Journal of Health Economics 90 (2023) 102773

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhe

Reducing socioeconomic health inequalities? A questionnaire study of majorization and invariance conditions☆

Kirsten I.M. Rohde [b,c,d], Tom Van Ourti [a,b,c,e,*], Amar Soebhag [b,c]

[a] Erasmus School of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
[b] Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
[c] Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
[d] Erasmus Research Institute of Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
[e] Erasmus Centre for Health Economics Rotterdam, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

ARTICLE INFO

Keywords:
Health inequality
Concentration index
Preference conditions
Lab experiment

ABSTRACT

We study the appeal of basic preference conditions that underpin health inequality indices, including the widely used concentration index. We did a lab experiment in which 349 respondents had to choose repeatedly between two policies that generated a distribution of income and health among five groups in society. We found stronger support for preference conditions that focus on inequality in the marginal distribution of health (and income) than for preference conditions that favor reduced correlation between both dimensions. Respondents' choices were more in line with the principle of income related health transfers when policies did not affect the ranking of groups in terms of health. Respondents also expressed more concern about the correlation between income and health when health was expressed as a shortfall rather than an attainment. Support for the preference conditions was unaffected when all groups in society experienced the same absolute or relative health change.

1. Introduction

Reducing inequalities in health is an important challenge for public policy. In the US, for example, inequalities in life expectancy have been increasing during this century (Chetty et al., 2016). These inequalities were more stable in Europe (Bosworth, 2018) but have intensified during the COVID-19 pandemic throughout the West (Williamson et al., 2020; Decoster et al., 2021; Schwandt et al., 2022). Estimated health inequality trends may depend on the way health inequalities are measured (Wagstaff and Van Doorslaer, 2004). A simple approach measures inequalities in health and considers the full extent of all health differences (Le Grand, 1987; Gakidou et al., 2000; van Raalte and Caswell, 2013). The more popular approach in health economics ignores health differences unrelated to income, and focuses on the distribution of health across a measure of socioeconomic status such as income (Wagstaff et al., 1991; Van Doorslaer et al., 1997). Other features of health inequality indices — such as whether these are sensitive to the bounded nature of health, whether they are sensitive to absolute or relative health differences, or how much focus is placed on the poor versus the rich (Erreygers, 2009a; Erreygers and Van Ourti, 2011; Wagstaff, 2011; Kjellsson and Gerdtham, 2013; Khaled

☆ Erasmus Research Institute of Management, Netherlands and the Erasmus Initiative Smarter Choices for Better Health provided financial support. The internal review board of the Erasmus Research Institute of Management granted approval to conduct this experiment (IRB application 2018-04 Rohde). Part of the work was undertaken while Tom Van Ourti was a visiting professor at the Milken Institute of Public Health of the George Washington University. The authors would like to thank Owen O'Donnell, Eddy Van Doorslaer and seminar participants at several universities and conferences for useful comments and suggestions. We have no conflicts of interest to disclose, and all errors are our own.

* Corresponding author.
E-mail addresses: rohde@ese.eur.nl (K.I.M. Rohde), vanourti@ese.eur.nl (T. Van Ourti), soebhag@ese.eur.nl (A. Soebhag).

https://doi.org/10.1016/j.jhealeco.2023.102773
Received 24 October 2022; Received in revised form 15 April 2023; Accepted 24 May 2023
Available online 13 June 2023

et al., 2018) – might also impact estimated health inequality trends. The choice of health inequality index, and thus the choice of monitoring framework, may depend on its appeal to people. This paper assesses the appeal of different basic preference conditions that underpin different health inequality indices.

We developed a questionnaire that asks respondents to choose repeatedly between two policies as if they were giving advice to policy makers of a hypothetical country. Each policy generated a distribution of income and health among five groups in society. Each question was designed to measure support for one of the preference conditions underlying different health inequality measures. These questions can be used to assess which preference conditions are most consistent with our respondents' ideas about good public policy in the domain of health and income (Amiel and Cowell, 1992; Gaertner and Schokkaert, 2012).

We consider two types of preference conditions: four inequality and three invariance conditions. The inequality conditions concern extensions of the unidimensional Pigou–Dalton transfer principle, which states that a mean-preserving reduction in the spread of an attribute reduces inequality. The principle of income-related health transfers underlies popular measures of income-related health inequalities such as the concentration index (and its variants) and the slope and relative indices of inequality (Wagstaff et al., 1991; Mackenbach and Kunst, 1997; Harper and Lynch, 2006; O'Donnell et al., 2008, 2016). This principle states that a mean preserving transfer of health from a socioeconomically better- to worse-off person reduces inequality (Bleichrodt and Van Doorslaer, 2006). We also consider the more stringent downward-positional-transfer-sensitivity condition which imposes a stronger preference for income-related health transfers when these transfers occur at the bottom versus the top of the income distribution (Wagstaff, 2002; Makdissi and Yazbeck, 2014; Khaled et al., 2018). Both inequality conditions combine features of uniform and correlation increasing majorization which have been proposed in the literature on multidimensional inequality measurement. Uniform majorization imposes that inequality in well-being is reduced when the same mean preserving reduction in spread is applied to health and income simultaneously (Kolm, 1977; Tsui, 1995). The principle of correlation increasing majorization is in line with the intuition that inequality rises when the correlation between health and income increases (Atkinson and Bourguignon, 1982; Tsui, 1999). The three invariance conditions describe scenarios under which orderings based on the inequality conditions should remain unchanged. Unit invariance and translation invariance require that changes in the average levels of health and income should not affect inequality orderings provided everyone faces respectively the same relative or absolute change (Tsui, 1995; Gajdos and Weymark, 2005). The mirror condition states that inequality orderings should not depend on representing health as a gain (e.g. health level minus minimum health) or a loss (e.g. maximum health minus health level) (Erreygers, 2009a,b; Lambert and Zheng, 2011; Erreygers and Van Ourti, 2011).

We administered a questionnaire to 349 students and found that support for uniform majorization (UM) is stronger than support for the principle of income related health transfers (PIRHT), which in turn is stronger than support for correlation increasing majorization (CIM). Translation and unit invariance were strongly supported, suggesting that support for CIM, PIRHT, and UM does not depend on average income or health. Support for CIM and PIRHT was stronger when health was expressed in terms of shortfalls than in terms of attainments, in particular for transfers at the bottom of the income distribution. Apart from political orientation, support for these principles did not depend on background characteristics of respondents. Moreover, we did not find any evidence that transfers of health or income at the bottom of the distribution get stronger support than transfers at the top of the distribution.

This paper adds to the empirical literature on preference conditions underlying health inequality indices (a.o. Bleichrodt et al., 2012; Tarroux, 2015). One of our contributions lies in showing that support for PIRHT and CIM is higher when presenting societal groups in terms of their income rank (from poor to rich) rather than in terms of their level of health (from unhealthy to healthy). We also empirically confirm – as argued by Bleichrodt and Van Doorslaer (2006) – that PIRHT is more acceptable when it does not affect the ranking of groups in terms of health. Hence, PIRHT is arguably more acceptable when applied to a society where health and income are strongly correlated, making it less likely that a given health transfer will change the ranking in terms of health. Since the workhorse tool of health inequality measurement – rank-dependent inequality measures (Wagstaff et al., 1991; Erreygers and Van Ourti, 2011; O'Donnell et al., 2016) – satisfies PIRHT, this also means that these measures receive more normative support when health and income correlate strongly. We also show that support for CIM and PIRHT is crucially correlated with respondents claiming to only focus on the affected groups when choosing between different policies. This is not unexpected as it is in line with the implied separability assumptions underlying PIRHT, CIM and UM. However, about 35 percent of our respondents consider the non-affected groups as well, substantially lowering support for the preference conditions underlying health inequality measurement tools.

Our paper also relates to the literature that estimates aversion to income-related health inequality using parametric assumptions about the underlying social welfare function (Hurley et al., 2020; Hardardottir et al., 2021). Under this approach, conditions such as PIRHT and CIM are assumed to hold, and the extent of inequality aversion is estimated. It allows for an assessment of the downward positional transfer sensitivity condition that we consider in this paper, and additionally imposes precise functional forms for the difference in social weights attached to the poor versus the rich by the social planner (Yitzhaki, 1983; Wagstaff, 2002). We found no evidence indicating that support for CIM or PIRHT would differ when health transfers take place at the bottom versus the top of the income distribution. This is in line with Hurley et al. (2020) and Hardardottir et al. (2021) who found social weights to be on average close to those imposed by the standard concentration index.

The remainder of the paper proceeds as follows. Section 2 provides more background of the inequality and invariance conditions studied in this paper, and Section 3 describes the questionnaire and the experimental approach. Section 4 presents our findings; and Section 5 discusses the main implications of these findings.

2. Background

We consider a social planner who has to choose between allocations of income and health. Society consists of n groups of individuals. Every individual in group i has income level y_i and health level h_i. An *allocation* (y, h) is a $2 \times n$ matrix, with the ith column being (y_i, h_i).

We assume that the social planner has preferences $\succcurlyeq$ over allocations that can be represented by a *social welfare function* W, such that for all allocations (y, h) and (x, l) we have

$$(y, h) \succcurlyeq (x, l) \iff W(y, h) \geq W(x, l).$$

Strict preference is denoted by $\succ$, and indifference by $\sim$. Preference conditions on $\succcurlyeq$ generate specific functional forms of W, which translate into specific inequality measures. We will consider two types of preference conditions, which we will label *inequality* and *invariance* conditions.

2.1. Inequality conditions

2.1.1. Correlation increasing majorization (CIM)

In this paper we are interested in whether and to what extent decision makers care about the distribution of health across income groups. One potential intuition is that inequalities in health and income reinforce each other when income and health are positively correlated, while inequalities are 'canceled out' under negative correlation. A preference condition that captures this idea, is correlation increasing majorization (Tsui, 1999; Gajdos and Weymark, 2005).

Allocation (y, h) is obtained from (x, l) through a *correlation increasing transfer* if there are two groups i, j such that $y_i = \min\{x_i, x_j\}$, $h_i = \min\{l_i, l_j\}$, $y_j = \max\{x_i, x_j\}$, and $h_j = \max\{l_i, l_j\}$, $(y_m, h_m) = (x_m, l_m)$ for all $m \notin \{i, j\}$, and $(y, h) \neq (x, l)$ and (y, h) is not simply a permutation of the groups in (x, l). Thus, (y, h) is obtained from (x, l) through a *correlation increasing transfer* if for two groups i, j we rearrange their income and health in such a way, that group i always gets the worse income and health. *Correlation increasing majorization (CIM)* holds if $(x, l) \succ (y, h)$ whenever (y, h) is obtained from (x, l) through a sequence of correlation increasing transfers.

2.1.2. Principle of income-related health transfers (PIRHT)

CIM is implied by the principle of income related health transfers, the main condition underlying the concentration index that is widely applied for the measurement of income-related health inequality (Wagstaff et al., 1991; Kakwani et al., 1997; Bleichrodt and Van Doorslaer, 2006). The *principle of income-related health transfers (PIRHT)* holds if a transfer of health from a richer to a poorer group improves social welfare, as long as it does not change the ranking of groups in terms of income. PIRHT agrees with CIM that correlation increasing transfers are undesirable: both PIRHT and CIM imply that transfers that increase the correlation between income and health, without affecting the marginal distributions of income and health, are socially not preferred. CIM only concerns transfers that change the correlation between income and health without affecting the marginal distributions of income and health. The transfers that are considered by CIM thereby always involve a re-ranking of groups in terms of health or income, depending on which dimension we consider as given. PIRHT, however, also considers transfers that do not imply such a re-ranking. It, for instance, requires that any change in the spread of health improves social welfare, as long as health is transferred from the rich to the poor. Thereby, PIRHT is a stronger condition than CIM.

An alternative way to think about health inequalities, is to think of them as arising from mean-preserving increases in the spreads of health and income simultaneously. Uniform majorization is a preference condition that captures this intuition.

2.1.3. Uniform majorization (UM)

Allocation (y, h) is obtained from (x, l) through a uniform majorization if (y, h) is obtained by multiplying (x, l) by a bistochastic matrix and it is not a simple permutation of the groups. *Uniform majorization (UM)* holds if $(y, h) \succ (x, l)$ whenever (y, h) is a uniform majorization of (x, l) (Tsui, 1999; Gajdos and Weymark, 2005).

2.1.4. Downward positional transfer sensitivity (DPTS)

UM and CIM are independent conditions (Tsui, 1999). We will systematically assess the robustness of the support for UM and CIM. First of all, for both UM and CIM we will test whether the given transfers in income and health are perceived to be more valuable when taking place at the bottom than at the top of the income distribution. We thereby also test for *downward positional transfer sensitivity (DPTS)*, a condition that requires the concern about income-related health transfers in PIRHT to be stronger when these transfers occur at the bottom rather than the top of the income distribution (Wagstaff, 2002; Makdissi and Yazbeck, 2014; Khaled et al., 2018). More precisely, the combination of two income-related health transfers that are identical in terms of both the amount of health transferred and the number of steps on the income ladder separating the transferring groups, is inequality reducing when one transfer is from a poor to a poorer group and the other from a rich to a richer one. PIRHT imposes that the former transfer is inequality reducing, while the latter increases inequality. DPTS then imposes that the former gets more weight than the second because the transfers occur more towards the bottom of the income distribution. While PIRHT captures a first-order aversion to inequality, DPTS captures a second-order aversion to inequality.

We will also test several invariance conditions that play an important role in inequality measurements.

2.2. Invariance conditions

Average levels of income and health may matter for inequality measurement. We consider two preference conditions that impose stability on attitudes towards inequalities when average levels of health and/or income change: translation invariance and unit invariance. Translation invariance is required for absolute inequality indices, while unit invariance is required for relative inequality indices (Erreygers and Van Ourti, 2011).

We will also consider another important invariance property, which relates to the framing of health in terms of attainments or shortfalls and is often referred to as the mirror property (Erreygers, 2009a,b; Lambert and Zheng, 2011). It requires inequality orderings to be independent of framing health in terms of gains or losses, which is a natural requirement for many health measures that have a finite upper and lower bound.

2.2.1. Translation invariance

Translation invariance refers to the idea that the level of inequality is unaffected by a change in the average level of health or income that leaves the absolute differences between groups in that dimension unchanged. It is implied by absolute inequality indices. By $(y+\alpha, h)$ we denote the distribution that results from (y, h) by adding the constant α to all income levels. By $(y, h+\beta)$ we denote the distribution that results from (y, h) by adding the constant β to all health levels. *Income translation invariance* holds if for all α we have that $(y, h) \succcurlyeq (x, l)$ implies $(y+\alpha, h) \succcurlyeq (x+\alpha, l)$. *Health translation invariance* holds if for all β we have that $(y, h) \succcurlyeq (x, l)$ implies $(y, h+\beta) \succcurlyeq (x, l+\beta)$. *Translation invariance* holds if both income and health translation invariance hold. Translation invariance is equivalent to the translatability condition of Tsui (1995). As Gajdos and Weymark (2005) discussed, this version of translation invariance, which they call strong translatability, allows for a different constant to be added to different dimensions. This is appropriate in our setting where the two dimensions are measured on a different scale. Yet, in a setting where the dimensions would be measured on the same scale, for instance, tax expenditures and expenditures on health insurance, it may be more appropriate to only allow for the same constant to be added to both dimensions simultaneously.

2.2.2. Unit invariance

Unit invariance, which is implied by relative inequality indices, refers to the idea that the level of inequality is unaffected by a change in the average level of health or income that leaves the relative differences between groups within the dimension unchanged. By $(\lambda y, h)$ we denote the distribution that results from (y, h) by multiplying all income levels by the constant λ. By $(y, \kappa h)$ we denote the distribution that results from (y, h) by multiplying all health levels by the constant κ. *Income unit invariance* holds if for all $\lambda > 0$ we have that $(y, h) \succcurlyeq (x, l)$ implies $(\lambda y, h) \succcurlyeq (\lambda x, l)$. *Health unit invariance* holds if for all $\kappa > 0$ we have that $(y, h) \succcurlyeq (x, l)$ implies $(y, \kappa h) \succcurlyeq (x, \kappa l)$. *Unit invariance* holds if both income and health unit invariance hold.

Our definition of unit invariance allows for a different multiplication factor for health and income. In a setting where the two dimensions are measured on the same scale, it may be more appropriate to require the factor to be the same for both dimensions (Tsui, 1995; Gajdos and Weymark, 2005).

2.2.3. Mirror invariance

Health is a good example of a dimension that can equally naturally be described in terms of attainments (gains) or shortfalls (losses). Consider, for instance, an individual who suffers from backpain for 20 days per year on average. We can say that in a year this person has 20 days with backpain, but we can as well say that this person has 345 days without backpain. Both framings represent the same objective level of health. It is therefore normatively appealing to require preferences of the social planner to be independent of whether health is framed in terms of attainments or shortfalls. This condition, which we will call *mirror invariance*, is a desirable condition for any dimension that is represented by a bounded variable (Erreygers, 2009a,b; Lambert and Zheng, 2011; Bosmans, 2016). Mirror invariance is implied by the stronger mirror property of Erreygers and Van Ourti (2011).

Evidence from behavioral economics and psychology suggests that mirror invariance will not hold in practice. Preferences have been shown to depend on framing in terms of gains or losses. The reflection effect, for instance, is the finding that people have opposite risk attitudes for gains than for losses (Kahneman and Tversky, 1979). Thus, people who are risk averse for gains, may well be risk seeking for losses. This reflection effect suggests that attitudes to inequalities, of both social planners and the people affected by their decisions, are likely to depend on whether health is framed in terms of attainments or shortfalls. Thus, while framing should not matter normatively, in practice it may matter for decision makers and for people in society who are affected by these decisions. Our questionnaire study will shed light on the extent to which framing may matter in the context of health inequality measurement.

Finally, mirror invariance and unit invariance are generally incompatible (Erreygers and Van Ourti, 2011; Lambert and Zheng, 2011). Thus, it is theoretically impossible for the conditions of mirror invariance and unit invariance to hold simultaneously for *all possible* allocations. If we find support for both mirror and unit invariance in our questionnaire this will only hold for the distributions used and cannot hold in general.

Table 1
Example of two distributions.

Group	1	2	3	4	5
Income (Euro)	6000	10 000	13 000	18 000	32 000
Days without back pain	315	320	325	330	335
Group	1	2	3	4	5
Income (Euro)	6000	10 000	13 000	18 000	32 000
Days without back pain	315	324	325	326	335

3. Questionnaire

We designed a questionnaire to measure the support for the aforementioned inequality and invariance conditions in the context of health and income. Respondents were asked to imagine a small hypothetical country, called Alfaland, which consists of five groups. Every group consists of an equal number of individuals. Within each group, all individuals have the same income and health. All respondents were asked 38 questions where they had to choose between two distributions of income and health in Alfaland. We told them that these distributions are the result of two different policy scenarios.[1] Respondents were asked to choose which of the two policies they would support if they were asked to advise the government of Alfaland. We chose to let respondents reveal their attitudes to inequalities by giving them choices between policies rather than asking them explicitly which policy they consider more unequal. This is consistent with the normative approach that derives inequality indices from preferences of social planners (Bleichrodt et al., 2012; Hurley et al., 2020; Hardardottir et al., 2021).

We gave our respondents the role of advisor to the government of Alfaland to let them take the perspective of an impartial policy maker (Amiel et al., 2009; Almås et al., 2020; Konow, 2009; Tarroux, 2015). We did not implement the alternative of asking them to decide in which society they preferred an imaginary grandchild to be born (Johansson-Stenman et al., 2002), as it can induce respondents to consider the situation of only their own income group (Bleichrodt et al., 2012). In this study we wanted respondents to consider the entire distributions of income and health from the perspective of an impartial social policy maker, and therefore chose the Alfaland formulation.

We considered income in terms of yearly net income per person (in euro) in 2019. Health was expressed in terms of days without backpain per year. A person with backpain was described to

- have moderate problems in walking about,
- have moderate problems doing usual activities (e.g., work, study, housework, family or leisure activities), and
- have moderate pain or discomfort

(The EuroQoL Group, 1990). We chose backpain, because this condition is relatively easy to imagine and need not be chronic (Attema et al., 2012). Moreover, in practice suffering from backpain often does not impact one's ability to work and related income. We also wanted to make sure that both income and health would be expressed according to a similar time frame. Therefore, we chose to express them in terms of yearly amounts and numbers of days per year.

A previous and related study (Bleichrodt et al., 2012) conceptualized income as a flow (monthly income) and health as a stock (life expectancy) to study attitudes towards socioeconomic health inequalities. A potential drawback of mixing flows with stocks is that it may induce respondents to compute lifetime income by multiplying health and income. This would confound our measurements, because it would be unclear which dimensions respondents consider in their decisions: only health and income as they are presented to them, or also lifetime income. Redistributions in monthly income and life expectancy that change the dispersion in each of these dimensions without affecting means (such as for UM), often will lead to a change in mean lifetime income. Respondents might not favor such redistributions because of the associated efficiency loss (Bosmans et al., 2015). Similarly, reducing the correlation between income and health may weaken the support for CIM as it lowers total lifetime wealth. We wanted to avoid this potential problem as much as possible and therefore chose to express both monetary (yearly amounts) and health outcomes (number of days) in terms of flows. Moreover, the results of an extensive pilot study (Online Appendix Section 2) indicate that the presentation in terms of stocks or flows does not affect support for CIM, PIRHT, and UM significantly in our setting.

Distributions were presented in tables.[2] One question would, for example, ask to choose between the two tables in Table 1. As is common in experimental studies, we did not allow for indifference. In case of indifference, respondents were asked to make a choice, in order to discourage them from reporting indifference merely to reduce cognitive effort (Nowlis et al., 2002). In every question, the distributions were presented in random order, so that if all subjects were indifferent and would, for instance, choose the first distribution presented to them, we would expect 50% of them to choose one distribution, and the others choosing the other one. Every choice was presented on a separate screen, the two distributions were presented one above the other, and the order of questions was randomized.

[1] Section 1 of the Online Appendix describes the actual wording used in the questionnaire. The question wording discouraged participants to think about the exact mechanisms and processes underlying the policies that lead to the final distributions of health and income (Schokkaert and Devooght, 2003; Le Clainche et al., 2015).

[2] In Bleichrodt et al. (2012) we displayed distributions both in tables and in graphs, but as most subjects focused on the tables, we chose not to display distributions in graphs in this paper.

Table 2
Questions to test for CIM, PIRHT, and UM.

	Policy A					Policy B				
	1	2	3	4	5	1	2	3	4	5
CIM_b	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 320	10 000 315	13 000 325	18 000 330	32 000 335
CIM_m	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 315	10 000 330	13 000 325	18 000 320	32 000 335
CIM_t	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 315	10 000 320	13 000 325	18 000 335	32 000 330
$PIRHT_b$	6000 316	10 000 319	13 000 325	18 000 330	32 000 335	6000 318	10 000 317	13 000 325	18 000 330	32 000 335
$PIRHT_{b,nr}$	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 317	10 000 318	13 000 325	18 000 330	32 000 335
$PIRHT_m$	6000 315	10 000 322	13 000 325	18 000 328	32 000 335	6000 315	10 000 326	13 000 325	18 000 324	32 000 335
$PIRHT_{m,nr}$	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 315	10 000 324	13 000 325	18 000 326	32 000 335
$PIRHT_t$	6000 315	10 000 320	13 000 325	18 000 331	32 000 334	6000 315	10 000 320	13 000 325	18 000 333	32 000 332
$PIRHT_{t,nr}$	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 315	10 000 320	13 000 325	18 000 332	32 000 333
UM_b	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	7600 317	8400 318	13 000 325	18 000 330	32 000 335
UM_m	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 315	13 200 324	13 000 325	14 800 326	32 000 335
UM_t	6000 315	10 000 320	13 000 325	18 000 330	32 000 335	6000 315	10 000 320	13 000 325	23 600 332	26 400 333

3.1. Stimuli

The questionnaire consisted of 38 choices between distributions and several background questions at the end. We constructed 8 different versions of the questionnaire. This section first presents the questions for one of the versions, and subsequently explains how the other versions of the questionnaire were derived.

CIM

We asked 3 questions to test for CIM, as summarized in Table 2. Each question involved a choice between a baseline distribution and another distribution derived from the baseline by a correlation increasing majorization. The income levels in the baseline distribution (Policy A) were based on actual household incomes in the Netherlands adjusted for household composition, rounded to numbers that are relatively easy to process cognitively. The average number of days without backpain was equal to the average number reported in the Netherlands. Moreover, as health and income are negatively correlated, a higher income implied more days without backpain in our baseline distribution.

CIM_b considers a correlation increasing majorization between groups 1 and 2, who are at the bottom (b) of the income distribution. Similarly, CIM_m and CIM_t consider a correlation increasing majorization between groups 2 and 4 and between groups 4 and 5, who are at the middle (m) and the top (t) of the income distribution. CIM predicts that respondents will choose Policy B in each of these questions. Comparing responses in CIM_b and CIM_t allows us to assess whether a similarly sized transfer of health is valued more when it takes place at the bottom than at the top of the income or health distribution.

PIRHT

From the three CIM questions we derived six questions to test for PIRHT, as summarized in Table 2. $PIRHT_b$ is derived from CIM_b by considering a health transfer from income group 2 to 1, which satisfies two properties. First, the health transfer is smaller than the health transfer implied by CIM_b. Second, the health transfer implies that the ranking of the two groups in terms of health is reversed, similarly as in CIM_b: in Policy A group 2 is in better health than group 1 and in Policy B group 2 is in worse health than group 1. $PIRHT_m$ and $PIRHT_t$ are derived from CIM_m and CIM_t in a similar way, with the sizes of the transfers proportional to those of the CIM questions. Comparison of health transfers at the top and bottom of the income distribution allows shedding light on support for DPTS with stronger support for $PIRHT_b$ than $PIRHT_t$ being in line with DPTS.

$PIRHT_b$, $PIRHT_m$, and $PIRHT_t$ each consider health transfers that change the ranking of groups in terms of health. From these questions we derived $PIRHT_{b,nr}$, $PIRHT_{m,nr}$, and $PIRHT_{t,nr}$ which satisfy two conditions: (1) their health transfers do not affect the ranking in terms of health (*nr* denotes no reranking), and (2) their health transfers are equal to the ones in $PIRHT_b$, $PIRHT_m$, and $PIRHT_t$ respectively. PIRHT predicts that for all six questions Policy B is preferred.

Table 3
Summary of questions to test for translation and unit invariance.

(a) Translation invariance			
$TI_{CIM,i,-}$	$= CIM_m -$ €3160	$TI_{UM,i,-}$	$= UM_m -$ €3160
$TI_{CIM,i,+}$	$= CIM_m +$ €632	$TI_{UM,i,+}$	$= UM_m +$ €632
$TI_{CIM,h,-}$	$= CIM_m -$ 65 days	$TI_{UM,h,-}$	$= UM_m -$ 65 days
$TI_{CIM,h,+}$	$= CIM_m +$ 13 days	$TI_{UM,h,+}$	$= UM_m +$ 13 days
$TI_{CIM,h,gl}$	$= CIM_m -$ 220 days	$TI_{UM,h,gl}$	$= UM_m -$ 220 days
(b) Unit invariance			
$UI_{CIM,i,-}$	$= CIM_m \times$ 0.8 income	$UI_{UM,i,-}$	$= UM_m \times$ 0.8 income
$UI_{CIM,i,+}$	$= CIM_m \times$ 1.04 income	$UI_{UM,i,+}$	$= UM_m \times$ 1.04 income
$UI_{CIM,h,-}$	$= CIM_m \times$ 0.8 days without BP	$UI_{UM,h,-}$	$= UM_m \times$ 0.8 days without BP
$UI_{CIM,h,-,sf}$	$= CIM_m \times$ 0.8 days with BP	$UI_{UM,h,-,sf}$	$= UM_m \times$ 0.8 days with BP
$UI_{CIM,h,+}$	$= CIM_m \times$ 1.04 days without BP		
$UI_{CIM,h,+,sf}$	$= CIM_m \times$ 1.04 days with BP		
$UI_{CIM,h,gl}$	$= CIM_m \times$ 105/325 days without BP	$UI_{UM,h,gl}$	$= UM_m \times$ 105/325 days without BP
$UI_{CIM,h,gl,sf}$	$= CIM_m \times$ 105/325 days with BP	$UI_{UM,h,gl,sf}$	$= UM_m \times$ 105/325 days with BP

UM

We asked three questions to test for UM, as summarized in Table 2. Question UM_b involves the same health transfer as Question $PIRHT_{b,nr}$, and income is transferred according to the same convex combination as health. Similarly, questions UM_m and UM_t involve the same health transfers as questions $PIRHT_{m,nr}$ and $PIRHT_{t,nr}$ with income transferred according to the same convex combination as health. Thus, the UM questions can fairly be compared to the PIRHT questions, and we can compare support of transfers in the bottom and top of the income distribution. In questions UM_b and UM_t the transfers do not affect the income rankings of the groups. In question UM_m, however, the income ranking of groups 2 and 3 is reversed. UM predicts that Policy B is preferred in all three questions.

Translation invariance

We asked 10 questions to test for translation invariance, as summarized in Table 3a and shown in Table A.6 in Appendix. We tested for translation invariance applied to both CIM and UM. In particular, the reference questions are CIM_m and UM_m. Hence, we study support for correlation increasing and uniform majorizations when they involve transfers from a group above to a group below the median and analyze how this support changes when average income or health changes. If respondents satisfy translation invariance, then their choices in the TI_{CIM} and TI_{UM} questions should be the same as in CIM_m and UM_m respectively. Income translation invariance was tested using both a decrease ($TI_{CIM,i,-}$ and $TI_{UM,i,-}$) and an increase ($TI_{CIM,i,+}$ and $TI_{UM,i,+}$) in income. Similarly, health translation invariance was tested using both a decrease ($TI_{CIM,h,-}$ and $TI_{UM,h,-}$) and an increase ($TI_{CIM,h,+}$ and $TI_{UM,h,+}$) in the number of days without backpain. The decrease in income and number of days without backpain was set equal to 20% of average income or days without backpain, i.e. €3160 and 65 days. The increase in income and number of days without backpain was set equal to 4% of average income and days without backpain, i.e. €632 and 13 days.[3]

We asked two additional questions to test for health translation invariance. The numbers of days without backpain in the distributions $TI_{CIM,h,gl}$ and $TI_{UM,h,gl}$ were constructed such that they equal the number of days *with* backpain in $TI_{CIM,h,-}$ and $TI_{UM,h,-}$ but in a different order. As will be explained later, we will have versions of the questionnaire where health is expressed in terms of days without backpain (attainment-framing), as well as versions where health is expressed in terms of days with backpain (shortfall-framing). If respondents' attitudes towards inequalities depend solely on the numbers they see, and not on the meaning of these numbers and if they consider both dimensions separately, we for instance expect similar choices for $TI_{CIM,h,-}$ in the attainment-framing as for $TI_{CIM,h,gl}$ in the shortfall-framing. Note that similar questions for income would not make sense as income is not a bounded variable.

Unit invariance

Unit invariance was also tested by applying it to both CIM and UM, with the reference questions being CIM_m and UM_m. Table 3b summarizes the unit invariance questions that can be found in Tables A.7 and A.8 in Appendix. Income unit invariance was tested using both a decrease ($UI_{CIM,i,-}$ and $UI_{UM,i,-}$) and an increase ($UI_{CIM,i,+}$ and $UI_{UM,i,+}$) in income.

Health is a bounded variable in our study, so unit invariance can be tested in two ways here. Unit invariance can apply either to the number of days *without* backpain, or to the number of days *with* backpain. Respondents are expected to satisfy at most one of these two versions of unit invariance. Health unit invariance was tested using a decrease applied to the number of days without backpain ($UI_{CIM,h,-}$ and $UI_{UM,h,-}$) as well as a decrease applied to the number of days with backpain ($UI_{CIM,h,-,sf}$ and $UI_{UM,h,-,sf}$, with sf denoting shortfall). Health unit invariance was also tested using an increase applied to the number of days without backpain ($UI_{CIM,h,+}$ and $UI_{UM,h,+}$) as well as an increase applied to the number of days with backpain ($UI_{CIM,h,+,sf}$ and $UI_{UM,h,+,sf}$, with

[3] Section 2.2 discusses that testing for income and health translation invariance separately is appropriate when income and health are measured on a different scale.

sf denoting shortfall). For decreases and increases in income and health we applied the same percentages of average income and health as for the translation invariance questions: 20% decrease and 4% increase.[4,5,6]

Questions $UI_{CIM,h,gl}$ and $UI_{UM,h,gl}$ equal the distributions in CIM_m and UM_m with the number of days without backpain multiplied by 105/325. Thus, $UI_{CIM,h,gl}$ and $UI_{UM,h,gl}$ give the same average number of days with backpain as $TI_{CIM,h,gl}$ and $TI_{UM,h,gl}$. Questions $UI_{CIM,h,gl,sf}$ and $UI_{UM,h,gl,sf}$ equal the distributions in CIM_m and UM_m with the number of days *with* backpain multiplied by 105/325.

Mirror invariance

Health can be expressed in terms of attainments/gains (number of days without backpain) or shortfalls/losses (number of days with backpain). We constructed both a health-attainments and health-shortfalls version of our questionnaire, where the health-shortfalls version has equal numbers of days with backpain as the health-attainments version. If respondents satisfy mirror invariance, we expect the same responses in the attainments as in the shortfalls version of the questionnaire. We therefore tested mirror invariance *between* respondents. Testing for mirror invariance within respondents would have required each respondent to answer questions in terms of attainments as well as shortfalls, which might encourage respondents to respond consistently and thereby lead to an overestimation of the support for this property, and which is also cognitively demanding.

Versions

In policy A of the questions to test for CIM the difference in health between subsequent income groups is 5 days without backpain. This difference could be perceived to be relatively small and could lead respondents to mainly focus on the income dimension. We therefore constructed an additional version of the questionnaire with larger differences in terms of health. The CIM and PIRHT questions of this version are summarized in Tables A.9–A.11 in Appendix. Policy A in the CIM questions had the same maximum number of days without backpain (335) as in the version with smaller health differences, but the differences between subsequent groups were multiplied by 6 such that they amount to 30 days, i.e. approximately 1 month. Policy B was also constructed by multiplying the transfer in the original version by 6. All other PIRHT-questions were derived from the original version in a similar way. The UM and unit invariance tests were derived from the CIM tests as described for the original version. The translation invariance tests were derived from the unit invariance tests, such that they involve the same percentage changes in average health, as described above.

Since the version with large health differences has a lower average number of days without backpain (275 instead of 325), it is also informative on the incompatibility between mirror and unit invariance: when average health lies exactly at the mid interval of minimum (365 days with backpain) and maximal health (0 days with backpain), both invariance conditions are compatible for all possible health-income allocations with mid-interval mean health. When average health is closer to the extreme values of health, the incompatibility more likely shows up in particular income-health allocations such as ours. Comparing both versions could thus inform us on which invariance condition receives more support (Lambert and Zheng, 2011; Erreygers and Van Ourti, 2011).

In the example in Table 1, health is expressed as a gain, the income-rows are presented above the health-rows and the five groups are ranked based on the top row, i.e. from poorest (group 1) to richest (group 5). In the corresponding version of the questionnaire, we always ranked the groups in Alfaland from small to large in terms of income. We also have versions where health is always presented above income and where groups are ranked on health — from few to many days without backpain for health attainments, from many to few days with backpain for health shortfalls. Normatively, ranking distributions on income or health should not matter as they represent the exact same distributions. In practice, however, responses could be sensitive to the ranking. For CIM, for instance, a given transfer can be viewed as a health transfer from a higher to a lower income group in the income-ranked version while being viewed as an income transfer from a better to a poorer health group in the health-ranked version. If respondents have different attitudes towards income transfers than towards health transfers, then the ranking may have an influence on responses.

To summarize, from both the version with smaller health differences and the version with larger health differences, we constructed a health attainments and a health shortfalls version with groups ranked on income or health. Thus, we have 8 versions in total:

- health attainments - income-ranked - small differences (*AIR*)
- health attainments - health-ranked - small differences (*AHR*)
- health shortfalls - income-ranked - small differences (*SIR*)
- health shortfalls - health-ranked - small differences (*SHR*)
- health attainments - income-ranked - large differences (*AIRL*)
- health attainments - health-ranked - large differences (*AHRL*)
- health shortfalls - income-ranked - large differences (*SIRL*)
- health shortfalls - health-ranked - large differences (*SHRL*)

[4] The numbers of days without backpain in $UI_{CIM,h,-,sf}$, for instance, equal $365-(365-x)\times 0.8$, where x is the numbers of days without backpain in CIM_m.
[5] Questions $UI_{UM,h,+}$ and $UI_{UM,h,+,sf}$ were dropped due to an error in the health distribution in policy B.
[6] Testing for income and health unit invariance separately is appropriate when income and health are measured on a different scale, see also Section 2.2.

Background questions

The last few questions in the questionnaire were background questions and two debriefing questions. In the first debriefing question, respondents were asked whether (and why), when making their choices, they focused only on the columns of the tables that differed between the policies, or whether they also considered the other columns. If they consider only the columns that differ, then their choices will imply separability between groups of individuals. The second and open-ended debriefing question inquired about any rules used to make the choices. Next, we asked for gender, age, study program, and number of days with backpain experienced in the last year. We also asked respondents to rate their political views (using the American National Election Studies 10-point scale: very left to very right),[7] family's income 10 years ago relative to the average income in society 10 years ago (on a 5-point scale: relatively low to relatively high), and expected own income in 10 years relative to the average income in society in 10 years (on a 5-point scale: relatively low to relatively high). Finally, we asked whether they are familiar with the following concepts: gini-index, lorenz-curve, and/or concentration index.

Implementation

We recruited 349 students of Erasmus University Rotterdam to participate in our study.[8] They were invited to the lab for a session that was expected to last 40 min and they were paid €10 for their participation. We ran 17 sessions in total, spread over 6 days in the spring of 2018. In each session, respondents were randomly assigned to one of the versions of our questionnaire.

4. Results

Table A.13 in Appendix summarizes the background characteristics of the 349 respondents. It shows that 54% of them are female and their average age is 22 years. Around 89% of the respondents are or were enrolled in an economics or business program (*econbus* = 1; 0 otherwise). 82% of the respondents experienced at least one day of backpain in the year prior to the experiment (*BP-1-365-days* = 1; 0 otherwise), and 23% experienced more than 28 days of backpain (*BP-29-365-days* = 1; 0 otherwise). We summarized political views into a variable *right*, which equals one if a subject gave an answer from 7 to 10 on the political views scale and zero otherwise. Past family income, relative to average income, was perceived to be in the highest category by around 10% of respondents (*highpastincome* = 1 if answer 5 on the 5-point scale; 0 otherwise), and future income is expected to be in the highest category by 21% (*highfutureincome* = 1 if 5 on the 5-point scale; 0 otherwise). Many respondents, 78%, are familiar with the Gini index and/or the Lorenz curve (*gini-lorenz* = 1; 0 otherwise), and only 16% are familiar with the concentration index (*conc-index* = 1; 0 otherwise) reflecting that the Gini index and Lorenz curve are part of the curriculum in secondary education in the Netherlands, but not the concentration index. Almost 65% of the respondents applied group-independence when making their choices, implying that they only focused on the groups that differed between the two policies when making a decision (*indep* = 1; 0 otherwise).

Balancing tests confirmed absence of systematic differences in average background characteristics across questionnaire versions (see Table A.12 in Appendix), and there were no differences in terms of exclusive focus (or not) on columns affected by the policies. Time required to read the instructions was similar between versions indicating that all versions were equally difficult to understand.

We analyzed the responses to the choice questions using a seemingly unrelated regression (SUR) framework, where the choices in each question – coded as 1 if consistent with the relevant preference condition and 0 otherwise – are modeled as a linear probability model and estimated jointly to allow for correlation between different choice questions. A SUR model with question-specific-constants only, informs whether the proportion of choices consistent with the preference conditions differs significantly from 50%, which would be observed if respondents chose randomly. We assess the impact of framing in terms of attainments or shortfalls, having small or large health differences, and policies ranked on income or health, by extending the question-specific-constants-only SUR model with dummies accounting for different versions of the questionnaire.[9] We additionally assess the importance of respondent's background characteristics by estimating a SUR model that lets the choices also depend on background characteristics.

4.1. CIM

Panel A of Fig. 1 shows that respondent's choices do not deviate significantly from 50% and that questions involving transfers at the top, middle, and bottom of the income distribution receive similar support (p-value of the joint equality test equals 0.97). When investigating individual response patterns, we see that these findings are driven by individual heterogeneity rather than by random responses: 38% of the respondents are always consistent and 30% are always inconsistent with CIM.

This individual heterogeneity is partly driven by presenting policies as ranked on income versus ranked on health (see coefficient of IR in Table 4). When policies are ranked on income (IR equals 1), support for CIM_b, CIM_m and CIM_t is significantly larger than 50% (for example, support for CIM_b equals $0.632 = 0.393 + 0.067 * 0.499$ (share of respondents with $S = 1$ for health

[7] https://electionstudies.org/.

[8] The required number of observations was based on a power calculation for the difference in support between two versions of the questionnaires for just one of the 38 questions. This indicated that 43 participants per version are required to have a minimum detectable effect (MDE) of 30 percentage points (pp) – 65p vs 35pp – implying 344 observations in total. The MDE using the 349 observations for testing question-specific deviations from the random response pattern of 50% is 7.5 pp.

[9] The other randomizations – order of the 38 choice questions, order of the policies within each choice question –, the interview date, the time required to read the instructions and the time required to respond to each choice questions did not affect the response patterns in a systematic way.

Table 4
The impact of framing in terms of attainments or shortfalls, having small or large health differences, and policies ranked on income or health on respondents' choices.

(a)

	CIM_b	CIM_m	CIM_t	$PIRHT_b$	$PIRHT_m$	$PIRHT_t$	$PIRHT_{b,nr}$	$PIRHT_{m,nr}$	$PIRHT_{t,nr}$	UM_b	UM_m	UM_t
S	0.067	0.096	0.107	0.217	0.148	0.090	0.125	0.096	0.045	0.078	0.073	0.050
	(0.201)	(0.070)	(0.043)	(0.000)	(0.004)	(0.083)	(0.012)	(0.063)	(0.359)	(0.120)	(0.164)	(0.285)
L	0.049	0.043	0.020	0.037	0.038	0.096	0.074	0.079	0.098	0.017	−0.013	0.058
	(0.348)	(0.413)	(0.701)	(0.463)	(0.468)	(0.065)	(0.139)	(0.126)	(0.046)	(0.741)	(0.804)	(0.209)
IR	0.181	0.142	0.130	0.240	0.182	0.171	0.206	0.165	0.217	−0.001	0.084	−0.007
	(0.001)	(0.007)	(0.014)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.982)	(0.109)	(0.887)
const	0.393	0.399	0.416	0.312	0.381	0.404	0.443	0.433	0.489	0.627	0.524	0.700
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349	349
R^2	0.040	0.031	0.029	0.107	0.057	0.047	0.069	0.044	0.066	0.007	0.013	0.008
p	0.002	0.012	0.017	0.000	0.000	0.001	0.000	0.001	0.000	0.469	0.210	0.429

(b)

	TI_{CIM} $_{i,-}$	TI_{CIM} $_{i,+}$	TI_{CIM} $_{h,-}$	TI_{CIM} $_{h,+}$	TI_{CIM} $_{h,gl}$	UI_{CIM} $_{i,-}$	UI_{CIM} $_{i,+}$	UI_{CIM} $_{h,-}$	UI_{CIM} $_{h,-,sf}$	UI_{CIM} $_{h,+}$	UI_{CIM} $_{h,+,sf}$	UI_{CIM} $_{h,gl}$	UI_{CIM} $_{h,gl,sf}$
S	0.028	−0.001	−0.041	−0.012	0.011	−0.007	0.022	−0.035	0.005	0.005	−0.018	0.011	−0.018
	(0.433)	(0.985)	(0.295)	(0.759)	(0.779)	(0.861)	(0.564)	(0.357)	(0.896)	(0.889)	(0.629)	(0.786)	(0.643)
L	0.038	−0.014	0.003	−0.026	−0.026	0.026	0.020	−0.059	−0.008	−0.020	−0.008	0.009	0.015
	(0.285)	(0.722)	(0.931)	(0.524)	(0.515)	(0.490)	(0.597)	(0.113)	(0.830)	(0.609)	(0.831)	(0.822)	(0.700)
IR	0.005	0.011	0.004	−0.024	0.033	0.016	0.022	0.010	0.028	0.039	0.016	0.011	0.039
	(0.879)	(0.794)	(0.909)	(0.553)	(0.396)	(0.667)	(0.560)	(0.786)	(0.479)	(0.309)	(0.658)	(0.784)	(0.310)
const	0.842	0.836	0.859	0.859	0.830	0.836	0.819	0.898	0.830	0.836	0.870	0.830	0.830
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349	349	349
R^2	0.005	0.001	0.003	0.003	0.004	0.002	0.003	0.010	0.002	0.004	0.001	0.001	0.004
p	0.616	0.978	0.773	0.837	0.749	0.875	0.813	0.326	0.905	0.726	0.923	0.978	0.707

(c)

	TI_{UM} $_{i,-}$	TI_{UM} $_{i,+}$	TI_{UM} $_{h,-}$	TI_{UM} $_{h,+}$	TI_{UM} $_{h,gl}$	UI_{UM} $_{i,-}$	UI_{UM} $_{i,+}$	UI_{UM} $_{h,-}$	UI_{UM} $_{h,-,sf}$	UI_{UM} $_{h,gl}$	UI_{UM} $_{h,gl,sf}$
S	0.051	0.016	−0.030	0.032	−0.013	0.091	0.050	0.016	0.004	−0.019	−0.036
	(0.249)	(0.719)	(0.523)	(0.469)	(0.786)	(0.044)	(0.277)	(0.724)	(0.927)	(0.688)	(0.442)
L	−0.079	0.025	−0.067	0.042	−0.027	0.007	0.001	0.036	−0.010	0.002	−0.015
	(0.076)	(0.569)	(0.152)	(0.352)	(0.561)	(0.880)	(0.981)	(0.412)	(0.827)	(0.970)	(0.740)
IR	0.016	−0.007	−0.054	−0.047	−0.002	0.022	0.005	−0.007	−0.053	0.004	−0.002
	(0.716)	(0.878)	(0.255)	(0.296)	(0.972)	(0.618)	(0.918)	(0.879)	(0.236)	(0.928)	(0.970)
const	0.785	0.774	0.814	0.763	0.768	0.711	0.728	0.763	0.808	0.763	0.780
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349
R^2	0.013	0.001	0.011	0.007	0.001	0.012	0.003	0.002	0.004	0.001	0.002
p	0.206	0.923	0.289	0.473	0.937	0.230	0.754	0.843	0.691	0.982	0.871

Note: Point estimates and p-values (in parentheses) obtained from SUR models with dummies accounting for different versions of the questionnaire. The three panels result from one single SUR estimation. S equals 1 when presented in terms of shortfalls, and equals 0 in case of attainments. L equals 1 when health differences are large and 0 when health differences are small. IR equals 1 when policies are ranked on income, and 0 when ranked on health. p shows the p-value of the F-test that all coefficients are jointly equal to zero.

presented in shortfalls) + 0.049 ∗ 0.496 (share of respondents with $L = 1$ for large health differences) + 0.181).[10] When policies are ranked on health (IR equals 0), support does not differ significantly from 50%. Support for CIM did not depend on the magnitude of health differences (L equals 1 for large health differences and 0 for small health differences). Framing in terms of attainments (S equals 0) or shortfalls (S equals 1), also for PIRHT, UM, and translation and unit invariance, is discussed in Section 4.5.[11]

[10] Out of 349 respondents, 174 respondents faced a questionnaire in terms of shortfalls and 173 in terms of large health differences.

[11] Conditioning on the interactions between all three aspects of framing (ranking on income or health, magnitude of health differences, and attainments or shortfalls), did not affect the coefficient estimates for ranking by income or health in a meaningful way.

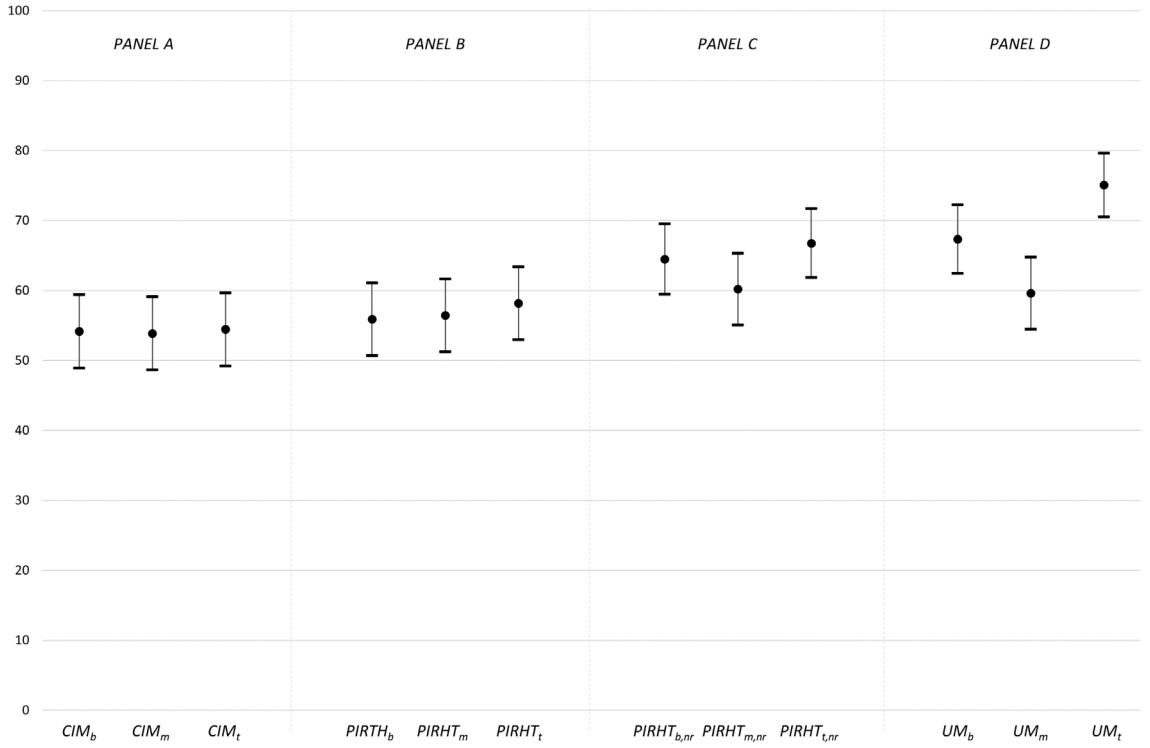

Fig. 1. Proportion responding in line with CIM, PIRHT and UM. Note: Point estimates and 95% Confidence Intervals obtained from a constant-only SUR of linear probability models of all 36 choice questions.

4.2. PIRHT

Panels B and C of Fig. 1 summarize the support for PIRHT. The percentage of choices consistent with PIRHT is significantly larger than 50% for each question. Moreover, we find that support for PIRHT is larger than for CIM, especially when there is no reranking, suggesting that PIRHT receives more support because it can accommodate transfers without reranking. We find no systematic difference between support for transfers at different parts of the income distribution (p-values of joint equality test equal 0.07 and 0.63 for PIRHT without and with reranking respectively). Individual response patterns indicate that around 45% of the respondents choose very consistently with PIRHT (5 to 6 choices consistent with PIRHT). Finally, support for PIRHT is strongly reduced to around 50% when policies are ranked by health rather than income, while the magnitude of health differences has no systematic impact (see Table 4).

4.3. UM

Support for UM is summarized in Panel D of Fig. 1. The percentage of choices consistent with UM is significantly larger than 50% in all three questions. There is significantly more support for UM in UM_t than in UM_m, and more support in UM_b than in UM_m, suggesting that inequality reducing transfers implied by UM are most preferred among the best-off and least preferred around the median. The finding that support is lowest for transfers around the median may have to do with the fact that UM_m involved a re-ranking in terms of health between the second and the third income groups. For CIM and PIRHT we similarly found that transfers that involve re-rankings in terms of health get lower support. We also find 70% of the respondents choosing at least 2 out of 3 times in line with UM, while the share persistently rejecting UM (8%) is low. Ranking of policies on income or health and magnitude of health differences both had no impact on support for UM.

4.4. Translation and unit invariance

Figs. A.2 and A.3 in Appendix show the proportions of respondents who chose the same policy as in the corresponding CIM_m and UM_m questions, i.e. consistent with translation and unit invariance. Even though CIM_m was supported by only slightly more than 50% of the respondents, the level of consistency in the corresponding translation and unit invariance questions was very high (between 83% and 88%) and did not differ significantly between questions ($p = 0.68$). Moreover, within-respondent support for choosing consistently did not depend on whether the same proportional transformation was applied to health attainment or health

shortfall (compare the questions with and without sf-extension in Table 3b). The same pattern emerges for consistencies between UM_m and the corresponding translation and unit invariance questions, with between 74% and 79% of consistent responses, a p-value of 0.70 for differences between questions, and within-respondent support not depending on applying proportional transformations to health attainments or shortfalls.[12] In addition, consistency with CIM_m and UM_m did not depend on ranking by income or health and not on the magnitude of health differences.

4.5. Mirror invariance

Table 4 shows that we generally find more support for CIM and PIRHT when health is expressed as a shortfall rather than as an attainment. Support for UM is also comparatively higher for shortfalls, but insignificantly so. For PIRHT, the differences between framing become more pronounced, in terms of size as well as significance, when moving from transfers at the top to transfers at the bottom of the income distribution, but do not depend on the magnitude of health differences or on ranking by income and health (see footnote 11).[13]

4.6. Group-independence

Respondents were asked whether they applied *group-independence* when making their choices, i.e. whether they focused only on the groups that differed between the two policies when making a decision. Group-independence is thus in accordance with CIM, PIRHT and UM, which all assume that the non-affected groups should not matter when choosing between policies. Table 5 shows that support of CIM and PIRHT depends on support for group-independence (*indep* equals 1).[14] Among those applying group-independence, support for CIM and PIRHT is around two thirds for each question, while support among those not applying group-independence is around one third for CIM and PIRHT with reranking, and 50% for PIRHT without reranking. Group-independence only matters for UM when there is a re-ranking in terms of health between income groups, i.e. for UM_m. Group-independence did not matter for translation and unit invariance. The responses to the open question which motivated students to reflect on their reasons for supporting (or not) group-independence were vague and not informative.[15]

4.7. Demographics

The estimates in Tables A.14–A.16 in Appendix reveal the association between the background characteristics and support for CIM, PIRHT, and UM.[16,17] Of all background variables (gender, age, study program, political views, past family income, future personal income, number of days with backpain during the last year, and familiarity with gini-index, lorenz-curve, or concentration-index), only political orientation was systematically related to support for the preference conditions, with respondents with a more rightish political orientation supporting CIM, PIRHT, and UM_m, the only UM question with reranking, less. We did a similar analysis for all translation and unit invariance questions and found no association between choices and background characteristics.

5. Conclusion and discussion

Our results show support for UM and for PIRHT, with more than 50% of the responses being consistent with these preference conditions. Support for PIRHT was stronger when presenting policies as income-ranked and when the transfer of health did not involve a re-ranking of groups in terms of health, while support for CIM was only confirmed when policies were presented as income-ranked. Support for UM did not depend on presenting policies as income- or health-ranked. Higher rates of consistency with CIM, PIRHT and UM_m were observed among respondents focusing only on the groups for whom income or health changed.

---

[12] We could not reject that attitudes towards inequalities depend solely on the numbers, i.e. responses to $TI_{CIM,h,-}$ under the attainment scenario do not differ significantly from those to $TI_{CIM,h,gl}$ under the loss scenario and vice versa; and the same holds true for UM.

[13] We also tested whether mirror and unit invariance are incompatible for the allocations used in our questionnaire with respect to UM_m, but not CIM_m as mirror was rejected for CIM. We find no evidence that the incompatibility more likely occurs when health differences are small (average health = 40 days of backpain) compared to large (average health = 90 days of backpain), which would be expected as average health is closer to the mid-interval level of health (182.5 days) in the latter case: the interaction terms between shortfalls/attainments and large/small health differences for the 6 unit invariance questions for UM are jointly insignificant in our SUR model (p-value = 0.121). This is in line with the finding in Section 4.4 that within-respondent support for choosing consistently with CIM_m and UM_m did not depend on proportional transformations to health attainments or health shortfalls.

[14] The proportion of respondents who claimed to have applied group-independence did not differ between versions of the questionnaire (see Table A.12 in Appendix) and was also unrelated to the background characteristics (p-value of the F-test of a regression of all background characteristics on group-independence equals 0.27).

[15] Some responses did not relate to group-independence but indicated that respondents were focusing on the mechanisms underlying the correlation between income and backpain, such as 'richer individuals can afford the care/support required to deal with backpain' or 'healthier individuals are more productive and should have higher incomes'.

[16] As background characteristics did not differ significantly between versions of the experiment, we did not control for the different questionnaire versions in these regressions (see Table A.12 in Appendix).

[17] The responses to the second open-ended debriefing question, which inquired about any decision rules used, were generally vague such as trying to "balance out", "reduce inequality", "reduce gaps".

Table 5
The association between group-independence and respondents' choices.

(a)

	CIM_b	CIM_m	CIM_t	$PIRHT_b$	$PIRHT_m$	$PIRHT_t$	$PIRHT_{b,nr}$	$PIRHT_{m,nr}$	$PIRHT_{t,nr}$	UM_b	UM_m	UM_t
indep	0.290	0.310	0.307	0.291	0.288	0.302	0.224	0.183	0.160	0.094	0.161	0.051
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.073)	(0.003)	(0.291)
const	0.355	0.339	0.347	0.371	0.379	0.387	0.500	0.484	0.565	0.613	0.492	0.718
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349	349
R^2	0.077	0.089	0.087	0.079	0.077	0.086	0.050	0.032	0.026	0.009	0.025	0.003

(b)

	TI_{CIM} $_{i,-}$	TI_{CIM} $_{i,+}$	TI_{CIM} $_{h,-}$	TI_{CIM} $_{h,+}$	TI_{CIM} $_{h,gl}$	UI_{CIM} $_{i,-}$	UI_{CIM} $_{i,+}$	UI_{CIM} $_{h,-}$	UI_{CIM} $_{h,-,sf}$	UI_{CIM} $_{h,+}$	UI_{CIM} $_{h,+,sf}$	UI_{CIM} $_{h,gl}$	UI_{CIM} $_{h,gl,sf}$
indep	−0.003	−0.045	−0.057	−0.054	−0.049	−0.027	−0.069	−0.022	−0.007	−0.010	−0.059	−0.040	−0.073
	(0.925)	(0.279)	(0.163)	(0.201)	(0.235)	(0.503)	(0.085)	(0.574)	(0.868)	(0.796)	(0.123)	(0.325)	(0.069)
const	0.879	0.863	0.879	0.863	0.871	0.871	0.895	0.871	0.847	0.855	0.903	0.871	0.895
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349	349	349
R^2	0.000	0.003	0.006	0.005	0.004	0.001	0.009	0.001	0.000	0.000	0.007	0.003	0.010

(c)

	TI_{UM} $_{i,-}$	TI_{UM} $_{i,+}$	TI_{UM} $_{h,-}$	TI_{UM} $_{h,+}$	TI_{UM} $_{h,gl}$	UI_{UM} $_{i,-}$	UI_{UM} $_{i,+}$	UI_{UM} $_{h,-}$	UI_{UM} $_{h,-,sf}$	UI_{UM} $_{h,gl}$	UI_{UM} $_{h,gl,sf}$
indep	−0.004	0.001	−0.042	0.041	0.047	0.070	0.048	−0.033	−0.004	−0.003	−0.032
	(0.923)	(0.986)	(0.397)	(0.378)	(0.337)	(0.138)	(0.323)	(0.472)	(0.923)	(0.958)	(0.508)
const	0.782	0.790	0.766	0.750	0.718	0.726	0.726	0.806	0.782	0.758	0.774
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349
R^2	0.000	0.000	0.002	0.002	0.003	0.006	0.003	0.002	0.000	0.000	0.001

Note: Point estimates and p-values (in parentheses) obtained from SUR models with a dummy for group-independence (*indep* equals 1 when the respondents applied group-independence, and 0 otherwise). The three panels result from one single SUR estimation.

We found strong support for translation and unit invariance, indicating that choices did not depend on changes in average levels of health or income that left absolute or relative differences between groups unaffected. Mirror invariance, however, was rejected for CIM and PIRHT, but not for UM.

Our data revealed four additional insights.[18] First, re-ranking of groups in society matters greatly for support of CIM, PIRHT and UM. The re-ranking due to health transfers that is implicit in CIM and may occur for PIRHT and UM, lowers support for CIM, PIRHT and UM. While re-ranking always occurs for CIM, our findings indicate that PIRHT and UM are acceptable to a larger pool of respondents when the correlation between income and health is strong. A firmer correlation lowers the potential for re-ranking as the richer individual is more likely to stay healthier than the poorer individual after transferring health from rich to poor. The importance of health re-ranking is in line with the lack of support for PIRHT reported in Bleichrodt et al. (2012) where 11 of the 12 tests for PIRHT included rank reversals. It is also in line with Tarroux's (2015) finding of income re-ranking reducing the support of UM. Our results extend the importance of re-ranking to the domain of health and to preference conditions on the correlation between income and health such as PIRHT and CIM.

Second, ranking policies in terms of income or health also mattered. While the ranking variable should not matter from a normative perspective as they represent the exact same distributions of health and income, in practice responses revealed a stronger support for PIRHT and CIM when groups were ranked by income. This could derive from interpreting the case of ranking by income as representing transfers of health and the case of ranking by health as representing transfers of income, which would then imply that our results show that respondents are more supportive of health than income transfers. Whether this is indeed what drives this result, remains a topic for future research.

Third, our results also point to the importance of only focusing on those groups for whom income or health changed. Respondents who claimed to make decisions with this in mind had a stronger support for CIM, PIRHT and UM_m, 35 percent of our respondents did not respond in line with group-independence (see Table A.12 in the Appendix) and group-independence was unrelated to any of the background characteristics included in our questionnaire.

[18] A SUR model with group-independence, income versus health ranking, and political orientation as sole regressors showed almost identical coefficients as those reported in Table 4. Combined with group-independence and political orientation showing no differences across questionnaire version (see Table A.12 in Appendix), this suggests that these are three complementary explanations for support for CIM, PIRHT, and UM.

Fourth, support for CIM, PIRHT and UM was consistently higher for respondents with more leftist political orientation. This is in line with Hardardottir et al. (2021) who reported less aversion to income-related health inequalities among those with rightist political views.

We also found that support for CIM and PIRHT does not depend on transfers/permutations occurring at the bottom or top of the income distribution. Responses were thus not consistent with preference conditions, such as downside positional transfer sensitivity (Wagstaff, 2002; Makdissi and Yazbeck, 2014; Khaled et al., 2018), that put more weight on health transfers at the bottom (versus top) of the income distribution. Instead, response patterns were in line with the weighting scheme of the concentration index (Wagstaff et al., 1991) that weighs health transfers independently from where they occur in the income distribution. This is in line with Hardardottir et al. (2021) who parametrically estimate the level of inequality aversion to be on average close to that of the standard concentration index.

A final result concerns the importance of expressing health as a shortfall (days with backpain) compared to an attainment (days without backpain). We found more support for CIM and PIRHT when health is expressed as a shortfall. The difference in support for PIRHT between shortfall and attainment was more prominent for transfers among the poorest groups. These findings are inconsistent with those of Attema et al. (2019) who study the impact of framing in terms of attainments or shortfalls for multivariate risk preferences for longevity and wealth. Like us, they found correlation aversion to be sign-dependent, i.e. different between outcomes expressed in terms of gains or losses. Unlike us, however, they found support for correlation aversion to be stronger for gains rather than for losses. Two major differences between their study and ours, however, are that their gain and losses framings did not involve the same objective outcomes, whereas the framing in our study did; and that our study considered correlation aversion between dimensions over different groups from the perspective of a social planner, while Attema et al. (2019) studied attitudes towards correlation between dimensions over different states of nature from the individual perspective.

Overall, our results suggest that the presentation of policies influences support for CIM and PIRHT, but not for UM. Support for CIM and PIRHT was significantly stronger when policies were presented as income-ranked and when health was expressed as shortfall. From a normative perspective, this sensitivity to presentation is undesirable. Ideally, policy makers should be impartial and insensitive to presentation of policies. Our results show that policy makers may not satisfy this insensitivity. Moreover, even if they are, public support for their policies will depend on presentation. This calls for more research to study the exact drivers of this sensitivity and to assess whether and how it can be reduced. Sensitivity to framing in terms of attainments/gains or shortfalls/losses is well-known from the behavioral economics literature, where reference points have been shown to be influential drivers of behavior. Sensitivity to income- versus health-ranking is not well-known and its drivers are also not well-established. As discussed, one possibility is that an income-ranking presentation induces decision makers to interpret the choice as redistributing health, while health-ranking may induce an interpretation of redistributing income. An avenue for future research is to study to which extent this interpretation is indeed driving the sensitivity, or whether other explanations are more plausible. Sensitivity to income- versus health-ranking and to framing in terms of attainments/gains or shortfalls/losses would, for instance, suggest that choices may depend on the policy presented first – and therefore interpreted as the 'base-case'. Yet, we found that the order in which policies were presented did not affect choices in a systematic way.

There are several limitations to our study. First of all, the decisions made in our questionnaire were hypothetical. We thought that implementing real incentives would not be feasible for the purposes of our study. One could, of course, create a small society in the lab and ask decision makers to decide for this society. Then we would need two policies, each with five groups. Decision makers would be randomly allocated to one of the 5 groups of the policy they preferred and get paid according to the outcome for the allocated group. This would work in a uni-dimensional setting but would become much more complicated in our multidimensional setting. For any multidimensional setting, we would need one additional outcome dimension next to payment. Instead of health, one could think of effort, time in the lab, or non-monetary goods such as snacks. However, effort, time in the lab, and snacks can in principle be generated by income, which could lead respondents to reduce the multi-dimensional problem to a uni-dimensional one. Moreover, we were explicitly interested in health, which is much harder or even impossible to implement with real incentives.

A second limitation of our study is the homogeneous sample of respondents. All of them were student. One reason not to administer our questionnaire to a representative sample of the population is that our questions are cognitively demanding. Administering our questions to a representative sample of the population would have required us to reduce the number of questions and possibly extend the instructions. For this study we wanted to test as many of our conditions as possible within rather than between respondents. We therefore aimed for administering our questionnaire in the lab with highly educated participants, which we chose to be students. Moreover, students in economics and related fields are the future policy makers who will be faced with these decisions between policies. An interesting avenue for future research is to administer a similar survey to highly educated professionals, including policy makers, to assess the robustness of our results. Moreover, administering a similar survey to a general population sample will shed light on the extent to which education-level matters.

Finally, we should acknowledge that our findings are only internally valid for the chosen baseline distribution for CIM, PIRHT and UM. Our aim was to test preference conditions rather than providing parameter estimates. Thus, we did not estimate inequality aversion parameters by imposing particular functional forms for the underlying social welfare functions as is common in the literature (e.g. Hurley et al., 2020; Hardardottir et al., 2021), and instead focused on the underlying non-parametric preference conditions. The acceptability of (general) preference conditions is arguably less sensitive to the chosen baseline distribution than (specific) parameter estimates. We could also reject response patterns being driven by different heuristics. Moreover, counterbalancing of the presentations of the distributions as ranked by income or health mitigated responses reflecting consistent focus on the health distribution only. While this possibility could not entirely be excluded for PIRHT, it could be rejected for CIM as the marginal distributions of income and health were identical in both policies. Most importantly, we are not worried that respondents found the

differences in health levels between income quintiles irrelevant. Choices between policies did not depend on framing in terms of large versus small health differences, and it is thus unlikely that respondents reduced the choice problems to one dimension by only focusing on the income dimension. We also avoided reasoning in terms of lifetime income which could occur when combining stock measures for health, such as life expectancy, with flow measures for income. We leave it to future studies to assess the robustness of our results. One avenue for future research is to consider different types of health outcomes.

CRediT authorship contribution statement

Kirsten I.M. Rohde: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Tom Van Ourti:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. **Amar Soebhag:** Data curation, Investigation, Software.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

See Tables A.6–A.16, Figs. A.2 and A.3.

Table A.6
Questions to test for translation invariance.

	Policy A					Policy B				
	1	2	3	4	5	1	2	3	4	5
$TI_{CIM,i,-}$	2840 315	6840 320	9840 325	14 840 330	28 840 335	2840 315	6840 330	9840 325	14 840 320	28 840 335
$TI_{CIM,h,-}$	6000 250	10 000 255	13 000 260	18 000 265	32 000 270	6000 250	10 000 265	13 000 260	18 000 255	32 000 270
$TI_{CIM,i,+}$	6632 315	10 632 320	13 632 325	18 632 330	32 632 335	6632 315	10 632 330	13 632 325	18 632 320	32 632 335
$TI_{CIM,h,+}$	6000 328	10 000 333	13 000 338	18 000 343	32 000 348	6000 328	10 000 343	13 000 338	18 000 333	32 000 348
$TI_{CIM,h,gl}$	6000 95	10 000 100	13 000 105	18 000 110	32 000 115	6000 95	10 000 110	13 000 105	18 000 100	32 000 115
$TI_{UM,i,-}$	2840 315	6840 320	9840 325	14 840 330	28 840 335	2840 315	10 040 324	9840 325	11 640 326	28 840 335
$TI_{UM,h,-}$	6000 250	10 000 255	13 000 260	18 000 265	32 000 270	6000 250	13 200 259	13 000 260	14 800 261	32 000 270
$TI_{UM,i,+}$	6632 315	10 632 320	13 632 325	18 632 330	32 632 335	6632 315	13 832 324	13 632 325	15 432 326	32 632 335
$TI_{UM,h,+}$	6000 328	10 000 333	13 000 338	18 000 343	32 000 348	6000 328	13 200 337	13 000 338	14 800 339	32 000 348
$TI_{UM,h,gl}$	6000 95	10 000 100	13 000 105	18 000 110	32 000 115	6000 95	13 200 104	13 000 105	14 800 106	32 000 115

Table A.7
Questions to test for unit invariance for CIM.

	Policy A					Policy B				
	1	2	3	4	5	1	2	3	4	5
$UI_{CIM,i,-}$	4800 315	8000 320	10 400 325	14 400 330	25 600 335	4800 315	8000 330	10 400 325	14 400 320	25 600 335
$UI_{CIM,h,-}$	6000 252	10 000 256	13 000 260	18 000 264	32 000 268	6000 252	10 000 264	13 000 260	18 000 256	32 000 268
$UI_{CIM,h,-,sf}$	6000 325	10 000 329	13 000 333	18 000 337	32 000 341	6000 325	10 000 337	13 000 333	18 000 329	32 000 341
$UI_{CIM,i,+}$	6240 315	10 400 320	13 520 325	18 720 330	33 280 335	6240 315	10 400 330	13 520 325	18 720 320	33 280 335
$UI_{CIM,h,+}$	6000 328	10 000 333	13 000 338	18 000 343	32 000 348	6000 328	10 000 343	13 000 338	18 000 333	32 000 348
$UI_{CIM,h,+,sf}$	6000 313	10 000 318	13 000 323	18 000 329	32 000 334	6000 313	10 000 329	13 000 323	18 000 318	32 000 334
$UI_{CIM,h,gl}$	6000 102	10 000 103	13 000 105	18 000 107	32 000 108	6000 102	10 000 107	13 000 105	18 000 103	32 000 108
$UI_{CIM,h,gl,sf}$	6000 349	10 000 350	13 000 352	18 000 354	32 000 355	6000 349	10 000 354	13 000 352	18 000 350	32 000 355

Table A.8
Questions to test for unit invariance for UM.

	Policy A					Policy B				
	1	2	3	4	5	1	2	3	4	5
$UI_{UM,i,-}$	4800 315	8000 320	10 400 325	14 400 330	25 600 335	4800 315	10 560 324	10 400 325	11 840 326	25 600 335
$UI_{UM,h,-}$	6000 252	10 000 256	13 000 260	18 000 264	32 000 268	6000 252	13 200 259	13 000 260	14 800 261	32 000 268
$UI_{UM,h,-,sf}$	6000 325	10 000 329	13 000 333	18 000 337	32 000 341	6000 325	13 200 332	13 000 333	14 800 334	32 000 341
$UI_{UM,i,+}$	6240 315	10 400 320	13 520 325	18 720 330	33 280 335	6240 315	13 728 324	13 520 325	15 392 326	33 280 335
$UI_{UM,h,gl}$	6000 102	10 000 103	13 000 105	18 000 107	32 000 108	6000 102	13 200 105	13 000 105	14 800 105	32 000 108
$UI_{UM,h,gl,sf}$	6000 349	10 000 350	13 000 352	18 000 354	32 000 355	6000 349	13 200 352	13 000 352	14 800 352	32 000 355

Table A.9
Questions to test for CIM, PIRHT, and UM with large health differences between groups.

	Policy A					Policy B				
	1	2	3	4	5	1	2	3	4	5
CIM_b	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 245	10 000 215	13 000 275	18 000 305	32 000 335
CIM_m	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 215	10 000 305	13 000 275	18 000 245	32 000 335
CIM_t	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 215	10 000 245	13 000 275	18 000 335	32 000 305
$PIRHT_b$	6000 221	10 000 239	13 000 275	18 000 305	32 000 335	6000 233	10 000 227	13 000 275	18 000 305	32 000 335
$PIRHT_{b,nr}$	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 227	10 000 233	13 000 275	18 000 305	32 000 335
$PIRHT_m$	6000 215	10 000 257	13 000 275	18 000 293	32 000 335	6000 215	10 000 281	13 000 275	18 000 269	32 000 335
$PIRHT_{m,nr}$	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 215	10 000 269	13 000 275	18 000 281	32 000 335
$PIRHT_t$	6000 215	10 000 245	13 000 275	18 000 311	32 000 329	6000 215	10 000 245	13 000 275	18 000 323	32 000 317
$PIRHT_{t,nr}$	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 215	10 000 245	13 000 275	18 000 317	32 000 323
UM_b	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	7600 227	8400 233	13 000 275	18 000 305	32 000 335
UM_m	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 215	13 200 269	13 000 275	14 800 281	32 000 335
UM_t	6000 215	10 000 245	13 000 275	18 000 305	32 000 335	6000 215	10 000 245	13 000 275	23 600 317	26 400 323

Table A.10
Summary of questions to test for translation invariance with large health differences between groups.

$TI_{CIM,i,-}$	$= CIM_m -$ €3160	$TI_{UM,i,-}$	$= UM_m -$ €3160
$TI_{CIM,i,+}$	$= CIM_m +$ €632	$TI_{UM,i,+}$	$= UM_m +$ €632
$TI_{CIM,h,-}$	$= CIM_m -$ 55 days	$TI_{UM,h,-}$	$= UM_m -$ 55 days
$TI_{CIM,h,+}$	$= CIM_m +$ 11 days	$TI_{UM,h,+}$	$= UM_m +$ 11 days
$TI_{CIM,h,gl}$	$= CIM_m -$130 days	$TI_{UM,h,gl}$	$= UM_m -$ 130 days

Note: Average health in CIM_m equals 275. The reduction of 55 days is 20% of 275, and the increase of 11 days equals 4% of 275.

Table A.11
Summary of questions to test for unit invariance with large health differences between groups.

$UI_{CIM,i,-}$	$= CIM_m \times$ 0.8 income	$UI_{UM,i,-}$	$= UM_m \times$ 0.8 income
$UI_{CIM,i,+}$	$= CIM_m \times$ 1.04 income	$UI_{UM,i,+}$	$= UM_m \times$ 1.04 income
$UI_{CIM,h,-}$	$= CIM_m \times$ 0.8 days without BP	$UI_{UM,h,-}$	$= UM_m \times$ 0.8 days without BP
$UI_{CIM,h,-,sf}$	$= CIM_m \times$ 0.8 days with BP	$UI_{UM,h,-,sf}$	$= UM_m \times$ 0.8 days with BP
$UI_{CIM,h,+}$	$= CIM_m \times$ 1.04 days without BP		
$UI_{CIM,h,+,sf}$	$= CIM_m \times$ 1.04 days with BP		
$UI_{CIM,h,gl}$	$= CIM_m \times$ 145/275 days without BP	$UI_{UM,h,gl}$	$= UM_m \times$ 145/275 days without BP
$UI_{CIM,h,gl,sf}$	$= CIM_m \times$ 145/275 days with BP	$UI_{UM,h,gl,sf}$	$= UM_m \times$ 145/275 days with BP

Table A.12
Balancing tests: mean of the background questions per questionnaire version.

Variable	*AIR*	*AHR*	*SIR*	*SHR*	*AIRL*	*AHRL*	*SIRL*	*SHRL*	*Diff*
female	0.533	0.568	0.581	0.477	0.512	0.465	0.698	0.455	0.300
age	21.29	21.11	21.58	21.27	22.19	22.33	21.581	22.455	0.361
econbus	0.911	0.886	0.930	0.886	0.837	0.860	0.930	0.841	0.756
BP-1-365-days	0.800	0.773	0.884	0.727	0.814	0.884	0.860	0.841	0.529
BP-29-365-days	0.244	0.136	0.256	0.227	0.279	0.209	0.233	0.227	0.802
right	0.267	0.364	0.256	0.364	0.349	0.326	0.209	0.341	0.642
highpastincome	0.156	0.136	0.047	0.068	0.140	0.023	0.093	0.159	0.073
highfutureincome	0.289	0.250	0.140	0.273	0.186	0.140	0.233	0.182	0.499
gini	0.622	0.500	0.512	0.568	0.628	0.535	0.535	0.432	0.629
lorenz	0.622	0.705	0.721	0.705	0.721	0.651	0.674	0.591	0.873
conc-index	0.178	0.136	0.186	0.136	0.209	0.140	0.140	0.159	0.979
indep	0.689	0.500	0.651	0.659	0.674	0.581	0.698	0.705	0.519
time	130	139	135	134	147	130	129	143	0.534
N	45	44	43	44	43	43	43	44	

Note: The 8 questionnaire versions are abbreviated with the following acronyms: AIR (health attainments – income ranked – small differences), AHR (health attainments – health-ranked – small differences), SIR (health shortfalls – income-ranked – small differences), SHR (health shortfalls – health-ranked – small differences), AIRL (health attainments – income-ranked – large differences), AHRL (health attainments – health-ranked – large differences), SIRL (health shortfalls – income-ranked – large differences), SHRL (health shortfalls – health-ranked – large differences). The column *Diff* shows the *p*-value of the F-test for joint orthogonality of each variable across all questionnaire versions. *time* denotes the time required to read the instructions.

Table A.13
Descriptives of the background questions.

Variable	*Mean*	*StdDev*	*Min*	*Max*
female	0.536		0	1
age	21.72	3.28	18	54
econbus	0.885		0	1
BP-1-365-days	0.822		0	1
BP-29-365-days	0.226		0	1
right	0.309		0	1
highpastincome	0.103		0	1
highfutureincome	0.212		0	1
gini-lorenz	0.779		0	1
conc-index	0.160		0	1
indep	0.645		0	1

Table A.14
The association between background characteristics and respondents' choices.

	CIM_b	CIM_m	CIM_t	$PIRHT_b$	$PIRHT_m$	$PIRHT_t$	$PIRHT_{b,nr}$	$PIRHT_{m,nr}$	$PIRHT_{t,nr}$	UM_b	UM_m	UM_t
female	0.008	0.012	−0.004	0.020	−0.075	−0.058	−0.040	0.027	−0.011	0.007	0.031	−0.004
	(0.886)	(0.830)	(0.940)	(0.709)	(0.173)	(0.289)	(0.449)	(0.621)	(0.837)	(0.901)	(0.564)	(0.941)
age	−0.003	0.000	0.005	−0.009	0.004	0.008	0.001	0.001	−0.001	−0.001	−0.002	−0.003
	(0.699)	(0.952)	(0.518)	(0.252)	(0.658)	(0.319)	(0.866)	(0.922)	(0.911)	(0.917)	(0.786)	(0.693)
econbus	0.039	0.010	0.116	0.291	0.052	0.114	0.055	0.077	0.107	0.048	−0.033	−0.014
	(0.659)	(0.250)	(0.183)	(0.001)	(0.551)	(0.193)	(0.510)	(0.381)	(0.197)	(0.563)	(0.701)	(0.860)
right	−0.145	−0.243	−0.227	−0.153	−0.168	−0.149	−0.185	−0.063	−0.149	−0.047	−0.140	−0.081
	(0.014)	(0.000)	(0.000)	(0.008)	(0.004)	(0.011)	(0.001)	(0.281)	(0.007)	(0.402)	(0.016)	(0.113)
high-past-income	0.058	0.036	−0.088	−0.108	−0.030	0.033	0.196	−0.009	0.133	0.109	0.005	0.027
	(0.517)	(0.683)	(0.324)	(0.220)	(0.740)	(0.709)	(0.022)	(0.923)	(0.118)	(0.203)	(0.958)	(0.730)
high-future-income	0.101	0.010	−0.009	0.029	0.098	0.033	0.001	0.086	−0.004	0.006	0.078	0.024
	(0.131)	(0.876)	(0.895)	(0.655)	(0.140)	(0.617)	(0.992)	(0.197)	(0.955)	(0.921)	(0.240)	(0.678)
BP-1-365-days	0.049	−0.049	−0.036	−0.101	−0.019	−0.037	0.031	−0.003	−0.040	−0.018	−0.056	−0.079
	(0.498)	(0.488)	(0.617)	(0.156)	(0.795)	(0.611)	(0.657)	(0.964)	(0.562)	(0.792)	(0.430)	(0.213)
BP-29-365-days	−0.113	−0.131	−0.009	−0.024	−0.092	−0.028	−0.097	0.042	−0.037	0.010	−0.062	−0.035
	(0.088)	(0.044)	(0.890)	(0.709)	(0.161)	(0.668)	(0.124)	(0.519)	(0.553)	(0.869)	(0.341)	(0.546)
gini-lorenz	−0.038	−0.056	−0.038	−0.014	−0.014	−0.009	0.057	0.026	0.086	0.037	−0.011	0.103
	(0.572)	(0.395)	(0.569)	(0.838)	(0.834)	(0.895)	(0.366)	(0.696)	(0.172)	(0.563)	(0.866)	(0.078)
conc-index	0.012	0.050	0.024	0.036	−0.023	−0.023	0.066	0.002	0.033	−0.143	−0.038	0.010
	(0.875)	(0.497)	(0.744)	(0.622)	(0.751)	(0.752)	(0.351)	(0.980)	(0.638)	(0.042)	(0.601)	(0.883)
const	0.604	0.609	0.466	0.640	0.563	0.415	0.566	0.476	0.599	0.654	0.759	0.835
	(0.010)	(0.008)	(0.045)	(0.005)	(0.016)	(0.075)	(0.011)	(0.041)	(0.007)	(0.003)	(0.001)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349	349
R^2	0.034	0.067	0.054	0.064	0.042	0.032	0.055	0.012	0.039	0.022	0.030	0.024
p	0.286	0.007	0.035	0.010	0.136	0.343	0.033	0.949	0.184	0.682	0.391	0.587

Note: Point estimates and p-values obtained from SUR models. The estimates in Tables A.14, A.15, and A.16 were obtained from a single SUR estimation. p shows the p-value of the F-test that all coefficients are jointly equal to zero. *gini-lorenz* equals 1 if the respondent indicated to be familiar with the gini-index and/or the lorenz-curve, and 0 otherwise.

Table A.15
The association between background characteristics and respondents' choices.

	TI_{CIM} $_{i,-}$	TI_{CIM} $_{i,+}$	TI_{CIM} $_{h,-}$	TI_{CIM} $_{h,+}$	TI_{CIM} $_{h,gl}$	UI_{CIM} $_{i,-}$	UI_{CIM} $_{i,+}$	UI_{CIM} $_{h,-}$	UI_{CIM} $_{h,-,sf}$	UI_{CIM} $_{h,+}$	UI_{CIM} $_{h,+,sf}$	UI_{CIM} $_{h,gl}$	UI_{CIM} $_{h,gl,sf}$
female	0.006	0.025	−0.034	−0.002	−0.020	0.022	−0.054	0.026	0.022	0.036	−0.022	0.006	0.007
	(0.871)	(0.545)	(0.400)	(0.967)	(0.623)	(0.580)	(0.167)	(0.505)	(0.588)	(0.365)	(0.568)	(0.881)	(0.869)
age	0.006	−0.004	−0.008	−0.003	0.001	0.001	−0.003	0.001	0.002	−0.001	−0.003	−0.006	−0.004
	(0.277)	(0.544)	(0.199)	(0.630)	(0.892)	(0.838)	(0.653)	(0.826)	(0.790)	(0.836)	(0.625)	(0.357)	(0.537)
econbus	−0.002	−0.030	0.044	−0.002	0.004	−0.011	0.095	0.056	0.007	0.144	−0.006	−0.018	−0.020
	(0.973)	(0.655)	(0.494)	(0.976)	(0.951)	(0.864)	(0.127)	(0.365)	(0.917)	(0.024)	(0.917)	(0.777)	(0.753)
right	0.014	0.023	0.013	−0.054	−0.003	0.033	−0.030	−0.049	−0.008	−0.008	0.014	−0.006	0.043
	(0.730)	(0.605)	(0.759)	(0.231)	(0.949)	(0.434)	(0.471)	(0.243)	(0.853)	(0.848)	(0.730)	(0.892)	(0.320)
high-past-income	0.014	0.066	0.093	0.005	0.111	−0.000	0.071	0.013	0.008	0.012	0.037	−0.071	0.076
	(0.818)	(0.331)	(0.159)	(0.937)	(0.096)	(0.996)	(0.264)	(0.836)	(0.905)	(0.856)	(0.548)	(0.281)	(0.244)
high-future-income	0.021	−0.003	−0.063	0.018	−0.074	0.035	−0.099	−0.001	−0.035	0.004	−0.041	−0.013	−0.010
	(0.631)	(0.952)	(0.203)	(0.719)	(0.137)	(0.461)	(0.036)	(0.976)	(0.471)	(0.928)	(0.369)	(0.786)	(0.844)
BP-1-365-days	0.035	0.010	−0.033	0.028	−0.019	0.012	0.058	−0.048	0.003	−0.009	0.030	0.009	0.011
	(0.470)	(0.860)	(0.534)	(0.608)	(0.728)	(0.816)	(0.258)	(0.349)	(0.960)	(0.857)	(0.543)	(0.866)	(0.827)
BP-29-365-days	0.039	0.049	0.037	0.053	−0.021	0.079	0.033	0.095	0.042	0.059	0.056	0.036	0.050
	(0.377)	(0.325)	(0.448)	(0.297)	(0.676)	(0.094)	(0.475)	(0.042)	(0.382)	(0.221)	(0.217)	(0.459)	(0.298)
gini-lorenz	0.079	0.073	0.054	0.080	0.046	0.115	0.062	0.050	0.159	0.056	0.083	0.075	0.045
	(0.074)	(0.146)	(0.266)	(0.119)	(0.355)	(0.016)	(0.190)	(0.295)	(0.001)	(0.243)	(0.073)	(0.126)	(0.352)
conc-index	0.019	0.097	0.076	0.036	0.082	0.017	0.101	0.054	0.069	0.024	0.054	0.058	0.043
	(0.704)	(0.082)	(0.163)	(0.525)	(0.135)	(0.754)	(0.053)	(0.302)	(0.203)	(0.660)	(0.288)	(0.285)	(0.430)
const	0.632	0.825	0.956	0.807	0.804	0.688	0.757	0.749	0.652	0.675	0.834	0.909	0.862
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	349	349	349	349	349	349	349	349	349	349	349	349	349
R^2	0.019	0.028	0.034	0.020	0.024	0.031	0.057	0.027	0.046	0.027	0.028	0.022	0.018
p	0.753	0.450	0.299	0.751	0.605	0.358	0.026	0.497	0.095	0.482	0.452	0.665	0.787

Note: Point estimates and p-values obtained from SUR models. The estimates in Tables A.14, A.15, and A.16 were obtained from a single SUR estimation. p shows the p-value of the F-test that all coefficients are jointly equal to zero.

Table A.16
The association between background characteristics and respondents' choices.

	TI_{UM} $_{i,-}$	TI_{UM} $_{i,+}$	TI_{UM} $_{h,-}$	TI_{UM} $_{h,+}$	TI_{UM} $_{h,gl}$	UI_{UM} $_{i,-}$	UI_{UM} $_{i,+}$	UI_{UM} $_{h,-}$	UI_{UM} $_{h,-,sf}$	UI_{UM} $_{h,gl}$	UI_{UM} $_{h,gl,sf}$
female	0.018	−0.028	0.002	−0.006	−0.053	0.021	0.052	0.009	0.007	−0.020	0.037
	(0.703)	(0.537)	(0.969)	(0.899)	(0.275)	(0.655)	(0.275)	(0.844)	(0.875)	(0.677)	(0.434)
age	−0.010	−0.002	−0.004	0.006	−0.001	−0.010	0.002	0.001	−0.004	−0.006	0.010
	(0.163)	(0.785)	(0.609)	(0.372)	(0.888)	(0.159)	(0.825)	(0.908)	(0.543)	(0.405)	(0.158)
econbus	0.022	0.039	0.021	0.039	0.056	0.048	0.117	0.105	0.087	−0.082	−0.045
	(0.764)	(0.590)	(0.790)	(0.601)	(0.472)	(0.524)	(0.125)	(0.153)	(0.235)	(0.287)	(0.561)
right	0.046	0.087	0.014	0.041	−0.020	0.063	0.026	0.005	0.017	−0.051	0.041
	(0.351)	(0.070)	(0.791)	(0.401)	(0.700)	(0.204)	(0.604)	(0.921)	(0.727)	(0.316)	(0.425)
high-past-income	0.101	0.086	0.002	−0.004	0.081	−0.022	−0.071	0.049	−0.212	0.025	−0.004
	(0.178)	(0.243)	(0.980)	(0.962)	(0.306)	(0.774)	(0.365)	(0.515)	(0.005)	(0.749)	(0.962)
high-future-income	0.033	0.033	0.036	0.019	0.043	0.064	0.114	0.021	0.065	0.015	0.044
	(0.551)	(0.552)	(0.549)	(0.732)	(0.461)	(0.262)	(0.049)	(0.711)	(0.242)	(0.798)	(0.445)
BP-1-365-days	−0.028	0.017	−0.016	−0.043	−0.024	−0.054	−0.031	−0.020	−0.084	0.031	−0.057
	(0.644)	(0.771)	(0.804)	(0.481)	(0.705)	(0.380)	(0.619)	(0.734)	(0.161)	(0.617)	(0.360)
BP-29-365-days	−0.041	−0.050	0.026	−0.123	−0.024	0.085	0.030	−0.029	−0.052	−0.062	0.109
	(0.461)	(0.362)	(0.656)	(0.027)	(0.677)	(0.131)	(0.601)	(0.595)	(0.342)	(0.283)	(0.057)
gini-lorenz	0.061	−0.020	−0.040	0.049	0.040	−0.054	−0.045	−0.034	−0.012	−0.022	−0.023
	(0.275)	(0.714)	(0.501)	(0.385)	(0.492)	(0.345)	(0.441)	(0.538)	(0.824)	(0.709)	(0.691)
conc-index	0.028	−0.018	0.106	0.033	0.066	0.071	0.045	0.080	0.078	0.084	0.078
	(0.657)	(0.769)	(0.109)	(0.596)	(0.313)	(0.262)	(0.485)	(0.196)	(0.203)	(0.194)	(0.228)
const	0.910	0.785	0.812	0.613	0.721	0.959	0.611	0.696	0.871	0.973	0.557
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)	(0.006)
N	349	349	349	349	349	349	349	349	349	349	349
R^2	0.030	0.026	0.011	0.030	0.020	0.028	0.025	0.014	0.045	0.017	0.031
p	0.411	0.520	0.959	0.407	0.730	0.469	0.581	0.904	0.105	0.834	0.385

Note: Point estimates and p-values obtained from SUR models. The estimates in Tables A.14, A.15, and A.16 were obtained from a single SUR estimation. p shows the p-value of the F-test that all coefficients are jointly equal to zero.

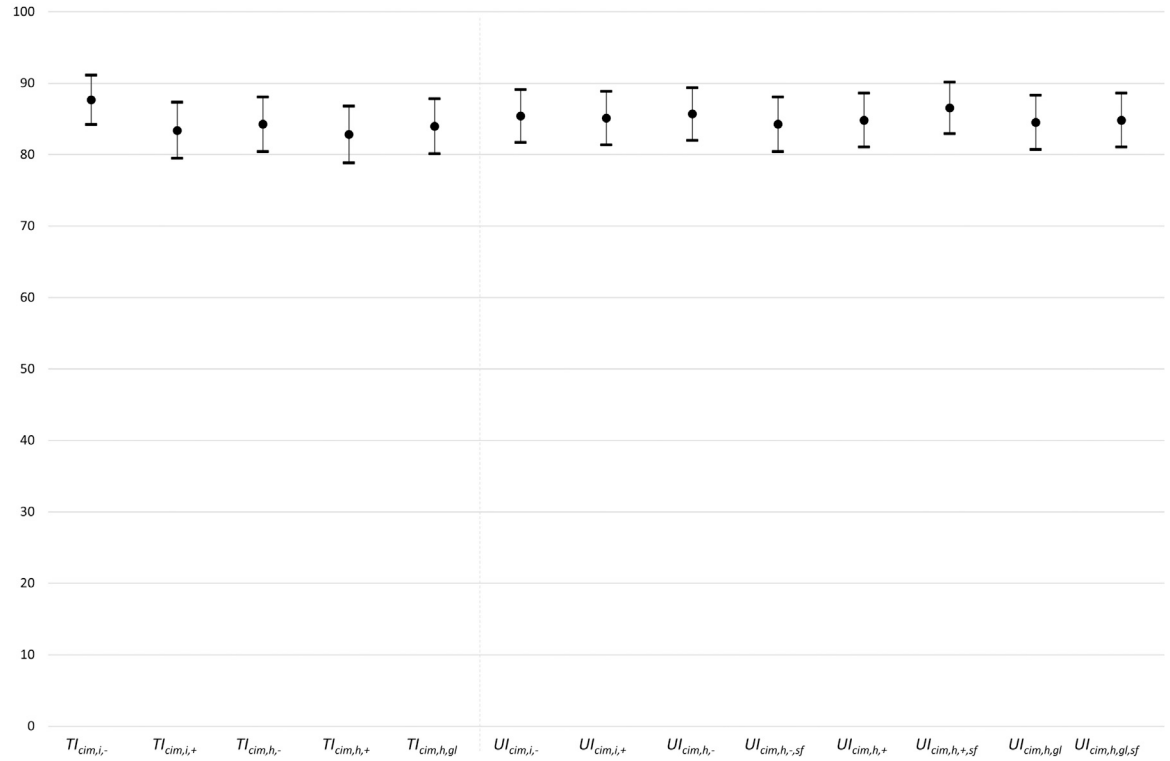

Fig. A.2. Proportion of respondents satisfying translation and unit invariance for CIM_m.
Note: Point estimates and 95% Confidence Intervals obtained from a constant-only SUR of linear probability models of whether the choices correspond to that of CIM_m.

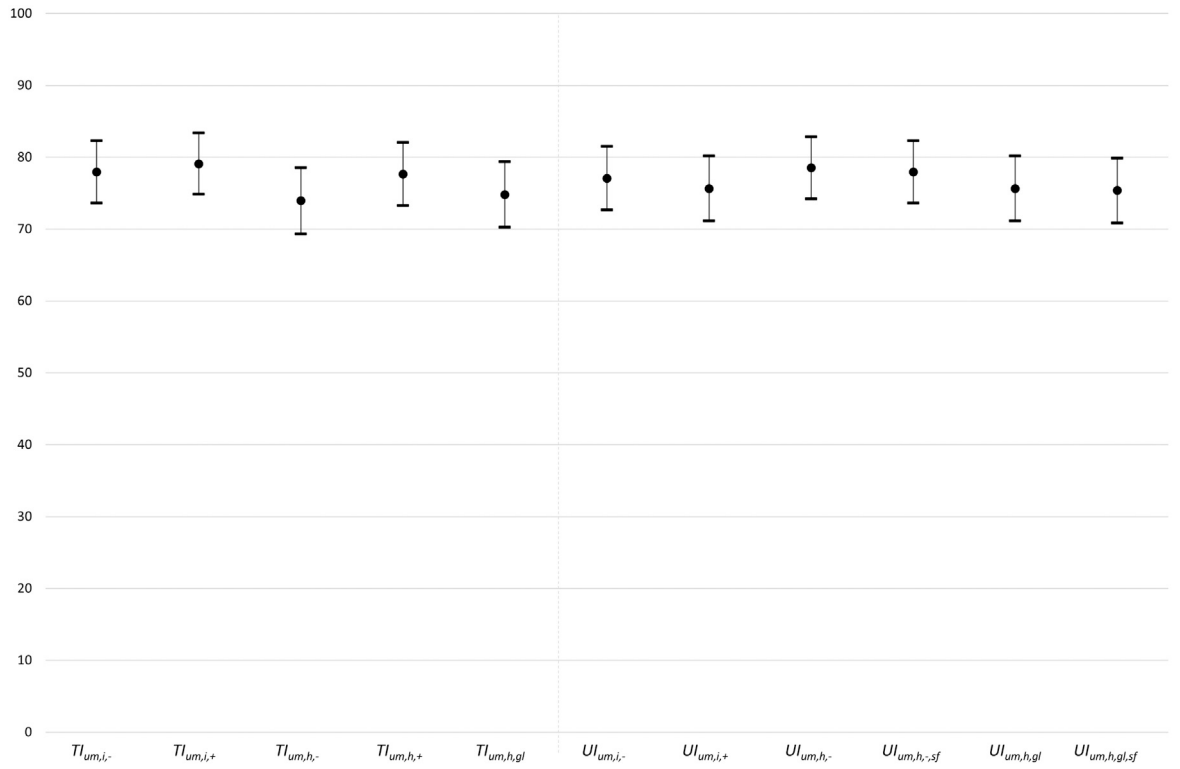

Fig. A.3. Proportion of respondents satisfying translation and unit invariance for UM_m.
Note: Point estimates and 95% Confidence Intervals obtained from a constant-only SUR of linear probability models of whether the choices correspond to that of UM_m.

Online appendix

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jhealeco.2023.102773.

A data repository including data and software allows replication of all findings and is available at https://doi.org/10.17632/k26f8rh58k.1.

References

Almås, Ingvild, Cappelen, Alexander W., Tungodden, Bertil, 2020. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than scandinavians? J. Polit. Econ. 128 (5), 1753–1788.

Amiel, Yoram, Cowell, Frank A., 1992. Measurement of income inequality: experimental test by questionnaire. J. Public Economics 47 (1), 3–26.

Amiel, Yoram, Cowell, Frank A., Gaertner, Wulf, 2009. To be or not to be involved: a questionnaire-experimental view on Harsanyi's utilitarian ethics. Soc. Choice Welf. 32, 299–316.

Atkinson, A.B, Bourguignon, F., 1982. The comparison of multi-dimensioned distributions of economic status. Rev. Econom. Stud. 49 (2), 183–201.

Attema, Arthur E., Bleichrodt, Han, Wakker, Peter P., 2012. A direct method for measuring discounting and QALYs more easily and reliably. Med. Decis. Making 32, 583–593.

Attema, Arthur E., l'Haridon, Olivier, Kuilen, Gijs van de, 2019. Measuring multivariate risk preferences in the health domain. J. Health Econ. 64, 15–24.

Bleichrodt, Han, Rohde, Kirsten, Ourti, Tom Van, 2012. An experimental test of the concentration index. J. Health Econ. 31 (1), 86–98.

Bleichrodt, H, Van Doorslaer, E., 2006. A welfare economics foundation for health inequality measurement. J. Health Econ. 25, 945–957.

Bosmans, Kristof, 2016. Consistent comparisons of attainment and shortfall inequality: a critical examination. Health Econ. 25, 1425–1432.

Bosmans, Kristof, Decancq, Koen, Ooghe, Erwin, 2015. What do normative indices of multidimensional inequality really measure? J. Public Econ. 130, 94–104.

Bosworth, Barry, 2018. Increasing disparities in mortality by socioeconomic status. Annu. Rev. Public Health 39, 237–251.

Chetty, Raj, Stepner, Michael, Abraham, Sarah, Lin, Shelby, Scuderi, Benjamin, Turner, Nicholas, Bergeron, Augustin, Cutler, David, 2016. The association between income and life expectancy in the United States, 2001–2014. JAMA 315 (16), 1750–1766.

Decoster, André, Minten, Thomas, Spinnewijn, Johannes, 2021. The income gradient in mortality during the Covid-19 crisis: Evidence from Belgium. J. Econ. Inequal. 19, 551–570.

Erreygers, Guido, 2009a. Correcting the concentration index. J. Health Econ. 28 (2), 504–515.

Erreygers, Guido, 2009b. Can a single indicator measure both attainment and shortfall inequality. J. Health Econ. 28 (4), 885–893.

Erreygers, Guido, Van Ourti, Tom, 2011. Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: [a] recipe for good practice. J. Health Econ. 30, 685–694.

Gaertner, Wulf, Schokkaert, Erik, 2012. Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice. Cambridge University Press.

Gajdos, Thibault, Weymark, John A., 2005. Multidimensional generalized gini indices. Econom. Theory 26, 471–496.

Gakidou, E., Murray, C., French, J., 2000. Defining and measuring health inequality. Bull. World Health Org. 78 (1), 42–54.
Hardardottir, Hjördis, Gerdtham, Ulf-G, Wengström, Erik, 2021. Parameterizing standard measures of income and health inequality using choice experiments. Health Econ. 30, 2531–2546.
Harper, S., Lynch, J., 2006. Measuring inequalities in health. In: Oakes, J., Kaufman, J. (Eds.), Methods in Social Epidemiology. Jossey-Bass, San Francisco.
Hurley, Jeremiah, Mentzakis, Emmanouil, Walli-Attaei, Marjan, 2020. Inequality aversion in income, health, and income-related health. J. Health Econ. 70, 102276.
Johansson-Stenman, Fredrik Carlsson, Daruvala, Dinky, 2002. Measuring future grandparents' preferences for equality and relative standing. Econom. J. 112, 362–383.
Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–292.
Kakwani, Nanak, Adam, Wagstaff, Van Doorslaer, Eddy, 1997. Socioeconomic inequalities in health: Measurement, computation, and statistical inference. J. Econometrics 77 (1), 87–103.
Khaled, M.A., Makdissi, P., Yazbeck, M., 2018. Income-related health transfers principles and orderings of joint distributions of income and health. J. Health Econ. 57, 315–331.
Kjellsson, Gustav, Gerdtham, Ulf-G, 2013. On correcting the concentration index for binary variables. J. Health Econ. 32 (3), 659–670.
Kolm, Serge-Christophe, 1977. Multidimensional egalitarianisms. Q. J. Econ. 91 (1), 1–13.
Konow, James, 2009. Is fairness in the eye of the beholder? An impartial spectator analysis of justice. Soc. Choice Welf. 33, 101–127.
Lambert, Peter, Zheng, Buhong, 2011. On the consistent measurement of attainment and shortfall inequality. J. Health Econ. 30, 214–219.
Le Clainche, C., Wittwer, J., 2015. Responsibility-sensitive fairness in health financing: Judgments in four European countries. Health Econ. 24 (4), 470–480.
Le Grand, Julian, 1987. Inequalities in health. Some international comparisons. Eur. Econ. Rev. 31, 182–191.
Mackenbach, J.P., Kunst, A.E., 1997. Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. Soc. Sci. Med. 44, 757–771.
Makdissi, P., Yazbeck, M., 2014. Measuring socioeconomic health inequalities in presence of multiple categorical information. J. Health Econ. 34, 84–95.
Nowlis, S.M., Kahn, B.E., Dhar, R., 2002. Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. J. Consum. Res. 29 (3), 319–334.
O'Donnell, O., O'Neill, S., Van Ourti, T., Walsh, B., 2016. Conindex: Estimation of concentration indices. Stata J. 16 (1), 112–138.
O'Donnell, O., Van Doorslaer, E., Wagstaff, A., Lindelow, M., 2008. Analyzing Health Equity using Household Survey Data: A Guide To Techniques and their Implementation. World Bank, Washington.
Schokkaert, E., Devooght, K., 2003. Responsibility-sensitive fair compensation in different cultures. Soc. Choice Welf. 21 (2), 207–242.
Schwandt, Hannes, Currie, Janect, von Wachter, Till, Kowarski, Jonathan, Chapman, Derek, Woolf, Steven H., 2022. Changes in the relationship between income and life expectancy before and during the COVID-19 pandemic, california, 2015–2021. JAMA 328 (4), 360–366.
Tarroux, Benoît, 2015. Comparing two-dimensional distributions: a questionnaire-experimental approach. Soc. Choice Welf. 44, 87–108.
The EuroQoL Group, 1990. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 6 (3), 199–208.
Tsui, Kai-yuen, 1995. Multidimensional generalizations of the relative and absolute inequality indices: The atkinson-kolm-sen approach. J. Econom. Theory 67, 251–265.
Tsui, Kai-yuen, 1999. Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation. Soc. Choice Welf. 16, 145–157.
Van Doorslaer, Eddy, Wagstaff, Adam, Bleichrodt, Han, Calonge, Samuel, Gerdtham, Ulf-G, Gerfin, Michael, Geurts, José, Gross, Lorna, Häkkinen, Unto, Leu, Robert E., O'Donnell, Owen, Propper, Carol, Puffer, Frank, Rodriguez, Marisol, Sundberg, Gun, Winkelhake, Olaf, 1997. Income-related inequalities in health: some international comparisons. J. Health Econ. 16 (1), 93–112.
van Raalte, Alyson A., Caswell, Hal, 2013. Perturbation analysis of indices of lifespan variability. Demography 50, 1615–1640.
Wagstaff, Adam, 2002. Inequality aversion, health inequalities and health achievement. J. Health Econ. 21 (4), 627–641.
Wagstaff, Adam, 2011. The concentration index of a binary outcome revisited. Health Econ. 20 (10), 1155–1160.
Wagstaff, A., Paci, P., Van Doorslaer, E., 1991. On the measurement of inequalities in health. Soc. Sci. Med. 33 (5), 545–557.
Wagstaff, Adam, Van Doorslaer, Eddy, 2004. Overall versus socioeconomic health inequality: a measurement framework and two empirical illustrations. Health Econ. 13, 297–301.
Williamson, Elizabeth J., Walker, Alex J., Bhaskaran, Krishnan, Bacon, Seb, Bates, Chris, Morton, Caroline E., Curtis, Helen J., Mehrkar, Amir, Evans, David, Inglesby, Peter, Cockburn, Jonathan, McDonald, Helen I., MacKenna, Brian, Tomlinson, Laurie, Douglas, Ian J., Rentsch, Christopher T., Mathur, Rohini, Wong, Angel Y.S., Grieve, Richard, Harrison, David, Forbes, Harriet, Schultze, Anna, Croker, Richard, Parry, John, Hester, Frank, Harper, Sam, Perera, Rafael, Evans, Stephen J.W., Smeeth, Liam, Goldacre, Ben, 2020. Factors associated with COVID-19-related death using opensafely. Nature 584, 430–436.
Yitzhaki, S., 1983. On an extension of the gini inequality index. Internat. Econom. Rev. 24 (3), 617–628.