participants	244	325	244
Observations	7320	9750	7320

Note: Regressions of health shares, $\tilde{h}_{irs} = h_{irs}/\sum_i h_{irs}$, participant s gives to individual i in round r of Treatment B_1, B_2 or B_3 on indicators of sex, income and smoking, respectively. Participant random effects models. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

"No Difference" is the reference category. We estimate analogous bivariate regressions of health shares to poorer individuals (Treatment B_2) and non-smokers (Treatment B_3) on the respective responses. The estimates from these regressions, shown in Table D4, confirm that participants' perceptions of their priorities are consistent with their actual allocations in relation to non-health characteristics. Those who reported distributing more health to females, poorer individuals and non-smokers did indeed allocate larger health shares to those individuals.

Table D5 shows regressions of health shares, $\tilde{h}_{irs}$, to females in Treatment B_1 (column (1), the poorest individual in each round of Treatment B_2 (column (2)) and non-smokers

Table D4: Actual and Perceived Prioritisation of Females, Poorest and Non-Smokers

	Health share to		
	Female (=Y)	Poorest (=Y)	Non-smoker (=Y)
	(1)	(2)	(3)
Perceive giving more to (ref. No difference)			
- X	-0.0014	-0.0698***	0.0150
	(0.011)	(0.026)	(0.036)
- Y	0.0258***	0.0466***	0.1140***
	(0.006)	(0.007)	(0.008)
- Don't Know	0.0108***	0.0091	-0.0014
	(0.002)	(0.020)	(0.020)
Constant	0.3348***	0.3369***	0.3398***
	(0.001)	(0.002)	(0.003)
Participants	243	325	242
Observations	3660	3250	2420

Note: Regressions of health shares, $\tilde{h}_{jrs} = h_{irs}/\sum_i h_{irs}$, where i is an individual in round r, j is a specific individual defined as female, poorest and a non-smoker in columns (1), (2) and (3) corresponding to Treatments B_1, B_2 or B_3, respectively, and s is a participant. The independent variables are responses to the question "In general, did you try to distribute more Health to X or Y?", where X and Y were "males" and "females", "richer individuals" and "poorer individuals", and "smokers" and "non-smokers" in Treatments B_1, B_2 and B_3, respectively. Answer options were: "More to X", "More to Y", "No Difference", or "Don't Know". "No Difference" is the reference category. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in Treatment B_3 (column (3)) on participants' characteristics. The allocations participants make to individuals distinguished by sex, income and smoking are not associated with their own reported sex, income and smoking behavior.

D.3 Responsibility Beliefs

Table D6 shows the association between participants' responsibility beliefs and the average health share allocated to the poorest individual in each round of Treatment B_2 and non-smokers in each round of Treatment B_3, each averaged over all rounds of the respective treatment. Participants who gave higher health shares to the poorest tend to believe that

Table D5: Regressions of Health Shares on Participants' Characteristics

	Health share to		
	Female (1)	Poorest (2)	Non-Smoker (3)
Participants' characteristics			
- Female	0.0004	-0.0059	0.0112
	(0.002)	(0.007)	(0.013)
- Income (£'000)	-0.0000	0.0001	0.0005
	(0.000)	(0.000)	(0.000)
- Cigarettes per day	0.0003	-0.0007	0.0003
	(0.000)	(0.001)	(0.002)
Constant	0.3378***	0.3564***	0.4049***
	(0.003)	(0.008)	(0.015)
Participants	218	297	218
Observations	3300	2970	2180

Note: Regressions of health shares, $\tilde{h}_{irs}$, to females in Treatment B_1 (column (1)), the poorest individual in each round of Treatment B_2 (column (2)) and non-smokers in Treatment B_3 (column (3)) on participants' reported sex, income and cigarettes smoked per day. Participant random effects models. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

people are less responsible for the income they earn. Those who gave higher health shares to non-smokers tend to believe that people are more responsible for their smoking behavior.

Table D6: Regressions of Responsibility Beliefs on Health Shares

	Responsible for	
	Income (1)	Smoking (2)
Health Share: Poor	-0.3856*	
	(0.233)	
Health Share: Non-Smokers		0.4400***
		(0.095)
Constant	0.7192***	0.7048***
	(0.081)	(0.047)
Participants = Observations	325	244

Note: Participant-level linear regressions of elicited beliefs about responsibility for income earned (1) and smoking behavior (2) on health shares to poorest individual in each round of Treatment B_2 (1) and non-smokers in each round of Treatment B_3 (2), in each case averaged over all rounds of the respective treatment for each participant. Beliefs re-scaled to 0-1, with a larger value indicating greater responsibility. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D7 shows estimates from regressing participants' responsibility beliefs on their own characteristics. Participants with higher (reported) incomes tend to believe that people are more responsible for the income they earn. There is no significant association between beliefs about responsibility for smoking and own smoking behavior.

Table D7: Regressions of Responsibility Beliefs on Participants' Characteristics

	Responsible for	
	Income (1)	Smoking (2)
Female	-0.0734	0.2850
	(0.266)	(0.248)
Income (£'000)	0.0110***	0.0074
	(0.003)	(0.006)
Cigarettes Smoked Per Day	0.0420	-0.0194
	(0.030)	(0.037)
Constant	5.4651***	8.6090***
	(0.219)	(0.306)
Participants = Observations	297	218

Note: Participant-level linear regressions of elicited beliefs about responsibility for income earned (1) and smoking behavior (2) on participants' sex (female=1), income (£'000) and smoking behavior (number of cigarettes smoked per day). Beliefs re-scaled to 0-1, with a larger value indicating greater responsibility. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D8 shows estimates from regressing responsibility beliefs on participant-level estimates of preferences (see Appendix E.6). These regressions, unlike those of Table D6, separate the association of responsibility beliefs with estimates of the weight parameter for the respective non-health characteristic, β_k, from their association with estimated inequality aversion, ε. Participants who give a higher priority to the health of poorer individuals (larger β_2) tend to believe that people are less responsible for the income they earn. Participants who give a higher priority to the health of non-smokers (larger β_3) tend to believe that people are more responsible for their smoking behavior. Responsibility beliefs are not significantly associated with inequality aversion.

Table D8: Regressions of Responsibility Beliefs on Participant-Level Preferences

	Responsible for	
	Income (1)	Smoking (2)
Log Inequality Aversion, $\ln(\hat{\varepsilon})$	-0.1124	-0.0859
	(0.076)	(0.083)
Weight Parameter, $\hat{\beta}_k$	-1.0257**	0.6841***
	(0.512)	(0.179)
Constant	7.1554***	8.0822***
	(0.564)	(0.296)
Participants = Observations	324	243

Note: Participant-level linear regressions of elicited beliefs about responsibility for income earned (1) and smoking behavior (2) on participant-level estimates of log inequality aversion, $ln(\hat{\varepsilon})$, and weight parameters, $\hat{\beta}_k$. The parameter estimates used in columns (1) and (2) are estimated separately using data from Treatment B_2 and B_3, respectively. Beliefs re-scaled to 0-1, with a larger value indicating greater responsibility. Robust standard errors in parentheses. p-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix E Additional Parametric Results

E.1 Robustness to not Pooling over Treatments

The top four rows of Table E1 give estimates of the inequality aversion parameter and the equity weight parameters obtained from analysis of data from each treatment separately. Data are pooled over all rounds of a treatment for each participant and then over participants, but data are not pooled over treatments. Each log-likelihood is as (C5) without summation over treatments $t \in T$, giving treatment specific estimates of ε and, for each variant of Treatment B, β_k. The bottom row reproduces the estimates from the top panel of Table 1 obtained from pooling over treatments.

Estimates of the inequality aversion parameter are consistent across treatments, although the point estimate obtained using Treatment B_2 data is slightly larger in comparison with the estimate obtained from data pooled over all treatments. Each equity weight parameter estimate obtained using only data from the respective Treatment B is highly consistent with the respective estimate obtained when data are pooled over all treatments.

Table E1: Preference Parameters Estimated from each Treatment Separately

	Inequality	Equity Weight			
	Aversion	Sex	Income	Smoking	Obs.
	ε	β_1	β_2	β_3	
Treatment A (Anonymous)	1.420				17,070
	[1.35 , 1.49]				
Treatment B_1 (Sex)	1.403	1.021			7,320
	[1.31 , 1.53]	[0.995 , 1.05]			
Treatment B_2 (Income)	1.568		1.080		9,750
	[1.46 , 1.69]		[1.06 , 1.10]		
Treatment B_3 (Smoking)	1.429			1.356	7,320
	[1.33 , 1.54]			[1.33 , 1.39]	
All Treatments (Pooled)	1.450	1.021	1.078	1.383	41,460
	[1.41 , 1.50]	[0.99 , 1.05]	[1.06 , 1.10]	[1.34 , 1.42]	

Note: Maximum likelihood estimates in first four rows obtained using data from the treatment specified in the row heading. Within each the respective treatment, data are pooled over all rounds and across all participants. Each log-likelihood is as (C5) without summation over treatments. For each variant of Treatment B, the weights are specified as in (4) and (5). The bottom row reproduces the estimates from the top panel of Table 1, which are obtained from pooling data over all treatments and maximising log-likelihood (C5). In brackets are 90% bootstrap confidence intervals (percentile method, 1000 replications). Number of participants is 569 for Treatment A and Pooled, 244 for Treatments B_1 and B_3, and 325 for Treatment B_2.

E.2 Robustness to Sample Selection and Weights

In the main analysis, we drop 66 participants who gave incorrect answers to at least 3 out of 5 questions designed to test comprehension after a tutorial. We drop another 20 participants with incomplete item response on the covariates required for the sample weights. Sample weights are applied to improve balance between Samples 1 and 2 (see Appendix B.2). Here, we test robustness of estimates to not excluding participants with poor comprehension of the task or incomplete item response, and not applying the sample weights.

Table E2 gives our main estimates in row (1). Row (2) shows estimates obtained with the same sample but without the application of sample weights. All estimates are robust to this change. Row (3) gives estimates without application of weights and without excluding those with low comprehension or incomplete item response. The estimates of the preference parameters remain robust. The estimate of the precision parameter decreases, indicating lower precision when using data that includes participants with low comprehension of the task.

Table E2: Robustness to Application of Sample Weights and Restrictions

	Sample Weights	Sample Restricted	Inequality Aversion ε	Equity Weight Sex β_1	Equity Weight Income β_2	Equity Weight Smoking β_3	Precision σ	Obs.
(1)	YES	YES	1.450	1.021	1.078	1.383	7.846	41,460
			[1.41, 1.50]	[0.99, 1.05]	[1.06, 1.10]	[1.34, 1.42]	[7.49, 8.26]	
(2)	NO	YES	1.426	1.021	1.068	1.379	7.613	41,460
			[1.38, 1.47]	[0.99, 1.05]	[1.05, 1.09]	[1.35, 1.41]	[7.25, 8.00]	
(3)	NO	NO	1.420	1.011	1.068	1.368	7.009	47,520
			[1.38, 1.46]	[0.99, 1.04]	[1.05, 1.09]	[1.34, 1.40]	[6.72, 7.33]	

Note: Row (1) gives the main estimates (Table 1, top panel) obtained with application of sample weights and sample restricted by excluding participants answering at least 3/5 comprehension questions incorrectly. In row (2), sample weights are not applied. In row (3), the sample is not restricted. Estimation method is as described in notes to Table 1. In brackets are 90% bootstrap confidence intervals (percentile method, 1000 replications). Number of participants is 569 in rows (1) and (2), and 655 in row (3).

E.3 Learning and Fatigue Effects

Here, we assess whether estimates are robust to the order in which treatments were completed and whether repetition of the tasks over multiple rounds affected estimates, possibly through learning or fatigue.

All participants completed Treatment A before Treatment B. With Sample 2, we randomised the order of Treatment B_1 (sex) and Treatment B_3 (smoking). If there was learning in the course of completing one treatment, then the order may affect the estimates.

The estimates shown in the top panel of Table E3 are obtained using all data (from all participants) from treatments A and B_2. The estimates in row (1a) additionally use data from treatments B_1 and B_3 obtained from participants who completed B_1 before B_3. The estimates in row (1b) use treatments B_1 and B_3 data from participants who completed B_3 before B_1. Estimates of the preference parameters are very similar irrespective of the order in which B_1 and B_3 were completed. Precision is a little higher when B_1 was completed first.

Table E3: Checks for Learning and Fatigue Effects

	Inequality	Equity Weight				
	Aversion ε	Sex β_1	Income β_2	Smoking β_3	Precision σ	Obs.
Treatment order:						
(1a) B_1-B_3	1.462	1.015	1.080	1.375	8.456	30,660
	[1.41 , 1.52]	[0.98 , 1.06]	[1.06 , 1.10]	[1.33 , 1.42]	[7.98 , 9.05]	
(1b) B_3-B_1	1.443	1.027	1.078	1.395	7.528	30,300
	[1.39 , 1.50]	[0.99 , 1.07]	[1.06 , 1.10]	[1.34 , 1.45]	[7.13 , 7.99]	
Data from:						
(2a) First Rounds	1.328	1.145	1.110	1.497	8.006	3,414
	[1.21 , 1.48]	[0.999 , 1.31]	[1.05 , 1.18]	[1.32 , 1.69]	[7.02 , 9.56]	
(2b) Last Rounds	1.442	1.095	1.114	1.391	7.263	3,414
	[1.30 , 1.62]	[0.96 , 1.26]	[1.04 , 1.19]	[1.27 , 1.53]	[6.20 , 8.84]	

Note: Top panel estimates obtained from all data from treatments A and B_2, plus: row (1a), data from treatments B_1 and B_3 from participants who completed B_1 before B_3; and, row (1b), data from treatments B_1 and B_3 from participants who completed B_3 before B_1. In the bottom panel, row (2a) gives estimates obtained using only data from the first round of each session. Row (2b) gives estimates using only data from the last round of each session. Otherwise, estimation is as described in the notes to Table 1. In brackets are 90% bootstrap confidence intervals (percentile method, 1000 replications). Number of participants is 569.

To assess whether learning or fatigue affected participants' choices over rounds within each session, we compare estimates obtained using data only from the first round of each session — row (2a) in the bottom panel of Table E3 — with estimates using data only from the last round of each session — row (2b) in the bottom panel of Table E3.[18] There is little difference between preference parameter estimates. The largest difference is in the point estimate of β_1, but neither estimate is significantly different from 1. Precision does appear

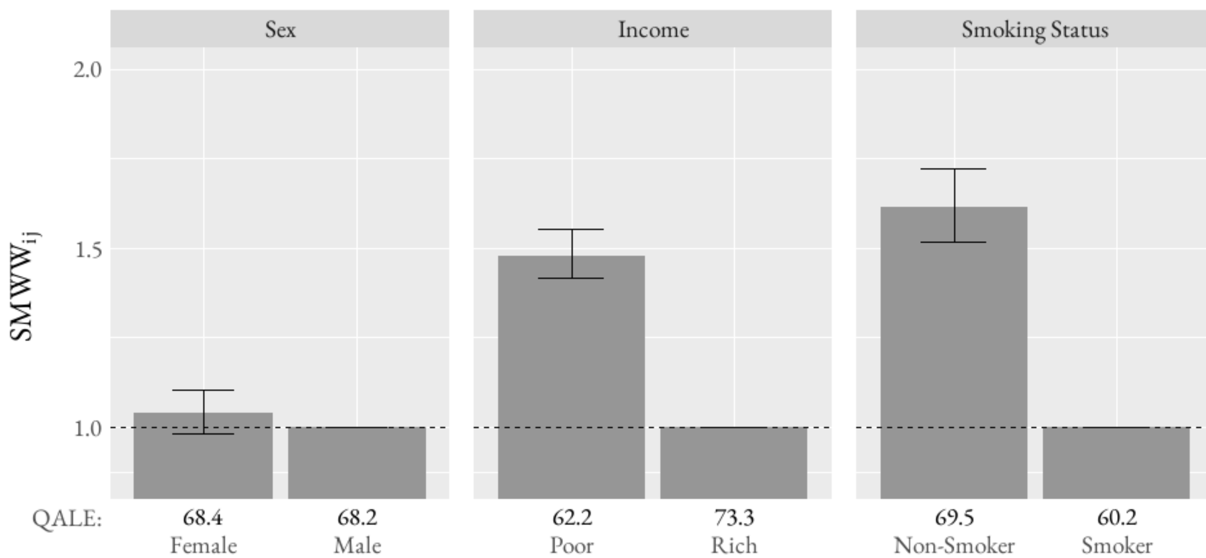
[18] In both samples, participants completed Treatment A in the first session. In Sample 1, participants completed Treatment B_2 in the second session. In Sample 2, participants completed Treatments B_1 and B_3, with a random order of these two treatments.

to drop somewhat when using data from the last rounds, which would be consistent with fatigue dominating over learning. But the confidence intervals overlap substantially.

E.4 Relative Social Marginal Welfare Weights within each Characteristic

Figure 3 shows relative social marginal welfare weights (SMWW) between the health of individuals differing in multiple characteristics. Figure E1 shows relative SMWW between individuals who differ in only one dimension (at a time).

Figure E1: Relative Social Marginal Welfare Weights within each Characteristic



Note: Estimates of (3). The weights are given by (4) and (5). The reference categories (j) for sex, income and smoking are male, rich and smokers, respectively. Parameter estimates are given in the top panel of Table 1. Health (h_i) is Quality Adjusted Life Expectancy (QALE), is shown below x-axis for each subgroup and is obtained from external estimates (see Appendix F). Interval lines show 90% bootstrap confidence intervals (percentile method, 1000 replications).

E.5 Goodness-of-Fit and Precision

Table E4 gives estimates of the precision parameter (σ) and goodness-of-fit measures for three models. Model (1) imposes constant weights ($\omega_i = 1/N \, \forall_i$), model (2) allows weights to vary over non-health characteristics but without sensitivity to responsibility beliefs — our main estimates in the top panel of Table 1 — and model (3) allows for responsibility-sensitive equity weights — our main estimates in the bottom panel of Table 1. The mean squared errors and log-likelihoods show that data fit improves on going from model (1) to (2) to (3) and Akaike Information Criteria confirm that this holds even after penalising the addition of parameters. Precision is slightly highest for model (2) but only marginally and confidence intervals overlap.

Table E4: Goodness-of-Fit and Precision Parameters

	Precision		Goodness-of-Fit			
Model	σ	[90% CI]	Mean Squared Error	Log-Likelihood	AIC	Obs.
(1)	7.436	[7.10 , 7.79]	0.0420	13933.2	-27864.5	41,460
(2)	7.846	[7.49 , 8.26]	0.0394	14757.4	-29504.7	41,460
(3)	7.832	[7.09 , 8.24]	0.0392	14814.7	-29615.3	41,460

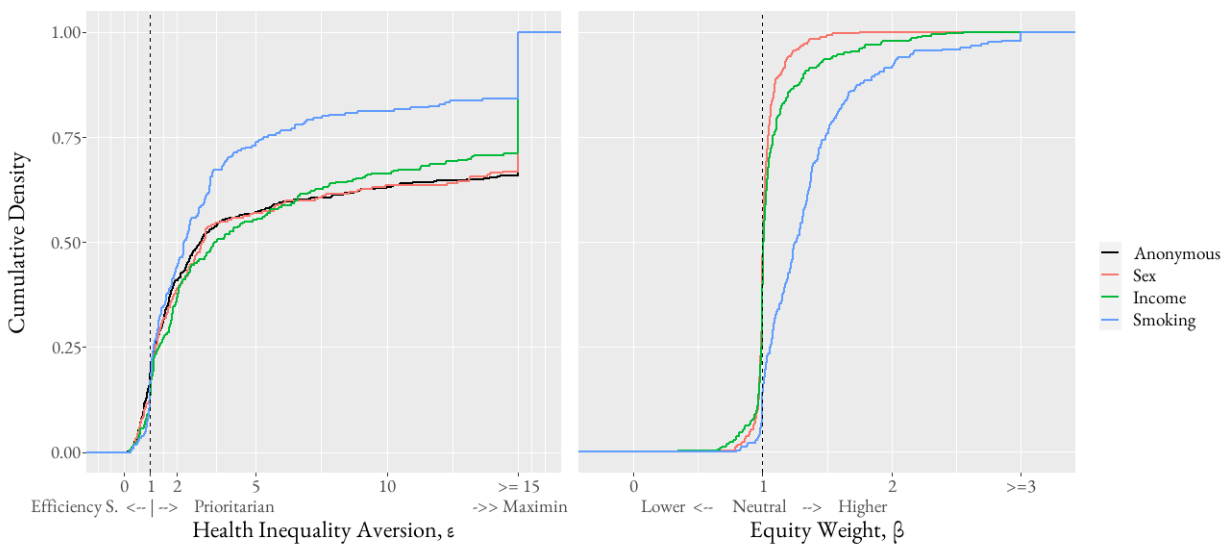
Note: Parameter estimates and statistics from maximum likelihood estimation explained in Appendix C. Model (1) imposes constant weights, $\omega_i = 1/N \, \forall_i$. In Model (2), weights are given by (4) and (5). In Model (3), weights are given by these two equations, plus (6). Preference parameter estimates of Models (2) and (3) given in the top and bottom panels of Table 1, respectively. Mean squared error is, $MSE = \frac{1}{n} \sum_s \sum_t \sum_r^{10} \sum_i^3 (\tilde{y}_{irts} - \tilde{y}^*_{irts})^2$, where $\tilde{y}_{irts}$ is the share of resources participant s allocates to individual i in round r of treatment t and $\tilde{y}^*_{irts}$ is the optimal resource share. The log-likelihood (LL) is the value of equation (C5) in Appendix C evaluated at the estimates of the respective model. Akaike Information Criterion is, $AIC = 2d - 2LL$, where d is the number of parameters of the respective model. Number of participants is 569.

E.6 Participant-Specific Preferences

Figure E2 shows the distribution of estimated participant-level preferences. The left panel shows ε, the health inequality aversion parameter, from the anonymous and information treatments. The right panel shows β, the equity weights, from information treatments.

In the anonymous treatment the median estimate for ε is 2.91, indicating a substantial degree of health inequality aversion. There is, however, extensive heterogeneity between participants: 14.2% can be classified as *Efficiency Seeking* ($0 < \hat{\varepsilon} \leq 0.9$), 8.44% as *Cobb-Douglas* ($0.9 < \hat{\varepsilon} < 1.1$), 43.23% as *Prioritarian* ($1.1 \leq \hat{\varepsilon} < 15$) and 34.09% as *Maximin* ($\hat{\varepsilon} \geq 15$). Estimates are similar for sex and income treatments, but we generally find fewer participants who have high levels of inequality aversion for smoking treatments.

Figure E2: Distribution of Health Inequality Aversion and Equity Weights

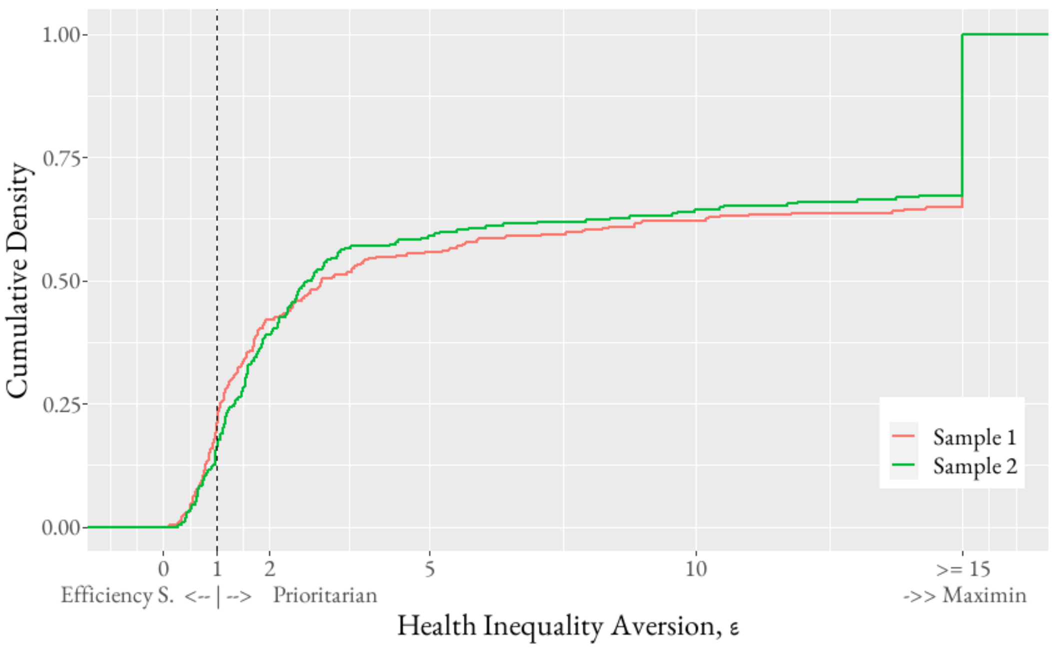


Note: Number of Observations and Participants is 17,070 and 569, respectively, for the anonymous treatment; 7,320 and 244 for the sex and smoking treatments, and 9,750 and 325 for the income treatment.

In the right panel, for smoking we see positive weights for almost all participants, this means that priority is given to those who smoke fewer cigarettes. For sex and income, there is more heterogeneity. Part of the sample are pro-female or pro-poor ($\beta > 1$), part are neutral ($\beta = 1$) and a smaller number are pro-male and pro-rich ($\beta < 1$).

Figure E3 shows two distributions of participant-specific estimates of the health inequality aversion parameter—one for Sample 1 (Treatment B_2 - income) and the other for Sample 2 (Treatments B_1 & B_3 - sex and smoking). The fact that there is very little difference between the distributions further demonstrates the good balance between the two samples.

Figure E3: Distributions of Inequality Aversion Parameter obtained from Sample 1 and 2



Note: Sample 1 is given Treatment A (Anonymous) and Treatment B_1 (Income). Sample 2 is given Treatment A and Treatments B_2 (Sex) and B_3 (Smoking). The figure shows the distribution of estimates of the health inequality aversion parameter, ε, obtained from Treatment A observations from the respective sample: respective number of Observations and Participants is 9,750 and 325 for Sample 1, and 7,320 and 244 for Sample 2.

Appendix F Quality Adjusted Life Expectancy

To calculate relative social marginal welfare weights (SMWW) for subgroups defined by sex, income and smoking, we combine our estimates of equity weight and health inequality aversion parameters with external estimates of Quality Adjusted Life Expectancy (QALE).

Love-Koh et al. (2023) estimate QALE at birth (in 2017-18) for sub-populations defined by sex and quintile groups of an Index of Multiple Deprivation (IMD) — a composite of income, employment, education and housing — for small geographical areas of England. For male and female QALE, we use that study's respective estimates for all England. We use the study's sex-specific QALE estimates for the first and fifth IMD quintile groups to approximate QALE for the groups we refer to as poor and rich. These are the QALE estimates we use to calculate the SMWW by sex and income groups shown in the bottom panel of Figure 4 and the left and middle panels of Figure E1.

Estimates of QALE by smoking status for the United Kingdom (UK), or even England, are not available. Xu et al. (2021) estimate QALE at birth by sex and smoking status in the United States. We combine these estimates with official estimates of smoking prevalence in England by sex and IMD quintile groups (Office of National Statistics, 2023) to estimate QALE for smokers and for non-smokers (by sex and IMD quintile). In doing this, we ensure that the sub-population QALE obtained from the weighted averages of the estimated QALE of smokers and non-smokers are consistent with the respective sub-population estimates in Love-Koh et al. (2023). We use these estimates of QALE by smoking status in the calculation of the SMWW for smokers and non-smokers shown in the bottom panel of Figure 4 and the right panel of Figure E1). We use the QALE estimates by sex, income (top and bottom IMD quintile group) and smoking in the calculations of the multidimensional SMWW given in Figure 3.

These QALE estimates are limited in three respects. First, QALE estimates for the UK are not available, and so we rely on estimates from England, where most of our sample is located, and the United States. Second, we proxied QALE by income with QALE by an index of small-area multiple deprivation (only partly determined by income). Third, while the estimates capture variation in the sex difference in QALE by income (deprivation) and by smoking, they do not allow the difference in QALE between smokers and non-smokers to vary by income, and vice versa.

Notwithstanding these limitations, the external QALE estimates, combined with our preference parameter estimates, allow approximation of welfare weights for interesting sub-populations. Applications of the weights in evaluations of health policies could test sensitivity to the QALE estimates.