What kind of inequality do you prefer? Evaluating measures of income and health inequality using choice experiments*

Hjördis Hardardottir[†], Ulf-G Gerdtham[‡] and Erik Wengström[†]

Abstract

When measuring inequality using conventional inequality measures, ethical assumptions about distributional preferences are often implicitly made. In this paper, we ask whether the ethical assumptions underlying the concentration index for income-related inequality in health and the Gini index for income inequality are supported in a representative sample of the Swedish population using an internet-based survey. We find that the median subject has preferences regarding income-related inequality in health that are in line with the ethical assumptions implied by the concentration index, but put higher weight on the poor than what is implied by the Gini index of income inequality. We find that women and individuals with a poorer health status put higher weight on the poor than men and healthier individuals. Ethically flexible inequality measures, such as the s-Gini index and the extended concentration index, imply that researchers have to choose from a toolbox of infinitely many inequality indices. The results of this paper are indicative of which indices (i.e. which parameter values) reflect the views of the population regarding how inequality should be defined.

Keywords: Socioeconomic inequality in health; Extended concentration index; S-Gini index; Distributional preferences; Income inequality

JEL Codes: I14; D9; D31; D63.

*The authors would like to thank the participants of the 15th CNEE workshop at the Department for Food and Resource Economics, University of Copenhagen, and the seminar participants at the Institute of Economic Studies at the University of Iceland for their useful comments. We are grateful for financial support from the Swedish Research Council (dnr 2014-646). The Health Economics Program (HEP) at Lund University also received core funding from the Government Grant for Clinical Research ("ALF") and Region Skåne (Gerdtham). Any remaining errors are the authors' own.

[†]Department of Economics, Lund University

[‡]Health Economics Unit, Department of Clinical Sciences, Lund University; Department of Economics, Lund University; Centre for Economic Demography, Lund University

1 Introduction

The importance of understanding the causes and consequences of inequality in income and health is hard to overstate; both income and health are important components of well-being. The increasing income gap in many countries around the world has been widely discussed, and differences in health across socioeconomic statuses have caught the attention of researchers. There are many reasons behind this large interest. For instance, inequality has been linked to malfunctioning political systems, lack of economic freedom, and lower economic growth (Persson & Tabellini, 1994; Stiglitz, 2013; Piketty, 2014). In addition, inequality has been considered a driver of environmental degradation, crime, poor health, and a range of other societal problems (Pickett & Wilkinson, 2009, 2015; Baland et al., 2007). It should be pointed out that many of these findings have been disputed and are associative rather than causal in nature (see, e.g., Forbes, 2000; Snowdon, 2010; Scheve & Stasavage, 2017). However, beyond these instrumental effects of inequality, there are also clear normative reasons to care about it. Most people simply find excessive inequality undesirable on moral grounds. In order to understand the causes and consequences of inequality and track its development over time, a transparent definition and measure of inequality is a prerequisite. Given the many faces of inequality, there is not one correct answer to the question of how it should be defined and measured. A closer look at the most common inequality measures reveals that they build on implicit ethical assumptions about how inequality is defined.

Two of the most common measures of income inequality are the Gini index for income inequality and the concentration index for income-related inequality in health.[1] Both measures implicitly assign weights over the underlying distribution that are symmetric around the median. Accordingly, both indices can be written as a sum of weighted health or income shares in which the weights depend on the rank of each individual in the underlying distribution such that the weight assigned to an individual of percentile rank 0.5 - x is the negative of the weight assigned to an individual of percentile rank 0.5 + x with $x \in [0, 0.5]$. This is problematic as there is no clear ethical argument for using the given weight structure. As shown by Donaldson & Weymark (1980) and Yitzhaki (1983) for the Gini index and by Wagstaff (2002) for the concentration index, one way to address this problem is to introduce a parameter to the index that determines the weight structure underlying the inequality measure. Altering the weights has been shown to have

[1] Hereafter, we will use income-related inequality in health synonymously with socioeconomic inequality in health.

important implications for how countries are ranked in terms of inequality (Wagstaff, 2002).

While the ethically flexible s-Gini developed by Donaldson & Weymark (1980) and Yitzhaki (1983) and the extended concentration index developed in Wagstaff (2002) both define a theoretical framework to address the lack of ethical flexibility in their original versions, they do not answer the question of which parametrization is the most appropriate to use. This is an ethical question that, to our knowledge, has not yet been addressed. Arguably, the parametrization used when inequality is measured using the s-Gini or the extended concentration indices should reflect the general views of the population as to how inequality should be defined and measured. Therefore, in this paper, we elicit preferences regarding the parametrization of the s-Gini for income inequality and the extended concentration index for income-related health inequality in a representative sample of the Swedish population. This implies that we estimate the weighting scheme each participant in our survey perceived as being appropriate when inequality is measured. For example, it might well be the case that people generally find it appropriate to give a higher weight to poor people than to rich people when inequality is measured.

Our results are indicative about which parametrization of the two indices best represents the view of how inequality should be defined and should therefore be of interest to academics and policymakers wishing to measure inequality in a way that is consistent with how the population thinks inequality should be defined. We are also one of the first studies to compare the health and income domains when it comes to views about inequality.[2] Given the close connection between the extended concentration index for income-related health inequality and the corresponding s-Gini for income inequality, it is interesting to see whether preferences regarding income-related inequality in health differ from preferences regarding income inequality.

We find that the ethical assumptions underlying the concentration index for income-related inequality in health are more or less in line with the median views of our representative sample. Thus, the symmetric weight structure of the concentration index is, hence, validated. However, for income inequality, we find that the median subject prefers to put higher weights on the poorer part of the distribution relative to the richer part than what is implied by the symmetric weight structure of the Gini index.

Although the focus of our study is on preferences regarding weighting schemes over the income distribution rather than inequality aversion, it re-

[2] Alessón & Tsuchiya (2014) compared inequality aversion and risk aversion in the domains of health and income. However, their setup is not comparable to ours. First, they studied inequality in losses, not in the distribution as a whole. Secondly, they studied pure inequality in health, while we study income-related health inequality.

lates to previous research on inequality aversion in the health and income domains, measured as the willingness to reduce inequality at the cost of efficiency.[3] The methods we use to elicit our ethical parameters are similar to the methods used in the literature to measure inequality aversion. However, our ethical parameters differ in one important aspect from measures of inequality aversion. While measures of inequality aversion measure the price people assign to inequality in terms of efficiency, our measure is neutral to efficiency and only captures how people weight different income groups when assessing inequality. Contrary to our results, in the literature that measures inequality aversion, people are generally found to be more inequality averse in the health domain than in the income domain. In the health domain, previous studies found that people are generally willing to give up considerable efficiency in order to reduce inequality. Dolan & Tsuchiya (2011) and Robson et al. (2017) quantified these preferences by measuring the Atkinson's ε and discovered a value of $\varepsilon = 10.95$ and $\varepsilon = 28.9$, respectively. Similar studies that investigated inequality aversion over income found a value for Atkinson's ε between 0.1 and 3 (Amiel et al., 2002; Carlsson et al., 2005; Pirttilä & Uusitalo, 2009). However, none of the papers that studied inequality aversion compared preferences in the health and income domains.[4]

Furthermore, our paper also relates to the literature on the heterogeneity of distributional preferences. Previous research has found distributional preferences and fairness ideals to vary with observable variables such as age, race, education level and education type (see for example Bellemare et al., 2008; Corneo & Fong, 2008; Cappelen et al., 2015; Fisman et al., 2015; Li et al., 2017). At the same time, Fisman et al. (2017) found that socioeconomic outcomes only predict a small share of the observed heterogeneity in preferences regarding the trade-off between equality and efficiency, while political preferences were shown to relate tightly to preferences regarding the equality-efficiency tradeoff. Our paper adds to this literature by studying how preferences regarding the weighting of different income groups relates to a large battery of variables that capture socio-economic status, life outcomes as well as attitudes and preferences. We find that our estimated ethical

[3] In addition to developing the extended concentration index, Wagstaff (2002) introduced an index of health achievement, which depends negatively on the extended concentration index and positively on the average level of health in society. The health achievement index hence captures a tradeoff between efficiency and equality, as does the Atkinson index of inequality. However, to our knowledge, no study has yet used the health achievement index as a basis for estimating inequality aversion.

[4] Alessón & Tsuchiya (2014) compared aversion to inequality in health losses and income losses between the income and health domains. Although their approach was slightly different from the other studies discussed here, in line with the aforementioned papers, the authors found that people were more averse to inequality in health than income inequality.

parameters are strongly related to survey-reported attitudes towards inequality and political views. We also report that women and individuals in poor health (self-reported and BMI) put higher weight on the poor than men and individuals in good health.

The paper is organized as follows. Section 2 illustrates the theoretical background of the study. In Section 3, the data is presented, and the methods are described. Section 4 presents the results of the paper, and Section 6 concludes.

2 Theoretical background

The Gini index and the concentration index are the two most widely used inequality measures for income and health, respectively. Both indices assume that transfers at the top of the distribution (e.g. a transfer from the richest person to the second richest person) reduce inequality as much as if the same transfers happen at the bottom of the distribution (e.g. from the second poorest person to the poorest person). Equalizing the two types of transfers implies that when inequality is measured with the Gini coefficient or the concentration index, an implicit ethical assumption is made. This ethical dilemma has been addressed by the extended concentration index developed by Wagstaff (2002) and in the s-Gini index developed by Donaldson & Weymark (1980) and Yitzhaki (1983) which allow for a more flexible weighting structure.

In the domain of income inequality, the s-Gini is defined as (written as the sum of weighted incomes):

$$G(y,\nu) = \frac{1}{\mu}\int_0^1 [1-\nu(1-p)^{\nu-1}]y(p)dp, \tag{1}$$

where $y(p)$ is the income of the individual with percentile rank p, and μ is the average income in the society.

In the domain of income-related health inequality, the extended concentration index is analogous to the s-Gini index and defined as (again written as the sum of weighted health measures):

$$C(h,\nu) = \frac{1}{\mu_h}\int_0^1 [1-\nu(1-p)^{\nu-1}]h(p)dp, \tag{2}$$

where $h(p)$ stands for the health of the individual with percentile rank p in the underlying distribution of income (or socio-economic status, more generally) and μ_h is the average of the health variable in the society. Given the two-dimensional nature of the concentration index, the concentration

curve, which corresponds to the Lorenz curve in the one-dimensional setting, can be above the 45-degree line, yielding values of the concentration index that are on $[-1, 0]$. This occurs when the measure of health (or ill-health) decreases with income. When health increases with income, the index takes values on $[0, 1]$ as in the standard one-dimensional case, since it is reduced to the Gini index.

The ν parameter allows for flexibility in the weighting structure in order to accommodate other ethical assumptions aside from the symmetric-weights assumption of the standard Gini and concentration indices. In the s-Gini and the extended concentration indices, the weights always sum to zero, with negative weights at the bottom of the underlying distribution and positive weights at the top of the distribution. The location at which the weights turn positive depends on the value of ν, which is defined on $]1, \infty[$[5] As the value of ν gets closer to 1, inequality matters less and less, and in the limit, when $\nu \to 1$, the value of the s-Gini and the extended concentration index approaches zero for all distributions. When $1 < \nu < 2$, the point at which the weights turn from negative to positive is above the median. The case when $\nu = 2$ corresponds to the conventional Gini and concentration indices in which the weights are symmetric around the median of the underlying distribution. When $\nu > 2$, the point at which the weights turn from negative to positive is below the median. The positioning of the point where the weights turn from positive to negative can be interpreted as the place in the distribution below which a rank-preserving payment or health increment decreases inequality. A rank-preserving transfer of income or health from a richer person to a poorer person will always decrease inequality as long as $\nu > 1$. However, the size of the inequality reduction varies with ν. Figure 1 illustrates the weights for the s-Gini and the extended concentration index for various values of the ν parameter. As can be seen in the figure, both the point at which the weights switch from negative to positive and the relative size of the weights depend on ν.

In theory, assuming that the size of the society studied approaches infinity, the standard Gini and concentration indices are defined on $[0, 1]$ and $[-1, 1]$, respectively. However, when applied to data, the bounds of the two indices are dependent on the size of the population. As can be seen from Figure 1, the smaller the sample studied, the lower weights in absolute terms that the richest and the poorest individuals receive (because the percentile rank of

[5]There is a generalization of the s-Gini in which ν also takes values on the range $]0, 1[$ (Gisbert et al., 2009). In this generalization, as ν decreases from 1, more and more weight is put on the richest individuals in society, and at the limit when ν approaches zero, all weight is put on the richest person. As will be discussed in Section 3, when we elicit the value of ν, we include one category for individuals with $\nu < 1$.

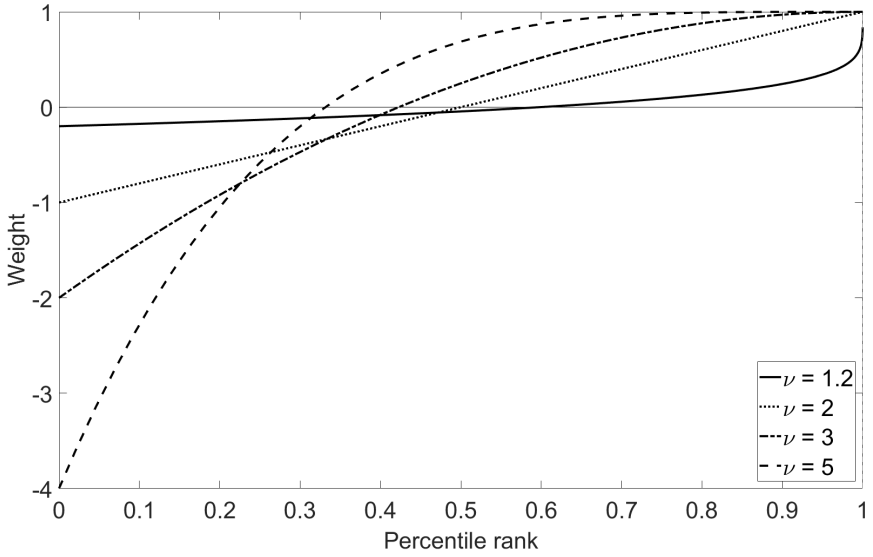

Figure 1: Weights as a function of percentile rank for different values of ν

these individuals differs from 1 and 0, respectively, when the sample is small). Since the maximum value of the index is precisely the value of the index when all income or health is accumulated at the richest person in society, the upper bounds of both the Gini index and the concentration index coincide with the weight attached to the richest person in society. Similarly, the lower bound of the concentration index corresponds to the value the index takes when all health is accumulated at the poorest person in society. However, the lower bound of the Gini index is always zero, irrespective of the sample size. In addition, regarding the concentration index, the definition of the health variable (in particular its upper and lower bounds) has been shown to determine the bounds of the index (Erreygers, 2009a,b; Wagstaff, 2009; Kjellsson & Gerdtham, 2013; Kjellsson et al., 2015).

The same reasoning holds when moving from the standard Gini and concentration indices to the s-Gini and the extended concentration indices. The bounds of the indices are defined by the weights attached to the richest and the poorest person. Since these weights depend on ν (see Figure 1), the bounds also depend on ν. While this is only an issue for the s-Gini index when the sample size is limited (since the weight attached to the richest person is lower than one only when the sample size is limited), in the case of the extended concentration index, the lower bound of the index is dependent on ν even when assuming an infinite sample. This can be seen in Figure 1 from the fact that the intercept of the weighting function with the y-axis (which

corresponds to the weight attached to the poorest person when the sample size is infinite) decreases with ν.

In this study, we assume that our proxy of health, life expectancy, is an unbounded variable that weakly increases with income.[6] Given these assumptions, the extended concentration index coincides with the s-Gini index.

Regarding the interpretation of the ν parameter, a high value of the ν parameter implies that high weight is put on the poorest and the share of the population to which a rank-preserving payment or health increment reduces inequality is reduced. Conversely, a low value of the ν parameter implies that relatively high weight is put on the richest and the share of the population to which a rank-preserving payment or health incremen reduces inequality is increased. Hence, a person that is concerned with reducing inequality at the bottom of the distribution rather than reducing inequality at the top has a high ν value, and a person that is more concerned with reducing inequality at the top of the distribution rather than at the bottom has a low ν value.

The value of ν can be put in relation to the Rawlsian theory of justice. A Rawlsian individual who follows the maximin rule will only care about the income or health of the worst off and hence will have an infinitely large ν value. On the other end of the spectrum, a health or income maximizer will be indifferent between all societies in the survey, since the average income or level of health is held constant across all societies presented. This illustrates the difference between the weighting parameter ν and more traditional measures of inequality aversions, such as Atkinson's ε. Atkinson's ε measures the extent to which individuals are willing to trade inequality with efficiency. When $\varepsilon \to \infty$, increasing inequality is infinitely expensive in terms of efficiency. Hence, all weight is put on the poorest individual in a Rawlsian manner. Thus, the interpretations of $\nu \to \infty$ and $\varepsilon \to \infty$ are similar. However, an individual with $\varepsilon \to 0$ places all weight on increasing efficiency and none on decreasing inequality in a health or income maximizing manner. While this individual would be indifferent between the societies in our survey where the average level of income or health is always the same, an individual with $\nu \to 1$ would place all (positive) weight on the richest person in society, caring only about inequality at the very top.

Duclos (2000) offered an alternative interpretation of the ν-value for the s-Gini. Given the structure of the weights of the s-Gini and the value the

[6]Some might argue that life expectancy is a bounded variable, given that, at least today, there is an upper bound on how old people can get. However, we argue that since there is no defined maximum age, it is more natural to treat life expectancy as an unbounded variable. Moreover, when eliciting the ν parameter, health always increases with income. In addition to making the parameter elicitation simpler, empirical evidence has indicated that health indeed increases with income.

s-Gini takes, $G(\nu)$ equals the expected relative deprivation (as a proportion of mean income) of the most deprived individual in a group of $\nu - 1$ persons. Duclos also proposed a thought experiment to assess reasonable values of the ν parameter. Making use of the fact that the s-Gini corresponds to a class of social welfare functions, a leaky bucket experiment as described in Okun (1975) could be implemented: Assume a rank-preserving transfer from a person with rank p_j in the distribution to a person with a lower rank p_i, where a certain share of the transfer "leaks out" due to, e.g., transaction costs or taxes. For a given transfer, the size of the leakage that is tolerated (such that social welfare does not decrease after the transfer) could be translated into a ν-value. In the thought experiment, Duclos concluded that a value of ν that lies in the interval $[1, 4]$ is reasonable, since higher values imply very high tolerance for leakage. Despite the experiment suggested by Duclos, a leaky bucket experiment to estimate individual values of the ν parameter has not been implemented.

3 Data and methods

The data were collected through an internet-based survey with the help of the Swedish survey company Enkätfabriken. Out of the 2340 invited to participate in the survey, 1060 subjects answered the survey, of which 923 completed the entire survey. The group of participants that completed the entire survey was representative of the Swedish population in the dimensions of gender, age (in the span of 18-65 years), and geographical location (NUTS 2). The participants received 10 SEK[7] in compensation for their participation in the survey, which they could choose to receive as a payment or to donate to charity. Prior to the main data collection, two pilot surveys were conducted. Furthermore, we excluded the subjects that were in the quickest 20% answering the survey, which eliminated all subjects who answered the entire survey in less than 9 minutes.[8] This reduced our sample to 755 participants. The median age of the respondents in the final sample was 41.8 years, and 52% of the respondents were women. Summary statistics of the main background variables can be found in Table 1.

The survey consisted of two parts that elicited (i) the ν parameter of the s-Gini for income inequality and (ii) the ν parameter of the extended

[7] 10 SEK equals 1$.

[8] In Section 7.3 of the Online Appendix, we perform a robustness analysis of this cut-off and find that our results are not sensitive to the cut-off at nine minutes. In Section 7.6 of the Online Appendix, we compare the background variables of the final sample to the excluded subjects.

Table 1: Summary statistics

		No. Obs	Mean	SD	Min	Max
Age		753	41.8	13.6	17	77
BMI		734	25.8	4.8	14.5	44.8
Has Kids		755	0.54	-	0	1
CR score		755	1.6	1.3	0	4
Health		755	3.5	1	1	5
Risk tolerance		755	6.2	2.2	1	11
Attitude to health inequality		755	7.2	2.5	0	10
Attitude to income inequality		755	7.9	2.3	0	10
Education:		755	%			
	Primary school	55	7.3			
	High school	392	51.9			
	University	308	40.8			
Sex	Man	361	47.8			
	Woman	389	51.5			
	Neither/non specified	5	0.7			

Notes: The variable "Health" mesures self-assessed health with the question "In your opinion, how is your health condition?" Possible answers are 1= very bad, 2=bad, 3=ok, 4=good, and 5=very good. The variable "Risk tolerance" is measured with the question "How willing or unwilling are you to take risk" where 0=Very unwilling to take risks and 10= Very willing to take risks. The two variables that measure attitudes to health and income inequality are based on the questions "How important is it to reduce income (health) inquality" where 0="Not at all important" and 10="Very important". CR score is the number of correct answers out of 6 on the cognitive reflection questions. Abnormal answers for height and weight have been removed (height below 112 cm and above 400 cm, one observation with weight = 133568 kg), as well as two abnormal answer for age (age=3 and age=198). In addition, there were 6, respectively 9 observations, where data is missing on height and weight.

concentration index for income-related inequality in health. The order in which the subjects answered the two parts was randomized. In addition, information about socioeconomic background, health status, preferences, and attitudes to redistribution was collected. Lastly, the subjects answered a series of cognitive reflection questions.

To elicit ν, the subjects were asked to imagine that they were consultants for the government of an imaginary country called Alfaland. Their task was to advise the government regarding the choice between two policy reforms that would result in two different distributions of health or income. For each parameter, the subjects answered five questions in which they chose between two resulting societies, society A and society B, described in terms of the average income (health) of the poorest third, the middle third, and the richest third of the society. In Figure 2, an example of a choice screen is presented. The average level of health or income was the same in all societies. Society A was constant in all questions, but society B varied from question to question. A consistent subject would shift only once from choosing the B society to choosing the A option, thus allowing us to determine an interval for the ethical parameter for the subject. For assessing the value of ν for income inequality, the average income of the three income groups varied between the B societies. For assessing the value of ν for income-related inequality in health, life expectancy of the three income groups was used as a measure of health and was varied between the B societies.

Before answering the main survey, the participants answered six questions that tested their comprehension of the survey setup. The questions consisted of reading a table similar to the tables used to present the Alfaland societies and answering questions about its content. The comprehension questions can be found in Section 7.7 of the Online Appendix.

We used life expectancy as our measure of health for two main reasons. First, it is a measure that is easy to understand, which was an important factor in our survey setting. Secondly, it is a continuous measure that has no defined upper bound. This allowed us to create variations in the health profiles of the societies in the survey as well as to disregard issues regarding measuring income-related inequality in health of bounded variables (Erreygers, 2009a,b; Wagstaff, 2009; Kjellsson & Gerdtham, 2013; Kjellsson et al., 2015). Both the levels of life expectancy and income used in the survey were chosen so that the resulting societies would not be too far from Swedish reality. Data on life expectancy from the Public Health Agency of Sweden and data on income distribution from Statistics Sweden were used for reference.

The reason we used an imaginary country in which the subjects acted as a consultant to the government was to induce choices that reflected the participants' opinions about which society was desirable, irrespective of their

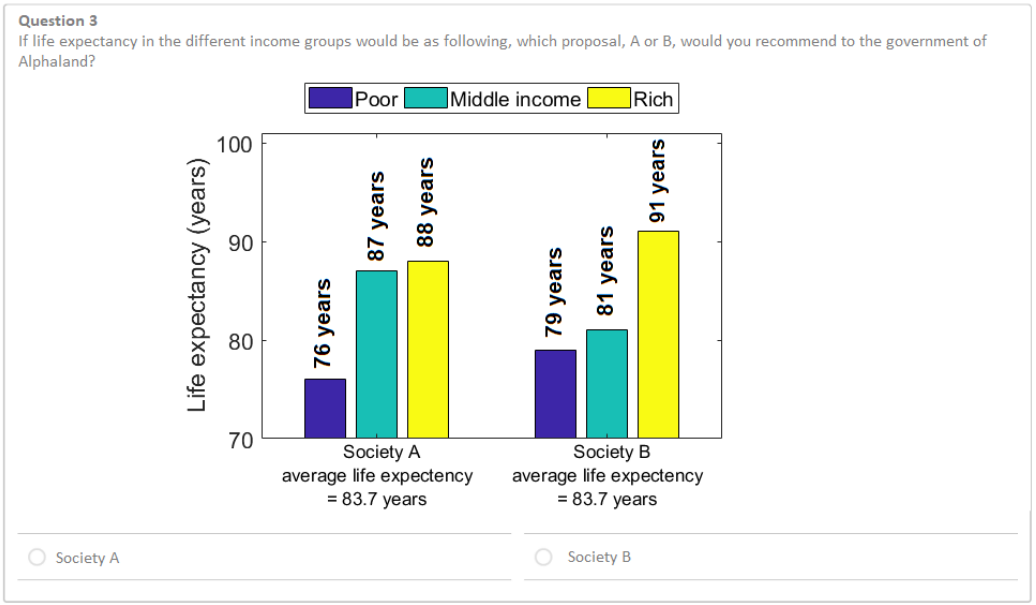

Figure 2: A screenshot from the survey eliciting ν for health (translated from Swedish)

own position in it. That is, we wanted participants to choose the society that they thought would generate the highest social welfare. Making participants choose between hypothetical societies minimized concerns that they would choose societies in which their own health or income would be maximized.

Regarding the definition of the s-Gini index (G) and the extended concentration index (C), a subject with the ethical parameter ν would be indifferent between societies A and B in terms of income inequality if the two societies had the same s-Gini index:

$$G(y_A, \nu) = G(y_B, \nu), \tag{3}$$

where y_A and y_B are the income distributions in the two societies.[9]. Similarly, a subject with the ethical parameter ν would indifferent between two societies, A and B, in terms of income-related inequality in health when

$$C(h_A, \nu) = C(h_B, \nu), \tag{4}$$

[9]We used the discrete version of the s-Gini and the extended concentration index presented in Wagstaff (2002), in which fractional ranks are used instead of percentile ranks to define the weights. Due to the low number of groups in our societies (only three income groups), we use the small-sample weights presented in Erreygers et al. (2012) when calculating the value of ν for which $A \sim B$ to correct for small-sample bias.

where h_A and h_B are the distributions of health over income in the two societies.

To control for the possibility that the initial health/income level of the poor and the rich, relative to the middle class in the B societies, would affect the elicited ν value, we had two versions of the survey. The only difference between the two versions was the marginal increase in health/income from the poorest third to the middle third compared to the marginal increase in health/income from the middle third to the richest third in the initial B society. The participants were randomly assigned to one of the two versions for both the income and health parts of the survey. In Tables 2 and 3, the A and B societies used in both versions of the survey, as well as the ν values that yielded indifference between A and B, are listed. As can be seen from the two tables, the ν value, which implies indifference between society A and society B, increases as we go from society B1 to society B5. This implies that the subjects with a higher ν value shifted later from society B to society A because they put higher weight on the poor relative to the rich than the subjects with a lower ν value.

Table 2: Societies used to elicit ν for income

ν income version 1	Min. income	Mean income	Max. income	ν-value if indifferent between A and B
Society A	14000	28000	35000	
Society B1	16500	25000	35500	<1
Society B2	16000	25000	36000	1.24
Society B3	15500	25000	36500	2
Society B4	15000	25000	37000	3
Society B5	14500	25000	37500	4.64
ν income version 2	Min. income	Mean income	Max. income	ν-value if indifferent between A and B
Society A	14000	28000	35000	
Society B1	18000	23000	36000	<1
Society B2	17000	23000	37000	1.53
Society B3	16000	23000	38000	2.56
Society B4	15500	23000	38500	3.25
Society B5	15000	23000	39000	4.21

The last part of the survey consisted of questions on the subjects' socioeconomic background and health as well as questions on preferences, values, and attitudes from the World Value Survey and from Falk et al. (2018). We also asked participants to answer a set of four cognitive reflection questions, borrowed from Toplak et al. (2014).

In particular, we collected information about self-reported health, active lifestyle, and workout habits. In addition, we asked subjects to report their political attitudes on a left-to-right scale, their opinions about redistribution,

Table 3: Societies used to elicit ν for health

ν health, version 1	Poor - life exp.	Mid - life exp.	Rich - life exp.	ν-value if indifferent between A and B
Society A	76	87	88	
Society B1	78.5	84	88.5	<1
Society B2	78	84	89	1.24
Society B3	77.5	84	89.5	2
Society B4	77	84	90	3
Society B5	76.5	84	90.5	4.64
ν health, version 2	Poor - life exp.	Mid - life exp.	Rich - life exp.	ν-value if indifferent between A and B
Society A	76	87	88	
Society B1	80.5	81	89.5	<1
Society B2	80	81	90	1.24
Society B3	79	81	91	2
Society B4	78	81	92	3
Society B5	77	81	93	4.64

their opinions about sources of success, and their views on the role of the government. We also asked subjects about their attitudes towards reducing inequality in both health and income, attitudes to punishment, and whether they give to charity. Lastly, we asked about their time and risk preferences. All the posed questions are listed in Section 7.7 of the Online Appendix.

The study was pre-registered in the AEA RCT registry. Furthermore, it has been approved by the local ethical research committee in Sweden (Etiksprövningsnämnden in Lund).

4 Results

In this section, we present our results in two steps. First, we provide an overview of the estimates of ν for the income domain, ν_{income}, and the health domain, ν_{health}. Second, we explore individual heterogeneity by relating the ν measures to a wide range of background variables using interval regressions. Throughout the results section, we only include consistent answers with one switching point between society B and society A. This leads to a reduction of our sample by around 28%, but at the end of this section, we present evidence suggesting that this has a very limited impact on our results.

4.1 General results

There are systematic and and significant differences in the attitudes towards inequality between the health and income domains. Figure 3 displays the distributions of the estimated parameter values for ν_{income} and ν_{health}. The

estimates presented in the figure clearly indicate that the distribution of ν values is skewed more towards lower values for health than for income. That is, the subjects generally put more weight on reducing the inequality among the poor in the income domain than in the health domain. Kolmogorov-Smirnov tests confirm that the data comes from different distributions (p-value<0.001) and that the distributions of both ν_{health} and ν_{income} are not uniform (p-value<0.001 for both distributions). We note that a relatively large share of the subjects reported answers that correspond to the two extremes of our scale for ν. However, the main difference between the income and the health domains remains when two extreme choices are excluded. In Section 7.5 of the Online Appendix, we address this issue further and conclude that the observed pattern does not seem to be driven by noise or to be an artifact of the survey design.

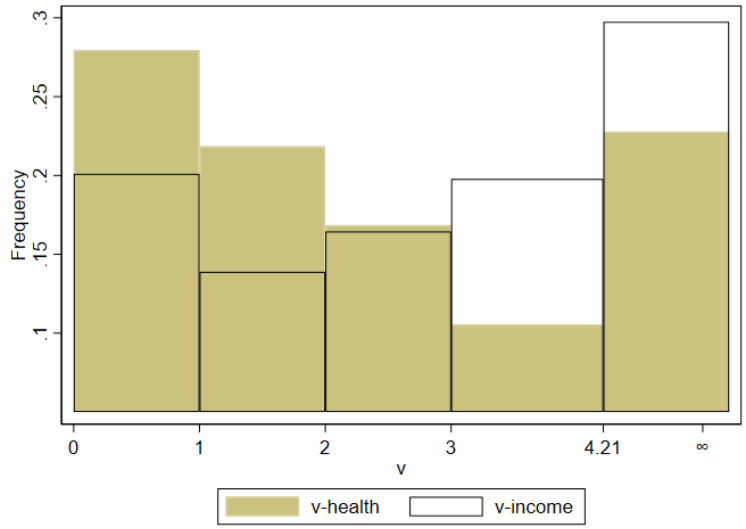

Figure 3: Distribution of ν for income and health

Notes: For ease of visualization, we used the same ν-intervals across elicitation tasks. To account for the fact that version 2 of the income elicitation task used slightly different ν intervals, we made reallocations between the bins assuming a uniform distribution. The detailed underlying distributions can be found in Tables 5 and 6.

It is also of interest to analyze whether the difference we observed in the aggregate also holds within subjects. For all the subjects who answered consistently for both ν_{health} and ν_{income}, we calculate the number of subjects with $\nu_{income} < \nu_{health}$, $\nu_{income} = \nu_{health}$, and $\nu_{income} > \nu_{health}$, respectively. We perform this comparison for the entire sample of subjects consistent in

both ν_{health} and ν_{income}. There were, in total, 423 subjects who answered consistently to both the income survey and the health survey. The results are presented in Table 4. From the table, it can be seen that when we study differences in ν_{income} and ν_{health} within subjects, it is more common for subjects to have ν_{income} larger than ν_{health}, than vice versa.

Table 4: The difference in estimated values of ν_{income} and ν_{health} within subjects.

	Percent (%)
$\nu_{income} - \nu_{health} < 0$	29.08
$\nu_{income} - \nu_{health} = 0$	26.95
$\nu_{income} - \nu_{health} > 0$	43.97

More details about the distributions of the estimated ν parameters are presented in Tables 5 and 6. Here, the responses are broken down by the two different survey versions. We can see that the median subject has a ν_{health} value in the interval $[2, 3]$ and a ν_{income} in the interval [3,3.25] (obtained by overlapping the median intervals for version 1 and version 2). Hence, our parameter estimates indicate that for health, the value 2 that corresponds to symmetric weights when the s-Gini or the extended concentration index is calculated is inside the obtained interval. On the other hand, for the parameter estimate for ν_{income}, our estimates do not include the value 2. This indicates that the implicit ethical principles governing the concentration index for income-related inequality in health are in line with the views of the median subject, whereas this is not the case for the Gini index for income inequality. Regarding the income domain, our results indicate that the median subject preferred to put more weight on inequality at the bottom of the distribution compared to what is implied by the Gini index.

Table 5: Parameter estimates of the ν parameter for health

	Version 1			Version 2			Total		
	Freq	**%**	**Cum %**	**Freq**	%	**Cum %**	**Freq**	%	**Cum %**
$\nu < 1$	66	24.91	24.91	85	30.91	30.91	151	27.96	27.96
$1 \leq \nu < 1.24$	22	8.30	33.21	36	13.09	44.00	58	10.74	38.70
$1.24 \leq \nu < 2$	31	11.70	44.91	29	10.55	54.55	60	11.11	49.81
$2 \leq \nu < 3$	44	16.60	61.51	47	17.09	71.64	91	16.85	66.67
$3 \leq \nu < 4.64$	31	11.70	73.21	32	11.64	83.27	63	11.67	78.33
$\nu \geq 4.64$	71	26.79	100.00	46	16.73	100.00	117	21.67	100.00
	N=265			N=275			N=540		

Table 6: Parameter estimates of the ν parameter for income

Version 1				Version 2			
	Freq	**%**	**Cum %**		**Freq**	**%**	**Cum %**
$\nu < 1$	44	17.46	17.46	$\nu < 1$	66	22.37	22.37
$1 \leq \nu < 1.24$	14	5.56	23.02	$1 \leq \nu < 1.53$	27	9.15	31.53
$1.24 \leq \nu < 2$	17	6.75	29.76	$1.53 \leq \nu < 2.56$	39	13.22	44.75
$2 \leq \nu < 3$	38	15.08	44.84	$2.56 \leq \nu < 3.25$	58	19.66	64.41
$3 \leq \nu < 4.64$	51	20.24	65.08	$3.25 \leq \nu < 4.21$	41	13.90	78.31
$\nu \geq 4.64$	88	34.92	100.00	$\nu \geq 4.21$	64	21.69	100.00
N=252				N=295			

Total

	Freq	**%**	**Cum %**
$\nu < 1$	110	20.11	20.11
$1 \leq \nu < 1.24$	14	2.56	22.67
$1 \leq \nu < 1.53$	27	4.94	27.61
$1.24 \leq \nu < 2$	17	3.11	30.71
$1.53 \leq \nu < 2.56$	39	7.13	37.84
$2 \leq \nu < 3$	38	6.95	44.79
$2.56 \leq \nu < 3.25$	58	10.60	55.39
$3 \leq \nu < 4.64$	41	7.50	62.89
$3.25 \leq \nu < 4.21$	51	9.32	72.21
$\nu \geq 4.21$	64	11.7	83.91
$\nu \geq 4.64$	88	16.09	100
N=547			

4.2 Heterogeneity of fairness perceptions

In order to explore the heterogeneity of our parameter estimates, we regress the ν-intervals on a large set of explanatory variables, including socioeconomic variables, survey-based variables on attitudes, beliefs, and preferences, and results on comprehension questions and cognitive reflection questions. Since the parameter estimates are an interval, we employ interval regressions. Furthermore, in the regressions, we exclude subjects who reported abnormal values for their age, weight, and height[10] as well as subjects who reported their gender to be neither man nor woman or who did not want to disclose their gender. In addition, there were six and nine observations, respectively, in which data on length and height was missing. This left us with a sample of 728 subjects.

The full set of coefficient estimates are presented in Tables 7-9 of the Online Appendix. We proceed by presenting the estimates in graphs. Figure 4 visualizes the coefficient estimates and corresponding confidence intervals of

[10]Two observations with abnormal ages were removed (age =3 and age=198), and 13 observations with abnormal heights were removed (11 observations with height≤112 cm. and two observations with height>400 cm.). One observation with abnormal weight was removed (133568 kg).

a set of socio-economic background variables and basic preference measures. The mean ν-estimates are relatively stable across different groups of the population, but a few noteworthy exceptions exist. Men preferred lower ν-values for both income and health, whereas subjects with a high BMI preferred higher ν-values. This implies that women and those with a high BMI placed more weight on inequality at the bottom than men and those with a lower BMI did. Education and cognitive reflection are significantly related to lower ν-values for income but not for health. These differential effects for income and health can be seen as reassuring, since they suggest that the measures are not generally biased due to some cognitively-demanding aspect of the elicitation format. If that was the case, we would have expected both measures to be affected to a similar extent. This view is also supported by the fact that we find no relationship between ν and performance on the pre-elicitation survey comprehension questions. We also note that the different versions of the elicitation format led to different choices of ν. Finally, we find that the survey-based estimate of risk aversion is not related to the ν estimates, indicating that risk aversion is not a driving force behind the attitudes towards ν.

Figure 5, displays the relationship between ν and self-reported health, workout habits and active lifestyle. We observe that self-reported health is not related to the parameter estimates. Active lifestyle and workout habits, on the other hand, are negatively related to ν_{health} but not to ν_{income}. The lack of correlation between self-reported health and ν_{health} can be seen as reassuring since it indicates that subjects do not mirror their own health status onto Alfaland and make choices focusing only on individuals with health-status that is similar to their own. Rather, the results suggest that subjects also count in the health of the inhabitants of Alfaland whose health is better or worse than their own.

The negative relationship between ν_{health} and active lifestyle and workout habits suggests that individuals who have an active lifestyle and work out attach a lower weight to those who are worse off than individuals who have a less active lifestyle and exercise less. In an additional specification not reported here, we included a control for self-reported health in the regressions for active lifestyle and workout habits. This did not qualitatively affect the relationship between active lifestyle and ν_{health} and between workout habits and ν_{health} , which indicates that these relationships are not driven by actual health status but by lifestyle choices. One potential interpretation is that people that lead an active life and exercise regularly consider health to be highly influenced by individuals' own lifestyle choices and behaviors, causing them to care less about reducing inequality at the lower end of the distribution. To investigate such potential mechanisms further, we make use of our

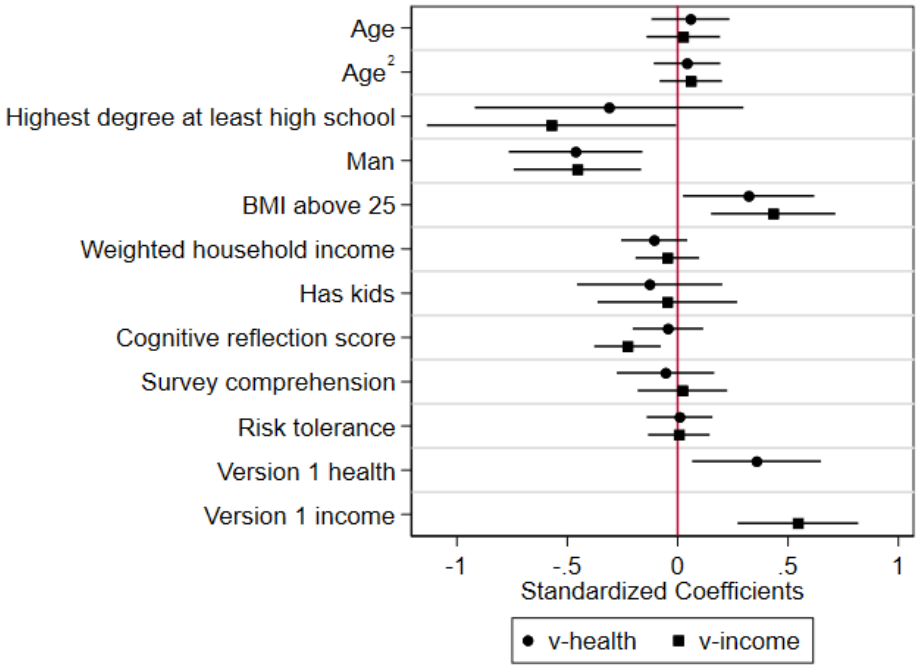

Figure 4: ν parameters regressed on background variables.

Notes: The figure is based on the coefficient estimates from the interval regression with the full set of coefficients presented in Table 7 in the Online Appendix. The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized, except for the dependent variables ν_{health} and ν_{income} and the dummy variables. The cognitive reflection score is defined as the number of correct answers on the four cognitive reflection questions. The survey comprehension is defined as the number of correct answers on the six survey comprehension questions. The questions used to elicit the cognitive reflection score, the pre-elicitation survey comprehension questions, and the question on risk tolerance can be found in Section 7.7 of the Online Appendix.

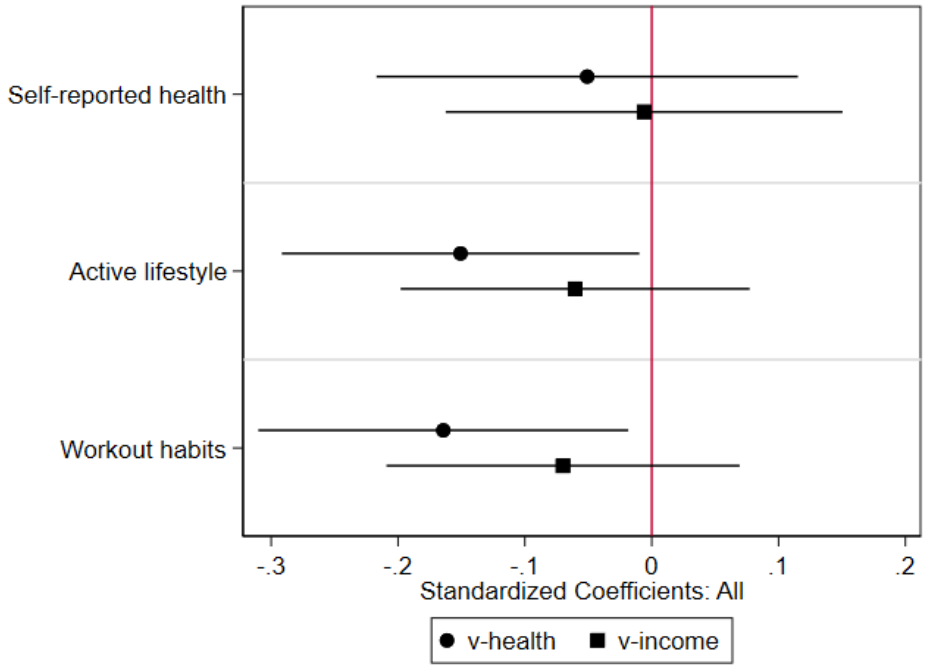

Figure 5: ν parameters regressed on individual health variables.

Notes: The figure is based on the coefficient estimates from the interval regression with the full set of coefficients presented in Table 8 in the Online Appendix. The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized, except for the dependent variables ν_{health} and ν_{income} In the regressions, we control for the set of background variables from Figure 4 and regress ν_{health} and ν_{income} on the health variables one at a time. The independent variables are self-reported health, active lifestyle, and workout habits, elicited on a scale from 1 to 4 (health and active lifestyle) and 1 to 6 (workout habits). The exact phrasing of the survey questions can be found in Section 7.7 of the Online Appendix.

large battery of attitudinal survey questions.

In Figure 6, we present regression estimates shedding light on the relationship between the ν parameters and attitudes towards politics and responsibility. In the figure, the ν values are regressed on each attitude variable one at a time (always with the same set of control variables, as in Figure 4). The first two variables capture attitudes toward reducing inequality, and the following three variables capture political attitudes. All of them are predictive of the value of ν for both health and income such that being less supportive of reducing inequality, more to the right in politics, less supportive of a reduction in the wage gap, and less in favor of an extensive social security system correlates with lower estimates of both ν_{health} and ν_{income}. These findings are quite intuitive and indicate that peoples' ideological orientation is related to their attitudes toward inequality. Furthermore, the subjects that were more willing to give to charity have, on average, higher estimates of ν for health and income. Finally, the results also point out that it does not seem to hold true that people who consider income and health outcomes to be mostly driven by individuals' choice of effort and lifestyle differ in their attitudes towards inequality. This goes against the argument of the previous paragraph that attitudes towards the sources of health could have driven the link between active lifestyle and workout habits and ν rather than health.

4.3 Selection effects

The above analysis uses only consistent answers in which the participants did at most one switch from society A to society B, but this reduction of the sample does not seem to affect results. The share of consistent answers was 71.5% for ν_{health} and 72.5% for ν_{income}, which seems relatively high given the general-population sample and the complexity of the task. Yet, a natural concern when subjects are eliminated from the data is that the representativeness of the original sample is lost. In the Online Appendix, we compare the distribution of age and gender of the subjects that responded consistently for both ν_{health} and ν_{income} to the distribution of age and gender in the Swedish population in 2016. This comparison shows that there is no detectable difference between the sample of consistent subjects and the distribution in Sweden in 2016 in the dimensions of gender and age (18-65 years). Thus, we conclude that the sample used in the analysis that follows is representative of the Swedish population in terms of gender and age. We also demonstrate that the entire sample of complete answers was representative in terms of the geographical location of the subjects. Unfortunately, we only have aggregated data on the geographical location of the subjects. Therefore, although the sample of consistent answers used in the analysis is

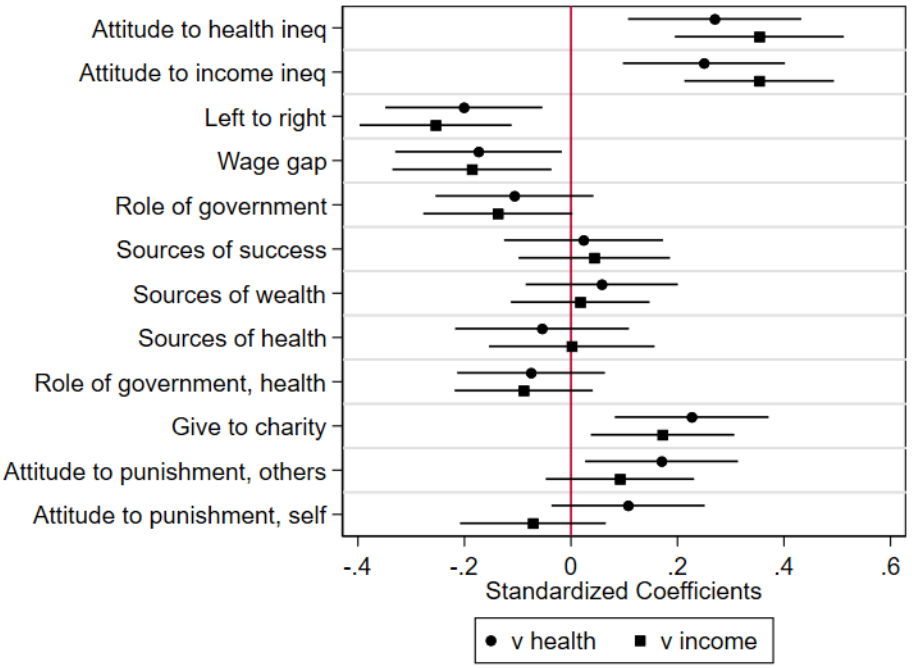

Figure 6: ν parameters regressed on background variables.

Notes: The figure is based on the coefficient estimates from the interval regression with the full set of coefficients presented in Table 9 in the Online Appendix. The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health}, ν_{income}. In the regressions, we control for the set of background variables from Figure 4 and regress ν_{health} and ν_{income} on the attitude variables one at the time. All attitude variables are defined on a scale from 0 or 1 to 10. The exact phrasing of the survey questions can be found in Section 7.7 of the Online Appendix.

heterogeneous in terms of geographical location, we cannot assert that it is representative of the Swedish population.

Another potential concern is that despite the fact that the sample we use is representative with respect to age and gender, it might be the case that answering in a consistent way is linked to other individual characteristics which are also correlated with the preferences towards inequality. In order to control for this, in the Online Appendix, we re-estimate the distribution of ν_{health} and ν_{income}, as well as the regressions presented in Figures 4-6 on a weighted sample. The weights are constructed as the inverse of the probability of being consistent based on each subject's background and attitude variables (the variables included in the regressions presented in Figures 4 and 6, respectively). Although our weights are based on extensive information on socioeconomic background as well as political views and attitudes towards inequality, our results show only very small differences between the results presented in Tables 5 and 6 and Figures 4-6 and the corresponding weighted results. This suggests that the results of the paper are not driven by selection in terms of observable background characteristics and attitudes.

5 Discussion

Taken together, our results show that people assess income inequality and income-related inequality in health differently, such that more weight is put on the poor part of the distribution when income inequality is assessed than when income-related inequality in health is assessed. This result goes against the general result of the experimental literature on inequality aversion in which estimated inequality aversion has been found to be higher for health than for income (Amiel et al., 2002; Carlsson et al., 2005; Pirttilä & Uusitalo, 2009; Dolan & Tsuchiya, 2011; Alessón & Tsuchiya, 2014; Robson et al., 2017). This result is rather surprising at first glance because it seems natural to be more averse to inequality in health than income. For example, Anand (2002) argued that inequality aversion in the health domain should be higher than in the income domain for two reasons. First, health is a special good with both an intrinsic and instrumental value, while income only has instrumental value. Secondly, while income inequalities might be justifiable (e.g. as a trade-off for higher overall efficiency), a similar argument is hard to make in the case of health inequality.

The fact that our estimates for ν are higher for income inequality than for income-related inequality in health is a result of our focus on eliciting how the subjects weighted different income groups when they assessed inequality, keeping average income or health constant. Consequently, we did not

measure how averse the subjects were to inequality but, rather, how they assessed it. Keeping this in mind, higher values for ν_{income} than ν_{health} suggest that subjects put relatively higher weights on the poor in the income domain than the health domain, or, equivalently, that subjects put relatively higher weights on the rich in the health domain than in the income domain. Thus, $\nu_{income} > \nu_{health}$ could be interpreted as a higher aversion to inequality at the top for health than for income. This view would make sense if, for instance, a subject did not like the idea that the rich can "buy health".

Although our measures of ν are not directly comparable to measures of individual inequality aversion of the Atkinson type, the issue of the dependence of parameter estimates and the method of elicitation is relevant for the elicitation of ν as well as for inequality aversion. In the literature that has elicited inequality aversion in the income domain using experimental methods, two main methods have been used. The first is a leaky bucket experiment, where aversion is measured as the tolerance for leakage in a transfer from a rich person to a poor person, and the second is a comparison of distributions or transfers, similar to the method used in this paper. Amiel et al. (2002) and Pirttilä & Uusitalo (2009) estimated inequality aversion from a survey using the leaky bucket design and found quite low values of inequality aversion. On the other hand, Carlsson et al. (2005) reported considerably higher values of inequality aversion using a survey in which distributions were compared. Additionally, Pirttilä & Uusitalo (2009) compared the results from the leaky bucket survey to a survey question in which two societies were compared and discovered that the latter type of question yielded much higher values of measured inequality aversion.

These results suggest that even in the case of ν, estimated values might be higher when elicited using a survey like ours in which societies are compared than when a leaky bucket approach is used to elicit ν. A possible explanation for this, as suggested by Carlsson et al. (2005), is that people are averse to the idea of redistribution, irrespective of the outcome. While the concept of redistribution from one given person to another is the core of the leaky bucket experiment, in our experiment, societies' post-redistributions were compared so that the exact extent of the redistribution was unknown. In light of the above suggestion in Duclos (2000) that a leaky bucket experiment is a possible method to elicit ν, eliciting ν with a leaky bucket approach and comparing the results to the outcome of this paper would be an informative next step in this line of research.

6 Conclusion

In this paper, we study attitudes towards inequality by estimating the ν parameter of the extended concentration index for income-related inequality in health and the s-Gini index for income inequality in a representative sample of the Swedish population. The use of inequality-measures that are ethically flexible creates the problem of infinitely many indices with different underlying ethical assumptions to choose from, each implying a different measure of inequality defined by the weights that different income groups obtain when inequality is measured. By estimating the ν parameter that determines the weighting structure underlying the s-Gini or the extended concentration index, our results provide an indication about which parameter values are reasonable to use when inequality is measured using the extended concentration index for income-related health inequality or the s-Gini for income inequality.

While the standard versions of the concentration index and the Gini index implicitly assume $\nu = 2$ and thus assume a weighting structure that is symmetric around the median of the income distribution, we find that the median respondent in our survey has an estimated value of ν in the interval [2,3] for income-related inequality in health and in the interval [3, 3.25] for income inequality. Since the estimated interval for ν_{health} contains the value 2, this indicates that the implicit ethical assumptions of the extended concentration have the support of the median participant in our survey. For the income domain, the value 2 is outside of the estimated range of ν_{income} for the median participant, suggesting that the median participant preferred to put higher weight on the poorer parts of the distribution when assessing inequality than what is implicitly assumed by the standard Gini index.

We relate our estimates of ν to socioeconomic variables as well as to self-reported health, lifestyle, preferences, and attitudes. Our results show that the individual estimates of ν relate strongly to political attitudes and views on inequality. We also note that our individual estimates of ν correlate with gender, BMI, education, and cognitive ability measured as performance on a set of cognitive reflection questions.

While the theoretical framework of the extended concentration index and the s-Gini that allow for ethically flexible inequality indices have been available for quite some time, this is the first attempt to estimate the ethical parameter ν. The results shed light on the appropriability of the ethical assumptions about the weights different income groups receive when income-related inequality is measured with the concentration index or income inequality is measured with the Gini index.

In the paper, we study only one specific type of weight structure over the

income distribution, the one implied by the extended concentration index and the s-Gini. Other indices with flexible weight structure over the income distribution include Erreyger's symmetric index (Erreygers et al., 2012), in which variations in the ethical parameter reflect the level of sensitivity to changes in health at the extremes of the income distribution, relative to the middle of the distribution. Estimating different weight structures would add to our knowledge about how people assess inequality and thus could be a next step in this line of research.

Moreover, the paper opens up the question of how preferences regarding inequality in health relate to preferences regarding income inequality. In a follow-up paper, we will study Atkinson's based inequality aversion in the health and income domains and how the two relate to each other and a set of background variables.

Another topic for future research is to examine how strongly the parameter estimates depend on the Swedish context of the study. In addition to the subject-pool being representative of the Swedish population, the different societies used in the elicitation of the parameters were chosen to be as close to Swedish reality (or potential reality) as possible. Investigating how the parameter estimates depend on the societies used in the elicitation, as well as performing a similar study in a different country, would be an interesting topic for future research.

References

Alessón, I. A. & Tsuchiya, A. (2014). Comparing inequality and risk aversion in health and income: An empirical analysis using hypothetical scenarios with losses. *Papeles de trabajo del Instituto de Estudios Fiscales. Serie economía*, (1), 3–30.

Amiel, Y., Creedy, J., & Hurn, S. (2002). Measuring attitudes towards inequality. *The Scandinavian Journal of Economics*, *101*(1), 83–96.

Anand, S. (2002). The concern for equity in health. *Journal of Epidemiology & Community Health*, *56*(7), 485–487.

Baland, J.-M., Bardhan, P., & Bowles, S. (2007). *Inequality, cooperation, and environmental sustainability*. Princeton University Press.

Bellemare, C., Kröger, S., & van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, *76*(4), 815–839.

Cappelen, A., Nygaard, K., Sørensen, E., & Tungodden, B. (2015). Social preferences in the lab: A comparison of students and a representative population. *The Scandinavian Journal of Economics*, *117*(4), 1306–1326.

Carlsson, F., Daruvala, D., & Johansson-Stenman, O. (2005). Are people inequality-averse, or just risk-averse? *Economica*, *72*(287), 375–396.

Corneo, G. & Fong, C. M. (2008). What's the monetary value of distributive justice. *Journal of Public Economics*, *92*(1-2), 289–308.

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

Donaldson, D. & Weymark, J. A. (1980). A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, *22*(1), 67 – 86.

Duclos, J.-Y. (2000). Gini indices and the redistribution of income. *International Tax and Public Finance*, *7*(2), 141–162.

Erreygers, G. (2009a). Correcting the concentration index. *Journal of Health Economics*, *28*(2), 504 – 515.

Erreygers, G. (2009b). Correcting the concentration index: A reply to wagstaff. *Journal of Health Economics*, *28*(2), 521 – 524.

Erreygers, G., Clarke, P., & Ourti, T. V. (2012). Mirror, mirror, on the wall, who in this land is fairest of all? Distributional sensitivity in the measurement of socioeconomic inequality of health. *Journal of Health Economics*, *31*(1), 257 – 270.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692.

Fisman, R., Jakiela, P., & Kariv, S. (2017). Distributional preferences and political behavior. *Journal of Public Economics*, *155*, 1–10.

Fisman, R., Jakiela, P., Kariv, S., & Markovits, D. (2015). The distributional preferences of an elite. *Science*, *349*(6254), aab0096.

Forbes, K. J. (2000). A reassessment of the relationship between inequality and growth. *American economic review*, *90*(4), 869–887.

Gisbert, F. J., Lasso de la Vega, C., & Urrutia, A. (2009). Generalizing the s-Gini family: Some properties. *Working papers = Documentos de trabajo: Serie AD, Nº. 16, 2009*.

Kjellsson, G. & Gerdtham, U.-G. (2013). On correcting the concentration index for binary variables. *Journal of Health Economics*, *32*(3), 659–670.

Kjellsson, G., Gerdtham, U.-G., & Petrie, D. (2015). Lies, damned lies, and health inequality measurements: understanding the value judgments. *Epidemiology*, *26*(5), 673.

Li, J., Dow, W. H., & Kariv, S. (2017). Social preferences of future physicians. *Proceedings of the National Academy of Sciences*, *114*(48), E10291–E10300.

Okun, A. M. (1975). *Equality and efficiency, the big tradeoff*.

Persson, T. & Tabellini, G. (1994). Is inequality harmful for growth? *The American Economic Review*, *84*(3), 600–621.

Pickett, K. E. & Wilkinson, R. G. (2009). *The Spirit Level: Why more equal societies almost always do better*. London, United Kingdom: Allen Lane/Penguin Group UK.

Pickett, K. E. & Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine*, *128*, 316–326.

Piketty, T. (2014). *Capital in the 21st century*. Cambridge, MA: Harvard University Press.

Pirttilä, J. & Uusitalo, R. (2009). A 'leaky bucket' in the real world: Estimating inequality aversion using survey data. *Economica*, *77*(305), 60–76.

Robson, M., Asaria, M., Cookson, R., Tsuchiya, A., & Ali, S. (2017). Eliciting the level of health inequality aversion in England. *Health Economics*, *26*(10), 1328–1334.

Scheve, K. & Stasavage, D. (2017). Wealth inequality and democracy. *Annual Review of Political Science*, *20*, 451–468.

Snowdon, C. (2010). *The spirit level delusion: Fact-checking the left's new theory of everything*. North Yorkshire, UK: Little Dice.

Stiglitz, J. E. (2013). *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York, NY: W. W. Norton & Company.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, *20*(2), 147–168.

Wagstaff, A. (2002). Inequality aversion, health inequalities and health achievement. *Journal of Health Economics*, *21*(4), 627 – 641.

Wagstaff, A. (2009). Correcting the concentration index: A comment. *Journal of Health Economics*, *28*(2), 516 – 520.

Yitzhaki, S. (1983). On an extension of the Gini inequality index. *International Economic Review*, *24*(3), 617–628.

7 Online Appendix

The Online Appendix is organized as follows. In Section 7.1, we present the regression tables that correspond to Figures 4, 5 and 6. In Section 7.2, we present comparisons of the distribution of gender and age in the sample of consistent answers, the whole sample, and the true distribution of age and gender in Sweden in 2016. In Section 7.3, we present a robustness analysis with regards to the exclusion of survey answers that took less than nine minutes. In Section 7.5, we present an analysis of the mechanisms behind the relatively large share of extreme answers. In Section 7.6, we compare the socio-economic background of participants with consistent answers when ν was elicited to the socio-economic background of participants with inconsistent answers. Finally, in Section 7.7, we list all the background questions included as well as the cognitive reflection questions and the pre-elicitation survey comprehension questions.

7.1 Regression tables

Here, we present the regression tables that correspond to the results presented graphically in the paper.

Table 7: Background variables. The table corresponds to Figure 4 in the paper.

	(1)	(2)
	ν_{health}	ν_{income}
Age	-0.0142	-0.0240
	(0.036)	(0.034)
Age^2	0.000228	0.000320
	(0.000)	(0.000)
Highest degree at least high school	-0.311	-0.571**
	(0.310)	(0.288)
Man	-0.462***	-0.454***
	(0.154)	(0.147)
BMI above 25	0.322**	0.433***
	(0.152)	(0.144)
Weighted household income	-0.0794	-0.0348
	(0.057)	(0.055)
Kids dummy	-0.127	-0.0471
	(0.168)	(0.161)
Cognitive reflection score	-0.0351	-0.181***
	(0.065)	(0.061)
Comprehension score	-0.0330	0.0133
	(0.068)	(0.062)
Risk tolerance	0.00399	0.00256
	(0.034)	(0.032)
Version 1	0.357**	0.545***
	(0.149)	(0.139)
N	535	540

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Integral regression. Coefficients are not standardized. Risk preferences are survey based and are defined between 0 to 10, where 10 implies maximum risk tolerance. Cognitive reflection score is the number of correct answers out of 4 on the cognitive reflection questions. Comprehension score is the number of correct answers out of 6 on the pre-elicitation survey comprehension questions. The phrasing of the questions used to elicit risk tolerance, the cognitive reflection questions and the survey comprehension questions can be found in section 7.7 of the Online Appendix.

Table 8: The relationship between self reported health, active lifestyle and workout habits, and the estimated ν for health and income. The table corresponds to Figure 5 in the paper.

	(1)	(2)	(3)	(4)	(5)	(6)
	ν_{health}	ν_{income}	ν_{health}	ν_{income}	ν_{health}	ν_{income}
Self-reported health	-0.0510	-0.00596				
	(0.085)	(0.080)				
Active lifestyle			-0.148**	-0.0593		
			(0.071)	(0.069)		
Workout habits					-0.113**	-0.0480
					(0.051)	(0.049)
N	535	540	535	540	535	540

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimated using interval regression. Coefficients are not standardized. The background variables from Table 7 are included as control variables in all regressions. The independent variables are self-reported health, active lifestyle and workout habits elicited on a scale from 1 to 4 (health and active lifestyle) and 1 to 6 (workout habits). The exact phrasing of the survey questions can be found in section 7.7 of the Online Appendix.

Table 9: Attitude variables. The table corresponds to Figure 6 in the paper.

	ν_{health}	ν_{income}
Attitude to health inequality	0.108***	0.142***
	(0.033)	(0.032)
Attitude to income inequality	0.0989***	0.140***
	(0.031)	(0.028)
Left to right	-0.0877***	-0.111***
	(0.033)	(0.032)
Reduce wage gap	-0.0658**	-0.0704**
	(0.030)	(0.029)
Role of government	-0.0411	-0.0532*
	(0.029)	(0.028)
Sources of success	0.00908	0.0169
	(0.029)	(0.028)
Sources of wealth	0.0250	0.00750
	(0.032)	(0.029)
Sources of health	-0.0220	0.000620
	(0.034)	(0.032)
Role of government, health	-0.0273	-0.0323
	(0.026)	(0.024)
Give to charity	0.0807***	0.0612**
	(0.026)	(0.024)
Attitude to punishment, others	0.0682**	0.0369
	(0.029)	(0.028)
Attitude to punishment, self	0.0400	-0.0266
	(0.027)	(0.026)
N	535	540

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimated using interval regression. Coefficients are not standardized. The background variables from Table 7 are included as control variables in all regressions. ν_{income} and ν_{health} are regressed on the attitude variables one at the time. The independent variables are survey based measures of attitudes elicited on a scale from 0 or 1 to 10. The exact phrasing of the survey questions can be found in section 7.7 of the Online Appendix.

7.2 Representativeness of the sample of consistent answers

In order to assert that the sample of complete answers used to estimate ν_{health} and ν_{income} is similar to the population in Sweden in 2016 in terms of age and gender[11], we compare the distribution of age and gender in the sample of consistent answers used in the analysis for both ν_{health} and ν_{income} to the true distribution of age and gender in Sweden in 2016. Additionally, we present the geographical distribution of the whole sample of complete answers and compare it to the true distribution in Sweden in 2016. Unfortunately, we only have aggregated information about the geographical distribution of subjects. Therefore, we cannot assert that the sample of consistent answers used in the analysis was representative of the Swedish population in terms of the geographical location of subjects.

Figures 7 and 8 illustrate the difference between the sample of consistent answers in ν_{health} and ν_{income}, respectively, used in the analysis (hence excluding subjects who answered in less than nine minutes) and the distribution of age in Sweden in 2016. The original sample of subjects that completed the survey was representative of subjects aged 18-65 years. Therefore, we drop 4 observations in which the reported age is above 65 years when we compare the sample of consistent answers to the distribution of age in Sweden.

We note that, compared to the Swedish population, for both ν_{health} and ν_{income}, the distribution of consistent answers is slightly skewed towards older subjects. We test the null hypotheses that the distributions of consistent answers used in the analysis differ from the age distribution in Sweden. The results from a Kolmogorov-Smirnov test of the equality of the distributions show that for both ν_{health} and ν_{income}, we cannot reject the null hypothesis that the distributions are the same (p=0.529 for ν_{health} and p=0.727 for ν_{income}), nor can we reject the null hypothesis that the sample of consistent answers used in the analysis contains larger values than the distribution of age in Sweden in 2016 (p=0.27 for ν_{health} and p=0.385 for ν_{income}).

The share of women and men in Sweden in 2016 was 50% women, 50% men. In the sample of consistent answers used in the analysis, the share of men was slightly lower: 48.8% in the sample where ν_{health} was consistent and 49.6% in the sample where ν_{income} was consistent. In order to investigate whether the gender distribution of the sample of consistent answers differed significantly from 50%-50%, we perform a test of equality of proportions. The results of this test show that the null hypotheses that the proportion of

[11]The distribution of age, gender, and geographical location in Sweden in 2016 was used when the whole representative sample was created.

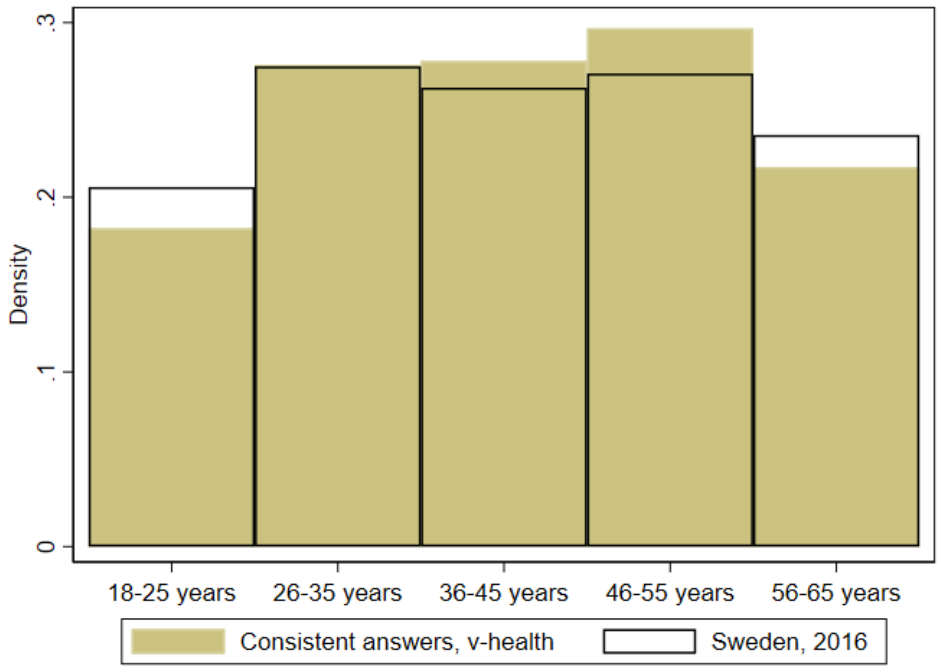

Figure 7: The distribution of age in the sample of subjects with consistent answers in terms of ν_{health} who use nine minutes or more to answer the whole survey, compared to the age distribution in Sweden in 2016.

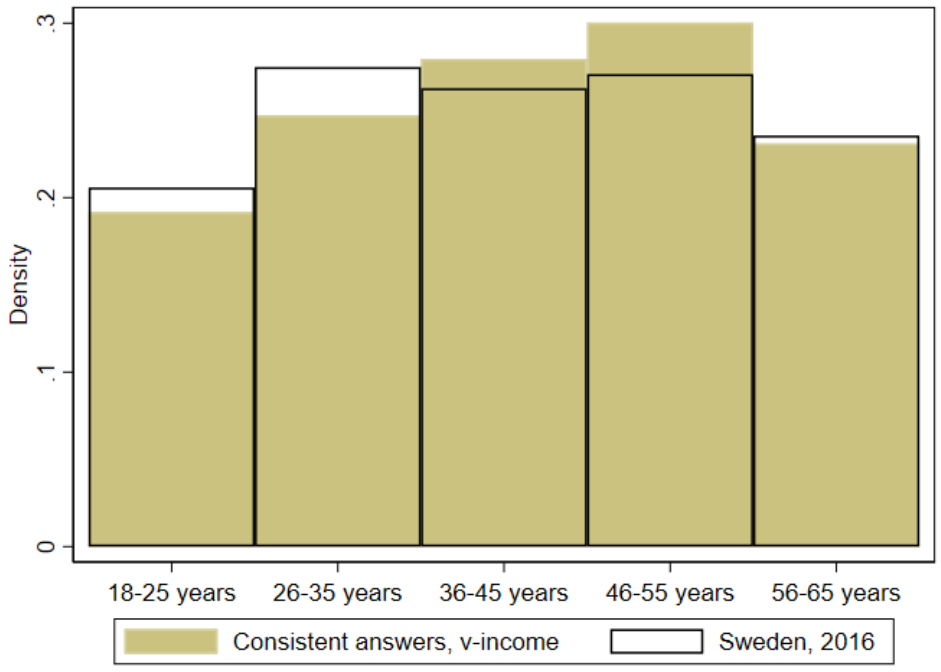

Figure 8: The distribution of age in the sample of subjects with consistent answers in terms of ν_{income} who use nine minutes or more to answer the whole survey, compared to the age distribution in Sweden in 2016.

men in the two consistent samples equals 0.5 cannot be refuted (p=0.6 for ν_{health} and p=0.86 for ν_{income}). Therefore, we conclude that the two samples of consistent answers used in the analysis of ν_{health} and ν_{income}, respectively, are representative of the Swedish population in terms of gender and age.

Although we only have data on the geographical location of subjects aggregated for the whole sample of consistent answers, in Table 10, we present the distribution of subjects in the whole sample according to the NUTS 2 standard and compare it to the same distribution in Sweden in 2016. From the table, it can be seen that the entire sample of complete answers very closely replicates the geographical distribution of inhabitants in Sweden in the year 2016. Even if we cannot assert the representativeness of the final sample of consistent answers in terms of the geographical distribution, the fact that the original sample was representative in the geographical dimension suggests that the final sample was heterogeneous in terms of the subjects' geographical location.

Table 10: The geographical distribution of inhabitants in Sweden in 2016 and the geographical distribution of subjects in the sample of whole, complete answers to the survey.

	Share in Sweden (%)	Share in whole sample (%)
Stockholm	23	21
Östra Mellansverige	17	17
Småland med öarna	8	9
Sydsverige	15	15
Västsverige	20	19
Norra Mellansverige	8	9
Mellersta Norrland	4	4
Övre Norrland	5	6

7.3 Robustness of cut-off at nine minutes

In order to study the robustness of our results regarding the exclusion of subjects that used less than nine minutes to answer the survey, we here present the main results of the paper without a time cut-off. In Table 11, we present the share of consistent answers in the nine minutes cut-off used in the paper compared to no cut-off in terms of answering time (but still only including subjects that answered the whole survey). From the table, a reduction in the share of consistent answers from around 72% to around 65% can be observed.

In Tables 12 and 13, we present the parameter estimates with no cut-off in terms of answering time. Comparing Tables 12 and 13 to Tables 5 and 6, only small differences can be seen in terms of the median parameter

Table 11: Number of consistent answers for the nine min. cutoff and without a cut-off.

	Nine min. cut-off		*No cut-off*	
	ν_{health}	ν_{income}	ν_{health}	ν_{income}
Total number of subjects	755	755	923	923
Number of consistent subjects	540	547	597	605
Share consistent	71.5%	72.5%	65%	65.5%

estimates. It is noteworthy that the bulk of the answers that were added when the cut-off at nine minutes was removed were answers in which either only society A was chosen or only society B was chosen.

Table 12: Parameter estimates of the ν parameter for health, no cut-off.

	Version 1			Version 2			Total		
	Freq	**%**	**Cum %**	**Freq**	%	**Cum %**	**Freq**	%	**Cum %**
$\nu < 1$	80	26.85	26.85	95	31.77	31.77	175	29.31	29.31
$1 \leq \nu < 1.24$	25	8.39	35.23	39	13.04	44.82	64	10.72	40.03
$1.24 \leq \nu < 2$	32	10.74	45.97	29	9.70	54.52	61	10.22	50.25
$2 \leq \nu < 3$	47	15.77	61.74	50	16.72	71.24	97	16.25	66.50
$3 \leq \nu < 4.64$	33	11.07	72.82	34	11.37	82.61	67	11.22	77.72
$\nu \geq 4.64$	81	27.18	100.00	52	17.39	100.00	133	22.28	100.00
N=298				N=299			N=597		

Figures 9 to 11 present coefficient estimates of the regressions presented in Figures 4 to 6, now including all subjects that finished the survey, irrespective of answering time. From the figures, no significant differences can be seen between the results in which there was a cut-off in terms of answering time and the results in which there was no cut-off.

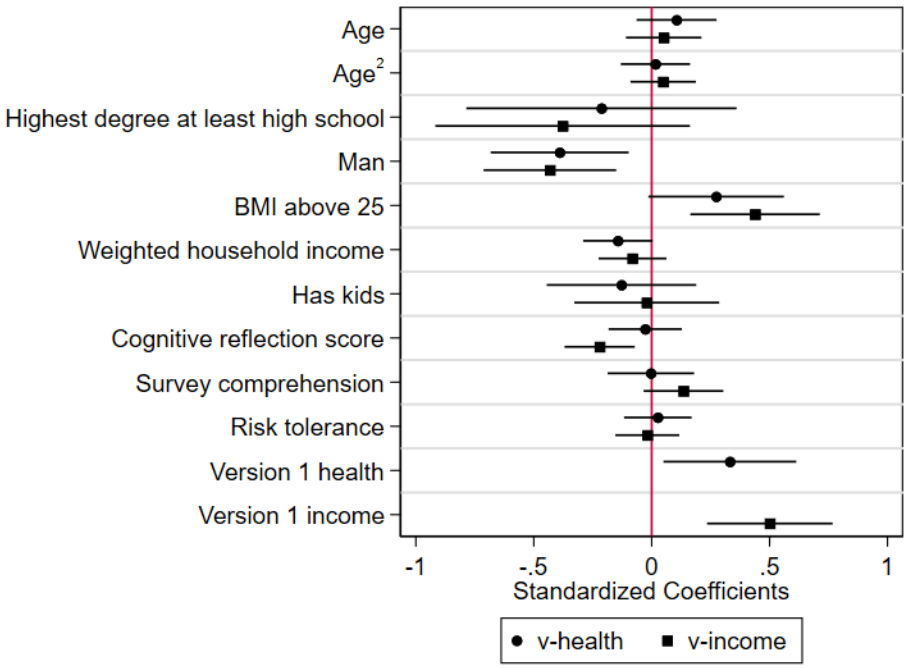

Figure 9: ν parameters regressed on background variables. The sample includes all subjects that finished the survey, irrespective of answering time.

Notes: The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health}, ν_{income} and the dummy variables. The cognitive reflection score is defined as the number of correct answers on the four cognitive reflection questions. The survey comprehension is defined as the number of correct answers on the six survey comprehension questions. The questions used to elicit the cognitive reflection score, the pre-elicitation survey comprehension questions, and the question on risk tolerance can be found in Section 7.7 of the Online Appendix.

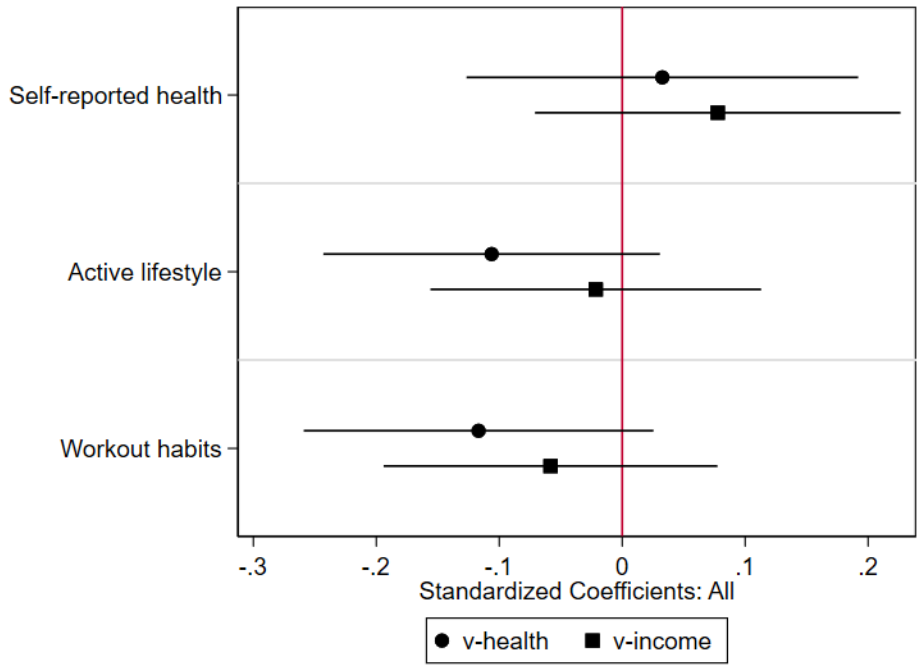

Figure 10: ν parameters regressed on individual health variables. The sample includes all subjects that finished the survey, irrespective of answering time.

Notes: The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health} and ν_{income} . In the regressions, we control for the set of background variables from Figure 4 and regress ν_{health} and ν_{income} on the health variables one at the time. The independent variables are self-reported-health, active lifestyle and workout habits elicited on a scale from 1 to 4 (self-reported health and active lifestyle) and 1 to 6 (workout habits). The exact phrasing of the survey questions can be found in section 7.7 of the Online Appendix.

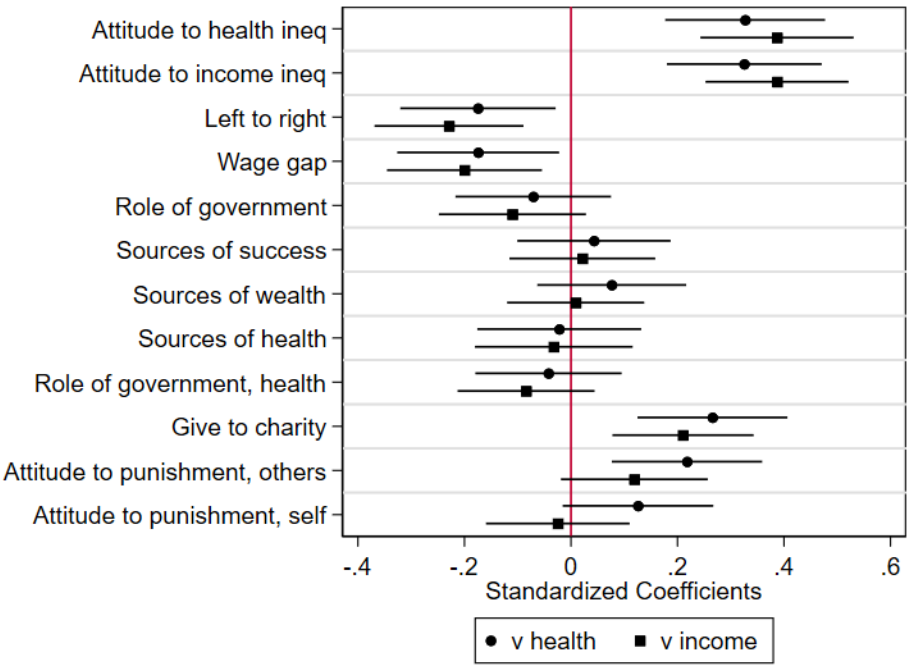

Figure 11: ν parameters regressed on attitude variables. The sample includes all subjects that finished the survey, irrespective of answering time.

Notes: The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health}, ν_{income}. In the regressions, we control for the set of background variables from Figure 4 and regress ν_{health} and ν_{income} on the attitude variables one at the time. All attitude variables are defined on a scale from 0 or 1 to 10. The exact phrasing of the survey questions can be found in section 7.7 of the Online Appendix.

Table 13: Parameter estimates of the ν parameter for income, no cut-off.

Version 1				Version 2			
	Freq	**%**	**Cum %**		**Freq**	**%**	**Cum %**
$\nu < 1$	56	19.72	19.72	$\nu < 1$	76	23.68	23.68
$1 \leq \nu < 1.24$	16	5.63	25.35	$1 \leq \nu < 1.53$	28	8.72	32.40
$1.24 \leq \nu < 2$	18	6.34	31.69	$1.53 \leq \nu < 2.56$	41	12.77	45.17
$2 \leq \nu < 3$	40	14.08	45.77	$2.56 \leq \nu < 3.25$	60	18.69	63.86
$3 \leq \nu < 4.64$	53	18.66	64.44	$3.25 \leq \nu < 4.21$	44	13.71	77.57
$\nu \geq 4.64$	101	35.56	100.00	$\nu \geq 4.21$	72	22.43	100.00
	N=284				N=321		

Total

	Freq	**%**	**Cum %**
$\nu < 1$	132	21.82	21.82
$1 \leq \nu < 1.24$	16	2.64	24.46
$1 \leq \nu < 1.53$	28	4.63	29.09
$1.24 \leq \nu < 2$	18	2.98	32.07
$1.53 \leq \nu < 2.56$	41	6.78	38.84
$2 \leq \nu < 3$	40	6.61	45.45
$2.56 \leq \nu < 3.25$	60	9.92	55.37
$3 \leq \nu < 4.64$	44	7.27	62.64
$3.25 \leq \nu < 4.21$	53	8.76	71.40
$\nu \geq 4.21$	72	11.90	83.31
$\nu \geq 4.64$	101	16.69	100
	N=605		

7.4 Possible selection in the sample of consistent answers

Although we do show in section 7.2 above that the sample of consistent answers cannot be distinguished from the Swedish population in terms of age and gender, it is possible that the subjects who are consistent differ from the subjects who are not consistent in dimensions other than age and gender. For example, it is possible that inconsistent subjects differ from consistent subjects in terms of background characteristics and/or political views and attitudes towards inequality. In that case, it is possible that the sample of inconsistent subjects has views on the weighting of income groups when inequality is measured that differs from the views of the sample of consistent subjects.

In order to control for this possibility, we re-estimate the results presented in Tables 5 and 6, as well as the results of the regressions presented in Figures 4-6, weighting each observation according to the inverse of the probability of being consistent on the health inequality questionnaire and the income inequality questionnaire, respectively, given the individual's background variables (the variables presented in Figure 4) and attitude variables (the variables presented in Figure 6). Hence, subjects with a low probability of being consistent get higher weights than subjects with a high probability of being consistent, given both their socio-economic background and their political views and attitudes towards inequality.

In order to create the weights for the estimates of ν_{health}, we estimate a probit regression with a dummy variable for being consistent in ν_{health} with the background and attitude variables as independent variables. We then predict the probability of being consistent in the health survey for each subject and finally define the individual weight as the inverse of this predicted probability. The procedure is identical in the case of ν_{income}.

In Tables 14 and 15, we present the ν estimates of the weighted sample. When comparing the two tables to the corresponding Tables 5 and 6 in the paper, we observe that weighting the data does not affect the distribution of ν estimates. In Figures 12-14, we present the results of the weighted regressions. Comparing the figures to the corresponding Figures 4-6 in the paper, we observe that the regression results with weighted data are very similar to the results in the paper.

7.5 Robustness analysis of extreme answers

Regarding Figure 3 in the paper, it is evident that a large share of responses for both ν_{health} and ν_{income} were responses in which either society A was

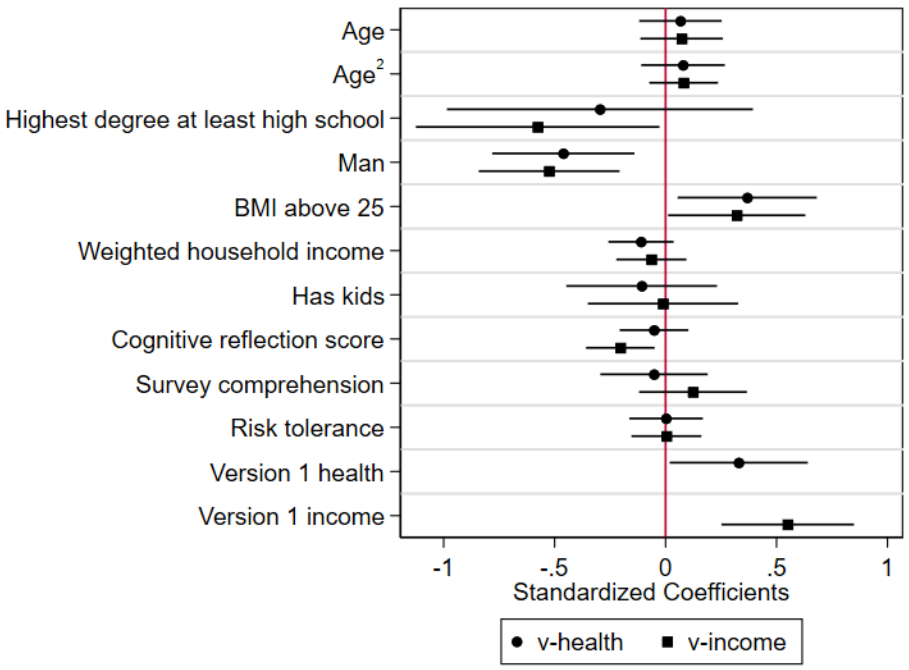

Figure 12: ν parameters regressed on background variables when subjects are weighted with the inverse of the probability of being consistent in answering the health- or the income survey based on their background and attitude variables.

Notes: The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health}, ν_{income} and the dummy variables. The cognitive reflection score is defined as the number of correct answers on the four cognitive reflection questions. The survey comprehension is defined as the number of correct answers on the six survey comprehension questions. The questions used to elicit the cognitive reflection score, the pre-elicitation survey comprehension questions, and the question on risk tolerance can be found in Section 7.7 of the Online Appendix.

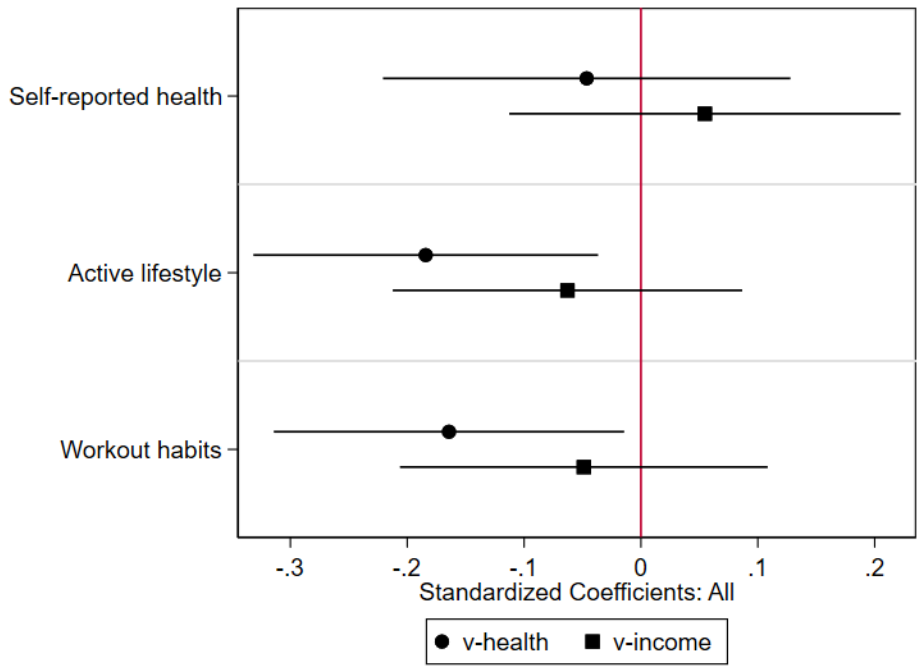

Figure 13: ν parameters regressed on individual health variables when subjects are weighted with the inverse of the probability of being consistent in answering the health- or the income survey based on their background and attitude variables.

Notes: The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health} and ν_{income} . In the regressions, we control for the set of background variables from Figure 4 and regress ν_{health} and ν_{income} on the health variables one at the time. The independent variables are self-reported-health, active lifestyle and workout habits elicited on a scale from 1 to 4 (self-reported health and active lifestyle) and 1 to 6 (workout habits). The exact phrasing of the survey questions can be found in section 7.7 of the Online Appendix.

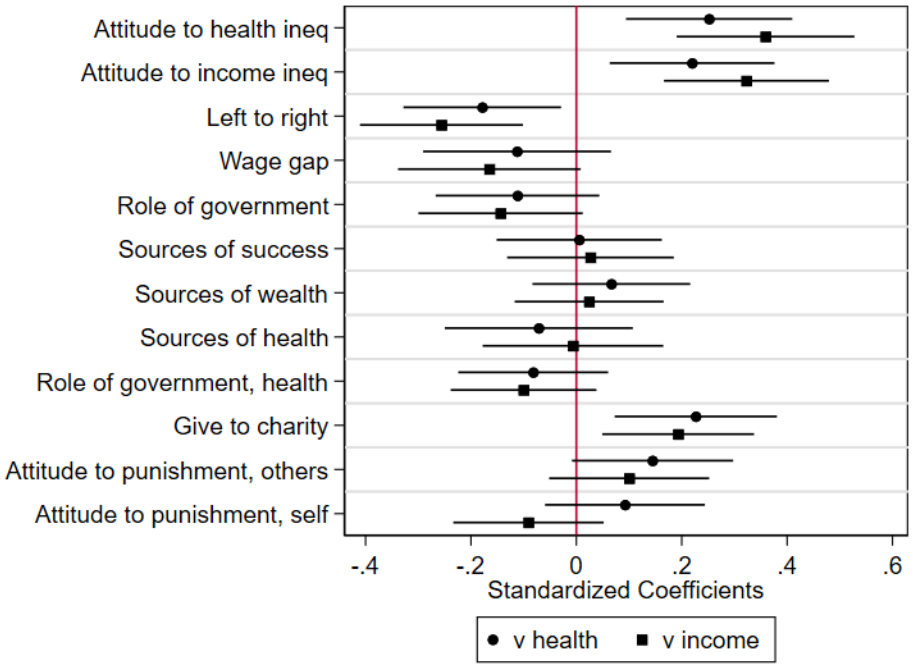

Figure 14: ν parameters regressed on attitude variables when subjects are weighted with the inverse of the probability of being consistent in answering the health- or the income survey based on their background and attitude variables.

Notes: The bars represent a 95% confidence interval for the coefficient estimates. The variables are standardized except from the dependent variables ν_{health}, ν_{income}. In the regressions, we control for the set of background variables from Figure 4 and regress ν_{health} and ν_{income} on the attitude variables one at the time. All attitude variables are defined on a scale from 0 or 1 to 10. The exact phrasing of the survey questions can be found in section 7.7 of the Online Appendix.

Table 14: Parameter estimates of the ν parameter for health when observations are weighted with the inverse of the probability of being consistent based on background and attitude variables.

	Version 1			Version 2			Total		
	Freq	**%**	**Cum %**	**Freq**	%	**Cum %**	**Freq**	%	**Cum %**
$\nu < 1$	129 781	26.11	25.70	166 101	30.98	30.98	222 786	28.20	28.20
$1 \leq \nu < 1.24$	44 379	8.93	34.27	69 106	12.89	43.87	88 327	11.18	29.38
$1.24 \leq \nu < 2$	50 764	10.21	45.05	54 618	10.19	54.06	83 764	10.60	49.99
$2 \leq \nu < 3$	83 908	16.88	62.18	80 343	14.99	69.04	128 719	16.29	66.28
$3 \leq \nu < 4.64$	53 643	10.79	72.50	56 842	10.60	79.65	88 431	11.19	77.47
$\nu \geq 4.64$	134 647	27.09	100.00	109 126	20.35	100.00	177 955	27.53	100.00
N	497 122			536 136			789 982		

Notes: The weights are the inverse of the probability of being consistent, multiplied with 1000 and then rounded to the nearest integer. The reason for this transformation is that the original weights (the inverse of the probability of being consistent) are defined on $[1.7, 4.4]$. Since frequency weights must be integers, the transformation is needed in order to apply the weights to the dataset.

Table 15: Parameter estimates of the ν parameter for income when observations are weighted with the inverse of the probability of being consistent based on background and attitude variables.

	Version 1				Version 2		
	Freq	**%**	**Cum %**		**Freq**	**%**	**Cum %**
$\nu < 1$	96 874	20.47	20.47	$\nu < 1$	132 91	22.19	22.19
$1 \leq \nu < 1.24$	22 613	4.78	25.25	$1 \leq \nu < 1.53$	49 787	8.31	30.50
$1.24 \leq \nu < 2$	32 369	6.84	32.08	$1.53 \leq \nu < 2.56$	69 631	11.63	42.13
$2 \leq \nu < 3$	60 317	12.74	44.83	$2.56 \leq \nu < 3.25$	109 066	18.21	60.34
$3 \leq \nu < 4.64$	88 779	18.76	63.58	$3.25 \leq \nu < 4.21$	85 228	14.23	74.57
$\nu \geq 4.64$	172 356	36.42	100.00	$\nu \geq 4.21$	152 287	25.43	100.00
N= 473 308				N=598 909			

	Total		
	Freq	**%**	**Cum %**
$\nu < 1$	170 629	21.91	21.91
$1 \leq \nu < 1.24$	17 573	2.26	24.16
$1 \leq \nu < 1.53$	38 234	4.91	29.07
$1.24 \leq \nu < 2$	24 338	3.12	32.20
$1.53 \leq \nu < 2.56$	53 100	6.82	39.01
$2 \leq \nu < 3$	48 178	6.19	45.20
$2.56 \leq \nu < 3.25$	80 801	10.37	55.57
$3 \leq \nu < 4.64$	56 679	7.28	62.85
$3.25 \leq \nu < 4.21$	68 011	8.73	71.58
$\nu \geq 4.21$	93 969	12.06	83.65
$\nu \geq 4.64$	127 360	16.35	100.00
N=778 872			

Notes: The weights are the inverse of the probability of being consistent multiplied with 1000 and then rounded to the nearest integer. The reason for this transformation is that the original weights (the inverse of the probability of being consistent) are defined on [1.74.4]. Since frequency weights must be integers the transformation is needed in order to apply the weights to the dataset.

always chosen or society B was always chosen. In order to rule out that these results were due to noise or an artifact of the survey design, we study the following possible explanations for the large share of extreme answers: (i) When there was a large gap between the rich and the middle class or the poor and the middle class, this gap was a focal point for the subjects, whose objective then became to minimize the gap. This was obtained by always choosing society A when the gap was large between the middle class and the rich and by always choosing society B when the gap was large between the poor and the middle class. (ii) Unconsidered answers were more probable to be such that society A was always chosen or society B was always chosen, suggesting that the observed pattern was driven by noise.

In order to study (i), we examine the two versions of the Alfaland survey for ν_{health} and ν_{income} and study the differences in the ν estimates between the two. For both health and income, the difference between version 1 and version 2 of the survey is that there was a smaller jump from the poor to the middle class in version 2 than in version 1, while the gap between the middle class and the rich was larger (see Tables 2 and 3 in the paper). It is possible that the relative income of the rich became the focal point, leading the respondents to focus on minimizing the gap between the middle class and the rich and hence always choose society A. In version 1 of the ν survey, the biggest "relative imbalance" was between the poor and the middle class in society A. This might have become a focal point, inducing respondents to focus on closing this gap and thus always choose society B. However, note that a respondent that was not affected by the focal points of the graphs and always wanted to reduce the gap between the middle class and the rich would always choose society A in version 1, just as in version 2. Hence, it must be that, at least partly independently of preferences, respondents were drawn to focus on a particular attribute of the distribution.

What can be noted for the ν estimates is that for both ν_{income} and ν_{health}, the estimates are sensitive to the framing of the questions in terms of the relative jump in income/health from the poor to the middle class or from the middle class to the rich. As can be seen in Tables 5 and 6, the share of respondents that had the lowest estimate of the ν-parameter (those who chose society A in all five questions) was considerably higher in variant 1 of the questionnaires, while the opposite was true for the share of respondents that had the highest estimate of the ν-parameter (those who choose society B in all five questions). Moreover, the differences between the two versions were larger in the share of only B answers than only A answers. For both ν_{income} and ν_{health}, there was about a five percentage point difference between the groups in the share of respondents choosing only society A. However, the difference between the only B answers was over 10 percentage points

for both health and income. This indicates that respondents tended to be more susceptible to focusing on the relative deprivation of the poor than the relative advantage of the rich.

It could also be the case that the only A and only B answers were results of unconsidered answers. In order to study (ii), we analyze the relationship between answering only A or only B and the individual score on a set of comprehension questions in which the understanding of the survey setup was tested. The results from logit regressions where a dummy variable for only A/only B answers is regressed on performance on the comprehension questions and the cognitive reflection questions are presented in Tables 16 and 17. In the case of only A answers, we find no correlation between always choosing society A and performance on the comprehension questions. However, in the case of only B answers, there is a negative and significant correlation between always choosing society B and performance on the comprehension questions. If we control for cognitive ability measured by the performance on a set of cognitive reflection questions, we no longer find a significant relationship between performance on the comprehension questions and the probability of always choosing society B. This indicates that subjects with low cognitive ability had lower scores on the comprehension questions and put higher weight on the poor, leading them to always choose society B. In order to control for the potential confound of low cognitive ability and low ν estimates driven by only B answers, we estimate the ν parameter for the subgroup of subjects with 6 out of 6 correct answers on the comprehension questions. The results are presented in Tables 18 and 19. The results show that the parameter estimates are within the same range, although slightly lower. The median value turns out to be the same in the case of ν_{health} in each of the two versions, but put together the median falls one interval lower ($\nu \in [1.24, 2]$). For ν_{income}, the median is now in $[2.56, 3]$, compared to $[3, 3.25]$ for the full sample. Hence, the difference between the ν_{income} and the ν_{health} estimates pertains, although the estimates are slightly lower.

Concerning the extreme answers in the sub-sample of high-scorers, we note that high-scorers were slightly more prone to choose society A in all five questions and slightly less prone to choose society B in all five questions than the sample as a whole presented in Tables 5 and 6. The pattern for the interior answers is similar to the whole sample. Hence, we draw the conclusion that it is differences in the share of only A answers compared to only B answers are behind the slightly lower parameter estimates for the sub-sample of high-scorers.

Table 16: The role of comprehension for the number of only A and only B answers, ν_{health}

	Only B, ν_{health}	Only B, ν_{health}	Only A, ν_{health}	Only A, ν_{health}
Comprehension score	-0.12	-0.053	0.001	-0.06
	(0.089)	(0.094)	(0.087)	(0.091)
Cognitive reflection score		-0.23**		0.18**
		(0.094)		(0.083)
N	515	515	515	515

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimated using logit regression. Comprehension score is the number of correct answers on the six pre-elicitation survey comprehension questions. Cognitive reflection score is the number of correct answers on the four cognitive reflection questions.

Table 17: The role of comprehension for the number of only A and only B answers, ν_{income}

	Only B, ν_{income}	Only B, ν_{income}	Only A, ν_{income}	Only A, ν_{income}
Comprehension score	-0.19**	-0.086	-0.063	-0.105
	(0.082)	(0.273)	(0.092)	(0.097)
Cognitive reflection score		-0.369***		0.133
		(0.089)		(0.091)
N	521	521	521	521

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimated using logit regression. Comprehension score is the number of correct answers on the six pre-elicitation survey comprehension questions. Cognitive reflection score is the number of correct answers on the four cognitive reflection questions.

Table 18: Parameter estimates of the ν parameter for health, only including high scorers on the comprehension questions (6/6 correct answers)

	Version 1			Version 2			Total		
	Freq	**%**	**Cum %**	**Freq**	%	**Cum %**	**Freq**	%	**Cum %**
$\nu < 1$	43	24.57	24.57	56	34.15	34.15	99	29.20	29.20
$1 \leq \nu < 1.24$	13	7.43	32	20	12.20	46.34	33	9.73	38.94
$1.24 \leq \nu < 2$	23	13.14	45.14	15	9.15	55.49	38	11.21	50.15
$2 \leq \nu < 3$	27	15.43	60.57	36	21.95	77.44	63	18.58	68.73
$3 \leq \nu < 4.64$	23	13.14	73.71	19	11.59	89.02	42	12.39	81.12
$\nu \geq 4.64$	46	26.29	100.00	18	10.98	100.00	64	18.88	100.00
	N=175			N=164			N=339		

Table 19: Parameter estimates of the ν parameter for income, only including high scorers on the comprehension questions (at least 6/6 correct answers)

Version 1				Version 2			
	Freq	**%**	**Cum %**		**Freq**	**%**	**Cum %**
$\nu < 1$	25	16.89	16.89	$\nu < 1$	45	23.44	23.44
$1 \leq \nu < 1.53$	8	5.41	22.30	$1 \leq \nu < 1.24$	21	10.94	34.38
$1.53 \leq \nu < 2.56$	11	7.43	29.73	$1.24 \leq \nu < 2$	26	13.54	47.92
$2.56 \leq \nu < 3.25$	30	20.27	50	$2 \leq \nu < 3$	38	19.79	67.71
$3.25 \leq \nu < 4.21$	32	21.62	71.62	$3 \leq \nu < 4.64$	26	13.54	81.25
$\nu \geq 4.21$	42	28.38	100.00	$\nu \geq 4.64$	36	18.75	100.00
N=148				N=192			
Total							
	Freq	**%**	**Cum %**				
$\nu < 1$	70	20.59	20.59				
$1 \leq \nu < 1.24$	8	2.35	22.94				
$1 \leq \nu < 1.53$	21	6.18	29.12				
$1.24 \leq \nu < 2$	11	3.24	32.35				
$1.53 \leq \nu < 2.56$	26	7.65	40				
$2 \leq \nu < 3$	30	8.82	48.82				
$2.56 \leq \nu < 3.25$	38	11.18	60				
$3 \leq \nu < 4.64$	26	7.65	67.65				
$3.25 \leq \nu < 4.21$	32	9.41	77.06				
$\nu \geq 4.21$	36	10.59	87.65				
$\nu \geq 4.64$	42	12.35	100.00				
N=340							

7.6 Comparing differences in background variables between consistent and inconsistent answers

In Tables 20 and 21, we present a comparison of the background variables of the final sample (consistent answers with an answering time of no less than nine minutes) compared to the sample of participants who completed the survey but were inconsistent in their answers. From the tables, it can be seen that participants who were inconsistent were generally younger, more risk-taking, and had a more negative view towards reducing inequality than the consistent participants.[12] In addition, and perhaps not surprisingly, the comprehension scores and scores on the cognitive reflection questions were lower for the inconsistent group than for the consistent group.

Table 20: Summary statistics for consistent and inconsistent subjects in the elicitation of ν for health, only including participants that completed the whole survey.

	Inconsistent		Consistent			
	mean	**sd**	**mean**	**sd**	**Difference**	**se**
Man	0.48	0.52	0.50	0.52	-0.02	(-0.49)
Age	38.66	13.45	41.61	13.62	-2.95**	(-3.26)
Education	2.27	0.62	2.36	0.60	-0.09*	(-2.29)
BMI	25.58	6.21	25.77	4.86	-0.19	(-0.49)
CR score	0.93	1.14	1.72	1.23	-0.80***	(-10.10)
Comprehension score	3.94	1.91	5.31	1.16	-1.37***	(-12.44)
Health	3.42	1.06	3.59	0.94	-0.16*	(-2.39)
Risk tolerance	6.71	2.23	6.09	2.18	0.63***	(-4.25)
Attitude to health inequality	7.84	2.66	8.89	2.28	-1.05***	(-6.25)
Attitude to income inequality	7.68	2.48	8.21	2.54	-0.53**	(-3.17)

Notes: The variable "Health" mesures self-assessed health with the question "In your opinion, how is your health condition?" Possible answers are 1= very bad. 2=bad. 3=ok. 4=good, and 5=very good. The variable "Risk preferences" is measured with the question "How willing or unwilling are you to take risk" where 0=Very unwilling to take risks and 10= Very willing to take risks. The two variables that measure attitudes to health and income inequality are based on the questions "How important is it to reduce income (health) inquality" where 0="Not at all important" and 10="Very important".

[12]Note that while the results regarding attitude towards inequality might have been driven by noise (since the average answer of the inconsistent participants was closer to the middle of the answering list, which would be the average if everybody answered randomly), the results regarding risk tolerance went in the opposite direction, suggesting that noise was not a driving factor behind the difference.

Table 21: Summary statistics for consistent and inconsistent subjects in the elicitation of ν for income, only including participants that completed the whole survey.

	Inconsistent		Consistent			
	mean	**sd**	**mean**	**sd**	**Difference**	**se**
Man	0.47	0.51	0.51	0.52	-0.04	(-1.22)
Age	38.09	13.08	41.97	13.76	-3.88***	(-4.32)
Education	2.27	0.61	2.35	0.60	-0.08*	(-2.05)
BMI	25.13	6.09	26.08	4.93	-0.95*	(-2.46)
CR score	0.94	1.14	1.71	1.24	-0.77***	(-9.72)
Comprehension score	3.95	1.91	5.28	1.19	-1.33***	(-11.97)
Health	3.50	1.04	3.53	0.97	-0.03	(-0.51)
Risk tolerance	6.71	2.20	6.10	2.20	0.61***	(-4.14)
Attitude to health inequality	7.54	2.63	9.08	2.19	-1.54***	(-9.31)
Attitude to income inequality	7.55	2.44	8.30	2.54	-0.75***	(-4.50)

Notes: The variable "Health" mesures self-assessed health with the question "In your opinion, how is your health condition?" Possible answers are 1= very bad. 2=bad. 3=ok. 4=good, and 5=very good. The variable "Risk preferences" is measured with the question "How willing or unwilling are you to take risk" where 0=Very unwilling to take risks and 10= Very willing to take risks. The two variables that measure attitudes to health and income inequality are based on the questions "How important is it to reduce income (health) inquality" where 0="Not at all important" and 10="Very important".

7.7 Questions on socio-economic status, health, attitudes and preferences

Here, we present the questions that were included in the survey (translated from Swedish). For the questions used in the analysis, the label is written in bold letters in front of the question.

Socio-economic status

- **Age:** What is your age in years? (*number*)
- **Woman:** What is your gender? (*woman / man / other or do not want to disclose*)
- **Highest degress at least high school:** What is your highest educational degree? (*primary school / high school / university*)
- **BMI above 25** (calculated from height and weight): How tall are you and how much do you weight? a) (*number*) cm. b) (*number*) kg.
- What is your civil status? Choose what best describes your living situation today. (*married / partnership / divorced / in a relationship but living apart / widow/widower / Single*)
- **Has kids:** Do you have any children? If so, how many? (*none / one child / two children / three children / four children / five or more children*)
- How many grown ups (20 years or older) live in your household? (*number*)
- How many children (19 years or younger) live in your household wholetime? (*number*)
- How many children (19 years or younger) live in your household parttime (half of the time or less)? (*number*)
- Choose the alternative that describes your background most correctly (*I am born in Sweden and I have at least one parent who is born here. / I am born in Sweden but my parents are not. / I am born in another country.*)
- **Weighted household income** (calculated from information on gross income of household and the number of persons living there): What

is the total gross-income of your household (income before tax deductions)? (*0-19999kr per month / 20000-29999 per month / 30000-39999 per month / 40000-49999 per month / 50000-59999 per month / 60000-69999 per month / 70000-79999 per month / 80000-89999 per month / 90000-99999 per month /100000 or more per month*)

Health

- **Self-reported health:** In your opinion, how is your health condition? (*very bad / bad / ok / good / very good*)
- Do you smoke? (*Yes, every day / Yes, sometimes / Only when partying / No, I have recently stopped or I am trying to stop / No, I have never smoked or I stopped more than 12 months ago*)
- **Workout habits:** How many times have you exercised for at least 15 minutes such that you get sweaty and short of breath in the last month? (*0-1 times per week / 2 times per week / 3 times per week / 4 times per week / 5 times per week / 6 times or more per week*)
- **Active lifestyle:** How much everyday physical activity do you get during a typical day? By everyday physical activity, we mean that you move (e.g. walk or bike) so that your pulse is elevated and the physical effort is such that you begin breathing through your mouth. (*0-14 minutes per day / 15-29 minutes per day / 30-59 minutes per day / 60 minutes or more per day*)

Preferences, attitudes and values

The following questions are borrowed from Falk et al. (2018)

- **Risk tolerance:** Please tell me, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.
- We now ask for your willingness to act in a certain way in four different areas. Please again indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

- How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?
- **Attitude to punishment, self:** How willing are you to punish someone who treats you unfairly, even if there may be costs for you?
- **Attitude to punishment, others:** How willing are you to punish someone who treats others unfairly, even if there may be costs for you?
- **Give to charity:** How willing are you to give to good causes without expecting anything in return?
-

The following questions are borrowed from the World Value Survey Sweden, 2011.

- **Left to right:**In politics one often talk about "left" and "right". How would you place your self on the scale below? (*0=left / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10=right*)
- How would you describe your views in the following questions? The number 1 indicates that you completely agree with the statement to the left and the number 10 that you completely agree with the statement to the right. If you consider your views are in between the two, you should choose the number that best describes your views.
  - **Wage gap:** Peoples wages should become more equal - We need bigger wage differences in order to encourage individual effort.
  - Private ownership in trade and industries should increase - Public ownership of trade and industries should increase.
  - **Role of government:** The government should take greater responsibility that everyone gets what they need - People should take more responsibility for their own situation.
  - **Sources of success:** In the long-run, hard work tends to lead to better living conditions - Hard work usually does not lead to success, success rather depends on luck and good connections.
  - **Sources of wealth:** You can only get rich at the expenses of others - Wealth can increase so that everybody gets better off.
  - **Sources of health:** In the long run, physical activity and good nutrition leads to better health and longer life - Physical activity

and good nutrition usually does not lead to better health and longer life, instead it depends on luck and genes.

- **Role of government, health:** The government should take bigger responsibility and encourage better habits regarding nutrition and physical activity (e.g. by introducing a tax on fat and sugar) - People should take bigger responsibility for their own nutrition and physical activity.

Additional questions on attitudes towards inequality, on a scale from 1 to 10

- **Attitude to income inequality:** How important is it to reduce inequality in income? (*1 = not at all important / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 = very important*)

- **Attitude to health inequality:** How important is it to reduce inequality in health across different income groups? (*1 = not at all important / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 = very important*)

Cognitive reflection questions

The cognitive reflection questions are borrowed from Toplak et al. (2014).

- If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? _ days (*correct answer = 4 days; intuitive answer = 9 days*)

- Johanna received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?_ students (*correct answer = 29 students; intuitive answer = 30 students*)

- A man buys a pig for 600 SEK, sells it for 700 SEK, buys it back for 800 SEK, and sells it finally for 900 SEK. How much has he made? _ kronas (*correct answer = 200 SEK; intuitive answer = 100 SEK*)

- Simon decided to invest 8000 SEK in the stock market one day early in 2013. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately, for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon has: a. broken even in the stock market, b. is ahead of where he began, c. has lost money (*correct answer = c, because the value at this point is 7000 SEK ; intuitive answer = b*)

Survey comprehension questions

Before accessing the survey, subjects answered six comprehension questions that aimed at eliciting their ability to retrieve information from the types of graphs used to present the Alfaland societies. Subjects were presented with the graph presented in Figure 15 and were asked to answer the following six questions (Everything is translated from Swedish).

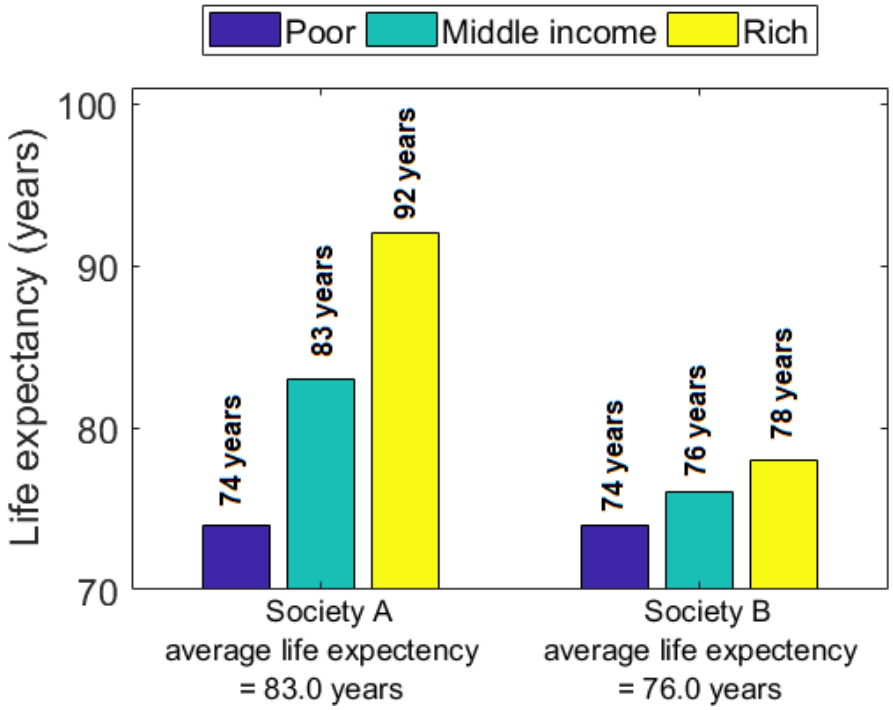

Figure 15: An example graph used in the pre-elicitation survey comprehension questions-
In this graph you see the distribution of life expectancy across different income groups in two societies, society A and society B. The average life expectancy for the poorest third (poor), the middle third (middle income) and the richest third (rich) in both societies is shown in the graph. Under the graph, you see the average life expectancy for the whole society.

- In which society does the poorest third have the lowest average life expectancy? (*Society A / Society B / The same*)
- In which society does the middle third have the highest average life expectancy? (*Society A / Society B / The same*)

- In which society is the difference between the average life expectancy of the richest third and the average life expectancy of the middle third the smallest? (*Society A / Society B / The same difference*)
- In which society is the difference between the average life expectancy of the middle third and the average life expectancy of the poorest third the biggest? (*Society A / Society B / The same difference*)
- In which society is the average life expectancy the lowest? (*Society A / Society B / The same*)
- How many years on average is the life expectancy of the richest third in society A? (*83 years / 78 years / 92 years*)