ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Health Economics

Public preferences over health distributions and beliefs on distributive justice

Name student: Maartje Tromm

Student ID number: 460330

Supervisor: Matthew Robson

Second assessor: Erwin Birnie

Date final version: 30-07-2024

Abstract

Preferences for health distribution have been shown to vary when applied to different groups in society. However, a clear understanding of the reasons behind those preferences and their differentiation between population segments is lacking from the literature. This study examines the public's health distribution preferences across groups defined by income, smoking status and sex. In addition, health distribution beliefs of the public for each of these groups are categorized into a list of principles of distributive justice based on philosophical literature and these categories' connection to the distribution direction is explored. For this purpose, experimental data of a UK sample (n = 655) is used to show how participants have aimed to distribute resources for health. Short open text responses (up to 82 words) explaining their choices are manually coded into the principle categories of distributive justice and three multinomial logit regression models demonstrate the predictive power of the categories on the preferred health distribution separately for the groups of Income, Smoking and Sex. These models are compared to three models in which the categories are replaced by several variables containing participant characteristics, like believed life expectancy of the groups and political orientations. The findings show that the majority of participants distributes equally between the rich and the poor or more to the poor, more to non-smokers and equally between males and females. For Income, Equality and Need are the most prominent principles, whereas for Smoking Responsibility is most often mentioned and for Sex primarily Equality. The categories possess great explanatory power regarding health distribution preferences, unlike the descriptive variables. These results underline the need for consideration of public beliefs on health distribution tailored to the subjects of specific health policies.

Table of contents

1 Introduction

Health inequalities in the UK affect life expectancy, quality of life and overall wellbeing (Marmot et al., 2020) and reflect broader issues of justice and equity. Inequality is a multidimensional concept, since it can point to, for instance, inequality of resources or inequality of outcomes, as well as inequality between different segments of society. The diversity of opinions around health inequalities stresses the complexity of the topic and its social relevance. For example, there are those who call for a redistribution of health recourses to mitigate socio-economic disparities, and others who emphasize the role personal responsibility and lifestyle choices play in health outcomes (Graham, 2009; Popay et al., 2010; Albertsen and Tsiakiri, 2023; Ruckert et al., 2017). Within the philosophical field of distributive justice the question of which inequalities are "unfair" is dealt with by defining a just distribution of goods within a community. The objective of this study is to shed light on the principles of distributive justice that steer public beliefs in the way health should be distributed. For this purpose, public preferences are examined for health resource distribution among different groups defined by income, smoking status and sex, in addition to the reasons behind those preferences. Understanding preferences and beliefs of the public is crucial for the development of health policy (Lynch and Gollust, 2010; Burnstein, 2003; Page and Shapiro, 2014; Jacobs, 1993, Smith et al., 2018).

Studies have shown the plurality of fairness beliefs of the public when it comes to inequality of health access versus health outcomes (Lynch and Gollust, 2010), and that of distribution preferences between different groups defined by income, sex or smoking status (Lynch and Gollust, 2010; Goto and mori, 2019; Norman et al., 2012; Robson et al., 2024a). The lack of uniformity of the findings across different countries calls for a separate analysis of preferences in the country of interest. For this paper data is analysed from an experiment by Robson et al. (2024b), in which distribution preferences are elicited among the UK public between the richest fifth and the poorest fifth of the UK population (Income), smokers and non-smokers (Smoking), males and females (Sex) and participants are asked for their reasons behind these preferences. This paper starts out by answering the question: *Does the public distribute more health towards the rich or the poor, smokers or non-smokers and men or women?*

The literature is not definitive on which principles of distributive justice are most valued by the public when it comes to health (Wilmot and Ratcliffe, 2002; Oedingen et al., 2019; Bae et al., 2023; Schwappach, 2002; Lim et all, 2012; Cookson and Dolan, 1999; Wikler, 2002; Stone, 2006). In each of these studies a limited number of possible principles is assessed. Differentiated use of principles is found across the different domains of health care, pensions and unemployment benefits (Van Hootegem et al., 2020), yet no previous research has established which principles are employed for health distribution across different groups. Focusing on this gap in the literature, a second research question is posed for each of the three distribution settings of Income, Smoking and Sex: *To which categories do public beliefs behind resource allocation for health belong?* This paper is unique in the analysis of survey answers of

a few sentences explaining participants' health distribution preferences. As will become clear, this kind of data has its limitations, but will also provide a distinct perspective, since it allows for the unconstrained beliefs of participants to be assessed and for any principle to be mentioned. Much of the literature relies on discrete choice experiments (DCEs) and multiple choice surveys, limiting the possible outcomes. In this study philosophical theory is linked to public beliefs on health distribution by allocating participants' answers to a newly created list of categories of principles of distributive justice. This list is based on the literature, though in its entirety it is one of a kind to accommodate the bespoke nature of the dataset and the goal of preserving the variety of possible answers.

Additionally, a multinomial logit regression model is outlined for testing the predictive power of the outcome categories on health distribution preferences. It is then compared to the fit of a similar model with participants' characteristics as independent variables to gauge the relative extent to which the categories of principles explain participants' choices. The third research question is therefore: *Is there a correlation between the category of health distribution beliefs and preferences for health distribution and if so, in which direction?*

Chapter 2 contains an overview of literature on health distribution preferences, principles of distributive justice, public beliefs of those principles and text analysis methods. Chapter 3 describes the experiment and variables used and the manual coding method for text analysis, precisely defines all of the outcome categories and lays out text mining methods for the creation of several descriptive figures of the text data. It concludes by outlining the regression models used. Chapter 4 shows the results of the distribution preferences, a visualization of the text data, the category outcomes and regressions. Chapter 5 discusses the findings, strengths, limitations and conclusions.

2 Theoretical background

2.1 Health distribution preferences

Health inequality is not a straightforward concept, for inequality can exist across different dimensions, such as inequality of resources and inequality of outcomes, and between different groups, like different income segments, ethnicities, males and females and even people with different behaviours. In the UK, aversion is stronger for health inequality related to income than for pure health inequality, where the only difference between individuals exists in the multipliers for health production or in unspecified diseases (McNamara et al., 2021; Robson et al., 2024a). A part of this difference was shown to be driven by factors outside of health, such as the financial inequality (McNamara et al., 2021). These findings concern aversion towards health outcomes, though other research suggests that the majority of the American public views inequality in access to health care as unfair, while only about a third perceives inequality in health states or health outcomes to be unfair (Lynch and Gollust, 2010). According to Ali et al. (2017) framing an inequality scenario at an individual level or at a population level makes little difference. Out of health and health care inequalities between four sets of social groups defined by gender, race, income and educational attainment, the American public finds inequality between the sexes the least unfair and between different income segments the most unfair (Lynch and Gollust, 2010).

Goto and Mori (2019) conducted a DCE among a Japanese and Korean sample, eliciting equity preferences for life expectancy gains. The compared attributes are sex, income, smoking status, healthy lifestyle, caring for family members, age and increase in life expectancy. Carers were most favoured to receive a program increasing their life expectancy, followed by non-smokers, low income individuals, males and then healthy individuals. Younger individuals were also significantly preferred. Both males and females in the sample preferred their own sex, though males did so more than females, which caused the overall choice to lean towards males. Since females are more often carers and carers were highly preferred to non-carers, the male favour might be counteracted. A similar study by Norman et al. (2012) with an Australian sample found no significant effect for sex, but carers, low income individuals and non-smokers were preferred. The effect was largest for non-smokers.

In light of these findings, the following hypotheses are formulated: The UK general public prefers to distribute more health to the poorest fifth of the population than to the richest fifth (H1a), more to non-smokers than to smokers (H1b) and the same to males and females (H1c).

2.2 Principles of distributive justice

This complexity in public perception underscores the need for an understanding of the principles of distributive justice. The literature on distributive justice is incredibly extensive and dates back, at least, to Aristotle. Deutsch (1975) attempted to make an overarching classification of principles of justice into three values: equity, equality and need. The principle of equity states that rewards or outcomes should be allocated to individuals based on the extent of their contributions. Therefore a distinction should be

made between inequalities arising from brute luck and from choice, acknowledging a certain responsibility people have for their own outcomes (Dworkin, 1981; Cohen, 1989; Arneson, 1989). The principle of equality on the other hand advocates for an equal distribution to everyone regardless of their contributions. There are many different interpretations of equality. Equality of outcomes signifies that everyone should end up with the same outcomes, while equality of resources aims to distribute resources evenly for the production of an outcome. Sen (1979) proposes the capability approach that focuses on equality of the capability to produce utility from resources rather than equality of the resources or outcomes themselves. Nussbaum (2000) extended the capability theory by creating a list of 10 central capabilities, one of which is bodily health. The principle of need is often associated with John Rawls's seminal work *A Theory of Justice* (1971). He proposes a difference principle that suggests a just distribution should benefit the least advantaged the most.

For the prioritization of organ allocation, the principles of equality, need and responsibility are considered as well as two others: utilitarianism and value for society (Wilmot and Ratcliffe, 2002; Oedingen et al., 2019). The principle of utility is described in the work of utilitarians like Jeremy Bentham and John Stuart Mill. It concerns the maximizing of total utility in society. Utility maximization is the traditional economic approach towards prioritization in health care through cost-effectiveness analyses, whereby effectiveness is primarily measured in QALY's. In the context of organ allocation, value for society means prioritizing those who have proven themselves to be more beneficial for society than others, through, for example, charity work. Value for society is in a different way often incorporated into cost-effectiveness analyses as well, by considering productivity losses due to illness as a cost (Setiawan, 2024). However, there are other possible externalities to certain distributions of health. Generally, more egalitarian societies enjoy all kinds of advantages over more unequal societies, other than the intrinsic justice of equality (Wilkinson, 2005). Furthermore, the allocation of health resources could be used as an incentive for certain health behaviours (Lunze and Paasche-Orlow, 2013). There are undoubtedly many more examples. Finally, a large number of studies shows the close relationship between inequality aversion and risk aversion (Kroll & Davidovitz, 1999; Kroll & Davidovitz, 2003; Carlsson et al., 2005), exemplary of how self-interest can play a role in distributing resources. People are, thereby, more likely to judge a distribution as fair when it is advantageous to them (Miller, 1992). Six principles for health distribution have now been described and serve as the main outcome categories: Equality, Need, Responsibility, Efficiency, Externalities and Self-interest. Five additional categories are employed to categorize answers that are not associated with any of these principles in particular: Don't know, Ambiguous, Fair, None and Empty. Their forming and definitions will be further illustrated in chapter 3, as well as hypotheses concerning distribution predictions per category.

2.3 Public beliefs of distributive justice

For the distribution of income and wealth, out of Equality, Need and Responsibility (equity) people primarily refer to equity or a combination of equity and need (Aalberg, 2003). In an overview of studies

about public views on distributive justice with a wide range of methods Miller (1992) also stresses these two principles. It is rare for people to mention one single principle to outline their beliefs on distributive justice (Kluegel & Mateju, 1995).

Beliefs about principles of health distribution may differ from those of the distribution of income, in light of the nature of health as a prerequisite for enjoying other types of goods (Daniels, 2008). Even though, for health distribution, Efficiency is probably the most widely used principle in practice, the evidence suggests that it is not the only principle endorsed by the public. According to Bae et al. (2023) a larger QALY gain did increase the probability of a treatment scenario in their discrete choice experiment to be chosen, but with a diminishing marginal value. In a preference study by Schwappach (2002) only a small number of subjects strictly distributes their budged for health based on QALY gain. Still, quality of life and life expectancy after treatment were found to be the most important attributes, followed by healthy lifestyle (Responsibility). A DCE by Lim et all (2012) indicates that prioritizing the most disadvantaged of society is supported by the majority, even when this was not the most efficient allocation in terms of health gain. Most also considered personal responsibility an important factor, though decreasing with the severity of disease. A study of small-group discussions regarding the prioritization order of four hypothetical patients found three principles to be broadly accepted by the public: A rule of rescue (Need), health-maximization (Efficiency) and reducing inequalities (Equality) (Cookson and Dolan, 1999). Two other principles were tested, namely a 'lottery' principle of not wanting to choose at all, which was not mentioned often and 'choisism' (Responsibility), which turned out to be controversial. Responsibility does seem to play a large role in aversion to health inequality (Lundell et al, 2013). Individuals with the conviction that people are responsible for their own health outcomes are more likely to perceive those health outcomes as fair (Wikler, 2002; Stone, 2006). Paradoxically, the naturalistic fallacy (Frankena, 1939) can make people believe a lack of responsibility for health outcomes makes them more fair as they are "natural". Van Hootegem et al. (2020) have found that there is a strong variation in preferences for just distributions across different welfare domains: in health care 82% prefers a distribution based on Equality, while this is only 58% for pensions and 52% for unemployment. Responsibility is found of less importance for health care than for pensions and unemployment and Need is preferred by almost one-fifth for the domain of unemployment, though only minimally for health care and pensions. Based on these studies, the following hypothesis is posed: The reasons behind health distribution differ when related to Income, Smoking or Sex (H2).

2.4 Text analysis methods

Automated text processing has become very advanced in recent years (Min et al., 2021) and obviously offers many advantages compared to traditional manual coding methods. However, a brief exploration of computerized text analysis methods unveiled that there are no premade algorithms fit to use on a dataset as bespoke as the one described in this paper. Furthermore, in order to create a custom algorithm or train one to categorise according to the specific definitions that will be further explained in this

chapter, training data is needed. This is problematic, because firstly, the dataset contains a limited number of observations and secondly, the training data would have to be categorized manually, negating some of the advantages of computerized text analysis.

The computerized text analysis method that is most often used when training data is not available is Dictionary Based Computerized Text Analysis (DBCTA) (Gentzkow et al., 2019). DBCTA uses a dictionary for each outcome category and simply determines the category based on the use of any of the dictionary words in a Document. A Document can be a whole article, though in the case of the described dataset one participant's text response would be a Document. There are many dictionaries available online for this purpose on various topics, yet again none of them match the category definitions for this paper. Therefore, custom dictionaries would have to be created. Nicolas et al. (2019) and Reid et al. (2023) explain DBCTA in detail as well as the creation of the dictionaries using R. This starts with finding seed words, from which synonyms are identified (WordNet), word embeddings (Word2Vec), inflections (Lexicon package in R) and conjugations (Verbix.com). Ideally, at least two human judges with knowledge about the parameters of the six categories would evaluate each word in the dictionaries and establish whether it aligns with the category definitions. The potential for this method was elaborately investigated, by manually coding a part of the sample and gauging the efficacy of different seed words. As the topic of the answers is quite complex and the differences between the different outcome categories often very subtle, it was difficult, if not impossible, to create a dictionary for each category that would not only accurately include all observations belonging to that category, but also exclude the ones that do not. Many of the subtleties namely do not lie in the use of different words, but rather in different orders and combinations of similar words. Finally, it was therefore decided that DBCTA could not code the answers as accurately as they could be manually coded. An advantage of DBCTA would have been the potential for cross-study comparison, since the created dictionaries could have been used again. However, this might not be relevant due to the niche focus of the dataset and the intended analysis.

3 Method

3.1 Experiment

Robson et al. (2024b) performed an experiment to derive social preferences for health distribution. A broadly representative sample of the UK adult population was recruited through Prolific and given a series of choice tasks where they were asked to distribute *resources* between three hypothetical individuals. Whereas their paper only concerns pure, income-related and income-caused health inequality, another sample was added to examine preferences for prioritization based on sex, smoking status and income, along with aversion to health inequality. The budget for each choice task was randomly chosen from the set $m \in \{180, 240, 360\}$ and had to be depleted. Each of the three hypothetical individuals would end up with a certain amount of *health*, depending only on the *resources* they were given and their personal *multiplier* of $d \in \{0.33, 0.5, 1\}$. The *multiplier* is a productivity factor that is the function turning *resources* into *health*: an individual with a *multiplier* of 0.5 who is given 180 of the budget produces a health of 90. Participants are informed that *health* is the number of life years adjusted for illness or disability (QALYs). Sliders are used for the division of *resources*, during which participants can see their allocation, health outcomes, resource and health gaps, the total health produced and the budged that is left. The experiment contains two treatments, A and B. In treatment A the participants do not receive any personal information about the individuals, while in treatment B they know the individuals' income, gender or smoking status. Treatment B_1 randomly assigns a biological sex to each individual from $x_i^1 \in \{Male, Female\}$, making sure at most two out of three individuals are of the same sex. Similarly, Treatment B_2 randomly allocates an income from $x_i^2 \in \{£5, 000, £10, 000, £25, 000, £50, 000, £100, 000\}$ and Treatment B_3 a smoking status from $x_i^3 \in \{Light\text{-}Smoker, Moderate\text{-}Smoker, Heavy\text{-}Smoker\}$. All participants first finish Treatment A and are then assigned to either Treatment B_2 (income) or both treatment B_1 (sex) and B_3 (smoking) subsequently, where the order is randomized.

After Treatments A and B participants are asked a range of questions concerning their own characteristics, health beliefs, distribution preferences and the reasons for their distribution. Only the responses to these questions are analysed in this paper.

3.2 Variables

The dataset initially contains five variables describing the characteristics of the sample. Cigs Per Day contains the number of cigarettes smoked per day by the participant. Female is a dummy with value 1 for female and 0 for male. Income contains a numerical value of the participant's pre-tax gross annual income. The variables Left-Right and Libertarian-Authoritarian hold the political orientation of participants on a scale from 0 to 10. For Left-Right 0 stands for extremely left and 10 for extremely right, just as for Libertarian-Authoritarian 0 means the participant believes to be extremely libertarian and 10 extremely Authoritarian.

For the Income section of the survey, health beliefs of participants are registered by asking them: "What do you think is the expected length of life in full health for the poorest fifth of the UK population?" and "What do you think is the expected length of life in full health for the richest fifth of the UK population?". The variables Health Belief Poor and Health Belief Rich contain the categorical answers to these questions, with [1 = "90 years or more", 2 = "85-90 years", 3 = "80-85 years", 4 = "75-80 years", 5 = "70-75 years", 6 = "65-70 years", 7 = "60-65 years", 8 = "55-60 years", 9 = "50-55 years", 10 = "45-50 years", 11 = "45 years or less", 12 = "Do not know", 99 = "Prefer not to say"]. The same questions are asked for Health Belief Smokers, Health Belief Non-smokers, Health Belief Males and Health Belief Females, but "the richest fifth of the UK population" is replaced by "someone in the UK who is a Heavy-Smoker for their entire life", "someone in the UK who is a Non-Smoker for their entire life", "someone in the UK who is Male" and "someone in the UK who is Female", respectively. The answer categories remain the same for all of these variables.

From the six Health Belief variables, six new variables were constructed. The categorical variables were turned into continuous variables by taking the lower number of the categories, taking on 90 for "90 years or more" and 40 for "45 years or less". These new variables are used in the analysis instead of the categorical variables, and will from now on be referred to as Health Belief Poor, Health Belief Rich, Health Belief Smokers, Health Belief Non-smokers, Health Belief Males and Health Belief Females. Afterwards, from these continuous variables another three new variables were constructed by subtracting Health Belief Poor from Health Belief Rich, subtracting Health Belief Smokers from Health Belief Non-Smokers and Health Belief Males from Health Belief Females. For only five observations the resulting variable Health Belief Difference (Income) is negative, for zero observations Health Belief Difference (Smoking) is negative and for twelve observations Health Belief Difference (Sex) is negative, showing that participants overwhelmingly believe that richer individuals, non-smokers and females live longer than their respective counterparts. The larger this believed gap is, the larger the values of the Health Belief Difference variables are.

Health Distribution (Income), Health Distribution (Smoking) and Health Distribution (Sex) are the three variables containing the participants choices regarding the resource distribution for health. For Health Distribution (Income) participants are asked: "In general, did you try to distribute Resources to give more Health to richer or poorer individuals?". For Health Distribution (Smoking) and Health Distribution (Sex), "richer or poorer individuals" is replaced by "Smokers or Non-Smokers" and "Males or Females". These are categorical variables with values [1 = "More to richer individuals", 2 = "More to poorer individuals", 3 = "No difference", 4 = "Do not know", 99 = "Prefer not to say"] for Income, [1 = "More to smokers", 2 = "More to non-smokers", 3 = "No difference", 4 = "Do not know", 99 = "Prefer not to say"] for Smoking and [1 = "More to males", 2 = "More to females", 3 = "No difference", 4 = "Do not know", 99 = "Prefer not to say"] for Sex.

After each of these variables, participants are asked "Why did you do this?". The answers to this question form the contents of the variables Health Distribution Reason (Income), Health Distribution Reason (Smoking) and Health Distribution Reason (Sex). These are string variables containing participant's answers of which the longest has 82 words.

For the manual coding of the data, separate dummy variables were created for all of the outcome categories, namely Equality (Income), Need (Income), Responsibility (Income), Efficiency (Income), Externalities (Income), Self-interest (Income), Fair (Income), Don't Know (Income), None (Income), Ambiguous (Income) and Empty (Income), and in a similar fashion as for the variables described before, (Income) is replaced with (Smoking) and (Sex) for the remaining two sections, totalling 33 category dummies.

3.3 Sample characteristics

Table 3.1 displays characteristics of the sample. The sample consists of 320 males and 335 females. The average income is 28566.55 and they smoke 1.03 cigarettes per day on average. 70 of them are smokers (they smoke more than 0 cigarettes per day) and the rest are non-smokers. The sample leans more left and libertarian. According to Robson et al. (2024b) the sample is representative of the UK adult population by age and sex.

Table 3.1 Sample descriptives

Variable	Observations	Mean	St. dev.	Min	Max
Female	655	0.51	0.50	0	1
Income	586	28566.55	251	2500	175000
Cigs per day	647	1.03	3.86	0	30
Left-Right	618	3.96	2.25	0	10
Libertarian-Authoritarian	612	3.69	1.95	0	10

3.4 Manual coding

For this study the method of Deductive or A Priori Coding was applied to the data. This method uses predetermined codes or categories based on existing literature and is suitable, since this research aims to test the specific hypothesis of which philosophical categories ground the UK general public's beliefs behind resource allocation for health (Saldana, 2016). Stata 17.0 was used for the manual coding of the answers contained in variables Health Distribution Reason (Income), (Smoking) and (Sex). For each of the outcome categories the corresponding dummy variable was assigned the value 1 for each observation with an answer that was deemed a part of that category. Initially answers were only coded in the predefined philosophical categories, Equality, Need, Responsibility, Efficiency, Externalities and Self-interest, as well as a few predetermined categories for answers not a part of any

of the aforementioned: Don't know, if a participant expresses not to be sure, Empty if the response is completely missing and None if it cannot be coded anywhere else. This way all answers are coded into at least one of the categories. During the coding process two issues arose, for which additional categories were created: Fair and Ambiguous. A substantial number of responses contained a notion of "fairness" without any additional explanation as to why something is fair or not. In order to retain more information from the data, these answers were not coded as None, but in their own Fair category. Additionally, various answers were found ambiguous between two categories. For the same reason, these answers were coded into both categories, along with the new Ambiguous category. Under "Category definitions" all categories are explained in more detail.

3.5 Text mining for figures

Benchimol et al. (2022) provide a detailed account of text mining techniques using R that was followed to create several figures from the text data, meant to give the reader a feel for the content of the responses, without having to read through all of them. From the text punctuation, numbers and stop words were removed and all letters were transformed to lower case, aiming to create uniform data devoid of meaningless terms. The remaining words were then stemmed: words were reduced to their root, so that words with similar meanings become the same. For example, "healthy" becomes "health" and "deserving" turns into "deserv". The cleaned data was mapped onto a Document Term Matrix (DTM), a vector space representation of the Corpus (all text data) in which each row represents a Document (a single response) and each column a word in the Corpus. Each cell in the matrix then contains the frequency of appearance of the corresponding word in the corresponding Document. These frequencies were weighted using a weighting scheme called term frequency-inverse document frequency or tf-idf. Equation (1), (2) and (3) show how this weighting is calculated. While tf(t) increases when a word is more frequently used in a single document, idf(t) decreases when it appears more often in the corpus as a whole. Tf-idf thereby performs better than a simple tf score at measuring the importance of a term in a document relative to the Corpus.

$$(1) \ \text{tf}(t) = \frac{\text{Number of times term } t \text{ appears in Document } D}{\text{Total number of terms in Document } D}$$

$$(2) \ \text{idf}(t) = \ln\left(\frac{\text{Total number of Documents}}{\text{Number of Documents containing term } t}\right)$$

$$(3) \ \text{tf} - \text{idf}(t) = \text{tf}(t) \times \text{idf}(t)$$

Each term thus has a different tf-idf score for each document, so to calculate the top terms across the entire corpus, the columns of the DTM were summed together. From the top terms word clouds were created including words with a minimum tf-idf score of 3 for the data in each variable; Health Distribution Reason (Income), (Smoking) and (Sex). Moreover, correlation maps were generated for each section, made up of words that appear in at least 15 documents. Lines mapping the words together indicate a Pearson correlation coefficient of at least 0.10. These minimum frequencies and correlations

for the word clouds and the correlation maps were essentially chosen through trial and error, ensuring the figures are comprehensible.

3.6 Category definitions

Below all principle categories are clearly defined to provide consistency within this paper and the manual categorization. This section concludes with a paragraph on *additional coding explanations*, providing further insight into the reasoning behind the allocation of responses that might be less straightforward, choices made for category definitions and illustrating the unsuitability of the data for DBCTA.

Equality

The category of "equality" contains beliefs based on the equal value that people intrinsically have and therefore the fairness of equal distribution. The answers contain key words like "equal", "the same" and "even". In this category no differentiation is made between different forms of equality, like equality of outcomes versus equality of resources and equality of health versus equality of welfare, since in most cases it is hard to say which of the two is meant by the participant.

Example answers:
"There should be no difference -we are all equal as human beings."
"Because everyone deserves health equality regardless of their life choices"

H3: Individuals who give an answer in the Equality category are expected to distribute the same amount of health to all groups.
In general an equal distribution is expected. Participants could distribute unequally if they perceive the current situation to be unequal and their distribution to be the equalizer. As poorer individuals, smokers and males generally have worse health than their counterparts, unequal distributions are expected to sway in their favour. On the other hand, if the argument is equality of welfare and females are perceived as disadvantaged in other aspects of life, the participant may distribute more to females.

Need

The principle of need focuses on distributing to those who are worst off. It differs from the principle of equality in the sense that complete equality of any attribute, whether it be outcomes or resources, is not required. It might be that the goal is not to achieve some sort of equality, but to at least let everyone reach a certain floor of health. Need also differs from efficiency, given that the most efficient allocation can be, but is not necessarily to the worst-off. The concept of need is easily found in answers that explicitly contain the word "need".

Example answers:
"because the poor need more help"
"it didnt seem to be a factor and health outcomes should be based on need"

H4: *Individuals who give an answer in the Need category are expected to distribute more to poorer individuals, smokers and males.*

Poorer individuals, smokers and males have a lower life expectancy and are therefore worse-off in terms of their health.

Responsibility

The Responsibility category consists of answers that either imply that people made a choice themselves and are responsible for that choice, or that there is a lack of responsibility and a situation is out of a person's control. Answers concerning different incomes contain the notion that poor people are not responsible for being poor and should therefore be compensated or that they are not responsible for their poor health, as this is caused by their lack of income. For smokers vs. non-smokers numerous participants report that smokers choose to smoke and are therefore responsible for their bad health, while some point out the addictive nature of smoking and therefore the lack of responsibility. For distributing between the sexes, a lack of responsibility for worse health is stated as a reason to allocate resources equally or more to the less healthy sex (mostly men), whereas sometimes this difference is regarded as "natural" and therefore unchangeable.

Example answers:

"Smoking is a choice that brings with it a wealth of health difficulties resources should be allocated to those individuals who do not smoke"

"People can't choose their gender"

H5: *Individuals who give an answer in the Responsibility category are expected to distribute more to the poor, more to non-smokers and the same amount to males and females.*

Arguments are expected to be made for both a responsibility and a lack of responsibility, so the distributions could go all directions. However, since the literature shows that for smoking attributions of responsibility are often made and this is not so much the case for income, it is expected that the lack of responsibility of the poor is emphasized more and the responsibility of smokers. No particular difference is expected between males and females as different reasonings under Responsibility can go both ways.

Efficiency

Answers explaining participants' aim to maximize total health are categorized as Efficiency. In the case of smokers vs. non-smokers, many answers point out that, as non-smokers live longer, they have longer to benefit from the allocated resources. This argument belongs to the Efficiency category, since it does not point out a disadvantage of non-smokers, but the fact that leaving resources unused is inefficient and that greater health benefits could be reaped by allocating them to non-smokers. This category is all about the maximizing of total *health*. When answers describe effects of health allocation on maximizing other facets of society, they are categorized as Externalities.

Example answers:

"I didn't care for their income just for their ability to use resources"

"i looked at the relative value of the resources. probably this would adversely affect the disabled."

H6: *Individuals who give an answer in the Efficiency category are expected to distribute the same amount of health to all groups.*

This hypothesis is formulated based on the orthogonality of multipliers, so in theory the greatest health benefits would be reaped under an equal distribution, even though examples of a different line of thought have already been shown.

Externalities

This category contains statements about the effects of different kinds of distributions that do not concern the health of the individuals in the experiment. These beliefs are either about the health of others in society that is affected through a secondary mechanism or effects that have nothing to do with health.

In theory the possibilities for externalities are endless. Individuals would distribute more resources to the rich if they believe rich people to contribute more to society in taxes or other endeavours. They would give more to the poor if their argument is that inequality has bad side effects, like criminal activity. Arguments can be made regarding the cost of smokers to society, either that they cost more or less. Second-hand smoke could also be an issue, encouraging the participant to give less resources to smokers. Even the distribution of resources as an incentive to smoke less can be seen as an externality, because it is not about the health of the individual him or herself, but about the health of society as a whole.

Example answers:

"They pay more taxes and contribute more to society"

"The world needs more females in power and for that they need health"

H7: *Individuals who give an answer in the Externalities category are expected to distribute the same amount of health to all groups.*

Because of Externalities possibilities in all directions, there is no reason to suspect a preference towards one of the groups.

Self-interest

Answers purely based on the idea that, if the experiment were real, the given distribution would benefit the participants themselves.

Example answers:

"Would like to live longer"

"i feel i relate to poorer people"

H8: *Individuals who give an answer in the Self-interest category are expected to distribute the same amount of health to all groups.*

All groups can be self-interested.

Don't know

This category contains all answers in which the writer makes clear not to be sure about their reason for distributing the resources the way they did.

Examples:

"To be honest the more this study went on near the end I got a bit confused and muddled in my head about all the figures and graphs. ?"

"I found it difficult to understand the principle of the multiplier as there were so many variables involved to consider like income individual characteristics and chance."

Fair

All answers mentioning "fairness" or "deservingness" (or a derivative word) and no further explanation as to why a certain distribution is more fair or a group more deserving were allocated to the category Fair. An answer that does include a reason belonging to one of the other categories, is categorized as such and not as fair. If not for the specific use of the concept of fairness and deservingness, these answers would be categorized as None.

Examples answers:

"I didnt think it was fair to take sex into account"

"I felt they deserved it more"

Ambiguous

Ambiguous contains answers that could be coded into either of at most two categories, but it is unsure which. When in doubt between more than two categories, an answer is coded as None. If an answer mentions multiple reasons and it definitely belongs to multiple categories, it is not Ambiguous. Answers that are coded as Ambiguous are also coded as the two possible other categories. The purpose of this category is to limit the loss of information that would follow from coding all Ambiguous answers as None. An example of an answer regarding Smoking is: "Good life choices should be rewarded". This was coded as Externalities, Responsibility and Ambiguous, as the category depends on the goal of rewarding good life choices. The goal could be justice (Responsibility) or benefitting society (Externalities).

Example answers:

Coded as Equality, Efficiency and Ambiguous:

"I think the resources should be shared to make the most or all peoples chances at good health."

Coded as Need, Externalities and Ambiguous:

"slightly more to females due to maternity health"

None

The None category contains all answers that do not fit into any of the other categories.

Examples answers:

"wealth should not be a factor"

"When it was tricky to balance the budget"

Empty

Responses in this category were completely empty, though these participants did enter their distribution preferences.

Additional coding explanations

Besides equality of health, participants could be concerned with equality of welfare. This could be the case when a participant distributes more health to an individual solely to compensate for their worse welfare. Although there were a few answers fitting this description (e.g.: "Life quality is worse for lower income in general so I wanted to reward them with greater health"), these were too few to warrant a separate category and for most answers the dimension that was meant by the writer could not be determined with certainty.

In some cases, the Need category is harder to distinguish from Equality and Efficiency. A type of answer also assigned to Need only contains a statement about the current state of society: rich people have access to all kinds of resources and poor people do not. Here it is not clear that the distribution of resources or health should be equal, so the answer is not coded as Equality. Instead, the rich are portrayed as better off as opposed to the poor who are worse-off. Some answers could also belong to either Need or Efficiency depending on the allocation that is made. For example, "Females live longer and are stronger" was written by one participant who allocated more resources to men. Here the statement seems to imply that females need the resources less. Another participant wrote: "females tend to live longer" and allocated more to females. This participant may have considered that the resources will be more efficiently used if given to females. This is the only case where the actual distribution made influenced the allocation to an outcome category, therefore slightly contributing to reverse causation in the regression analysis that follows in the Results chapter. The choice was made to minimally compromise the regression results this way rather than losing information about the outcome category shares.

A significant share of answers indicated to see the preferred distribution as "fair". This tells us nothing about why the participant thinks the distribution is fair, as all of the six categories could be considered a fair way of distribution by the participant. In addition, a group of people could deserve something because the current situation is unequal, because they are worse-off, because they earned it or because

of reasons other than health. There is a balance to be struck between giving the reader as much information as possible by creating more precise categories and overwhelming the reader with the number of categories. Due to the sheer number of answers concerning fairness, it was decided to group these together into a Fair-category.

Contrary to what the size of the None category might suggest, the answers do not indicate a seventh main principle category that was missed in the theoretical background. Rather, this allocation stems from a lack of precision in describing beliefs, resulting in indefinite conclusions. A large number of answers was similar in meaning to the following: "I did not look at gender", "I don't think wealth should affect their welfare". Possibly these individuals feel that the experiment is trying to steer them in a certain direction and therefore explicitly state the basis for not agreeing with that distribution. In any case, they do not specify the reason behind *their* distribution and therefore attributing their answer to any of the six main outcome categories would be merely speculative, even though the underlying sentiment might well belong to one of those.

These nuances of answers demonstrate the need for manual coding of the data and the shortcomings of Dictionary Based analysis, for in many instances the allocation could not be made accurately based on the flagging of a list of words.

3.7 Regression models

Regression models are used to identify the reasons behind participant's resource allocations to specific groups. Three multinomial logistic regression models (MNL) with robust standard errors were estimated with Health Distribution (Income), (Smoking) and (Sex) as the dependent variables in each and the categories of principles as independent variables. Health Distribution (Income), (Smoking) and (Sex) are three categorical variables that each contain one of four choices of distribution preferences. For Income, participants could indicate if they aimed to distribute health "More to poorer individuals", "More to richer individuals", "No difference" or "Do not know", of which the first three choices are included in the regression. Health Distribution (Smoking) and (Sex) store these same options concerning smokers and non-smokers, males and females. MNL is suitable as these variables have more than two categories that are not ordinal. The option "No difference" was used as the reference category. For the Income and Smoking regressions Equality, Need, Responsibility, Efficiency and Externalities are included as regressors and for the Sex regression Equality, Need, Responsibility and Efficiency. For the regression with Health Distribution (Income) as the dependent variable, the regression equation for category i (where $i \neq 3$) relative to the reference category, category 3, is given by equation (4). The equations for Smoking and Sex are the same, except all Income variables are replaced with those corresponding to the Sex and Smoking section (and Externalities (Sex) is excluded).

(4) $\log\left(\frac{P(Health\ Distribution\ (Income)=i)}{P(Health\ Distribution\ (Income)=3)}\right) = \beta_{i0} + \beta_{i1}\text{Equality (Income)} + \beta_{i2}\text{Need(Income)} +$

$\beta_{i3}\text{Responsibility (Income)} + \beta_{i4}\text{Efficiency(Income)} + \beta_{i5}\text{Externalities (Income)}$

To promote readability of the results (Wooldridge, 2010), the marginal effects of each independent variable on the probabilities of the different health distribution outcomes were obtained in Stata after running the MNL model. The marginal effect of an independent variable X_j on the probability of an outcome Y_i is given by equation (5). Here j indicates the category dummy and i the category of the dependent variable, for instance "More to poorer".

(5) $\text{ME}_{ij} = \frac{\partial P(Y=Y_i)}{\partial X_j}$

The estimated marginal effects are presented in the regression tables of the Results section according to equation (6).

(6) $\text{ME}_{ij} = \beta_{ij}$

Table 3.2 clarifies what the layout of the first regression table will therefore look like. Self-interest was left out from all three regressions as well as Externalities from the Sex regression, because the models could not be estimated accurately if these were included due to the low number of observations in these categories. Since there were also very few observations where the participants chose category 4, so "Do not know", for all three dependent variables, only the first three categories were considered in the regressions.

Table 3.2: Marginal effects

Independent variable	ME on Y_1	ME on Y_2	ME on Y_3
Equality	β_{11}	β_{21}	β_{31}
Need	β_{12}	β_{22}	β_{32}
Responsibility	β_{13}	β_{23}	β_{33}
Efficiency	β_{14}	β_{24}	β_{34}
Externalities	β_{15}	β_{25}	β_{35}

In the same way another three MNL's with robust standard errors were performed, using several descriptive variables as independent variables instead of the category dummies. Equation (7) depicts an example regression for the variables concerning Income. For the Smoking and Sex regressions the Health Distribution and Health Belief Difference variables were replaced by their counterparts and Health Belief Rich was replaced by Health Belief Non-smoker and Health Belief Female respectively.

(7) $\log\left(\frac{P(Health\ Distribution\ (Income)=i)}{P(Health\ Distribution\ (Income)=3)}\right) = \beta_{i0} + \beta_{i1}\text{Health Belief difference (Income)} +$

$\beta_{i2}\text{Health Belief Rich} + \beta_{i3}\text{Female} + \beta_{i4}\text{Income} + \beta_{i5}\text{Cigs Per Day} + \beta_{i6}\text{LeftRight} +$

$\beta_{i7}\text{LibertarianAuthoritarian}$

4 Results

4.1 Health Distributions

Figure 4.1 shows the chosen categories of the Health Distribution (Income), (Sex) and (Smoking) variables. Regarding the Income part of the experiment, a slight majority of participants chose to divide Health equally between poorer and richer individuals. Slightly less than half aimed to distribute more to the poor and only a small number gave more to the rich, thus H1a is not accepted. In contrast, less than a quarter of participants set out to distribute equally between smokers and non-smokers, while most tried to distribute more to non-smokers, even though the Health Belief for smokers is lower than for non-smokers. More information on the health beliefs for all groups is found in Appendix A. H1b predicted a distribution in favour of non-smokers and is therefore accepted. In the Sex-section of the experiment the largest number of people chose "No difference" compared to the Income- and Smoking-section. Of the remaining observations, more people aimed to distribute more to females than more to males. As the majority of participants distributes equally between males and females, H1c is accepted.
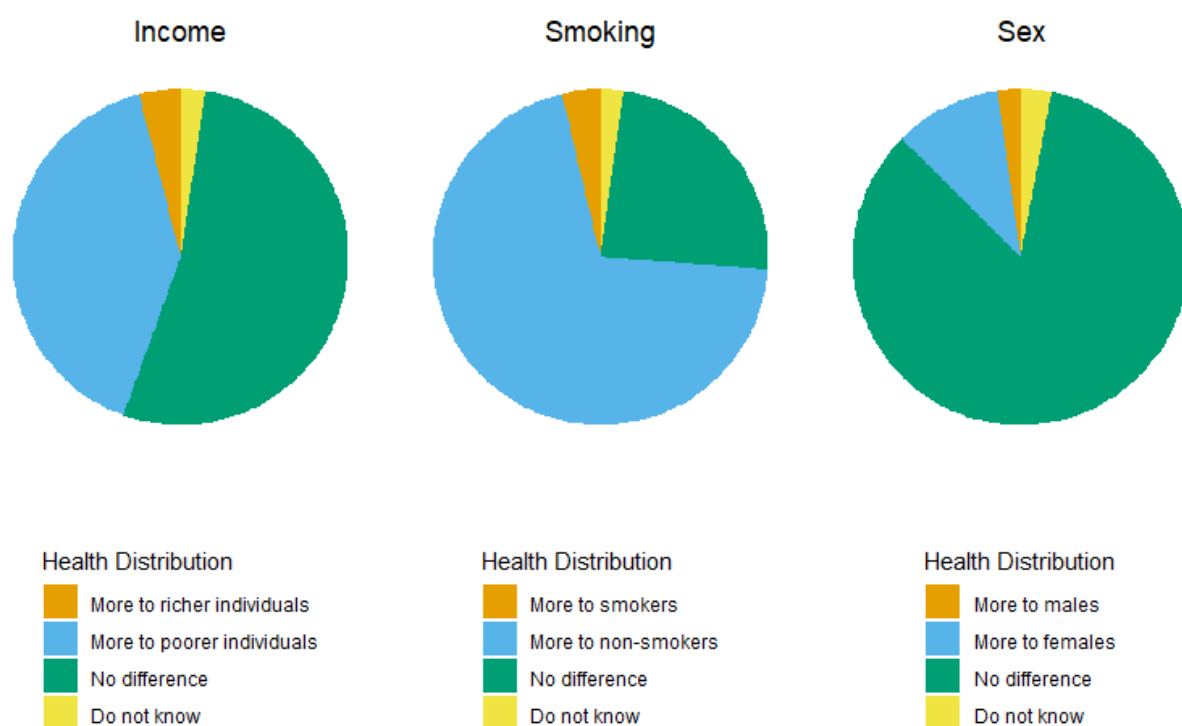
Figure 4.1 Preferences of Health Distribution (Income), Health Distribution (Smoking) and Health Distribution (Sex)

4.2 Visualisation of text data

Figure 4.2 contains three word clouds that provide a graphical representation of the word frequency in the responses contained by the Health Distribution Reason (Income), (Smoking) and (Sex) variables. The size of each word indicates its relative importance using the tf-idf frequency weighting.

The first word cloud for responses regarding the reasons behind the chosen Health Distribution of the Income-section prominently features the words "fair" and "equal", confirming the sentiment shown in figure 4.1, namely that most participants attempted to distribute health equally between the rich and the poor. These same words are most noteworthy in the responses of the Sex section, as seen in the third word cloud. Although "fair" is a key word in the Smoking responses too, most discussion seems to be around "smoke", "smoker" and "health", arguably suggesting what smoking does to health. Moreover, the words "choice" and "deserve" appear larger in this cloud, pointing to the believed responsibility the smoker has for their own (worse) health.

Income	Smoking	Sex

Figure 4.2 Word clouds of words of Health Distribution Reason (Income), (Smoking), (Sex) responses

Figures 4.3, 4.4 and 4.5 display correlation maps for the same responses in Health Distribution Reason (Income/Smoking/Sex) including the most frequent words according to the tf-idf frequency weighting. A line connecting two words indicates a Pearson correlation coefficient of at least 0.10. In these figures sentences take shape: in Figure 4.3 the words "distribute", "fair", "regardless" and "income" are connected on the left. These connections represent a common idea in the responses, for example: "I tried to give equal health to all individuals regardless of their income" and "Distribute evenly and fairly regardless of wealth". On the right of this figure the words "rich" and "poor" are connected to "need"

and "help". Although negations are not visible in this correlation map (words like "not" likely having gotten a low tf-idf frequency weighting, because of the high overall usage), seeing that Need is the second largest outcome category of the Income responses and that far more participants sought to distribute more Health to the poor than to distribute more to the rich, it might be concluded that these words represent the commonality of the notion that the poor need more help than the rich.
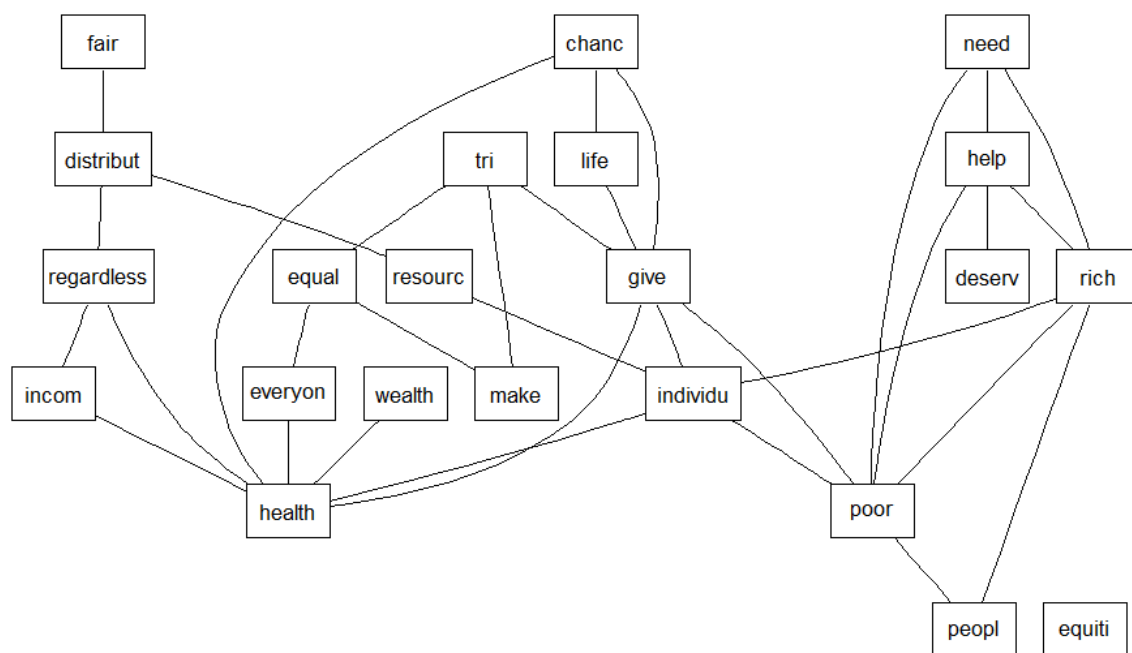
Figure 4.3 Correlation map of words in Health Distribution Reason (Income) responses

In Figure 4.4 the words "smoker", "deserv" and "health" are all interconnected, as well as "smoke", "health" and "choice", pointing again to answers regarding the responsibility of smokers for their health. It should be mentioned that different arguments might also be common, though not included in the figures if the language that is used to describe these arguments is more variable, the words therefore possibly not surpassing the frequency weighting threshold or the correlation threshold.
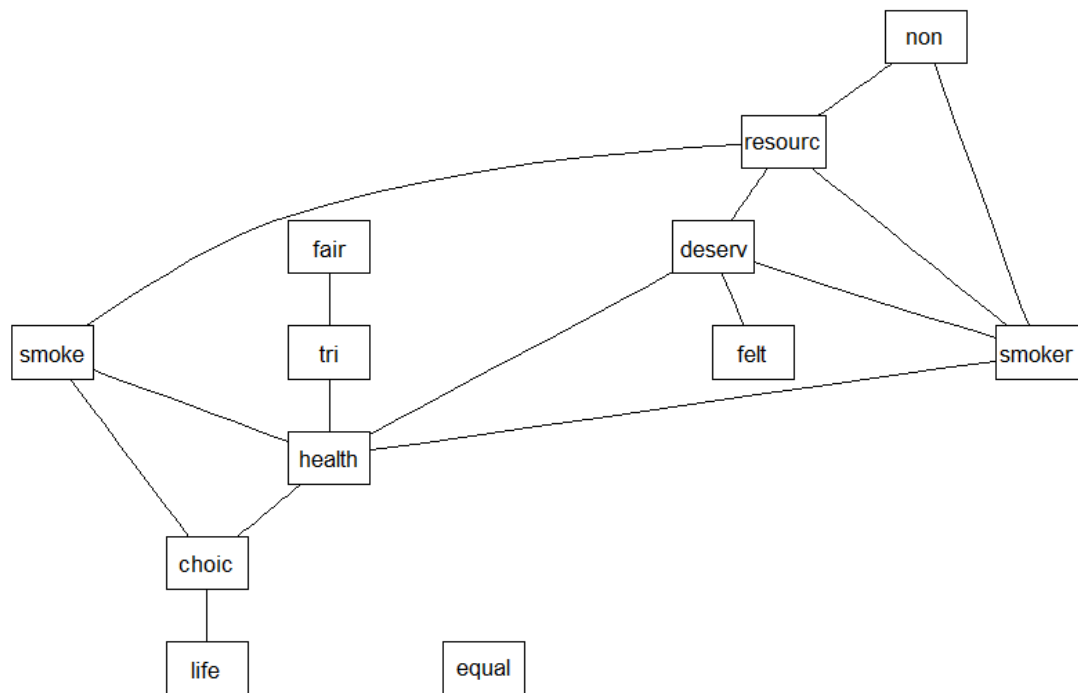
Figure 4.4 Correlation map of words in Health Distribution Reason (Smoking) responses

Finally, figure 4.5 demonstrates which words were often combined in the responses regarding the reasons behind the given Health Distribution between males and females. Some examples of clusters of words seen here are "sex", "didn't", "matter" and "distribute", "equal" and "reason", "discriminate", the latter probably depicting a view similar to the following response: "there is no reason to discriminate".

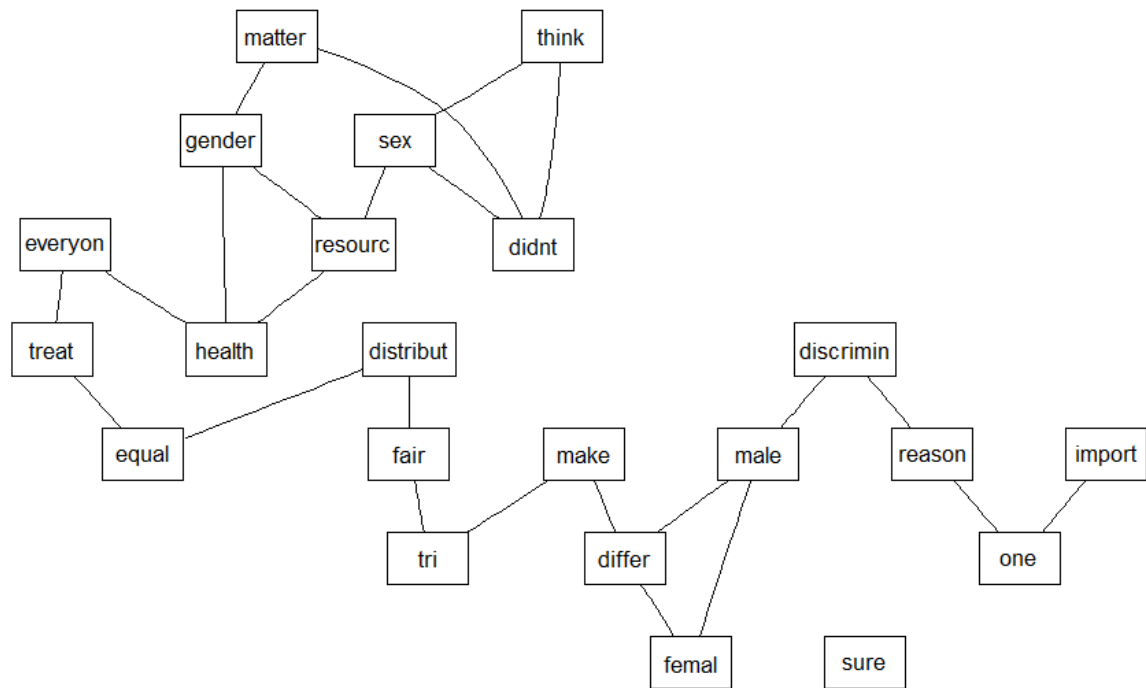*Figure 4.5 Correlation map of words in Health Distribution Reason (Sex) responses*

4.3 Principle category outcomes

Figure 4.6 displays the results of the manual coding into all outcome categories. The absolute counts the figure is based on can be found in Appendix B. The left column of figures shows the absolute category counts for Income, Smoking and Sex. For Income, Equality is by far the largest category, followed by Need. There is a stark difference with the category division of the Smoking data, where the number of responses in the Equality category is much lower, while Responsibility is the clear favourite. The number of answers coded as Efficiency and Externalities have also more than doubled. Of the reasons behind Health Distribution between the sexes, Equality is again the most common category. There is also a substantial number of answers categorized as None, far more than for Income and Smoking. A common theme in the None (Sex) category is that "sex is not a factor" in deciding the distribution. Similar to Equality, most participants whose answer was categorized as None distributed equally between the sexes. Self-interest is barely mentioned in any of the data sections. Considering these differences between Income, Sex and Smoking, especially in the categories Equality, Need and Responsibility, H2 is accepted: the reasons behind Health Distribution differ when related to Income, Smoking or Sex.
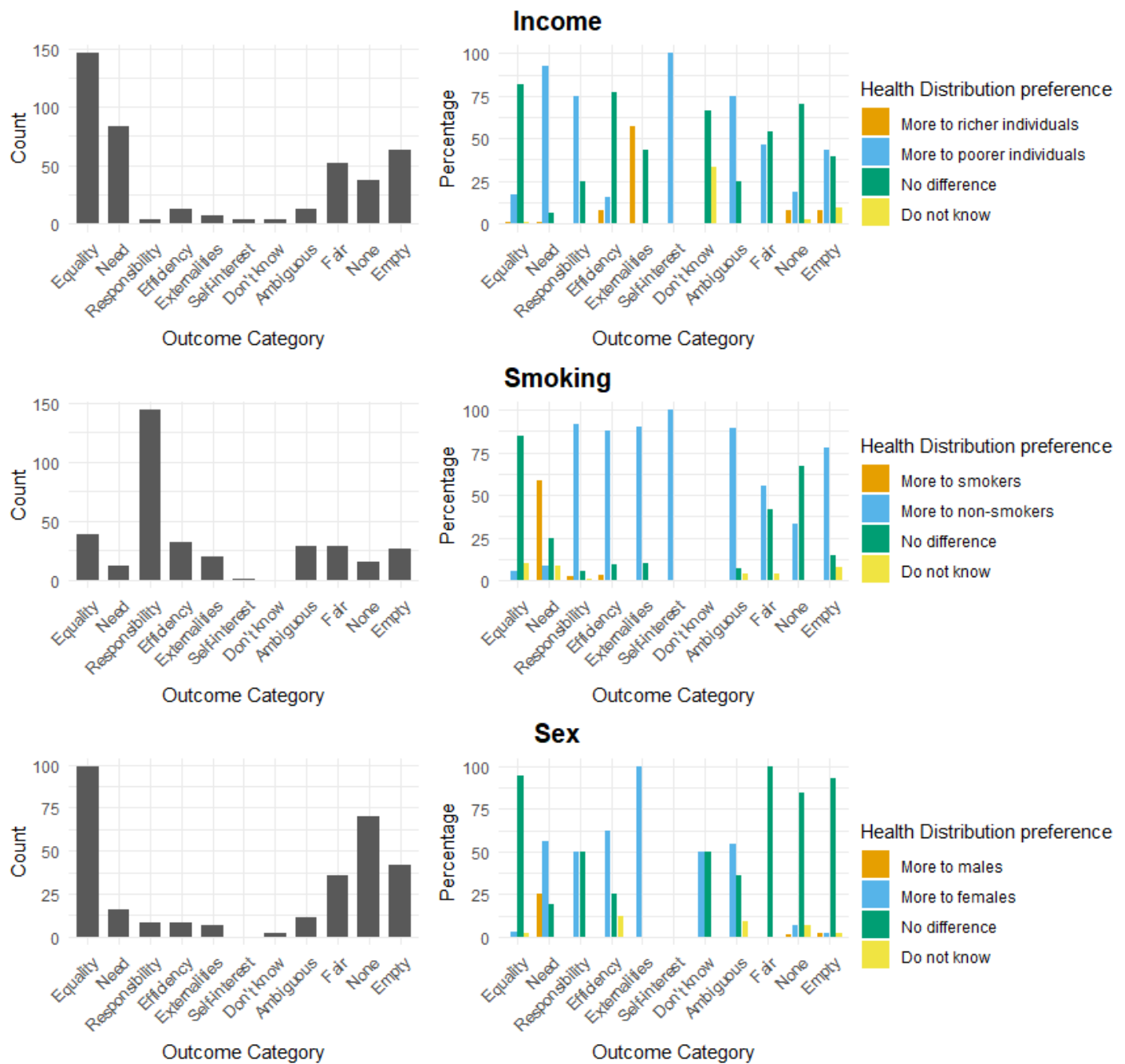
Figure 4.6 Absolute counts of observations in each category and relative percentages of Health Distribution per category

The figures in the column on the right of Figure 4.6 show the percentage of Health Distribution preferences in each category for Income, Smoking and Sex. Participants who's answer is coded as Equality predominantly distribute equally in all sections, in line with H3. As predicted by H4, among participants whose response is allocated as Need, distributing to the worst-off is most common, which is "More to poorer individuals" for Income and "More to smokers" for Smoking. Females are seen more often as in need of health than males, even though males are believed to have a lower life expectancy.

The distributions with regards to Responsibility differ, abiding by H5: related to Income, respondents primarily distribute more to poorer individuals (albeit based on only 4 observations), highlighting their lack of responsibility for their worse health, whereas related to Smoking nearly all respondents distribute more to non-smokers, emphasizing their choice to smoke and thus their responsibility for their worse health. Related to Sex, half of these respondents answered "More to females" and half "No difference". Consistent with H6 and the orthogonality assumption, participants who's answer was coded as Efficiency distributed equally among different income groups. However, this assumption holds neither when related to Smoking, as most participants distributed more to non-smokers, nor when related to Sex, where participants distribute more to Females most often. A common explanation is that the resources can be used more efficiently by these groups, since they live longer and are healthier anyways and the resources would be wasted on the unhealthy groups. This same logic was not seen applied to the poor on the other hand, so likely goes hand in hand with the use of other principles, like Responsibility. Contrary to H7, participants who gave a response coded as Externalities answered mainly "More to richer individuals" and "No difference" (Income), "More to non-smokers" (Smoking) and "More to Females" (Sex). The rare participants whose answers were allocated to Self-interest distributed "More to poorer individuals" and "More to non-smokers", inconsistent with H8. The next paragraph formally tests hypotheses H3 to H7. H8 cannot be tested, because of the lack of observations in the Self-interest category.

4.3 Regression results

Table 4.4 shows the regression results for the three MNL regressions with Health Distribution (Income/Smoking/Sex) as the dependent variable and several outcome categories as the independent variables. For Income the pseudo R^2 is 0.32, For Smoking 0.46 and for Sex 0.37. The pseudo R^2 calculated by Stata is McFadden's R^2, for which according to McFadden (1974), values between 0.2 and 0.4 represent an excellent fit. The chosen categories therefore explain a considerable amount of the variability found in the model.

Table 4.4 MNL regressions of principle categories on Health Distribution

Categories	Health Distribution			
A: Income	**More to richer** (1)	**More to poorer** (2)	**No difference** (3)	**Obs**
Equality	-0.0726* (0.0378)	-0.2061*** (0.0401)	0.2787*** (0.0385)	146
Need	-0.0184 (0.0351)	0.5017*** (0.0559)	-0.4833*** (0.0659)	83
Responsibility	-0.2618*** (0.0857)	0.3748** (0.1899)	-0.1129 (0.2040)	4
Efficiency	0.0288 (0.0402)	-0.3043 (0.1952)	0.2754 (0.1880)	13
Externalities	0.4870*** (0.1423)	-4.8071*** (0.3342)	4.3201*** (0.3343)	7
Pseudo R^2	0.3234	0.3234	0.3234	
Observations	15	153	201	
B: Smoking	**More to smokers** (4)	**More to non-smokers** (5)	**No difference** (6)	**Obs**
Equality	-0.2863** (0.1113)	-0.1436 (0.1183)	0.4299*** (0.0673)	39
Need	0.1055*** (0.0300)	-0.1635 (0.1141)	0.0580 (0.1076)	12
Responsibility	0.0138 (0.0171)	0.2036*** (0.0397)	-0.2174*** (0.0392)	145
Efficiency	0.0245 (0.0212)	0.1544 (0.0988)	-0.1789* (0.1028)	33
Externalities	-0.2398** (0.1052)	0.1948 (0.1215)	0.0450 (0.0864)	20
Pseudo R^2	0.4625	0.4625	0.4625	
Observations	10	190	65	
C: Sex	**More to males** (7)	**More to females** (8)	**No difference** (9)	**Obs**
Equality	-0.2583*** (0.0926)	0.0173 (0.0591)	0.2410*** (0.0859)	99
Need	0.0539*** (0.0197)	0.2274*** (0.0485)	-0.2813*** (0.0517)	16
Responsibility	-0.2702*** (0.0939)	0.2780*** (0.0600)	-0.0079 (0.0886)	8
Efficiency	-0.2658*** (0.0895)	0.3809*** (0.0567)	-0.1151 (0.0827)	8
Pseudo R^2	0.3705	0.3705	0.3705	
Observations	6	28	230	

Notes: This table shows the marginal effects for three multinomial logistic regressions performed in Stata 17.0. For the regression shown in Panel A, Health Distribution (Income) was used as the dependent variable and Equality (Income), Need (Income), Responsibility (Income), Efficiency (Income) and Externalities (Income) as the independent variables. Similarly, For the regression shown in Panel B the dependent variable is Health Distribution (Smoking) and the independent variables are Equality (smoking), Need (Smoking), Responsibility (Smoking), Efficiency (Smoking) and Externalities (Smoking). For Panel C the dependent variable is Health Distribution (Sex) and the independent variables are Equality (Sex), Need (Sex) Responsibility (Sex) and Efficiency (Sex). In all three panels the categories of the dependent variable are indicated by each column and the independent variables

are indicated by each row. The dependent variables are categorical variables with four categories of which three are included in each regression. The independent variables are all dummy variables, which equal 1 if the answer contained by Health Distribution Reason is coded into the category of the dummy. Pseudo R^2 shows McFadden's R^2 for each regression. The observations in the bottom row of each panel denote how many observations fall in each category of the dependent variable, while the observations in the column on the far right of each panel exhibit the number of observations for which the corresponding independent variable equals 1. Robust standard errors are placed in parentheses. *$p<0.10$, **$p<0.05$, ***$p<0.01$.

Panel A of table 4.4 shows the regression results for Health Distribution between individuals with different incomes, panel B for smokers and non-smokers and panel C for the sexes. Participants who gave an answer coded as Equality as the reason for their Health Distribution are more likely to opt for an equal distribution between groups defined by income (28.87 percentage points (pp), p=0.00), smoking status (42.99 pp, p=0.00) and sex (24.10 pp, p=0.01). They were significantly less likely to distribute more health to poorer individuals (20.61 pp, p=0.00), to smokers (28.63 pp, p=0.01) and to males (25.83 pp, p=0.01). H3 is therefore accepted for all three segments. Answers in the Need category are associated with lower probability of equal distribution between income groups (48.33 pp, p=0.00) and between the sexes (28.13 pp, p=0.00) and a higher probability of distributing more to poorer individuals (50.17 pp, p=0.00), to smokers (10.55 pp, p=0.00), to males (5.39 pp, p=0.01) and to females (22.74 pp, p=0.00). Since the poor and smokers are the disadvantaged populations in these scenarios in terms of health, these outcomes align with the predictions of H4. However, H4 is not accepted for Sex, because females are more preferred in the Need category than males, even though males are worse-off health-wise. That males have a lower health expectancy than females is believed by participants as well, as can be seen in Table A.1 in Appendix A. Participants whose reasons were categorised as Responsibility are more likely to ensure a distribution in favour of poorer individuals (37.48 pp, p=0.05), non-smokers (20.36 pp, p=0.00) and females (27.80 pp, p=0.00). This is a logical outcome if the answers in the Responsibility category mainly indicate a lack of responsibility of poorer individuals for their disadvantage and oppositely the choice of smokers for their unhealthy behaviour. Participants are less likely to distribute more to males (27.02 pp, p=0.00), aligned with the idea that males health have worse health than females and are responsible for that worse health. As no difference was predicted between males an females, H5 is accepted for Smoking and Income, but not for Sex. For answers in the Efficiency category individuals are less likely to distribute equally between smokers and non-smokers (17.89 pp, p=0.08), less likely to distribute more to males (26.58 pp, p=0.00) and more likely to distribute more to females (38.09 pp, p=0.00). Referring to the content of the responses, possibly more people believe females to live longer and therefore the resources to be more efficiently allocated to females. Even though, assuming orthogonality of multipliers, an equal distribution across all groups was expected, this is only true for Income and this result is not statistically significant (27.54 pp, p=0.14). H6 is thus not accepted for any of the segments and rejected for Sex. Individuals who gave a response coded as Externalities are more likely to answer "More to richer" (48.70 pp, p=0.00) and "No difference" (432.01 pp, p=0.00) and less likely to answer "More to poorer" (480.72 pp, p=0.00). For Smoking, they are only less likely to distribute more to smokers (23.98 pp, p=0.02). These findings are both partly in line with

H7, which predicted an equal distribution. The result for Smoking confirms perceived externalities of reducing the number of smokers, like the budget impact and second-hand smoking.

Appendix C contains regression results including observations marked as Ambiguous on the left and excluding those on the right. There are so few differences that it was decided for clarity to use all observations for the main regression tables. This shows that most findings in Table 4.4 are robust to some minor changes in coding. For only two variables the results significantly change: Responsibility (Smoking) and Efficiency (Smoking). The cause of this difference might be that 27 of the 29 Ambiguous observations were also coded as Responsibility and that 16 of 29 observations are ambiguous between Responsibility and Efficiency, and those are all of the Efficiency observations that are also Ambiguous.

Table 4.5 presents the marginal effects for the descriptive variables on the Health Distribution (Income), (Smoking) and (Sex) variables. Considering the pseudo R^2 of the three models, 0.04 for Income, 0.09 for Smoking and 0.09 for Sex, all three of them have a relatively low goodness of fit, as only a small percentage of the variation in the dependent variable is explained.

Panel A shows that for every unit increase in Health Belief Rich, individuals are 0.97 percentage points less likely to indicate that they would not differentiate between richer and poorer people for their health distribution ($p=0.03$). A 10 year increase in Health Belief Rich would therefore equate to an increased probability of almost 10%. Interestingly, the coefficient for Health Belief Difference in column (3) is not significant and positive. This suggests that a larger difference between the health beliefs for the rich and the poor does not affect the distribution choice, but a higher believed life expectancy overall does. Female, Income and Cigs Per Day seem to have no significant effect on this distribution. For the variables Left-Right and Libertarian-Authoritarian opposite effects are found. Every unit increase in Left-Right is associated with a 3.46 percentage points decrease in the probability of answering "More to poorer" ($p=0.02$) and a 3.25 percentage points increase in the probability of answering "No difference" ($p=0.03$). On the other hand, for every unit increase in Libertarian-Authoritarian, an individual has a 4.06 percentage points larger probability of answering "More to poorer" ($p=0.01$) and a 3.25 percentage points smaller probability of answering "No difference" ($p=0.01$). This result implies, unsurprisingly, that people both on the left side and the authoritarian side of the spectrum are more inclined to distribute more resources for health to poorer individuals. To illustrate, going from the far left to the far right of the spectrum would decrease this probability by 34.65 percentage points and going from extremely libertarian to extremely authoritarian would increase this probability by 40.61 percentage points, so quite substantial changes.

Table 4.5 MNL regressions of participant charachteristics on Health Distribution

Variables	Health Distribution		
A: Income	**More to richer (1)**	**More to poorer (2)**	**No difference (3)**
Health Belief Difference	0.0011	-0.0029	0.0018
	(0.0012)	(0.0033)	(0.0034)
Health Belief Rich	0.0024	0.0072	-0.0097**
	(0.0022)	(0.0044)	(0.0043)
Female	0.0428	0.1295	-0.0558
	(0.0291)	(0.0563)	(0.0564)
Income	0.0000	0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)
Cigs per day	-0.0000	-0.0033	-0.0033
	(0.0025)	(0.0068)	(0.0068)
Left-Right	0.0023	-0.0346**	0.0325**
	(0.0064)	(0.0148)	(0.0149)
Libertarian-Authoritarian	0.0006	0.0406**	-0.0412**
	(0.0078)	(0.0164)	(0.0168)
Pseudo R^2	0.0403	0.0403	0.0403
observations	15	153	201
B: Smoking	**More to smokers (3)**	**More to non-smokers (4)**	**No difference (5)**
Health Belief Difference	-0.0003	0.0190***	-0.0187***
	(0.0018)	(0.0053)	(0.0052)
Health Belief Non-smokers	0.0027	-0.0080*	0.0053
	(0.0020)	(0.0041)	(0.0038)
Female	-0.0213	0.1073*	-0.0860
	(0.0279)	(0.0589)	(0.0562)
Income	-0.0000	0.0000	0.0000
	(0.0000	(0.0000)	(0.0000)
Cigs per day	-0.0021	-0.0049	0.0070
	(0.0027)	(0.0061)	(0.0055)
Left-Right	0.0061	-0.0059	-0.0002
	(0.0058)	(0.0169)	(0.0158)
Libertarian-Authoritarian	-0.0027	0.0085	-0.0059
	(0.0080)	(0.0174)	(0.0156)
Pseudo R^2	0.0857	0.0857	0.0857
Observations	10	190	65
C: Sex	**More to males (7)**	**More to Females (8)**	**No difference (9)**
Health Belief Difference	0.0022	-0.0011	-0.0011
	(0.0024)	(0.0059	(0.0063)
Health Belief Females	0.0039*	-0.0032	-0.0007
	(0.0021)	(0.0021)	(0.0029)
Female	-0.0115	0.0560	-0.0444
	(0.0209)	(0.0445)	(0.0486)
Income	-0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
Cigs per day	-0.0421**	0.0026	0.0395**
	(0.0207)	(0.0054)	(0.0199)
Left-Right	-0.0070	-0.0118	0.0188
	(0.0061)	(0.0114)	(0.0126)
Libertarian-Authoritarian	0.0137	-0.0136	-0.0001
	(0.0098)	(0.0143)	(0.0169)
Pseudo R^2	0.0945	0.0945	0.0945
Observations	6	28	230

Notes: This table shows the marginal effects for three multinomial logistic regressions performed in Stata 17.0 Health Distribution (Income) was used as the dependent variable and Health Belief Difference (Income), Health Belief Rich, Female, Income, Cigs Per Day, Left-Right and Libertarian-Authoritarian as independent variables. The categories of the dependent variable are indicated by each column and the independent variables are indicated by each row. The dependent variable is a categorical variable with 4 categories of which 3 are included in the regression. Health Belief Difference and Health Belief Rich are continuous variables. Female is a dummy indicating 0 for Male and 1 for Female. Income and Cigs Per Day are continuous variables. Left-Right and Libertarian-Authoritarian are also continuous variables on a scale from 1 to 10. Pseudo R^2 shows McFadden's R^2. The observations in the bottom row of each panel denote how many observations fall in each category of the dependent variable. Robust standard errors are placed in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Panel B contains the coefficients for the regression with Health Distribution (Smoking) as the dependent variable. It shows a 1.90 percentage points increase in the probability of answering "More to non-smokers" for Health Belief Difference (Smoking) (p=0.00) and a 1.87 percentage points decrease in the probability of answering "No difference" (p=0.00). Paradoxically, a higher difference between the believed life-expectancies of smokers and non-smokers thus is associated with larger probability to distribute more health to the longer-living non-smokers. Perhaps individuals who dislike smoking (and smokers) the most also believe it is the most harmful. Furthermore, for every unit increase of Health Belief Non-Smokers, the probability to answer to distribute more to non-smokers decreases by 0.80 percentage points (p=0.05), suggesting that primarily a lower believed life-expectancy of poorer individuals is responsible for the previously found effect of Health Belief Difference. Again there are no significant effects found for Income and Cigs Per Day. Females do have a 10.73 percentage points higher probabiliy of answering to distribute more health to non-smokers (p=0.07). In this panel, political views also do not seem to significantly affect the indicated distribution choice.

In Panel C, marginal effects are displayed for the regression with Health Distribution (Sex) as the dependent variable. If Health Belief Females is increased by one unit, the probability of answering "More to Males" increases by 0.39 percentage points (p=0.06). Furthermore, for every extra ciggarette smoked per day, the probability of answering "More to males" decreases by 4.21 percentage points (p=0.04) and the probability of answering "No difference" increases by 3.95 percentage points (p=0.05). For Health Belief Difference, Female, Income, Left-Right and Libertarian-Authoritarian no significant effects are found.

Appendix D contains the results for several logit regressions with the categories as dependent variables and the same descriptive variables as above as the independent variables. As can be expected, there are not many significant correlations between these descriptives and the outcome categories.

5 Discussion and conclusion

5.1 Findings

Table 5.1 summarizes the research questions of this paper, hypotheses and findings. The results show that the preferred distribution of health by the UK public differs greatly based on the different characteristics of recipients: Income, Smoking and Sex. Most participants distributed health equally between the rich and the poor, though a large part also distributed more to the poor. Very few participants distributed more to smokers or males, even though they were both believed to have a lower life expectancy than non-smokers and females respectively. Between males and females an overwhelming majority chose to distribute health equally, while between smokers and non-smokers a similar majority preferred to distribute more health to non-smokers, thereby hypothetically increasing the already existing health gap. The fact that many participants distributed more to the poor, more to non-smokers and very few distributed more to males aligns with the literature, namely that people find health inequality between income groups more unfair than between the sexes (Lynch and Gollust, 2010) and that carers, non-smokers and low-income individuals are preferred for receiving life expectancy gains (Goto and Mori, 2019; Norman et al., 2012). With the different distributions, the reasons for those distributions differed as well. Equality was most often mentioned for the Income and Sex distributions, rather than Responsibility, which was the number one reason behind Smoking distributions. For Income and Sex, Responsibility was even one of the least mentioned reasons. Need is a substantially larger category for Income than for Smoking and Sex. Self-interest is a rarely found in any of the sections, contrary to the findings of Kroll & Davidovitz (1999), Kroll & Davidovitz (2003) and Carlsson et al. (2005). However, since in those studies participants are not explicitly asked whether they base their preferences on self-interest, it is possible that self-interest is an underlying factor for more people than those who are aware of it or are willing to admit it in their response. That Equality, Responsibility and Need are at the basis of most health distribution beliefs confirms results of Lim et al. (2012) and Cookson and Dolan (1999). The findings in the literature on the importance of Efficiency are mixed (Cookson and Dolan, 1999; Bae et al., 2023; Schwappach, 2002; Lim et al., 2012) and in this paper it is found to be mentioned scarcely, as are Externalities. The occurrence of both Efficiency and Externalities does increase for the Smoking section. Unlike the findings of Oedingen et al. (2019) and Kluegel and Mateju (1995), the reasons mentioned by most participants belong to only one category. It remains unclear if depending on the situation different principles are considered, or that in fact multiple principles are considered, but only the most prominent for the situation are mentioned. The fact that there are more observations in the Efficiency and Externalities category for Smoking than for Income and Sex suggests the former, because of the plausibility of implicit attributions of responsibility.

Table 5.1Research questions, hypotheses and conclusions

	A: Income	B: Smoking	C: Sex
Research question 1	*Does the UK general public distribute more health towards the rich or the poor, smokers or non-smokers and men or women?*		
H1 Result	More to the poor Majority No difference, also a large share More to the poor	More to non-smokers Large majority More to non-smokers	No difference Large majority No difference
Research question 2	*To which categories do the UK general public's beliefs behind resource allocation for health belong?*		
H2 Result	The reasons behind health distribution differ when related to Income, Smoking or Sex. Accepted. Responsibility and Efficiency are more prominent for Smoking, Need for Income.		
Research question 3	*What is the correlation between the category of health distribution beliefs and preferences for health distribution of the UK general public?*		
H3 Equality Result	No difference Accepted	No difference Accepted	No difference Accepted
H4 Need Result	More to the poor Accepted	More to smokers Accepted	More to males Rejected, more to females
H5 Responsibility Result	More to the poor Accepted	More to non-smokers Accepted	No difference Rejected, more to females
H6 Efficiency Result	No difference No significant effects	No difference No significant effects	No difference Rejected, more to females
H7 Externalities Result	No difference Positive significant effects for both More to richer and No difference	No difference Only significantly less to smokers	No difference Uncertain
H8 Self-interest Result	No difference Uncertain	No difference Uncertain	No difference Uncertain

Notes: The far left column indicates the research question, hypothesis and result. Column A contains all information regarding the Income section of the data, B for Smoking and C for Sex.

The categories have a high explanatory power over the direction of the preferred Health Distribution in all three segments of the data. The R^2's are considerably higher than those of the models using sample

characteristics as independent variables. There was also no notice of a substantial missing category among the responses. Thus, no significant argument has been found warranting a discouragement of future research using these categories in a comparable analysis.

Participants believe the difference in life expectancy to be largest between the different income groups and smallest between the sexes. A larger difference in health beliefs is only relevant for the Health Distribution between smokers and non-smokers.

5.2 Strengths and Limitations

The analysis of this paper provides a deeper understanding of the reasons behind people's health distribution preferences towards different groups of society and demonstrates that these reasons are highly dependent on the subjects of the distribution. The scope of previous studies is broadened by the creation of a new and more complete category list, enabled by the type of data that is used. Furthermore, this paper has contributed to the literature by showing that these categories have far more predictive power over health distribution preferences than several other relevant participant characteristics. These results can be extrapolated to the UK population as the sample is quite balanced on the political spectrum from left to right and libertarian to authoritarian, and effects of political orientation on the distribution choices are only found regarding Income. These effects become prominent once participants move from one extreme to another. If the distribution preferences for Income would be slightly different, conclusions about the differentiation found between distributions and distributive justice beliefs across different health groups still hold. Care should be taken for the extrapolation of these findings to other countries, as findings in the literature differ between countries and the effects of other environmental factors have not been assessed.

There are several limitations to these findings. Firstly, though the sample as a whole is large enough to obtain statistically significant results, once it is divided into the eleven categories, a few of the subsamples (Self-interest (Income), (Smoking), (Sex) and Externalities (Sex)) are too small to meaningfully assess their health distribution preferences. This could be solved by the use of a larger sample.

Secondly, this is a partly qualitative study and several components of the method are prone to subjectivity. Even though the forming of the categories was based on both philosophical theory and examples of similar studies, a different researcher might have ended up with slightly different categories to test the data to. It is challenging to prove or disprove the validity of the categories on their own, seeing that even philosophical experts do not all make the same overarching classification of principles of distributive justice. However, the regression results do advocate for this categorization system to be meaningful in this research.

Furthermore, manual coding is obviously subjective. Ideally, it is performed by multiple Human Judges. As time, resources and budget for this paper were highly limited, arranging for multiple people or even one more person knowledgeable about the topic to code the data was unfortunately not possible. To mitigate this problem, the choices made for the most common answers are explained in the methodology, so that the reader might decide for themselves whether they agree. Aside from deliberate choices, it is also possible that errors were made in the coding process through, for example, accidentally clicking the wrong category. This is not expected to have a large effect on the results, since Table A1 in Appendix A, from which Ambiguous observations are left out, suggest that a few minor changes in the coding should not change regression results very drastically, if there is not a pattern in the mistakes made (i.e. all of them belonging to the same category).

A slightly paradoxical element of the experiment is that the health that is distributed is supposed to be the final health outcome and multipliers are orthogonal, so a truly equal distribution of health would allocate the same amount of resources to the different individuals overall. This means that distributing the same resources to the poor and the rich would allow them the same overall health, as it would for the smokers and non-smokers and males and females. This concept might have been very counterintuitive for participants, given that many of them distributed more to the poor with the argument of equality and specifically mentioning that the rich enjoy all kinds of health advantages like private healthcare. Note that health advantages are mentioned, so this does not concern a compensation for welfare inequality. The clear set up of the experiment of advantaged versus disadvantaged individuals and the questions about participants' health beliefs might have fuelled this idea as well. For these reasons it is questionable whether the assumptions associated with orthogonality hold.

The distributions for Income, Smoking and Sex were tested as if they are completely separate groups in society, yet in reality they might be interrelated. If a participant believes poor people smoke more often or smokers are predominantly male, these beliefs may affect their distribution. In that case the differences in the results would probably be understated rather than overstated.

Furthermore, a limitation concerning the regression analysis is that there is one specific case of Health Distribution Reason answers for which the distribution choice was used as a basis for deciding which category the answer should belong to. This choice was deliberate, because of the conviction that indeed the distribution choice provides crucial additional information to the meaning of the argument. The resulting reverse causality in the regression models of Table 4.4 could have exaggerated the effects for the Need and Efficiency categories.

Lastly, Since the Ambiguous answers were also included in both of the categories they were ambiguous between, there could be a positive bias to a category size if there was a pattern in the Ambiguous answers. Again, Appendix A shows through the exclusion of Ambiguous observations that for most

categories this is not a major problem, except for the effects of Responsibility (Smoking) and Efficiency (Smoking), which may be untrustworthy.

Future research could mitigate subjectivity in the analysis in two main ways. One is to employ multiple human judges for manual coding of similar data. The other is to use supervised learning methods of text analysis (Grimmer and Stewart, 2013) on a sufficiently large sample of which a part is categorized by multiple human judges to train an algorithm that categorizes the rest. The lack of clarity of the responses that is found especially in the Ambiguous, Fair, None and Empty categories could be circumvented by conducting more in depth interviews or asking participants to give a more elaborate answer. Had there been in depth knowledge available of the intention behind these participants' answers, the category sizes might have differed through the accurate categorization of these answers in one of the six main categories. This difference would only be substantial if the outcome category is correlated with the lack of explanation in the answer, which in theory is possible. More detailed responses could thereby provide insight into the different forms of equality that are preferred. The distinction between equality of outcomes, resources and capabilities, for example, might be crucial for health policy choices and is a relevant topic for future research. Another interesting expansion of the research topic is the relationship between beliefs of principles of distributive justice and health policy preferences. By analysing the effects of these beliefs on the support for actual policies, rather than hypothetical health distributions, the gap from theory to reality could be further bridged. Finally, repeating a similar study in different political climates, but also cultural and social environments could provide more insight into how beliefs on distributive justice are affected by diverse surroundings.

5.3 Conclusion

To conclude, there are major differences found, both between the preferred health distribution between different income groups, smoking status and sexes, and between the reasons given for the health distributions of each group. The UK public appears most averse to health inequality between the rich and the poor, since they distribute more often to the poor than they distribute to non-smokers or males, who are also believed to have poorer health. The difference in health beliefs for these groups does not seem to play a major role in this discrepancy. More likely is that attributions of Need of the poor and Responsibility of smokers are at least a part of the cause. Several limitations could be solved by using a larger sample, more elaborate responses and an automated method of text analysis. The short answer format of the data did provide insight into what reasons the public comes up with when they are not limited by choices given. This insight confirms that the most important categories, Equality, Need and Responsibility, are covered in most discussed studies. Health policy makers should be aware that support for equality measures for different groups in society is most likely not consistent. Further research is necessary to inform on how different cultural and social group's beliefs on distributive justice vary.

References

Aalberg, T. (2003). *Achieving justice*. Brill Academic Pub. https://doi.org/10.1163/9789047402060

Albertsen, A., & Tsiakiri, L. (2023). Equality of Opportunity for Health: Personal Responsibility and Distributive Justice. In *Springer eBooks* (pp. 1–21). https://doi.org/10.1007/978-3-319-52269-2_86-1

Ali, S., Tsuchiya, A., Asaria, M., & Cookson, R. (2017). How Robust Are Value Judgments of Health Inequality Aversion? Testing for Framing and Cognitive Effects. *Medical Decision Making*, *37*(6), 635–646. https://doi.org/10.1177/0272989x17700842

Arneson, R. J. (1989). Equality and equal opportunity for welfare. *Philosophical Studies*, *56*(1), 77–93. https://doi.org/10.1007/bf00646210

Bae, E., Lim, M. K., Lee, B., Bae, G., & Hong, J. (2023). Public preferences in healthcare resource allocation: A discrete choice experiment in South Korea. *Health Policy*, *138*, 104932. https://doi.org/10.1016/j.healthpol.2023.104932

Burstein, P. (2003). The Impact of Public Opinion on Public Policy: A Review and an Agenda. *Political Research Quarterly*, *56*(1), 29. https://doi.org/10.2307/3219881

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

Cookson, R., & Dolan, P. (1999). Public views on health care rationing: a group discussion study. *Health Policy*, *49*(1–2), 63–74. https://doi.org/10.1016/s0168-8510(99)00043-3

Daniels, N. (2008). *Just health: Meeting health needs fairly.* Cambridge University Press.

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, *31*(3), 137–149. https://doi.org/10.1111/j.1540-4560.1975.tb01000.x

Dworkin, R. (1981). What is Equality? Part 2: Equality of Resources. *Philosophy & Public Affairs, 10*(4), 283-345.

Frankena, W. K. (1939). IV.—THE NATURALISTIC FALLACY. *Mind*, *XLVIII*(192), 464–477. https://doi.org/10.1093/mind/xlviii.192.464

Goto, R., & Mori, T. (2019). Comparison of Equity Preferences for Life Expectancy Gains: A Discrete Choice Experiment Among the Japanese and Korean General Public. *Value in Health Regional Issues*, *18*, 8–13. https://doi.org/10.1016/j.vhri.2018.05.004

Graham, H. (2009). *Understanding Health Inequalities*. McGraw-Hill Education.

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, *21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Jacobs, R. (1993). The health of nations. Cornell University press.

Kluegel, J. R., & Matějů, P. (1995). 9 Egalitarian vs. Inegalitarian Principles of Distributive Justice. In *De Gruyter eBooks* (pp. 209–238). https://doi.org/10.1515/9783110868944.209

Kroll, Y., & Davidovitz, L. (1999). Choices in Egalitarian Distribution: Inequality Aversion versus Risk Aversion. *Social Science Research Network*. tp://darp.lse.ac.uk/subjective/kroll_1999.pdf

Kroll, Y., & Davidovitz, L. (2003b). Inequality Aversion versus Risk Aversion. *Economica*, *70*(277), 19–29. https://doi.org/10.1111/1468-0335.t01-1-00269

Lim, M. K., Bae, E. Y., Choi, S., Lee, E. K., & Lee, T. (2012). Eliciting Public Preference for Health-Care Resource Allocation in South Korea. *Value in Health*, *15*(1), S91–S94. https://doi.org/10.1016/j.jval.2011.11.014

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

Lynch, J., Gollust, S.E. (2010). Playing Fair: Fairness Beliefs and Health Policy Preferences in the United States. *Journal of health, politics, policy and law, 35*(6), 849-87.

Marmot, M. (2020). *Health equity in England: the Marmot review 10 years on*. BMJ, 368, m693

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105-142). Academic Press.

McNamara, S., Tsuchiya, A., & Holmes, J. (2021). Does the UK-public's aversion to inequalities in health differ by group-labelling and health-gain type? A choice-experiment. *Social Science & Medicine*, *269*, 113573. https://doi.org/10.1016/j.socscimed.2020.113573

Miller, D. (1992). Distributive justice: What the people think. *Ethics*, *102*(3), 555–593. https://doi.org/10.1086/293425

Min, B., Ross, H., Sulem, E., Veyseh, A. P. B., Nguyen, T. H., Sainz, O., Agirre, E., Heinz, I., & Roth, D. (2021, 1 november). *Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey*. arXiv.org. https://arxiv.org/abs/2111.01243

Nicolas, G., X, B., & Fiske, S. (2019). Automated Dictionary Creation for Analyzing Text: An Illustration from Stereotype Content. *European Journal of Social Psychology*. https://doi.org/10.31234/osf.io/afm8k

Norman, R., Hall, J., Street, D., & Viney, R. (2012). EFFICIENCY AND EQUITY: a STATED PREFERENCE APPROACH. Health Economics, 22(5), 568–581. https://doi.org/10.1002/hec.2827

Nussbaum, M. C. (2000). *Women and human development: The capabilities approach.* Cambridge University Press.

Oedingen, C., Bartling, T., Mühlbacher, A., Schrem, H., & Krauth, C. (2019). Systematic review of public preferences for the allocation of donor organs for transplantation: Principles of Distributive Justice. *The Patient: Patient-Centered Outcomes Research*, *12*(5), 475–489. https://doi.org/10.1007/s40271-019-00363-0

Page, B. I., & Shapiro, R. (1983). Effects of Public Opinion on Policy. American Political Science Review, 77(1), 175–190. https://doi.org/10.2307/1956018

Popay, J., Whitehead, M., & Hunter, D. J. (2010). Injustice is killing people on a large scale--but what is to be done about it? *Journal Of Public Health*, *32*(2), 148–149. https://doi.org/10.1093/pubmed/fdq029

Rawls, J. (1971). *A theory of justice.* Harvard University Press.

Reid, S. W., McKenny, A. F., & Short, J. C. (2023). Synthesizing best practices for conducting Dictionary-Based Computerized Text Analysis research. In *Research Methodology in Strategy and Management* (pp. 43–78). https://doi.org/10.1108/s1479-838720220000014004

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

Robson, M., O'Donnell, O., & Van Ourti, T. (2024b). *Responsibility-Sensitive Welfare Weights for Health* (Tinbergen institute discussion paper TI 2024-045/V), tinbergen.nl/discussion-paper/6350/24-045-v-responsibility-sensitive-welfare-weights-for-health

Ruckert, A., Huynh, C., & Labonté, R. (2017). Reducing health inequities: is universal basic income the way forward? *Journal Of Public Health*, *40*(1), 3–7. ttps://doi.org/10.1093/pubmed/fdx006

Saldaña, J. (2016). *The coding manual for qualitative researchers*. https://lib.ugent.be/nl/catalog/rug01:002034125

Schwappach, D. L. (2002). Does it matter who you are or what you gain? an experimental study of preferences for resource allocation. *Health Economics*, *12*(4), 255–267. https://doi.org/10.1002/hec.713

Sen, A. (1979). Equality of what? In *S. McMurrin (Ed.), Tanner Lectures on Human Values* (Vol. 1, pp. 195-220). Cambridge University Press

Setiawan, E., Cassidy-Seyoum, S. A., Thriemer, K., Carvalho, N., & Devine, A. (2024). A Systematic Review of Methods for Estimating Productivity Losses due to Illness or Caregiving in Low- and Middle-Income Countries. *PharmacoEconomics*. https://doi.org/10.1007/s40273-024-01402-x

Stone, D. A. (2006). *The rhetoric of policy analysis: Health care reform*. In *Policy paradox: The art of political decision making* (pp. 45-76). W.W. Norton & Company.

Van Hootegem, A., Abts, K., & Meuleman, B. (2020). Differentiated Distributive Justice Preferences? Configurations of Preferences for Equality, Equity and Need in Three Welfare Domains. *Social Justice Research*, *33*(3), 257–283. https://doi.org/10.1007/s11211-020-00354-9

Wikler, D. 2002. Personal and Social Responsibility for Health. *Ethics and International Affairs 16*(2): 47–55.

Wilkinson, R. G. (2005). *The Impact of Inequality: How to Make Sick Societies Healthier*.

https://www.cdc.gov/pcd/issues/2006/jan/pdf/05_0156.pdf

Wilmot, S., & Ratcliffe, J. (2002). Principles of distributive justice used by members of the general

public in the allocation of donor liver grafts for transplantation: a qualitative study. *Health*

Expectations, *5*(3), 199–209. https://doi.org/10.1046/j.1369-6513.2002.00176.x

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.

Appendix A: Health beliefs

The results for the questions about health beliefs are displayed in Table A.1. The largest believed difference in life expectancy is the average of 18.11 years between the richest fifth of the UK population and the poorest fifth. This average difference is 14.92 years between heavy smokers and people who have been a non-smoker their entire life in the UK and 5.06 years between males and females in the UK.

Table A.1: Health beliefs of the rich, the poor, smokers, non-smokers, males and females

Variables	Mean	St. dev.	Min	Max
Health Belief Rich	80.42	7.09	45	90
Health Belief Poor	62.28	8.90	40	90
Health Belief Difference (Income)	18.11	9.70	-10	50
Health Belief Non-smokers	75.81	7.50	40	90
Health Belief Smokers	60.84	8.49	40	90
Health Belief Difference (Smoking)	14.92	6.53	0	35
Health Belief Females	76.57	8.26	40	90
Health Belief Males	71.52	8.32	40	85
Health Belief Difference (Sex)	5.06	4.30	-10	15

Notes: This table contains statistics of all variables concerning health beliefs. Health Belief Rich, Poor, Non-smokers, Smokers, Females and Males are continuous variables containing the believed life expectancy of these groups. Health Belief Difference (Income), (Smoking) and (Sex) are constructed by subtracting Health Belief Poor from Rich, Smokers from Non-smokers and Males from Females. The variables are indicated by the rows and the mean, standard deviation, minimum value and maximum value are indicated by the columns.

Appendix B: Absolute principle category counts

Table B.1 Absolute counts per category and Health Distribution for Income, Smoking and Sex

A: Income	Total observations	More to the rich	More to the poor	No difference	Don't know
Health Distribution	378	15	153	201	9
Equality	146	1	25	119	1
Need	83	1	77	5	0
Responsibility	4	0	3	1	0
Efficiency	13	1	2	10	0
Externalities	7	4	0	3	0
Self-interest	3	0	3	0	0
Don't know	3	0	0	2	1
Ambiguous	12	0	9	3	0
Fair	52	0	24	28	0
None	37	3	7	26	1
Empty	63	5	27	25	6

B: Smoking	Total observations	More to smokers	More to non-smokers	No difference	Don't know
Health Distribution	271	10	190	65	6
Equality	39	0	2	33	4
Need	12	7	1	3	1
Responsibility	144	3	132	8	1
Efficiency	32	1	28	3	0
Externalities	20	0	18	2	0
Self-interest	1	0	1	0	0
Don't know	0	0	0	0	0
Ambiguous	29	0	26	2	1
Fair	29	0	16	12	1
None	15	0	5	10	0
Empty	27	0	21	4	2

C: Sex	Total observations	More to males	More to females	No difference	Don't know
Health Distribution	272	6	28	230	8
Equality	99	0	3	94	2
Need	16	4	9	3	0
Responsibility	8	0	4	4	0
Efficiency	8	0	5	2	1
Externalities	7	0	7	0	0
Self-interest	0	0	0	0	0
Don't know	2	0	1	1	0
Ambiguous	11	0	6	4	1
Fair	36	0	0	36	0
None	70	1	5	59	5
Empty	42	1	1	39	1

Notes: This table contains the absolute counts of observations in each category and how many of those made which Health Distribution. The categories are all binary variables and the number of observations for which the category variable equals 1 is counted. The categories of the dependent variables Health Distribution (Income), (Smoking) and (Sex) are indicated by each column and the independent principle category variables by each row. Panel A contains the results for Income, panel B for Smoking and panel C for Sex.

Table B.1 contains the exact number of observations in each category, as well as the choices that were made for Health Distribution per category. The row on top of each panel depicts the numbers for the Health Distribution variables that were used for Figure 4.1. Underneath all outcome categories are listed with the total number of observations in each category and how many of those observations chose each answer for Health Distribution, used for Figure 4.6.

Appendix C: Exclusion of Ambiguous observations

Table C.1 MNL models regressing categories on Health Distribution variables with and without Ambiguous observations

Variables	With Ambiguous observations			Without Ambiguous observations		
Income	More to richer (1a)	More to poorer (2a)	No difference (3a)	More to richer (1b)	More to poorer (2b)	No difference (3b)
Equality	-0.0726*	-0.2061***	0.2787***	-0.0722*	-0.2219***	0.2942***
	(0.0378)	(0.0401)	(0.0385)	(0.0380)	(0.0432)	(0.0404)
Need	-0.0184	0.5017***	-0.4833***	-0.0222	0.4533***	-0.4311***
	(0.0351)	(0.0559)	(0.0659)	(0.0350)	(0.0597)	(0.0691)
Responsibility	-0.2618***	0.3748**	-0.1129	-0.1163*	2.2356***	-2.1193***
	(0.0857)	(0.1899)	(0.2040)	(0.0618)	(0.1889)	(0.1815)
Efficiency	0.0288	-0.3043	0.2754	0.0364	-0.3487	0.3123
	(0.0402)	(0.1952)	(0.1880)	(0.0424)	(0.2617)	(0.2471)
Externalities	0.4870***	-4.8071***	4.3201***	0.3635***	-3.2233***	2.8598***
	(0.1423)	(0.3342)	(0.3343)	(0.1017)	(0.2418)	(0.2434)
observations	15	153	201	15	144	198
Smoking	More to smokers (4a)	More to non-smokers (5a)	No difference (6a)	More to smokers (4b)	More to non-smokers (5b)	No difference (6b)
Equality	-0.2863**	-0.1436	0.4299***	-0.2740***	-0.1638	0.4378***
	(0.1113)	(0.1183)	(0.0673)	(0.1029)	(0.1250)	(0.0774)
Need	0.1055***	-0.1635	0.0580	0.3698***	-0.3525**	-0.0173
	(0.0300)	(0.1141)	(0.1076)	(0.1380)	(0.1607)	(0.1000)
Responsibility	0.0138	0.2036***	-0.2174***	0.2751**	0.0192	-0.2944***
	(0.0171)	(0.0397)	(0.0392)	(0.1149)	(0.1153)	(0.0613)
Efficiency	0.0245	0.1544	-0.1789*	0.2985**	-0.0542	-0.2443*
	(0.0212)	(0.0988)	(0.1028)	(0.1245)	(0.1625)	(0.1275)
Externalities	-0.2398**	0.1948	0.0450	-0.2610**	0.2721*	-0.0112
	(0.1052)	(0.1215)	(0.0864)	(0.1129)	(0.1511)	(0.1174)
Observations	10	190	65	10	164	63
Sex	More to males (7a)	More to females (8a)	No difference (9a)	More to males (7b)	More to females (8b)	No difference (9b)
Equality	-0.2618***	0.0414	0.2204**	-0.2534***	0.0297	0.2237**
	(0.0920)	(0.0445)	(0.0891)	(0.0963)	(0.0594)	(0.1000)
Need	0.0561***	0.1780***	-0.2342***	0.0596***	0.1347***	-0.1942***
	(0.0196)	(0.0416)	(0.0454)	(0.0194)	(0.0385)	(0.0388)
Responsibility	-0.2744***	0.2523***	0.0221	-0.2773***	0.2229***	0.0544
	(0.0947)	(0.0485)	(0.0941)	(0.0968)	(0.0498)	(0.0981)
Efficiency	-0.2547***	0.3107***	-0.0560	-0.1291***	0.9480***	-0.8189***
	(0.0860)	(0.0485)	(0.0861)	(0.0437)	(0.1916)	(0.1934)
Externalities	-0.0941**	0.9921***	-0.8980***	-0.1291***	0.9480***	-0.8189***
	(0.0392)	(0.1762)	(0.1782)	(0.0437)	(0.1916)	(0.1934)
Observations	6	28	230	6	22	226

Notes: This table shows the marginal effects for three multinomial logistic regressions performed in Stata 17.0. In columns 1-9a Ambiguous observations are included, whereas in columns 1-9b they are excluded. For the regression shown in Panel A, Health Distribution (Income) was used as the dependent variable and Equality (Income), Need (Income), Responsibility (Income), Efficiency (Income) and Externalities (Income) as the independent variables. Similarly, For the regression shown in Panel B the dependent variable is Health Distribution

(Smoking) and the independent variables are Equality (smoking), Need (Smoking), Responsibility (Smoking), Efficiency (Smoking) and Externalities (Smoking). For Panel C the dependent variable is Health Distribution (Sex) and the independent variables are Equality (Sex), Need (Sex) Responsibility (Sex) and Efficiency (Sex). In all three panels the categories of the dependent variable are indicated by each column and the independent variables are indicated by each row. The dependent variables are categorical variables with four categories of which three are included in each regression. The independent variables are all dummy variables, which equal 1 if the answer contained by Health Distribution Reason is coded into the category of the dummy. Pseudo R^2 shows McFadden's R^2 for each regression. The observations in the bottom row of each panel denote how many observations fall in each category of the dependent variable, while the observations in the column on the far right of each panel exhibit the number of observations for which the corresponding independent variable equals 1. Robust standard errors are placed in parentheses. *$p<0.10$, **$p<0.05$, ***$p<0.01$.

Table C.1 contains the results for the same regression model as Table 4.4, except with extra columns showing the marginal effects when Ambiguous observations are excluded. Most results are very similar in direction, only slightly differing in magnitude. Only for Responsibility (Smoking) and Efficiency (Smoking) the coefficients differ in sign. The cause of this difference might be that 27 of the 29 Ambiguous observations were also coded as Responsibility and that 16 of 29 observations are ambiguous between Responsibility and Efficiency, and those are all of the Efficiency observations that are also Ambiguous.

Appendix D: Effects of participant characteristics on principle categories

Six separate regular logistic regressions were run for the Income, Smoking and Sex section, totalling 18 regressions, to estimate the effect of the descriptive variables on the categories. The category variables Equality, Need, Responsibility, Efficiency, Externalities and Self-interest of (Income), (Smoking) and (Sex) were thus used as dependent variables and Health Belief Difference (Income), (Smoking) and (Sex), Health Belief Rich (for Income), Female, Income, Cigs Per Day, Left-Right and Libertarian-Authoritarian as the independent variables. The regression with Self-interest Sex as the dependent variable could not be estimated, due to the lack of sufficient observations, and was therefore omitted. Equation (1) shows an example equation that takes Equality (Income) as the dependent variable.

(1) $\log\left(\frac{P(\text{Equality (Income)}=1)}{1-P(\text{Equality (Income)}=1)}\right) = \beta_{i0} + \beta_{i1}\text{Health Belief Difference (Income)} +$

$\beta_{i2}\text{Health Belief Rich} + \beta_{i3}\text{Female} + \beta_{i4}\text{Income} + \beta_{i5}\text{Cigs Per Day} + \beta_{i6}\text{LeftRight} +$

$\beta_{i7}\text{LibertarianAuthoritarian}$

In the exact same way as described in Chapter 3, marginal effects were calculated and are displayed in Table D.1, D.2 and D.3.

In Table D.1, the marginal effects for the Income categories are shown. Every unit increase towards Authoritarian on the Libertarian-Authoritarian scale decreases the probability of the reasoning behind the Health Distribution (Income) being in the Equality category by 3.56 percentage points (p=0.03). For an increase from 0 to 10, so extremely Libertarian to extremely Authoritarian, this means an increase in this probability of 35.59 percentage points. There is also a significant effect of Libertarian-Authoritarian on the Need category, but positive. Individuals who are one unit more Authoritarian, are 2.63 percentage points more likely to give an answer belonging to Need (p=0.07). The Responsibility column shown the effect of Health Belief Difference (Income) on the probability of having the answer in Health Distribution Reason assigned to the Responsibility category. For every unit increase in Health Belief Difference (Income), the probability of this answer belonging to Responsibility decreases by 0.15 percentage points (p=0.06). Even if Health Belief Difference (Income) increases from 0 to 30 years, the probability only increases by 4.35 percentage points. Furthermore, the probability of an answer being in the Efficiency category increases by 0.45 percentage points (p=0.05) with each unit increase in the Health Belief for richer individuals.

Table D.1 Logistic regressions of participant characteristics on principle categories for Income

Variables	Categories					
	Equality	Need	Responsibility	Efficiency	Externalities	Self-interest
Health belief difference	-0.0012 (0.0034)	0.0008 (0.0028)	-0.0014* (0.0008)	-0.0007 (0.0009)	0.0007 (0.0006)	-0.0015 (0.0009)
Health belief rich	-0.0029 (0.0043)	0.0048 (0.0037)	0.0002 (0.0007)	0.0045* (0.0023)	-0.0003 (0.0009)	0.0015 (0.0009)
Female	-0.0315 (0.0547)	0.0157 (0.0481)	0.0015 (0.0153)	-0.0377 (0.0244)	0.0419 (0.0284)	-0.0115 (0.0166)
Income	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
Cigs per day	-0.0006 (0.0068)	-0.0070 (0.0066)	omitted	0.0011 (0.0013)	0.0004 (0.0017)	Omitted
Left-Right	0.0231 (0.0146)	-0.0163 (0.0139)	0.0022 (0.0042)	-0.0001 (0.0048)	0.0035 (0.0017)	-0.0041 (0.0034)
Libertarian-Authoritarian	-0.0356** (0.0168)	0.0263* (0.0145)	-0.0001 (0.0057)	-0.0013 (0.0060)	-0.0037 (0.0055)	0.0023 (0.0067)
Pseudo R^2	0.0225	0.0223	0.1183	0.1133	0.0931	0.1974
Observations	308	308	273	308	308	273

Notes: This table shows the marginal effects for six logistic regressions performed in Stata 17.0. For each regression a different principle category of Income is used as the dependent variable: Equality (Income), Need (Income), Responsibility (Income), Efficiency (Income), Externalities (Income) and Self-interest (Income). Health Belief Difference (Income), Health Belief Rich, Female, Income, Cigs Per Day, Left-Right and Libertarian-Authoritarian the independent variables in each regression. The dependent variable is indicated by each column and the independent variables are indicated by each row. The dependent variables are dummies that equal 1 if an answer is coded as such. Health Belief Difference and Health Belief Rich are continuous variables. Female is a dummy indicating 0 for Male and 1 for Female. Income and Cigs Per Day are continuous variables. Left-Right and Libertarian-Authoritarian are also continuous variables on a scale from 1 to 10. Pseudo R^2 shows McFadden's R^2. The observations in the bottom row of each panel denote how many observations fall in each category. Robust standard errors are placed in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table D.2 shows the same regressions for the questions about Smoking. There is a 0.28 percentage points (p=0.03) decrease in the probability of answering in the Need category associated with every unit increase in the difference between health beliefs for poorer and richer individuals. Secondly, the probability of giving an answer that belongs to the Externalities category decreases by 0.42 percentage points (p=0.05), when the difference in Health Beliefs for smokers and non-smokers increases by 1 unit.

Table D.2 Logistic regressions of participant characteristics on principle categories for Smoking

Variables	Categories					
	Equality	Need	Responsibility	Efficiency	Externalities	Self-interest
Health belief dif smoking	-0.0047	-0.0028**	0.0076	0.0039	0.0024	-0.0002
	(0.0043)	(0.0013)	(0.0054)	(0.0041)	(0.0022)	(0.0003)
Health Belief Non-smoker	0.0031	0.0001	-0.0053	-0.0026	-0.0042*	-0.0001
	(0.0028)	(0.0013)	(0.0042)	(0.0032)	(0.0022)	(0.0002)
Female	0.0020	0.0154	-0.0209	0.0613	0.0390	Omitted
	(0.0455)	(0.0274)	(0.0682)	(0.0469)	(0.0362)	
Income	0.0000	-0.0000	-0.0000	0.0000	0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Cigs per day	0.0020	-0.0006	-0.0122	-0.0036	0.0002	Omitted
	(0.0043)	(0.0022)	(0.0082)	(0.0051)	(0.0043)	
Left-Right	0.0046	-0.0000	-0.0109	0.0000	0.0105	-0.0009
	(0.0130)	(0.0047)	(0.0189)	(0.0146)	(0.0107)	(0.0013)
Libertarian-Authoritarian	-0.0105	-0.0116	-0.0025	0.0293	-0.0111	-0.0005
	(0.0135)	(0.0085)	(0.0223)	(0.0189)	(0.0111)	(0.0011)
Pseudo R^2	0.0187	0.0839	0.0206	0.0555	0.0600	0.0884
Observations	225	225	225	225	225	109

Notes: This table shows the marginal effects for six logistic regressions performed in Stata 17.0. For each regression a different principle category of Smoking is used as the dependent variable: Equality (Smoking), Need (Smoking), Responsibility (Smoking), Efficiency (Smoking), Externalities (Smoking) and Self-interest (Smoking). Health Belief Difference (Smoking), Health Belief Non-smoker, Female, Income, Cigs Per Day, Left-Right and Libertarian-Authoritarian the independent variables in each regression. The dependent variable is indicated by each column and the independent variables are indicated by each row. The dependent variables are dummies that equal 1 if an answer is coded as such. Health Belief Difference and Health Belief Non-smoker are continuous variables. Female is a dummy indicating 0 for Male and 1 for Female. Income and Cigs Per Da y are continuous variables. Left-Right and Libertarian-Authoritarian are also continuous variables on a scale from 1 to 10. Pseudo R^2 shows McFadden's R^2. The observations in the bottom row of each panel denote how many observations fall in each category. Robust standard errors are placed in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Regarding the questions relating Sex, table D.3 shows, first of all, that Females have a 13.95 percentage points lower probability of giving an answer that belongs to the Equality category compared to Males (p=0.03). is found that for every unit increase towards Authoritarian on the scale from 0 to 10, the probability of giving an answer belonging to the Efficiency category decreases by 1.12 percentage points (p=0.07). The change in probability can therefore never be more than 11.20 percentage points. Finally, the probability of the answer being in the Externalities category decreases by 0.23 percentage points for every unit increase in the Health Belief for Females (p=0.06). Coefficients for Self-interest could not be estimated, probably due to the low number of observations in this category.

Table D.3 Logistic regressions of participant characteristics on principle categories for Sex

| Variables | Categories | | | | |
	Equality	Need	Responsibility	Efficiency	Externalities
Health Belief Difference	-0.0063	-0.0030	0.0021	0.0040	-0.0007
	(0.0076)	(0.0038)	(0.0036)	(0.0041)	(0.0015)
Health Belief Female	-0.0020	-0.0007	0.0010	-0.0008	-0.0023*
	(0.0042)	(0.0014)	(0.0013)	(0.0020)	(0.0012)
Female	-0.1395**	0.0167	0.0181	0.0421	0.0401
	(0.0648)	(0.0301)	(0.0291)	(0.0345)	(0.0369)
Income	-0.0000	0.0000	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Cigs per day	-0.0046	0.0009	Omitted	Omitted	Omitted
	(0.0071)	(0.0029)			
Left-Right	0.0212	-0.0036	-0.0134	0.0051	0.0003
	(0.0180)	(0.0087)	(0.0086)	(0.0071)	(0.0037)
Libertarian-Authoritarian	0.0037	0.0079	-0.0050	-0.0112*	-0.0053
	(0.0213)	(0.0137)	(0.0089)	(0.0062)	(0.0070)
Pseudo R^2	0.0308	0.0186	0.1261	0.0795	0.1690
Observations	226	226	207	207	207

Notes: This table shows the marginal effects for six logistic regressions performed in Stata 17.0. For each regression a different principle category of Income is used as the dependent variable: Equality (Sex), Need (Sex), Responsibility (Sex), Efficiency (Sex) and Externalities (Sex). Health Belief Difference (Sex), Health Belief Female, Female, Income, Cigs Per Day, Left-Right and Libertarian-Authoritarian the independent variables in each regression. The dependent variable is indicated by each column and the independent variables are indicated by each row. The dependent variables are dummies that equal 1 if an answer is coded as such. Health Belief Difference and Health Belief Female are continuous variables. Female is a dummy indicating 0 for Male and 1 for Female. Income and Cigs Per Day are continuous variables. Left-Right and Libertarian-Authoritarian are also continuous variables on a scale from 1 to 10. Pseudo R^2 shows McFadden's R^2. The observations in the bottom row of each panel denote how many observations fall in each category. Robust standard errors are placed in parentheses. *p<0.10, **p<0.05, ***p<0.01.