

London School of Economics and Political Science

**Just health responsibility**

A comparative analysis focussing on the role of individual behaviour in relation to cancer and weight-control policy in German and US health care systems

**Harald Thomas Schmidt**

LSE ID: 200727771

A thesis submitted to the Department of Social Policy,  
London School of Economics,  
for the degree of Doctor of Philosophy

London, March 2012

## Declaration of Authorship

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I consider the work submitted to be a complete thesis fit for examination. I authorise that, if a degree is awarded, a paper and/or electronic copy of my thesis will be deposited in the British Library of Political and Economic Science and that, except as provided for in regulation 32 it will be made available for public reference and inter-library loan. I authorise that by submitting my thesis to the Library, the Library is permitted to digitise it to assist in the long-term preservation and accessibility of my thesis.

I authorise the School to supply a copy of the abstract of my thesis for inclusion in any published list of theses offered for higher degrees in British universities or in any supplement thereto, or for consultation in any central file of abstracts of such theses.

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I understand that if my thesis is not approved by the examiners, this declaration will become void.

## Abstract

This thesis seeks to examine the appropriate role of individual behaviour and responsibility in relation to cancer and weight-control policy in German and US health care systems. It contains six main parts.

The first describes and compares the ways in which personal responsibility features in law and policy in both countries. It analyses salient differences in underlying motivation and characterization and highlights ethical tensions that arise from these provisions and their implementation.

The second part reviews what established normative theories can do to address the issues that have been identified. It argues that these frameworks lack specificity and are ill-suited as a basis for policy in pluralist societies. It provides an analysis of different notions of the concept of personal responsibility, and makes a proposal for an overarching framework, adopting a procedural justice account that draws on work by Norman Daniels, Jim Sabin and Thomas Scanlon.

The third part systematically reviews survey literature on the proper role of personal responsibility and develops an instrument for semi-structured interviews with physicians and population-level surveys in the US and Germany. The instrument complements this earlier survey work and explores key themes that arose in the analysis of policy documents and the philosophical literature.

Based on this instrument, the fourth part analyses the findings from twenty semi-structured interviews with primary care physicians and oncologists in Berlin, Germany and Philadelphia, USA.

The fifth part presents findings from three population level surveys of 1,000 respondents each. Two surveys with identical instruments were conducted with non-probability samples (census-adjusted proportional quota sampling with regard to income) in Germany and the US, and one, using a subset of questions, was administered to a probability-based sample in the US. Findings are discussed comparatively between countries and in view of the interviews with physicians.

The last part concerns the policy implications of the analysis, and applies the framework proposed in the thesis to the case of colon cancer screening. It seeks to defend an incentive policy that attaches financial advantage to attending counselling on the advantages and disadvantages of colon cancer screening, building also on findings from the surveys, and interviews with physicians. The final chapter highlights a range of general policy implications for the evaluation and implementation of programmes seeking to incentivise personal responsibility.

## Contents

<b>List of Appendices.....</b>	<b>5</b>
<b>List of Tables and Boxes.....</b>	<b>7</b>
<b>List of Figures.....</b>	<b>9</b>
<b>List of Acronyms.....</b>	<b>11</b>
<b>Preface and Acknowledgements.....</b>	<b>12</b>
<b>Chapter 1: Introduction .....</b>	<b>15</b>
1.1 Background .....	15
1.2 Personal responsibility in the US and Germany .....	18
1.3 Research questions .....	20
1.4 The thesis structure .....	21
<b>Chapter 2: Responsibility, ‘carrots’ and ‘sticks’ in health care policy in Germany and the USA .....</b>	<b>25</b>
2.1 Introduction.....	25
2.2 On ‘carrots’ and ‘sticks’: when is an incentive an incentive?.....	26
2.3 Health responsibility in the German statutory health insurance .....	29
2.4 Health responsibility in the US employer-based insurance system .....	40
2.5 Conclusion.....	51
<b>Chapter 3: An ethical framework for evaluating health responsibility policies</b>	<b>55</b>
3.1 Introduction .....	55
3.2 Personal responsibility and behaviour change in the spectrum of coercion and ‘doing nothing’ .....	56
3.3 Personal responsibility: conceptual analysis.....	60
3.4 Personal responsibility as co-responsibility .....	68
3.5 Assessing the reasonableness of health responsibility policies.....	71
3.6 Conclusion.....	87
<b>Chapter 4: Incentivising individual behaviour: review of the survey and interview literature, and development of an instrument for new research .....</b>	<b>89</b>
4.1 Introduction .....	89
4.2 Key themes in relevant survey and interview work so far .....	91

4.3 Discussion of findings and development of a new survey instrument.....	111
4.4 Conclusion.....	117
<b>Chapter 5: Health responsibility: physicians' views.....</b>	<b>118</b>
5.1 Introduction and Aims.....	118
5.2 Methods.....	120
5.3 Findings.....	127
5.4 Discussion .....	156
5.5 Limitations .....	160
5.6 Conclusion.....	161
<b>Chapter 6: Health responsibility: views of the public.....</b>	<b>162</b>
6.1 Introduction and Aims.....	162
6.2 Methods.....	163
6.3 Findings.....	174
6.4 Discussion .....	220
6.5 Limitations .....	237
6.6 Conclusion.....	239
<b>Chapter 7: Policy implications: the case of colon cancer policy .....</b>	<b>241</b>
7.1 Introduction.....	241
7.2 Colon cancer prevention: denying the right not to think, but preserving the right to refuse treatment .....	242
7.3 Conclusion.....	260
<b>Chapter 8: Conclusion .....</b>	<b>262</b>
8.1 Introduction.....	262
8.2 Summary of findings.....	263
8.3 Policy implications.....	276
8.4 Limitations and next steps.....	285
8.5 Closing .....	290
<b>References .....</b>	<b>291</b>

## Appendices

### Chapter 2

Appendix 2A: Health-related activities incentivised for insurees of the Betriebskrankenkassen .....	307
Appendix 2B: Health care premiums of covered US workers by plan type, company size and employer/employee contribution .....	308
Appendix 2C: Cost shifting through wellness incentives: BeniComp® Advantage .....	309

### Chapter 3

Appendix 3A: Implications of different policy responses to the five groups problem.....	311
--	-----

### Chapter 4

Appendix 4A: Initial draft instrument for survey and semi-structured interviews .....	313
--	-----

### Chapter 5

Appendix 5A: Method of recruiting physicians via professional organisations, snowball sampling, and practice area profiling .....	319
Appendix 5B: Notes on use of NVIVO as computer-assisted qualitative data analysis software.....	321
Appendix 5C: development of code-list .....	322
Appendix 5B: Physician recruitment: email to German primary care physicians (version for oncologist adapted accordingly) .....	323
Appendix 5E: Physician recruitment: email to US oncologists .....	324
Appendix 5F: Physician recruitment: email to US primary care physicians, as edited and circulated by <i>Clinical Care Associates</i> .....	325
Appendix 5G: Physician interview guide (English version).....	326
Appendix 5F: Physician interview guide (German version).....	330
Appendix 5I: Initial code list.....	336
Appendix 5J: Final code list.....	338

**Chapter 6**

Appendix 6A: Changes in the survey instrument versus the guide used in physician interviews .....	342
Appendix 6B: Overview of the statistical significance of independent variables .....	344
Appendix 6D: Sample email from <i>Knowledge Network</i> to survey panel members about new available surveys .....	352
Appendix 6E: Overview of the effects of applying age-weights to sample Ger. ....	353

**Chapter 8**

Appendix 8A: A randomized controlled trial to study personalized incentives ..	355
--	-----

**Volume B** (on CD, see sleeve inside back cover): Statistical reports relating to the three surveys analysed in Chapter 6. CD also includes an electronic copy of the thesis.

## List of Tables and Boxes

### Chapter 2

Table 2.1: Mean health care costs among incentive program participants and their matched controls, 2006 (Evaluation of Bonus programmes of the *Techniker Krankenkasse*)

Table 2.3: Wellness incentives and cost-shifting options: impact on cost of coverage as passed on to consumers

Table 2.4: Principal arguments for and against promoting personal responsibility for health in policy

### Chapter 3

Table 3.1: Implications of different policy responses to the five groups problem

Box 3.1: The intervention ladder (*Nuffield Council on Bioethics*)

Box 3.2: Seven tests to evaluate the appropriateness of health responsibility policies

### Chapter 4

Box 4.1: MeSH terms for literature on empirical work relevant for questions around personal responsibility

Table 4.1: Quantitative findings from literature search on empirical work relevant for questions around personal responsibility

Table 4.2: Health behaviour attitudes/Germany: findings from a representative survey (Diederich/Schreier 2010)

Table 4.3: “Lifestyle” in relation to other factors regarding priority in treatment: findings from the 1998 Eurobarometer survey

Table 4.4: Should some groups of patients have higher or lower priority for treatment than others? Findings on changes in attitudes of participants of group discussions in the UK (Dolan, Cookson, and Ferguson 1999)

Table 4.5: Agreement of Norwegian physicians with statements about personal responsibility (Bringedal and Feiring 2011)

Table 4.6: Attitudes of UK doctors towards refusing treatment (Medical Hospital Doctor Magazine 2008)

### Chapter 5

Table 5.1: Code list of interviewed physicians and key attributes



## Chapter 6

Table 6.1: Numbers of respondents in samples US1 and Ger entering amounts for three types of incentive structures (overall and those entering amounts >\$1)

Table 6.2: Differences in the use participation and attainment incentives over time (in %) US1, Ger

Table 6.3: Differences between demographic groups (in \$ and as factors) between “false carrot” (FC) and “stick” (S), US1 and Ger

Table 6.4: Ranking of five scenarios exploring effort, desert and health behaviour in relation to weight control, population-level survey samples and physician interviews

Table 6.5: Significance of independent variables for past, current and future use of participation and attainment incentives in samples US1 and Ger

## List of Figures

### Chapter 2

Fig. 2.1: Companies' reasons for offering wellness incentive programmes

Fig. 2.2: The 'responsibility-gap' of unhealthy employees (Steve Burd)

### Chapter 4

Fig. 4.1: Priority in treatment/Germany: findings from a representative survey (Liedke 2009)

Fig. 4.2: Norwegian physicians' views on responsibility-related factors of relevance in priority setting decisions (Bringedal and Feiring 2011)

### Chapter 6

Fig. 6.1: Income distribution (in %): US and German population & US1 and Ger

Fig. 6.2: Health insurance coverage (in %): US population & US1 and US2

Fig. 6.3: Self-assessed health status (in %): US and German population & US1, US2, Ger

Fig. 6.4: Relationship of income and health in US1, US2 and Ger

Fig. 6.5: BMI distribution (in %): US & German population, US1, US2, Ger

Fig. 6.6: Incentive program use over time (in %): US1 and Ger

Fig. 6.7: Incentive use over time (in %), by income, US1 and Ger

Fig. 6.8: Attainment incentive use (in %) by health status, US1 and Ger

Fig. 6.9: Participation incentive use (in %), by health status, US1 and Ger

Fig. 6.10: Incentive use over time (in %), by BMI, US1 and Ger

Fig. 6.11: Percentage of respondents with different health status who rated five scenarios exploring effort, desert and health behaviour as unfair/completely unfair, US1 and Ger

Fig. 6.12: Percentage of respondents with different health status who rated five scenarios exploring effort, desert and health behaviour as fair/completely fair, US1 and Ger

Fig. 6.13: Percentage of respondents with different BMIs who rated five scenarios exploring effort, desert and health behaviour as unfair/completely unfair, US1 and Ger

Fig. 6.14: Percentage of respondents with different BMIs who rated five scenarios exploring effort, desert and health behaviour as fair/completely fair, US1 and Ger

Fig. 6.15: Levels (in \$) for three types of incentive structures, M and Mdn, gross and amounts >1\$, US1 and Ger

Fig. 6.16: Incentive levels (in \$), 3 types of structures, by age (US1 and Ger)

Fig. 6.17: Incentive levels (in \$), 3 types of structures, by gender (US1, Ger)

Fig. 6.18: Incentive levels (in \$), 3 types of structures, by income (US1, Ger)

Fig. 6.19: Incentive levels (in \$), 3 types of structures, by health status (US1, Ger)

Fig. 6.20: Incentive levels (in \$), 3 types of structures, by BMI (US1, Ger)

Fig. 6.21: Current use of participation and attainment incentives (in %), by lowest and highest income, health, and BMI groups (in %), US1 and Ger

Fig. 6.22: Intended future use of participation and attainment incentives (in %), by lowest and highest income, health, and BMI groups, US1 and Ger

## Acronyms

ACP: American College of Physicians  
 ANOVA: analysis of variance  
 BKK: Betriebskrankenkassen, Union of German employer-based sickness funds  
 BMA: British Medical Association  
 BMG: Bundesministerium für Gesundheit, German Ministry of Health  
 BMI: Body Mass Index  
 CAQDAS: computer-assisted qualitative data analysis software  
 CCA: Clinical Care Associates (primary care network of the University of Pennsylvania Health System)  
 CDC: Centers for Disease Control and Prevention  
 CDU: Christlich Demokratische Union Deutschlands, Christian Democrats  
 CSDH: Commission on Social Determinants of Health (of the WHO)  
 DHHS: US Department of Health and Human Services  
 DoL: US Department of Labor  
 DoT: US Department of Treasury  
 FDP: Freie Demokratische Partei, Free Democrats  
 G-BA: Gemeinsamer Bundesausschuss, Federal Joint Commission  
 GKV-FinG: Gesetz zur nachhaltigen und sozial ausgewogenen Finanzierung des Gesundheitswesens, Law for the sustainable and equitable financing of health care  
 GKV-WSG: Gesetz zur Stärkung des Wettbewerbs in der gesetzlichen Krankenversicherung, Law to strengthen competition among providers of statutory health-insurance  
 HDHP: High Deductible Health Plan  
 HIPAA: Health Insurance Portability and Accountability Act  
 HMO: Health Maintenance Organization  
 IOM: Institute of Medicine  
 MeSH: medical subject headings  
 NBGH: National Business Group on Health  
 OECD: Organisation for Economic Co-operation and Development  
 POS: Point Of Service plans  
 PPO: Preferred Provider Organization  
 QALY: Quality Adjusted Life Year  
 SGB V: Sozialgesetzbuch V, Book V of the German Social Security Code  
 SPD: Sozialdemokratische Partei Deutschlands, Social Democrats  
 WHO: World Health Organization

## **Preface and Acknowledgements**

The topic of personal responsibility for health raises a set of intricate issues in a range of different academic disciplines, including medicine, psychology, public health, health policy, law, political science, sociology, psychology and philosophy. As with most issues that demand an interdisciplinary approach, the analytic and explanatory powers of a single researcher are therefore limited. Here, the focus is on how to make progress with the oftentimes acrimonious debate around personal responsibility for health by examining what contribution philosophy and health policy analysis can make in clarifying the concept, and in assessing the appropriateness of specific rules or interventions. The philosophical approach is an applied one throughout, as I seek to set out a framework that specifies the areas in which justification is owed. This can help us decide when policies are acceptable and when it is reasonable to reject them, or to demand adjustments.

For better or worse, while I worked on this thesis I had plenty of opportunity to explore both theoretical and practical aspects of personal responsibility for health. Exercising in nearby Barnard Park in London before work I often saw a chubby boy walking to primary school. His breakfast, it seemed, consisted of a big bag of crisps that he dug into on his way through the park: we tended to look past each other casually, absorbed in our respective habits. Later, moving between the UK, the US and Germany, I had failed to keep up with dental check-ups, and narrowly avoided a painful root canal as a consequence. I initially missed the final appointment and was charged a substantial penalty when I came back, wondering, of course, whether this was reasonable or not. Chapter 4 was largely written with one hand, as I had broken the other when I fell off a branch doing pull-ups in New York's Central Park. As it happened, I was due to give a seminar on personal responsibility two days after the accident, and with my arm in a fresh plaster cast, at least we got to the topic instantly. The interviews with physicians were done after I had spent a year on a research fellowship in the US. Major health reform was passed then, accompanied by robust and, from a European perspective, oftentimes hard to comprehend, debate. I still remember walking inside a hospital in Berlin to meet with one of the oncologists for a research interview and being almost moved at the thought that no-one in Germany needed to fear personal bankruptcy due to medical bills. It did not

make me want to hug the patients smoking outside (some of them on intravenous drips), and of course, being German, the insight was not new to me. But in a relevant and new sense it made me appreciate how important universal and affordable access to healthcare is. Equally, it illustrated at a glance the point Thomas Scanlon made when he observed that in many situations where we are inclined to blame others, or wish to hold them accountable, our proper response ought instead to be: ‘there but for the grace of God go I’. Towards the final stage of the thesis, the irony was not lost on me that in one of the most busy periods I was not able to make as much time for exercise as I wished, nor enough time to cook or do other things that are more conducive to my overall wellbeing than generating p-values of survey findings, tracking down yet another insightful paper, or formatting tables. Among other things, what I take away from these experiences is that any framework seeking to address the question of personal responsibility needs to be compatible with the range of different ideas of a good life that people have; with the fact that opportunities of choice differ among people; and that talk about responsibility must be broader than just determining when we should blame and penalise people, as will hopefully become clear in the chapters that follow.

Much of the material included in this thesis has already been published, as is indicated in the first footnote of each chapter, which also lists conferences at which drafts of papers or chapters were presented, and individuals who assisted with technical questions that arose, with testing the instruments, or with discussions of issues that were insufficiently clear to me. Many friends and colleagues also provided most helpful guidance and advice in developing the initial approach of this thesis. Here I wish to thank again each of them: Johann Ach, Sandra Applebaum, Jorge Ayala, Tom Baldwin, Alan Balch, Ronal Barg, Gene Bishop, Jochen Breinlinger-O'Reilly, Berit Bringedal, James Brooke-Turner, Roger Brownsword, Johannes Bruns, Anna Bushan, Alena Buyx, Shawneequa Callier, Joseph Capella, Hanno Charisius, Zack Cooper, Tim Doran, Annette Dufner, Stefanie Ettelt, Nir Eyal, Ian Forde, Markus Grabka, Carmen Guerra, Rizwan Haq, Ingo Härtel, Peter Hasselblatt, Paul Henning, Ingo Höhr, Dirk Horenkamp-Sonntag, Karen Jochelson, Steven John, Martin Kaiser, Amit Kavnekar, Sabine Kies, Jeff Kullgren, Ian Forde, Tim Kennerly, Sabine Kies, Nicole Knaack, Jan Köser, Lore Korbei, Srinivas Kuruganti, Nilophar Lafrai, Karin Lange, Stefan Lenz, Kathy Liddell, Wolf-Dieter

Ludwig, Kristin Madison, Wendy Mansfield, Georg Marckmann, John McMillan, Claus-Dieter Middel, Stephanie Morain, Mathias Mrotzek, Axel Mühlbacher, Adam Oliver, Jens Niehoff, Sue Nelson, Charles Orellana, Engin Osmanoglou, Elfteria Panagiotou, Birte Pantenburg, Ursula Poch, Veronika Poch, Marianne Promberger, Lutz Rabe, Tim Rhode, Annette Rid, Ian Riner, Markus Rudolphi, Andrew Ryan, Bettina Schöne-Seiffert, Dorothea Schmidt, Lea Schmidt, Uwe Schmidt, Diana Seemann, Stephanie Stock, Julian Strauss, Stefan Subias, Humphrey Taylor, Sarah Thomson, Simon Tönsmeier, Andrea Troxel, Sridhar Venkatapuram, Andreas Vieth, Albert Weale, Dan Wang, Jürgen Wasem, Julia Weinmann, James Wilson, Jingsan Zhu, and the (anonymous) physicians who generously gave of their time, and frankly shared their views, in the interviews that formed part of this thesis.

Special thanks are due to David Ash, Norman Daniels, Scott Halpern, Julia Kreis, Dan Wikler, Kristin Voigt and Kevin Volpp, all of whom had substantial impact on the way I approached central issues in this thesis and beyond. I am deeply grateful to them. I must also thank the *Nuffield Foundation* for support in beginning this thesis, as well as the *Commonwealth Fund*, and the *Kolleg-Forschergruppe Normenbegründung in Medizinethik und Biopolitik* at the University of Münster, as their fellowships not only enabled me to dedicate more time to it, but also connected me with expert thinkers. The *Penn CMU Roybal P30 Center in Behavioral Economics and Health* enabled funding for parts of this research, which would otherwise have been impossible, through the *National Institute on Ageing of the US National Institutes of Health*. At the LSE I could not have been more lucky or privileged than to be guided by Julian Le Grand, Elias Mossialos, and Alex Voorhoeve, who provided just the right balance of hands-on and hands-off advice, support and expertise, invaluable, always. Finally, I am grateful, beyond measure, for her support, encouragement and steady companionship, to Joy Wang.

I dedicate this thesis to my family, and especially to my Grandfather, Herbert Holter. He read and wrote and played tennis and skied until his last year of life, dying at the age of 88. He would have been immensely pleased to know that this work is done, at least until there was a Mahler Symphony on the Radio, or Schnitzel mit Kartoffelsalat, or a good glass of Zweigelt.

New York, March 2012

## Chapter 1<sup>1</sup>

### Introduction

[T]he conditions in which people are born, live, and work are the single most important determinant of good health, or ill health; of a long and productive life, or a short and miserable one.

Margaret Chan (1947–), Director-General, World Health Organization

Medicine is a social science, and politics is nothing more than medicine on a grand scale.

Rudolf Virchow (1821–1902), physician, anthropologist and politician

I did it my way.

Frank Sinatra (1915–1998), singer

#### 1.1 Background

The debate around responsibility for health, like many long-standing controversies, has become somewhat polarised. At one end of the spectrum, commentators make what might be called a “get-real” argument (Burd 2009): they point to evidence that suggests, for example, that lifestyle choices about physical exercise and the consumption of food and alcohol clearly do affect health outcomes and account for up to 40% of the burden of morbidity and premature mortality (Schroeder 2007)<sup>2</sup>; that chronic diseases are on the rise (Scheller-Kreinsen, Blümel, and Busse 2009), that successful treatment depends considerably on patients’ compliance (Atreja, Bellam, and Levy 2005); or that peoples’ behaviour matters in relation to keeping appointments and other interactions with the health care system (Halpern, Bates, and Beales 2003). Since it can make sense to say that in all these areas people are responsible for their actions, proponents then often assert that people should also be held responsible. This may entail that they suffer a penalty or disadvantage in cases

---

<sup>1</sup> Text in this chapter draws on, and in part reproduces verbatim, material published in the following previous publications: Schmidt 2009d, 2011b.

<sup>2</sup> Drawing on prior work Steven Schroeder suggests for the US context that behavioural causes account for nearly 40% of all deaths, with the remaining causes genetics (30%), social circumstance (15%), inadequate health care (10%), and environmental factors (5%).



where they are judged to behave irresponsibly. Using such sanctions may also serve as an incentive for people to behave in the right way from the outset. In folk psychological terms such reasoning can fall on fertile ground. For the concept that one is responsible for the good and bad consequences of one's life, and that facing up to the choices one has made is a central part of what makes a life a good life, is deeply engrained in individual and autonomy centred societies (Jetten and Postmes 2006; McNamee and Miller 2009). Especially in the US, the notion of personal responsibility is integral to the concept of the American Dream, enshrined in the collective musical memory in Frank Sinatra's (and many of his epigones') rendition of *My Way*.

However, at the other end of the spectrum, commentators equally make a “get-real” argument when it comes to health responsibility. Here, it is stressed that the very concept of lifestyle choice can be cynical. The term seems to assume that it is equally easy for all to be healthy, with some simply choosing an unhealthy, as opposed to a healthy, lifestyle. But—leaving aside deeper philosophical debates about free will—there is much reason to doubt that this is the case. It is in this spirit that Margaret Chan cautioned that the environmental conditions in which people live are the single most important factor influencing health, when she presented the final report of the World Health Organization's (WHO) Commission on Social Determinants of Health (CSDH, Commission on Social Determinants of Health 2008).

The CSDH, as the WHO, has a global focus. Clearly the emphasis on the enabling conditions is of special relevance in this context. For there are stark contrasts in quality of life and life expectancy between developed and developing countries, and there is overwhelming historical evidence that traditional public health measures such as clean air and water, sanitation, and safe housing have a direct and highly significant impact on peoples' wellbeing. The German public health pioneer, physician, anthropologist and politician Rudolf Virchow was one of the first to emphasise this nexus, and work towards its implementation to improve health at the population level. In no small part, Virchow's work was prompted by his research regarding a typhus epidemic in a famine-ridden region of Prussia that he had been commissioned to carry out by the government. He was expected to make

recommendations for medical interventions and hygiene measures, but focused in his final report on the need for extensive social and political reform to improve health sustainably in the longer term.<sup>3</sup> For Virchow, the causes of the epidemic were as much due to poverty as to germs, and he argued that the most effective and sustainable way of responding to the outbreak was to improve social conditions.

Virchow's above-cited dictum that medicine cannot be reduced to a purely physical discipline, and that, conversely, political arrangements have very direct implications for people's health (Bauer A. 2005), is not only relevant for the living conditions in 19<sup>th</sup> century Prussia, and for developing countries lacking the ability to provide for peoples' basic needs. It is equally apt in relation to health inequalities in contemporary post-industrial developed countries, and the debate around personal responsibility for health.

Accordingly, researchers of social epidemiology and proponents of the social determinants of health view would argue that talking about a person's responsibility to maintain a healthy weight can be of relatively limited use in the case of a single unemployed teenage mother who grew up and lives in a deprived inner city borough with a high density of cheap fast food outlets, poorly maintained and unsafe parks, no affordable sports facilities, and so on. Holding her responsible for being overweight, by imposing some disadvantage or financial burden would be seen as a form of unacceptable victim-blaming. For not she, but the circumstances in which she lives, are seen as responsible for her poor weight. In the words of the CSDH report, and in line with Virchow's observations, instead of focusing on the causes of her poor health, more attention should be paid to the "causes of the causes" (Organization 2008).

Deciding on the appropriate balance of direct appeals to personal responsibility versus more indirect action at the level of the social determinants of health is of relevance to virtually any country with aspirations to operate effective and fair health and social policy frameworks.

---

<sup>3</sup> These included: full employment, higher wages, the establishment of agricultural co-operatives, universal education, and the disestablishment of the Catholic Church. Note that Virchow's travel to Silesia in February 1848 coincided with the emergence of revolutionary movements in France and Germany (Taylor and Rieger 1984).

## 1.2 Personal responsibility in the US and Germany

In this thesis, the empirical part focuses on policy in the US and Germany, which, for several reasons, has the potential to be particularly instructive. First, as will be described in more detail in the following chapter, the concept of personal responsibility features centrally in Germany's social health insurance system and is intricately linked to that of solidarity, versions of which also underlie many other European health care systems (Prainsack and Buyx 2011; Houtepen and ter Meulen 2000; Ashcroft, Campbell, and Jones 2000). By contrast, as will equally be explained in more detail, the US has a far more fragmented system. The majority of insurance coverage is provided in a market-based system by employers. As one consequence of the American credo in the individual pursuit of happiness, there is less of a societal expectation that citizens should consider the impact of their conduct on the wider community. Equally, there is less of an expectation that one is owed, or can rely on, assistance from the community. It can hence be instructive to analyse how two countries with very different background conditions address the question of incentivising individual behaviour in dealing with similar problems of rising levels of chronic conditions.

Second, while the nature of the problem posed by chronic conditions is indeed similar in both countries, it differs in magnitude. Levels of obesity in the US are approaching critical levels, with 34% of Americans being overweight in 2011, and 34% obese (Centers for Disease Control and Prevention 2011). In Germany an equal 34% are overweight, but only around half as many (13%) are obese (Statistisches Bundesamt 2011b). Given that the need to respond to the prevalence of overweight in the US is perhaps the most urgent of all post-industrial countries, and given that obesity is in the views of many a matter of personal responsibility, an analysis of the policies that are viewed as appropriate regarding personal behaviour is particularly timely.

Third, while the difference between the US and the German system is indeed in many ways stark, and the countries can be seen as positioned on opposite ends of a spectrum, in other ways there is also some noteworthy convergence. A major trend in the German (and broader European context) over the past decade or so is a stronger belief in using competition among both payers and providers of health care

to promote efficiency—therefore adopting a salient element of the American approach (Le Grand 1999, 2007; Wendt, Rothgang, and Helmert 2005). On the other hand, in a development that without exaggeration can be described as historical, recent US health reform has introduced the individual mandate. In effectively requiring each American to have health insurance, and putting in place structures that make coverage affordable through insurance exchanges and a number of other arrangements, the US have removed what health reform advocates regard as the US' “moral stain” (Emanuel 2012), and positioned the country closer to European states. While opponents are challenging the constitutionality of this measure before the Supreme Court at the time of writing, and while a political change in the 2012 Presidential elections would be likely to lead to severe and unpredictable disruptions, the signal emanating from this reform is a significant one and can indeed suggest some convergence between US and European health care systems (Callahan 2008). In this context it can then be instructive to review the ways in which the concept of personal responsibility is conceived of within them, given that it is itself beset by many of the controversies that are raised by the justifications of the different health care systems as a whole.

Lastly, in practical terms a comparison between the US and Germany is attractive because providers in both countries levy income-dependent individual insurance contributions. This method of financing enables financial incentives to have some leverage, as, among other forms, they can be provided as insurance discounts or surcharges, or differences in applicable co-payments. By contrast, a largely tax funded single payer system such as the UK's National Health Service generally has no such options, and co-payments and other cost-sharing play only a peripheral role. While one could compare, for example, cash incentives that are independent of insurance contributions, such a comparison would only represent one of several possible uses.

Practical, as well as political, cultural and normative grounds then make a comparison between the US and Germany particularly useful for a better understanding of the extent to which health care systems should adopt policies that incentivise personal responsibility and behaviour change at the level of the individual, in particular as a response to rising levels of chronic diseases. As will be

described in the outline of the thesis below, this comparison between the two countries will be made along four main axes: an analysis of the respective legal and policy framework (including formal reactions by key stakeholders); semi-structured interviews with physicians, and population level surveys; focusing on obesity and colon cancer prevention as conditions with considerable impact on health and health care budgets, yet distinct elements of individual control. The analysis will be guided by an analytic framework that builds on the philosophical literature on responsibility for health and a systematic review of the survey literature.

### **1.3 Research questions**

Overall, the central objective of this thesis is to explore in more detail the zone that lies between the two extreme ends of the continuum between social and individual responsibility for health. I contend that this is appropriate and necessary because the often encountered polarization is misguided and unhelpful for making progress in the debate around responsibility for health. I suggest that to some extent, the current situation is due to confusions about the way we typically use the concept of responsibility, and I propose a more nuanced approach that permits a debate about personal responsibility beyond the blame-game.

More specifically, I intend to address in a comparative manner the following research questions:

#### **Main research question:**

In promoting health and striving towards an efficient delivery of health care, to what extent should German and American health care systems adopt policies that incentivise personal responsibility and behaviour change at the level of the individual?

#### **Sub question 1:**

In what way do principal US and German health care policies seek to influence the behaviour of individuals regarding health maintenance and contribution to efficient services, and what are the underlying implicit and explicit values?

#### **Sub question 2:**

To what extent can the philosophical literature on personal responsibility and the justification of norms in pluralist societies help clarify the normative principles underlying the policies in the US and Germany, as identified in the analysis following Sub RQ 1, and offer a constructive way forward?

**Sub question 3:**

In the views of members of the public and physicians in the US and Germany: which aspects of cancer care and obesity policies focussing on individual behaviour and contribution to efficient services are reasonable, and which ones are not?

**1.4 The thesis structure**

Chapter 2 addresses the first sub-research question and describes the way in which principal US and German health care policies seek to influence the behaviour of individuals with regard to promoting health and contributing to efficient delivery and use of health care. It has been observed that two principal ways of conducting comparative health policy or health systems research would be (1) to analyse the extent to which two (or more) countries achieve specific stated goals (such as providing health care access to the entire population, or meeting certain thresholds of effectiveness in terms of producing health gains); or (2), to consider health systems as a whole, and compare the relative inherent structural and normative strengths and weaknesses (Busse 2002: 1). Throughout this thesis, the approach is a hybrid version of both.

While later chapters compare attitudes in both countries by reference to a common interview and survey instrument, Chapter 2 begins by providing background to the respective health systems and, in an inductive manner, seeks to identify the underlying implicit and explicit values supporting the respective appeals to personal responsibility. The chapter also provides an initial conceptual characterisation of the concept of ‘incentives’ and discusses in the abstract different ways in which incentives can be framed. The relevance of these framing options is then illustrated by a review of actual incentive policies in both countries. The chapter describes shared interests in using incentives as well as differences in approaches between the countries, and also contextualises evolving policy developments by summarising official position statements of professional medical associations.

Chapter 3 responds to the second sub-research question and explores to what extent the philosophical literature on personal responsibility and the justification of norms in pluralist societies can help clarify the normative principles underlying the

policies identified in Chapter 2, and resolves possible conflicts that may emerge between the rationales that have been identified there. The chapter begins by situating the use of incentives in the broader context of public health ethics and the scope and limitation of coercion in relation to health. In the following, I seek to show that the literature on the concept of personal responsibility for health is disjointed, as there is a wide range of different notions that feature in separate contributions. I suggest that none of these is sufficiently nuanced to be of practical use in planning, evaluating, and justifying health responsibility policies in contemporary pluralistic societies, and instead I synthesise the different strands into a workable concept of *health responsibility as co-responsibility*. I attempt to show how, drawing on the procedural justice framework of accountability for reasonableness, developed by Norman Daniels and James Sabin, (Daniels 1999) and Thomas Scanlon's contractualism (Scanlon 1998), the concept can be operationalised to provide a framework that specifies seven areas in which people who are interested in identifying mutually acceptable policies regarding personal responsibility ought to provide justification.

The first part of Chapter 4 provides a systematic review of the survey literature regarding personal responsibility for health, and focuses centrally on attitudes of members of the public and physicians. Building on these data, and on themes emerging from Chapters 2 and 3, an initial instrument as a basis for subsequent qualitative and quantitative research is developed, to explore in more detail the range of relevant reasons that need to be considered in relation to two specific areas: weight control and colon cancer prevention. These areas have been chosen for three main reasons. First, both overweight and colon cancer result in significant burdens of disease and there are widely held views that prevention can significantly reduce morbidity and associated cost. Second, while overweight is often regarded as being largely under the control of individuals, colon cancer is viewed more as a condition that strikes like bad luck, and in this sense those suffering from it are not seen as responsible in the same way that the overweight or obese are. Third, as Chapter 2 will show, Germany instituted a particularly controversial colon cancer prevention policy, giving rise to a range of complex issues, which have so far not been explored specifically in the survey literature.

Chapter 5 addresses the third sub-research question and describes the conduct of 20 semi-structured half-hour interviews with primary care physicians and oncologists in Berlin, Germany, and Philadelphia, USA. A refined version of the initial instrument developed in Chapter 4, the 13-item questionnaire concerned, among other things: the relationship of personal behaviour to other factors that can determine health; the adequacy of a prominent analogy made by proponents of higher health care premiums for overweight people, suggesting that they should be seen like risky drivers, who equally face higher car insurance premiums; the question of how high incentives for different formats should be; whether physicians would lie if a patient requests this in order to receive an incentive in relation to weight-control and common cancer prevention; and the acceptability of a new and unique obligation to think about one's health, that the current German colon cancer policy has established. The Chapter focuses on describing and discussing the reasons physicians gave in support and against these policies, and closes by contrasting similarities and differences between the American and German respondents, as well as between primary care physicians and oncologists.

Chapter 6 equally responds to the third sub-research question and concerns population-level surveys in Germany and the US, using an adapted version of the instrument that was used for the interviews in Chapter 5. For methodological and cost reasons, one survey each was fielded to non-probability samples of  $n=1,000$  in the US and Germany, and in a separate survey, a subset of five questions was fielded to a probability-based sample of  $n=1,115$  in the US. The findings are grouped under five main questions:

- To what extent do respondents use incentive programmes?
- To what extent do respondents agree with rewards and penalties?
- To what extent do respondents value opportunity of choice, motivation and effort?
- What levels of incentives are acceptable for particular incentive programmes?
- Do incentive programmes interfere with the doctor patient relationship?

The discussion focuses on key similarities and differences between the American and German samples; similarities and differences between views of survey



respondents and physician views as per Chapter 5; and winners and losers of incentive programmes.

Chapter 7 draws together the discussion from Chapter 2 (policy), 3 (ethics), 5 (physician interviews) and 6 (surveys) by applying the approach set out in Chapter 2 to the specific case of colon cancer prevention, illustrating what work the framework can, and cannot do. Finally, Chapter 8 provides a summary of the findings and new insights that have emerged in response to each of the research questions. The chapter argues that the framework presented in this thesis provides a novel and useful way of structuring the controversial debate around personal responsibility health, enabling more progress in policy and practice than has so far been the case. It then comments on policy implications that are of relevance to both the US and Germany, and emphasises the need to bring behavioural economics more effectively to bear on the Realpolitik of health policy; to carry out effective public and consumer engagement; and to conduct robust and suitably comprehensive evaluations of ongoing incentive programmes. Specific separate issues arising in the US and German policy context are also addressed, and the chapter, and the thesis, closes with limitations and next steps.

## Chapter 2<sup>1</sup>

### Responsibility, ‘carrots’ and ‘sticks’ in health care policy in Germany and the USA

#### 2.1 Introduction

The epidemiological and medical evidence regarding the burden of disease that can be attributed to personal behaviour clearly suggest that people’s individual actions play a significant role for health outcomes at the individual and population level, as noted in Chapter 1. Among other things, this raises the question of how to respond in law and policy. Not mincing his words, John H. Knowles, an outspoken critic of the American health care system, and viewed by some as a moderate Republican, left no doubt about his take on the matter when he wrote in an article in *Daedalus* in 1977 that:

[t]he cost of sloth, gluttony, alcoholic intemperance, reckless driving, sexual frenzy, and smoking is now a national, and not an individual, responsibility. This is justified as individual freedom—but one man’s freedom is another man’s shackle in taxes and insurance premiums. I believe the idea of a ‘right’ to health should be replaced by the idea of an individual moral obligation to preserve one’s own health—a public duty if you will. The individual then has the ‘right’ to expect help with information, accessible services of good quality, and minimal financial barriers.

Knowles does not really fit the standard political left-right pattern, in which the left usually abhors the concept and is in favour of universal health care, and the right favours personal responsibility, but is typically against health care for all: Knowles, by contrast, is interested both in greater emphasis on personal responsibility for health, and in universal health care for all Americans. His view also differs from traditional libertarian positions that would hold that one has the right to conduct one’s life however one wants, but is not owed assistance from others in cases where the resulting harm is the consequence of one’s choices. Knowles’ proposal did not translate directly into US policy. But the question of how to conceptualise the relationship of personal responsibility and risk-pooling—that may cushion some, or

---

<sup>1</sup> Text in this chapter draws on, and in part reproduces verbatim, material published in the following previous publications: Schmidt 2007a, 2007b, 2008a, 2008b, 2010; Schmidt, Stock, Gerber 2009; Schmidt, Stock, Doran, 2012. I am grateful for valuable discussion with Karen Jochelson and Joy Wang who reviewed drafts of the above papers and helped me better understand specific aspects discussed in them.

all of the negative effects of requiring treatment for conditions that may be said to be attributable to personal behaviour—is one any health care system has to answer, whether explicitly or implicitly.

In this chapter I seek to outline the ways in which two health care systems address the issue of personal responsibility for one's health: the German statutory health insurance system, and the employer-based insurance system in the USA. For both countries I first describe the overall framework that has been established in law and policy regarding the question of personal responsibility. In discussing the perspectives of important stakeholders, I then focus on two specific policies: incentives for cancer checkups in Germany, and wellness incentives for meeting Body Mass Index targets in the USA. This serves to identify the principal rationale behind these schemes and the concerns by those involved in their administration and use. The chapter also serves to prepare the ground for the discussion of ethical issues in the following Chapter 3, and the empirical work in chapters 4–6, which present a more in-depth analysis of attitudes of members of the public and physicians on particularly central aspects raised by personal responsibility policies. Before doing so, I begin with a brief conceptual clarification regarding the framing of health responsibility policies. This clarification is of relevance for the description as well as the evaluation of the policies that will be described later. For framing cannot be assumed to be simply a neutral background condition, but in many ways raises fairness issues by itself.

## **2.2 On 'carrots' and 'sticks': when is an incentive an incentive?**

There are a number of different ways in which personal responsibility can be promoted, as will be outlined in a more structured way in Chapter 3 (Section 3.2). On the one end of the spectrum, one could simply provide information about, for example, healthy eating. On the other, one might require people to exercise in a quasi-mandatory way, as in the case of the Chinese calisthenics programme, re-introduced in August 2010 in view of rising levels of obesity (Branigan 2010). Incentives, often called 'carrots', and disincentives or 'sticks', that form the centre of the discussion here, are two further means.

*Prima facie*, the distinction between ‘carrots’ and ‘sticks’ seems straightforward: one is a reward for behaviour that is judged to be positive, the other a penalty for behaviour that is judged to be negative. The use of incentives to influence individual behaviour has been implemented widely outside of health policy, especially for marketing purposes. For example, many airlines and supermarkets provide discounts for frequent flyers and shoppers in the hope that this will encourage customers to become more loyal. Levels of incentives are generally calculated in such a way that their cost is offset by customers’ more frequent purchases of goods and services. Programme enrolment is usually equally easy for all customers, and participation is often agreed at the point of purchase, with those not wishing to make use of the offer generally losing out on some benefits, such as reduced prices. On the whole, equal access, and the voluntary nature of the agreement means that no significant fairness issues are raised. However, problems can arise when this model is applied to the health care context. One of the main reasons is that it is not necessarily equally easy for all groups who are offered incentives to avail themselves of the opportunities. Depending on the size of the benefit, the question can arise whether losing out on a benefit can turn a ‘carrot’ approach into one that begins to have ‘stick’ character.

As the following discussion of real-world policy examples will show in more detail, the distinction between ‘carrots’ and ‘sticks’ is often far less clear than it seems. Much also depends on the way an incentive policy is *framed*. Conceptually, Dan Wikler set out one of the fundamental mechanism that is at work in implementing incentives as follows:

Suppose, for example, that the Government wants to induce the obese to lose weight, and that a mandatory national insurance plan is about to go into effect. The Government threatens the obese with higher premiums unless they lose their excess weight. Before the plan is instituted, however, someone objects that the extra charges planned for eager eaters make the plan coercive. No adequate justification is found. Instead of calling off the program, however, some subtle changes are made. The insurance scheme is announced with higher premiums than had been originally planned. No extra charges are imposed on anyone; instead, discounts are offered to all those who avoid overweight. Instead of coercion, the plan now uses positive incentives; and this does not require the kind of justification needed for the former plan (Wikler 1978: 330).

Let's assume that all those whose weight was in the range regarded as normal sign up for the discount rate, and all, or most of those who are overweight do not. Arguably, all plan members had an incentive to lose weight, and the switch in framing has transformed the policy from a 'stick' approach to a 'carrot' one—or has it?

I will not resolve this question here, but merely note that an answer is far from trivial. In the following, I will regard as the paradigm case of an incentive a situation where a population that has been offered certain services or goods at a specified price sees no changes in this baseline cost (leaving aside inflation-related adjustments and the like), but is offered a reduction in price, or another financial or in-kind benefit. I will call such cases 'genuine' incentives. In using this term, nothing is said about the ethical or other acceptability of the policy, it is intended merely to denote the mechanism that is at work. Because of the potentially misleading connotations, I will also not use the terms 'positive' or 'negative' incentives, which are commonly used in the literature (see, for example, Jochelson 2007). While, to some extent, these concepts also need to be seen as technical terms that seek to describe a mechanical aspect,<sup>2</sup> the adjectives 'positive' and 'negative' also imply a value judgment and may implicitly bias towards viewing an incentive scheme that works by offering a benefit as a good thing (or at least a better thing than a negative incentive programme). However, for much of the following analysis it is important to be open to the possibility that what might be described as a positive incentive may be bad (for at least some people) in some relevant way and equally, that a policy is not necessarily bad, just because it overtly uses penalties. Unless stated otherwise, I will therefore use a broad understanding of an incentive which encompasses both what is traditionally known as positive and negative incentives. On this view, knowing that one lives in a place that imposes the death penalty for murder can be as much an incentive not to kill someone as taxes can be incentives not to smoke, and a lump-sum of money may be to exercise regularly. It is now useful to review the characterisation of incentives to promote health and personal responsibility in law and policy.

---

<sup>2</sup> Namely, whether the incentive consists in the option of securing a benefit – in the case of a positive incentive – or in avoiding a penalty – in the case of a negative incentive.

### 2.3 Health responsibility in the German statutory health insurance

Germany's social health insurance system is governed by Book V of the German Social Security Code (SGB V—*Sozialgesetzbuch*) of 1988, revised most recently in 2007 under the *Gesetz zur Stärkung des Wettbewerbs in der gesetzlichen Krankenversicherung* (GKV-WSG—“Law to strengthen competition among providers of statutory health-insurance scheme”) and in 2010 under the *Gesetz zur nachhaltigen und sozial ausgewogenen Finanzierung des Gesundheitswesens* (GKV-FinG—“Law for the sustainable and equitable financing of health care”). SGB V regulates the provision of statutory health care through the *Gesetzliche Krankenkassen*, or sickness funds. The concept of solidarity centrally underpins the statutory insurance scheme in three different ways. First, by pooling contributions and expenditures without individual or gender-specific risk assessments, those who are able to work, especially healthy and young persons, support the less healthy and older ones. Second, due to income-tested contributions, the better off support the less well off. Third, those without families support those with dependents, as family members may be insured through the main insurance contributor. The norms specified in SGB V are binding for some 200 sickness funds that provide care for approximately 90% of the German population (the remainder being covered by private insurance). People have free choice of fund, and those with a gross annual income of less than €47,700 are obliged to be insured by one of the sickness funds, while higher-income earners may opt out and elect to be insured privately (Busse 2005; Weide 2005).

In terms of financing, until recently, employers and employees split insurance contributions evenly, each paying around 6–8% of the employee's gross salary. However, reforms enacted in 2010 led to a change in this arrangement. Employers had complained for some time that the high health insurance contributions had a stifling effect on their competitiveness in the European and world market, and the then recently elected industry-friendly coalition of Christian Democrats (CDU) and the Free Democrats (FDP) implemented reforms that capped employers' contributions at 7.3%, and increased contributions for employees to 8.2% of their gross income, with all future increases being shouldered by the employees through a newly introduced capitated fee and future increases in insurance contributions. Sickness funds often differ in their benefit packages, although a mandatory set of

core services that all funds need to provide is set out in SGB V. Services are not entirely free for all, and cost sharing is required for, among other things, prescription medicines, doctor visits and hospitalisation (Riesberg and Busse 2003; Gericke, Wismar, and Busse 2004).

### 2.3.1 *Legal framework—health responsibility*

Article 1 of SGB V has overarching function and is entitled “Solidarity and personal responsibility”. A characterization of both concepts is provided in the wording of the article itself which reads as follows:

In the spirit of a mutually supportive community [*Solidargemeinschaft*] the task of the statutory health insurance is to maintain, restore or improve health of the insured. The insured have co-responsibility for their health; through a health-conscious way of living, taking part in appropriately timed preventative measures [and] playing an active role in treatment and rehabilitation, they should contribute to avoiding illness and disability, and to overcoming the respective consequences. The statutory sickness funds are to assist the insured persons through the provision of information, advice and services, and should encourage a health-conscious way of living [my translation, HS<sup>3</sup>].

The principal characterization of solidarity and personal responsibility is that the community as a collective, and people individually, are “co-producers” of health (Forde and Raine 2008: 1695). The notion of co-responsibility has two important facets in this respect. First, it states that the “mutually supportive community” has a certain degree of responsibility for the health of each individual. In this sense, individuals are entitled to claims against the community for assistance (Dixon and Hermesse 2004: 171; Ashcroft, Campbell, and Jones 2000: 381; Houtepen and ter Meulen 2000: 330). Second, it also implies that the community has certain claims against individuals. Leaving prudential benefits aside, the appeal to staying healthy has the aim of containing overall expenditure and opportunity costs. For all care needs to be financed by the solidaristic community, and the underlying assumption is that cost can be reduced or at least contained if demands on the health care system are limited. Using services unnecessarily may also deprive another person in need of resources or medical attention, exacerbating resource allocation dilemmas. Article 2 SGB V on “necessity, cost-effectiveness, and personal responsibility” is unequivocal in stressing people’s obligations in this respect:

---

<sup>3</sup> Note that there is no official translation of SGB V. All translations therefore mine, HS.

Services ... are to be provided by the sickness funds with due respect to cost effectiveness [Wirtschaftlichkeitsgebot] ... and *insofar as the need for services is not attributable to the personal responsibility of the insured person*. [...] Sickness funds, service providers and insured persons must seek to ensure the clinical and cost-effectiveness of services, *which are only to be used insofar as necessary* [emphasis added, my translation, HS].

Article 2 SGB V raises the question of what exactly, in practice, the respective scope and limitation of solidaristic and personal responsibility should be. What forms will personal responsibility take: merely the demand for (co-)payments where the individual responsible for his own fate can pay? Or will there be a refusal of treatment when the person cannot pay for it by him or herself?

Article 52 SGB V sets out conditions under which statutory sickness funds may limit funding for services, and the health care reforms of 2007 included a noteworthy specification in this respect. In its pre-2007 version, Article 52 SGB V stated that insurers may demand a reasonable contribution to the costs of treatment if a person's need for health care is the result of engaging in a criminal activity. Solidarity and personal responsibility are interpreted in such a way that whoever harms the solidaristic community has lost the claim to having health care needs met at the community's exclusive expense.<sup>4</sup> In such cases, financing health care can become a matter of personal responsibility. However, this responsibility extends to the costs only: for solidarity still clearly requires the criminal person to be treated by the health care services. The 2007 reforms extended this principle to state that insurers may ask for equivalent contributions where people request treatment for complications arising from "cosmetic surgery, tattoos ... piercings," or another "non-medically indicated" measure. In its written justification of the bill, the Government argued that "since insured persons who decided to have [tattoos, cosmetic surgery, piercings] voluntarily exposed themselves to health risks, it is not appropriate to cover costs... through the collectively contributed funds" (Bundesregierung 2006: 300-301 [my translation, HS]). In addition to these provisions, there are a number of initiatives that are presented more explicitly as incentives for people to behave responsibly.

---

<sup>4</sup> Additionally, since sickness funds also administer sick-pay, the policy also states that otherwise applicable sickpay may be denied in part or entirely, or requested to be paid back, where it has already been paid.



The oldest programme is set out in Articles 55 and 56 SGB V and relates to dental check-ups. According to these articles, sickness funds must cover 50% of the costs for required dental replacements, with the other half falling to the insured person. Insurance providers must up their contribution by 20% of their initial payment, if, over a period of 5 years, adults have taken part in annual check-up programmes (and under 18-year-olds in biannual ones). If there are no gaps over the past 10 years this is increased to 30%. However, if an individual has not taken part in such check-ups, he or she has to pay his or her full 50%.

Since 2004, Article 65a SGB V, entitled “Bonus for health-conscious behaviour” provides that sickness funds may offer incentives (known as ‘bonuses’) to insured persons who participate in quality assured prevention, health-promotion, screening, and check-up programmes. Some of these programmes are specified elsewhere in SGB V, but in practice sickness funds incentivise a wide range of activities. An overview of activities with bonus awards offered by members of the *Betriebskrankenkassen* (BKK), which insure employees of large employers such as *Daimler*, *Siemens* or the postal service, is provided at Appendix 2A. Article 65a SGB V also includes a noteworthy provision on financial aspects of bonuses. While no maximum limit is specified for incentives for particular activities, paragraph 3 states that:

in the medium term, resources for bonus awards must result from cost-savings and efficiencies that result from participation in the bonus programmes. Sickness funds are required to report in regular intervals, at least every three years, to the relevant authorities about any savings that are made. Where no savings are made, bonuses may not be awarded for the respective activities.  
[my translation, HS]

The reporting requirement extends to cost only, and there is no requirement to provide information on health improvement, which can, at best, be inferred from data on health care usage.

Insurers differ in the ways in which they provide bonuses. One common type consists in giving reward points of a nominal value for taking part in certain measures. Usually points can be redeemed for items thought to be conducive to a healthy lifestyle, such as sports equipment, or health books. More recently, some insurers such as the *Techniker Krankenkasse*, one of the four largest insurers with 4m

contributing members (6m including dependents), have included iPods and vouchers for music downloads. Many providers also offer the option of redeeming the collected points for cash. In case of the *Techniker*, incentivised activities have a value of between €5 and €30 (with a mean of €10.48).<sup>5</sup> A four person family may redeem between 100 and 200€ per year through such schemes, and programmes of other insurers offer bonuses of a similar level (Bödeker, Friedel, and Friedrichs 2008; Stock, Schmidt, Buscher et al. 2010; Friedrichs 2009; Dreier and Linger 2006).

A more recent addition introduced under the 2007 reforms takes a slightly different form and provides incentives for limiting ones use of health care services. Article 53 SGB V on personalised health care plans (*Wahltarife*) provides that sickness funds may offer significantly reduced contributions (or lower co-payments, where required) to those agreeing to take part in schemes thought to reduce the burden of morbidity and costs for the sickness funds, such as managed care programmes. Sickness funds may also offer what amounts to ‘no-claim bonuses’: If the insured requires no primary care consultation leading to a prescription, or hospitalization over a year, reductions of up to 20% of the annual contributions may be granted, capped however at €600 maximum (or €900, where several different bonus plans are combined). Subsequent policy issued in 2011 stipulated that sickness funds intending to offer these programmes must demonstrate in advance through actuarial modelling that bonuses can be financed through efficiencies, in parallel to the above-cited cost-saving requirement pursuant Article 65a SGB V (Meißner 2011).

A further new policy introduced in Article 62 SGB V under the 2007 reforms concerns participation in cancer check-ups and compliance in treatment. This Article and subsequent guidance by the *Gemeinsame Bundesausschuss* (G-BA, Federal Joint Commission<sup>6</sup>) states that patients who suffer breast, colon, or cervical cancer will

---

<sup>5</sup> There are 31 measures, see: Techniker Krankenkasse [homepage on the internet]. “TK Bonusprogramme,” available in English from: [http://www.tk-online.de/centaurus/generator/tk-online.de/s07\\_english/dossiers/01\\_lexicon/tk\\_bonus\\_programme/tk\\_bonus\\_programme.html](http://www.tk-online.de/centaurus/generator/tk-online.de/s07_english/dossiers/01_lexicon/tk_bonus_programme/tk_bonus_programme.html) (accessed 20 August 2007)

<sup>6</sup> The G-BA is a self-governing body bringing together the payer community (the sickness funds), and health care providers (physicians, dentists, psychotherapists and hospitals) The body is tasked with determining the provision and reimbursement of pharmaceuticals, diagnostic and therapeutic procedures, medical devices and non-medical treatment for the statutory health insurance.

have to pay a maximum of 1% of their gross annual income as co-payments for treatments and medicines, provided they meet two conditions. First, they need to have attended counselling sessions on the advantages and disadvantages of the respective screens at the appropriate age. Second, they must not refuse treatment, should they require it. In case of non-compliance with either provision, the co-payment cap is 2%. Note, however, that both before and after introduction of this policy, the threshold for all other chronically ill was 1%.

In practical terms, all those insured<sup>7</sup> through the statutory sickness funds will be issued a ‘prevention-passport’ that documents their participation in counselling sessions. While people with severe mental diseases are exempt from the regulation, all others must attend counselling no later than two years after the recommended age for the respective disease to be eligible for the lower co-payments. Counselling is to be offered by medical professionals who are also qualified to carry out the respective screens. It is to be based on information brochures already available via the G-BA. Since this policy raises a number of interesting issues for the questions around the reasonable scope and limitations of personal responsibility that are of interest here, it will be discussed in somewhat more detail below in the consideration of the perspective of key stakeholders, as well as in the subsequent empirical chapters.

### 2.3.2 *Implicit and explicit rationales*

The different schemes outlined here have different rationales that may overlap:

- Improving health. In the most benevolent interpretation, programmes are simply intended to improve people’s health, and this notion is certainly emphasised in the written justification provided by the government in introducing the 2007 reforms, alongside the stress on individual’s personal responsibility to contribute to efficient services, as will be discussed in some more detail below (Section 2.3.3).
- Lower expenditure. There is the assumption that healthier people and those who limit their demand on the health care system over certain periods, will require less health care expenditure overall.

---

<sup>7</sup> Inclusion criteria: women born after 1 April 1987, and men born after 1 April 1962 – the point being that the policy should only apply to people who had the opportunity to chose to comply with it (or not) at the age at which screening is recommended.

- Competing for healthier customers. In a somewhat more indirect way, incentive programmes function similarly to schemes such as air miles or store loyalty cards (Müller 2003; Knaack 2008). If programmes are designed in such a way that they appeal in particular to the better off and healthy, incentives can help sickness funds attract and retain ‘good risk’ customers, who are likely to require less care, and contribute disproportionately more, as insurance contributions are income-dependent. In this way, incentive programmes may help insurers secure a competitive advantage. It is noteworthy that schemes such as the no-claims-bonuses (Article 53 SGB V) were introduced under the 2007 Reforms that signalled the aim of improving competition among sickness funds explicitly in the title.

Ideally, all three goals can be achieved, and the result of participation in bonus programmes is better health, lower health care cost, and improved competition between providers. However, tensions can clearly arise if sickness funds are guided more by one line of thought than by another. For example, Wolfgang Bödeker and colleagues noted that evidence on the initial evaluation of bonus programmes for the *BKKs* suggested that cash incentives are more effective than in-kind ones in motivating people to take up incentive schemes, although in-kind schemes provide the highest financial returns (Bödeker, Friedel, and Friedrichs 2008: 218). In principle, this may tempt insurers to focus on in-kind incentives that are less likely to promote behaviour change, but have more potential to reduce cost.

In a similar vein, the general emphasis on cost savings might turn out to be problematic. As noted, Article 65a SGB V assumes that healthier people are less costly to the health care system than sick ones, and specifies that bonuses may be paid only when savings are achieved.<sup>8</sup> Evidence from different evaluations indicated a return-on investment ratio of roughly 3:1 (Bödeker, Friedel, and Friedrichs 2008: 216; see also Baicker, Cutler, and Song 2010: 305; Stock, Schmidt, Büscher et al. 2010: 52). The normative implications of the cost-saving rationale will be taken up again in Chapter 3, Section 3.5.1. For now it can be observed that the assumption that bonus programmes will always lead to cost-savings can be problematic. Table

---

<sup>8</sup> It is interesting to note that the initial response from the sickness fund community to the requirement to demonstrate savings in three year intervals was more than skeptical, as it was feared that no such savings could be shown over such a short time frame (Bödeker, Friedel, and Friedrichs 2008).

2.2 shows the findings from an evaluation pursuant Article 65a SGB V of the *Techniker Krankenkasse*, a large sickness fund insuring approximately 9% of the German population. Overall, the mean difference for the sum of three cost categories (hospital spending, medication spending, and additional benefits) between matched pairs of bonus programme users and non-users over the three-year evaluation period amounted to €177 (£147) savings per insured enrolled in the incentive programme per year. If programme costs were taken into account, overall savings were reduced to €101 (£84). However, a subgroup analysis comparing the costs of participants and non-participants with no health care utilization in the year prior to introduction of the programme found differences in all three areas of spending, with those participating in the incentive programme having significantly higher costs overall (if lower, in both cases, than the mean cost in the complete intervention and control group). One implication of this finding is that it would seem odd to wind up programmes that have been shown to achieve behaviour change, simply because they may not have reduced costs: an unbalanced cost-saving requirement can therefore be problematic.

**Table 2.2: Mean health care costs among incentive program participants and their matched controls, 2006 (Evaluation of Bonus programmes of the *Techniker Krankenkasse*)**

	Total study population		Subgroup analysis: Participants without any health care costs in 2003	
	Incentive Programme Participants (Intervention group) (n=70,429)	Incentive Programme Non-Participants (Matched control group) (n=70,429)	Incentive Programme Participants (n=4,822)	Incentive Programme Non-Participants (n=4,822)
Hospital spending (2006 mean)	€469	€613	€576	€226
Medication spending (2006 mean)	€365	€413	€34	€26
Other medical benefits*(2006 mean)	€88	€81	€155	€108
Total**	€922	€1,107	€765	€360
*Other medical benefits include massage, physiotherapy, walking aides, etc. **Note: the amount of €177 (£147) mentioned in the text reflects the difference in the mean of the individual increase in cost in each cohort between baseline and the end of the study. This sum therefore differs from the aggregate of the three cost categories summarised here. Source: (Stock, Schmidt, Buscher et al. 2010; Schmidt, Stock, and Doran 2012)				

### 2.3.3 Stakeholder views

When Parliament passed the 2007 reforms, a heated debate raged over the issue of personal responsibility, and in particular over Articles 52 and 62 SGB V, concerning the chronically ill, and the question of whether the revised law might unduly burden

the chronically ill and force them to undergo screening. In introducing the bill, the government argued in its written justification of the provision, that those who benefit from expensive medical care had “special personal responsibilit[ies]” (Bundesregierung 2006: 305). Accordingly, the initial proposal envisaged that the lower co-payment threshold of 1% (which, as noted above, is the default threshold for all other chronically ill people) should be available only if women and men underwent the screens for colon, cervical and breast cancer at the recommended age (and subsequently in the appropriate intervals), and took an active part in complying with schemes such as disease management programmes. Otherwise they should face the 2% threshold (i.e.: twice the amount that all other chronically ill faced).

In the final debate on 2 February 2007, Brigitte Bender of the opposition Green party (Die Grünen) argued that this constituted a “a punishment... of severely ill people... in the form of increase in contributions... that is irrational.” (Deutscher Bundestag 2007: 8032). Gregor Gysi of the left (Die Linke) equally described the policy as a “penal fine” that he perceived as undermining the principle of solidarity, and as introducing a questionable principle of responsibility that could also be found in the new policy of requesting contributions to the treatment of people suffering from complications arising from piercings (as described above, regarding Article 52 SGB V, *ibid.*: 8015). By contrast, Elke Ferner, a member of the then governing Social Democrats (SPD), defended the policy, arguing “I do think that, if chronically ill are not willing to play an active part in treatment, then we do have to set appropriate incentives” (*ibid.* 8019), and her colleague Jens Spahn of the CDU, who formed a coalition government with the SPD, equally argued: “the bill says that if people do not come forward to the recommended checkups... they simply don’t get the reduced copay, and stay with the standard. So, it is not true that there is a ‘penal fine’, as you [the opposition] have been calling this provision over the recent weeks and months. It’s only that you don’t get a reduction. [...] Not getting a reduction is not the same as being made worse off [by this policy]” (*ibid.*: 8036, all translations mine, HS).

The provisions of Article 52 and 62 SGB V came to be nick-named “snitch-Articles” (Petzparagrafen), in some discussion fora of health professionals as they required physicians to inform sickness funds about people requiring treatment for

tattoo complications etc, and about non-compliant cancer-prevention behaviour. The provisions also received strong opposition from the medical profession. At the 2007 General Assembly of the *Bundesärztekammer* (BÄK, German Medical Association), delegates noted significant concerns about the concept of responsibility implied in Article 52 SGB V, and called for a critical debate about the role of physicians in relation to Article 62 SGB V, which, they feared, sought to assign them the role of “therapy-police” (“Therapiepolizei”) of the sickness funds (Bundesärztekammer 2007: 3).

The controversy could not be resolved in passing the law, and it was agreed that the precise content of the new requirement would be specified at a later stage by the G-BA. In its final report, the G-BA cautioned strongly against making co-payments conditional on actual participation in the respective screens and noted several areas of concern that spoke against more directive screening approaches. Further to the reservations expressed by the BÄK, these included an overall high ratio of numbers-needed-to-screen for each of the screens; the absence of evidence that screening programmes will lead to net savings; the importance of shared decision making and informed consent in view of the risk-benefit balance of the screens; and the “right not to know,” which was seen as widely accepted by ethicists, health professionals, and lawyers and backed up by the German Constitution (Gemeinsamer Bundesausschuss 2007), and perceived to be infringed if insured persons underwent screening simply to evade the higher co-payment.

Instead, the G-BA proposed the abovementioned policy, which leaves intact the fundamental architecture of establishing higher co-payment thresholds for the three types of cancer patients, but changed the requirement of undergoing screening to taking part in counselling sessions on the advantages and disadvantages of screening, and also relaxed the compliance requirement, as non-compliance would only be attested in cases where patients refused treatment—the latter shift was largely in response to opposition from physicians who viewed the possible policing function they were given under the initial law as incompatible with the doctor-patient relationship.

The then Minister for Health, Ulla Schmidt, welcomed the new regulation in her formal announcement, stating:

Early diagnosis enables better treatment. This is why the statutory sickness funds offer a range of screening and disease-prevention measures. I would like more people to make use of them, and the most recent reforms have created new incentives. It is important that all take more responsibility: for their own health, and also towards the community of people who jointly fund health care and ultimately pay for the treatment of chronic diseases. (Bundesministerium für Gesundheit 2007) [my translation, HS]

Arguably, as discussed initially in the review of structural issues around the concept of incentives, the policy still represents an incentive for people to consider the advantages and disadvantages of screens, and of refusing treatment should they require it. However, it is also clear that the approach is far from a genuine incentive: for, in principle, it could equally have been implemented by, for example, holding the 1% threshold constant for all chronically ill, and offering a 0.5% co-payment for those complying with the requirements.

In summary, then, salient features of the German situation are, first, the explicit reference to personal responsibility and its connection to solidarity. The characterization of these values in the provisions following Article 1 SGB V is that solidarity requires assistance for those with health care needs, but that such assistance is not unconditional. Several policies seek to prevent burdening the solidaristic community with costs that may be attributable to individuals' personal responsibility, even though it is unclear whether such a burden exists. A wide range of incentive schemes has been put in place, with the distinct, but often also overlapping aims of (1) promoting responsibility toward keeping oneself healthy, and regaining health in the case of sickness; (2) respecting the health of others (for example, in the case of bonuses for prenatal or childcare checkups, or vaccinations, which have a population-level benefit); and (3) contributing to cost-reduction and an efficient operation of the health care system (without direct cost-shifting among insurers). From the perspective of health care insurance providers, incentive programmes are also attractive as they may confer a competitive advantage in their effect on the composition of the risk pool. Some of the schemes have an overtly 'stick' character, while others are presented in the form of 'carrots'. However, the example of the provisions pursuant Article 62 SGB V regarding cancer prevention also illustrated that incentives can be framed in very different ways, and that 'carrots'



can also have a 'stick' character. In the responses of stakeholders to the 2007 reforms that introduced many of the more controversial new policies, the reaction of the medical profession is particularly noteworthy, both in view of the fact that they rejected aspects of the policy because of lack of insufficient evidence (as in the case of the Government's proposal to increase the number of people undergoing cancer screens under Article 62 SGB V) and their opposition to be given policing functions that they viewed as incompatible with their professional roles.

#### **2.4 Health responsibility in the US employer-based insurance system**

In contrast to Germany, the provision of health care in the USA is organised in a far less centralised fashion, and there is, as yet, no universal coverage, even though an individual mandate has been introduced under the 2010 health reforms. Overall, there is a mix of private and public provision of health care. Large employers often offer their own health insurance and use insurance companies mainly for administrative purposes, while designing their own benefit packages (Institute of Medicine 2012). Federal programmes such as *Medicaid* provide services for the least well off. People older than 65, and some who meet special criteria, are eligible for *Medicare* coverage. The *Veterans Health Administration* provides services for former military personnel. There are also significant differences in the way in which the different states provide services under Medicaid or Medicare, and while some view this diversity as an excellent opportunity to learn from different approaches, many view it as inequitable and problematic. The vast majority of Americans access health insurance through their employers, who may offer a range of different plans. A somewhat more detailed description of the way in which employers facilitate access to health care is necessary both because the system is significantly more complex than the German one, and because the impact and relevance of incentive schemes can be understood more clearly against this background.

In 2010, a total of 157 million Americans received health benefits through their employers. However, not all employers offer health insurance, and there are significant differences between large and small employers, the latter often leaving it up to employees to purchase coverage individually. In 2010, 34.6% of workers were employed in firms with more than 5,000 employees constituting 36.9% of those received coverage through their employers, whereas 8.3% of all workers were

employed by companies with 3 or less employees, constituting 5.5% of those receiving coverage. The average annual premiums in 2010 were \$5,049 for single coverage (with workers contributing on average 19% of this sum) and \$13,770 for family coverage (with workers' contributions amounting to 30%, on average), although there is significant variation around this figure. Seventeen per cent of covered workers have a single premium that is at least 20% higher than the average single premium, and 20% of covered workers have a single premium that is less than 80% of the average single premium (Kaiser Family Foundation 2010: 9, 69; Wallen and Williams 1982).

Price and benefit packages also depend significantly on the type of plan that employers offer. While 84% of employers offer only one type of plan (Kaiser Family Foundation 2010: 59), the four principal options are:

- Health Maintenance Organizations (HMOs: services are covered within specified networks only, and referrals are made within the network. This type usually has the lowest co-payments);
- Preferred Provider Organizations (PPOs: these have more flexibility and permit use of some out-of-network services, although patients usually need to pay co-insurance in such cases);
- Point Of Service plans (POSs: combining HMO and PPO features—an in-network physician is the primary provider, and out-of network services are covered as long as the primary care physician initiates a referral), and
- High Deductible Health Plans (HDHPs: these have generally the lowest premiums, but the highest deductibles, and most attractive for people of good health and/or in strained financial situations), see also Table 2.3.

While the differences in cost of coverage between the different plan options is relatively moderate, the amount workers contribute directly varies more significantly and can be almost twice as high on a HMO plan (\$1,028) compared to a HDHPs (\$632, see Appendix 2B). In the US employers therefore pay a much larger share of health care cost than in Germany, and co-payments and other forms of cost-sharing play a more prominent role, enabling the use of financial incentives that are coupled directly with the cost of coverage. To a significant extent, the employer-based health insurance system is governed by standard (insurance) market mechanisms of supply,

demand and risk-adjustment, although there is also some overarching regulation that imposes some limitations on the activities of insurance companies.

#### *2.4.1 Legal framework—health responsibility*

Not least because of the diversity in providers of health care, and the multitude of plans that may be offered in the private market and through employers, there is currently no explicit single set of norms that would specify responsibilities of health care users in different areas. However, similar to the German initiatives on health promotion and incentives, there is some federal guidance on the conditions under which particular types of incentives may be offered. The background to this regulation is the 1996 *Health Insurance Portability and Accountability Act* (HIPAA) which, among other things, sought to improve continuity of health insurance when individuals moved between providers because they changed employers (or lost employment). Prior to the act it was possible that someone who previously had coverage was denied insurance by another company, or charged a significantly higher premium, for example because of pre-existing conditions. HIPAA established that a group health plan may not demand higher premiums of individuals than of other “similarly situated” insurance holders, with comparisons based on a range of health-risk related factor such as medical history or disability.

But HIPAA also clarified that this did not prevent insurers from offering incentives for participation in health promotion and disease prevention programmes. Subsequently, the Departments of Labor (DoL), Treasury (DoT), and Health and Human Services (DHHS) issued guidance in 2006 to address the scope and limitations of this provision in guidance entitled *Nondiscrimination and wellness programs in health coverage in the group market* (in the following: Nondiscrimination rules, Departments of Labor, Treasury and Health 2006).

The *Nondiscrimination Rules* distinguished two different kinds of incentives. First, in the case of what can be called ‘participation-incentives’, a premium discount, rebate, or cash payment may be given simply for participating in a programme, for example one relating to weight-loss or smoking cessation. Second, in the case of what can be called ‘attainment-incentives’, a reimbursement may be

given for meeting certain health status targets, relating to risk factors such as Body Mass Index (BMI) or blood pressure. One important difference between these two types is that practically all people will be able to use participation-incentives, whereas many will not be able to use attainment programmes, whether for health, or other reasons. The Departments therefore considered it necessary to set a ceiling for the level of incentives that may be offered for attainment-incentives. Drawing on evidence from a leading benefit design consultant who suggested that premium discounts offered by insurers at the time ranged from around \$60 to \$480 annually in 1998, the departments calculated that this equated to approximately 3% or 23 % of the total cost of an employee's coverage (i.e., the employee's premium plus the employer's contribution). They subsequently specified three alternative thresholds of 10, 15 and 20% as incentive ceilings, that became the subject of a formal consultation (Departments of Health and Human Services 2001). In finalising the guidance the departments opted for the upper end of this scale, justifying as follows:

The final regulations provide that the amount of the reward may not exceed 20 per cent of the cost of coverage. The percentage limit is designed to avoid a reward or penalty being so large as to have the effect of denying coverage or creating too heavy a financial penalty on individuals who do not satisfy an initial wellness program standard that is related to a health factor. Comments from one employer and two national insurance industry associations requested that the level of the percentage for rewards should provide plans and issuers maximum flexibility for designing wellness programs. Comments suggested that plans and issuers have a greater opportunity to encourage healthy behaviors through programs of health promotion and disease prevention if they are allowed flexibility in designing such programs. The 20 per cent limit on the size of the reward in the final regulations allows plans and issuers to maintain flexibility in their ability to design wellness programs, while avoiding rewards or penalties so large as to deny coverage or create too heavy a financial penalty on individuals who do not satisfy an initial wellness program standard that is related to a health factor. (Depts of Labor, Treasury, Health 2006: 75018)

The question of the scope and limitation of wellness incentives, and of appropriate thresholds was revisited during health reform in 2009 and 2010. Both the House of Representative's *Affordable Health Care for America Act* (Section 112) and the Senate's *Patient Protection and Affordable Care Act* (Section 2705) included further proposals, with the latter proposing to increase the level of reimbursements for attainment-incentives to 30%, with the option of 50% for particular initiatives, subject to approval by the DoL, DoT and DHHS. Echoing earlier regulation, the proposal emphasised that for individuals for whom it is "unreasonably difficult due to a medical condition.... [or] medically inadvisable" to take part in the programmes,

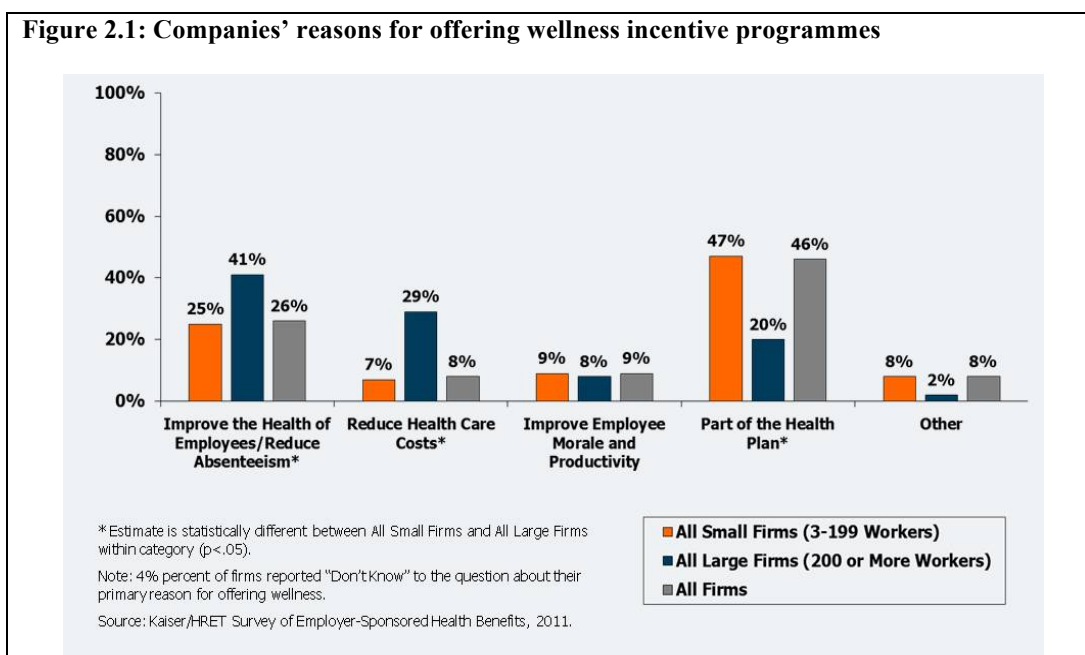
a reasonable alternative standard must be provided for reimbursements to be made, as long as plans were able to request the employee's physician to confirm the existence of a relevant condition. The Senate Bill's provisions passed unaltered into the final *Patient Protection and Affordable Care Act*, passed in March 2010.

In practice, there are several ways in which these provisions can be implemented. In the form of genuine incentives, contributions for insurance holders could be held constant, and those who are willing and able to seek to achieve attainment thresholds could be offered reductions up to 30 or 50% of the cost of coverage. In this case, cost for the incentive would need to be offset by savings resulting from healthier employees, or otherwise be absorbed by the employer. Alternatively, the model envisaged in Wikler's thought experiment (that could also be found in the German cancer prevention regulations pursuant Article 62 SGB V) could be used, by increasing contributions for all, and offering the option of returning to previous levels as an incentive. The *Nondiscrimination Rules* explicitly note that incentives may be implemented by imposing, in effect, differential premiums, as programmes may shift costs "from plan sponsors to participants who do not satisfy the standards, from participants who satisfy the standards to those who do not, or some combination of these." (Depts of Labor, Treasury, Health 2006: 75027). In this case the financing of incentives would result, at least in part, from cost-shifting, rather than gain-sharing. Depending on the type of health plan and the willingness of companies to risk controversy, in principle this provision can permit for substantial differences in cost of health care to employees, as Table 2.3 illustrates.

<b>Table 2.3: Wellness incentives and cost-shifting options: impact on cost of coverage as passed on to consumers</b>		
	30%	50%
Assuming the average cost of coverage of \$4,700, then incentives up to the following amounts may offered:	\$1,410	\$2,350
<i>Scenario A:</i> (genuine incentive) Initial cost of coverage remains unchanged. Incentive amount reduces overall cost of coverage	$\$4,700 - \$1,410 = \$3,290$	$\$4,700 - \$2,350 = \$2,350$
<i>Scenario B:</i> Initial cost of coverage changes: it is made up of previous cost + incentive amount. The incentive is to return to previous contribution levels	$\$4,700 + \$1,410 = \$6,110$	$\$4,700 + \$2,350 = \$7,050$

It is somewhat ironic that one of the major achievements of the 2010 health reforms was finally to end medical underwriting (further to the aforementioned restrictions already resulting from HIPAA), and that the same law should then, in principle, enable significant cost variation in a very similar vein under the guise of wellness incentives.

The extent to which employers will in fact make use of the new opportunities is not clear at this stage, although a marked increase over time can be noted. According to a survey by the National Business Group on Health, 36% of large employers (with at least 1,000 employees) used incentives in 2009. In 2011 this figure rose to 54%, and 80% plan to use incentives in 2012. The use of incentives as penalties over these time points has increased almost fivefold: levels more than doubled from 8% in 2009 to 19% in 2011, and then doubled again to an expected 38% in 2012 (National Business Group on Health and Towers Watson 2011). An earlier survey by the Group found that that 56% of large employers see wellness incentives as one of the top three priorities for curbing cost<sup>9</sup> (National Business Group on Health 2011), and a survey by the Kaiser Family Foundation indicates clear interest in using wellness programmes for health care cost control and improved productivity through reduced absenteeism (Figure 2.1).



<sup>9</sup> The other two being cost-sharing and consumer-directed plans.

While the discussion so far has been somewhat abstract and dealt mainly with what is legally possible, it is also instructive to consider a representative real world example that illustrates the concerns about cost-shifting are not merely armchair philosophy. The BeniComp® Advantage scheme, offered by the benefit design consultant BeniComp is intended to work with a HDHP, and its mechanics are set out clearly on the website, both for employers and employees, and is reproduced in full at Appendix 2C.

The information BeniComp provides for employers opens candidly:

What haven't you tried?... You've increased employee contributions, raised copays and coinsurance, ... We all have tried to shift the cost of health care with plan design changes and by having our employees share in the cost. It simply has not worked.

The company then praises its product which is alleged to be the “first plan that rewards employees for managing their own health”. The function is described as follows: “If an employee maintains or improves their health, BeniComp® Advantage offers a financial reward. For those employees who choose not to modify their lifestyle, they do not receive these rewards and will incur higher out of pocket cost.” The framing of “reward” used here follows directly Wikler’s thought experiment introduced in Section 2.2 and the company acknowledges that the plan “does shift cost, but only to those who have lifestyles that threaten controlled premiums for the rest of the population<sup>10</sup> [...] Either way, costs are allocated more fairly”.

An illustration then shows how the plan works. If employees previously had a deductible of \$500, this is increased to \$2,000. Employees can reduce this by achieving four attainment-incentives, relating to not smoking, and meeting BMI, cholesterol, and blood pressure targets.<sup>11</sup> Each has a value of \$500, and the information emphasises: “Unearned credits=employer savings!”, suggesting that savings of between 12–30% can result from better employee health and higher

---

<sup>10</sup> Regarding the population perspective it needs to be born in mind that this, of course, only refers to the population that is covered by employer-based insurance, and not to people outside of it, nor to people over 65, as employers typically only provide coverage until retirement at which point employees are covered by Medicare (which can lead employer-base plans to adopt a more short-term perspective regarding health promotion and cost-savings).

<sup>11</sup> “Blood Pressure: <120/80, Colesterol LDL: <100, Body Mass Index (BMI): <24.9. While these are established national guidelines, your employer may choose to modify these parameters”, source as below.

deductibles, as well as from employees “who choose other health care options”, i.e. leave the health plan. The programme may also lead to selection effects if prospective unhealthy employees are aware of it in the application phase, and decide to seek employment at a comparable company that does not operate this kind of incentive programme.

Under the heading “This seems aggressive. How do most employees react?” BeniComp suggest that employees understand that companies “cannot continue to provide rich benefits with plan expenses outpacing profits 5:1”. Furthermore, it is suggested to pitch the programme in an analogy to “auto insurance [:] those employees who are doing what they can to lower their risk of serious illness receive the equivalent of a ‘safe driver discount’. Employees understand this concept and receive it very well”.<sup>12</sup> BeniComp’s programme is fully compatible with the legal and regulatory framework, as the information materials also emphasise. Despite the far more obvious potential for inequalities in abilities to use programmes such as these, the US has no reporting requirement whatsoever, neither regarding actual cost-savings that may result from the initiatives, (as in the German case), nor on health improvement or user profiles in terms of health or income groups.

#### 2.4.2 Stakeholder views

Perhaps the most remarkable change in the US policy framework on health incentives is the most recent increase in levels of permissible incentives. It is useful to review the debate that led to it in some more detail. The provisions found their way into draft health reform legislation first in the Ensign-Carper amendment, passed in winter 2009/10 by a vote of 18:4 in the Senate Finance Committee’s work on *Americas Healthy Future Act* that was to become the *Patient Protection and Affordable Care Act* (Ensign 2008). The amendment was also known as the *Safeway-amendment* as, according to Sen. Carper (D-Del.), co-sponsor of the amendment, Sen. John Ensign (R-Nev.) is “a big advocate of the Safeway program” (Ross 2008; Hilzenrath 2010). Steve Burd, CEO of supermarket chain *Safeway*, had published an OpEd entitled “How Safeway Is Cutting Health-Care Costs: Market-

---

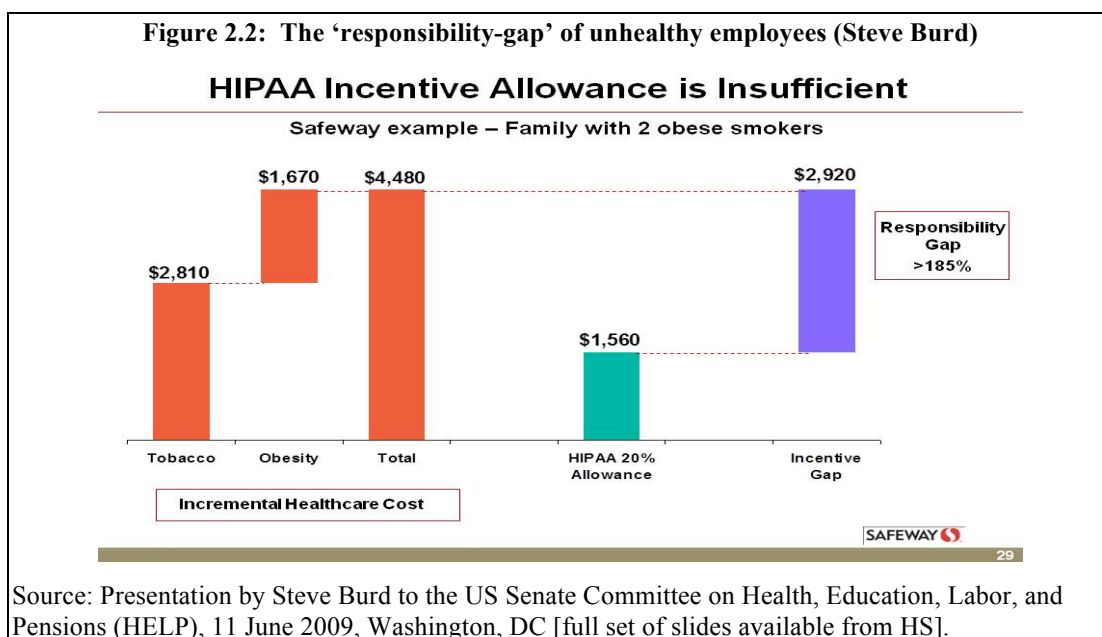
<sup>12</sup> All quotations from: BeniComp® Advantage, “Detailed overview”, available at: [http://www.benicompadvantage.com/index.php?option=com\\_content&view=article&id=8&Itemid=4](http://www.benicompadvantage.com/index.php?option=com_content&view=article&id=8&Itemid=4), accessed 18 February 2012



based solutions can reduce the national health-care bill by 40%” in the *Wall Street Journal* in June that year, claiming that the wellness programme implemented by his company had helped keep health care expenditure constant over the previous four years by varying premiums by as much as \$780 for individuals and \$1,569 for families (Burd 2009). In his commentary Burd also set out the line of argument that can be found in BeniComp’s explanation of the company’s *Advantage* programme:

For decades, driving behavior has been correlated with accident risk and has therefore translated into premium differences among drivers. Stated somewhat differently, the auto-insurance industry has long recognized the role of personal responsibility. As a result, bad behaviors (like speeding, tickets for failure to follow the rules of the road, and frequency of accidents) are considered when establishing insurance premiums. Bad driver premiums are not subsidized by the good driver premiums.

Burd also received considerable attention from policymakers in Washington, and discussed his insights, among other occasions, in a presentation 11 June 2009 to the US President’s Domestic Policy Council, in which he reiterated many of the points made in the *Wall Street Journal* OpEd. Criticising the permissible premium variation under HIPAA, he also set out what he termed the ‘responsibility-gap’, arguing that the incremental health care costs of a family with two obese smokers amounted to a total of \$4,480, of which HIPAA only allowed employers to recoup \$1,560, see Figure 2.2.



Burd’s view is frank, but also somewhat at odds with the declared intention of wellness incentives to focus on health promotion, as well as with the DoL’s, DHHS’ and DoP’s emphasis that the initial 20% threshold was acceptable, as otherwise there

was the risk of incentives having “the effect of denying coverage or creating too heavy a financial penalty” as stated in the above-cited section from the *Nondiscrimination Rules*.

Unsurprisingly, Burd’s intervention attracted criticism from several quarters. First, on the grounds that he has so far not been forthcoming with actual evidence for the cost-reductions that he claims have been achieved through *Safeway’s* programmes. David Hilzenrath, a journalist with the *Washington Post* argued that the case showed how “the untested claims of interest groups can take on a life of their own and shape national policy” (Hilzenrath 2010). Hilzenrath reviewed *Safeway* policy documents and interviewed senior staff, finding that health care costs had in fact not been kept flat for four years. Costs had dropped in 2006 by 12.5 % as a result of the company overhauling its benefits, according to *Safeway* Senior Vice President Ken Shachmut, one of Hilzenrath’s interlocutors. However, this decline had nothing to do with the company’s incentive programme, which had not been implemented until 2009, according to Shachmut, and after 2006 costs began to climb again. According to company spokesman Brian Dowling it would also be difficult for premium incentives to have such an impact on the company’s overall workforce of about 200,000, because the Health Measures programme had been open to only 28,000 of whom 17,000 to 18,000 had enrolled, mainly general office workers rather than store personnel (who are covered by union contracts).<sup>13</sup>

Regardless of similar concerns that had been circulating among informed circles as health reform was being debated, the *Safeway Amendment* still received much support, including from medical leaders such as the Cleveland Clinic’s chief executive, who acknowledged that more evidence would be helpful, but equally suggested that “the lack of empirical third party data does not preclude moving forward” (Hilzenrath 2010). At the same time, a sign-on letter to members of Congress authored by the *American Heart Association* and endorsed by more than one hundred health care advocacy groups including *AARP* (formerly the American Association of Retired Persons), the *American Cancer Society-Cancer Action Network*, the *American Diabetes Association*, the *American Public Health*

---

<sup>13</sup> Hilzenrath points out that “Burd’s assertions about flatlining costs pertained to those workers eligible for Healthy Measures, as did the projected 8.5 per cent increase for 2009.” (Hilzenrath, *ibid.*)

*Association, Consumers Union, and the Obesity Action Coalition* urged Congressmen and women to abandon the amendment because of concerns that cost-shifting can be used to penalise sicker employees; because there was lack of clarity of what constitutes a “reasonably designed” wellness programme; and because the ‘alternative standard’ is only provided for people with a medical condition (and not others who may have reasonable concerns, for example about sharing their health information with people other than their doctor, see: American Heart Association 2009).

A position paper on wellness incentives issued in September 2010 by the *American College of Physicians (ACP)* equally expressed concern about approaches along the lines envisaged by Burd. The ACP supported “the use of positive incentives to motivate behavior change” but also cautioned that any discriminating potential needed to be avoided, and that the “incentive structure must not penalize individuals by withholding benefits for behaviors or actions that may be beyond their control” (American College of Physicians 2010: 1-2). Regarding possible impact on the doctor-patient relationship the ACP noted that incentives:

should support appropriate patient autonomy and participation in decision making, including the right to refuse treatment, without punitive consequence [and that programmes] should be designed to recognize and support the physician’s ethical duty to provide care, the physician’s ethical responsibility to discuss all appropriate care options with the patient in a culturally sensitive manner and the physician’s professional obligation to make recommendations on the basis of medical merit. (American College of Physicians 2010: 1-2)

While relevant stakeholder groups therefore do not oppose the principle of using incentives for health improvement, the reforms eventually introduced under the Patient Protection and Affordable Care Act are viewed as problematic due to lacking evidence that would support the change, as well as because of concerns regarding the impact on vulnerable populations and the doctor-patient relationship.

#### *2.4.3 Implicit and explicit rationales*

In the USA, as in Germany, there is a clear interest by policymakers and health care professionals in using wellness programmes and incentives for health promotion, and Figure 2.2 showed that one third of large employers think that they are a suitable means for this purpose, with the added benefit of reducing absenteeism and improving productivity. Almost as many think that such initiatives can help reduce

health care expenditure (see also NBGH 2010). In principle, cost-reduction may be achieved in different ways. One option would be that better health will lead to lower health care usage, as noted regarding the potentially problematic German requirement, that bonuses may only be paid, if savings result from the programmes (Section 2.3.2). Another option, emphasised, for example by BeniComp, who advise that their Advantage scheme will result in “immediate” savings, is to shift cost to employees, or to motivate high-cost enrollees to seek coverage elsewhere. Although American employees have far higher contributions than their German counterparts, employers still cover the larger part of health care costs, which may explain some of the interest in using incentives to shift cost back to employees. Another, perhaps more implicit, and more value-laden rationale for such initiatives may be found in an appeal to fully embrace the ethos of ‘live free or die’. As Schachmut, *Safeway’s* Senior Vice President observed: “I have no problem with a smoker having a 10-pack-a-day habit and killing him or herself... I mean, it's a personal choice. It's a free country. I just don't want to have to pay the health-care costs of that personal choice... and the same thing is true for obesity” (Hilzenrath 2010).

## **2.5 Conclusion**

In part for overlapping reasons, in part for distinct ones, individual-level incentives are of considerable interest to policymakers and different stakeholders in Germany and the USA. There is universal agreement that they may be suitable tools to promote better health and efficiency of health care services in principle, although their acceptability depends considerably on the specific details of their implementation.

In the German context, goal conflicts can arise as incentives can be used by sickness funds not just for health promotion or efficiency purposes, but also to secure a competitive advantage over other providers in attracting ‘good-risk’ insurees. The fact that in the American context employers are far more directly involved in the provision of health care introduces a different kind of potential goal conflict, as outlined above, leading to higher willingness to shift cost and achieve reductions in absenteeism. In both countries health care professionals observed that incentives had the potential to interfere with the doctor patient relationship, with German doctors in particular opposing being given a policing function.

A major difference between the German and the US approach is that while, *prima facie*, both countries intend to use the schemes for health promotion, the levels of incentives differ significantly and are far lower in Germany. A more conceptual difference that is however directly related to issues around the permissible level of incentives is that US policy recognises that attainment-incentives are generally not open to all, whereas this is more likely the case with participation-incentives. Such a distinction is not recognised in Germany, where both types are mixed (see Appendix 2A). Equally, while US policy requires that alternative standards be provided to those who (for medical reasons) are unable to meet attainment-incentive targets, there is no such provision in Germany. While Germany imposes a reporting requirement on the financial impact of incentives, this is not required in the US, and neither country requires evidence that incentive programmes lead to health improvement.

The framing of incentives plays a significant role in both countries, and a number of different approaches can be distinguished. The implementation of the German cancer policy pursuant Article 65 SGB V was officially introduced as an incentive intended to assist responsible behaviour, but was widely criticised as a penalising approach, and not perceived as a merely benevolent encouragement. The perception of a penalty had less to do with the absolute size of the difference (2% instead of 1% of one's annual gross income), which is unlikely to result in financial hardship, but with the mechanics of creating the incentive by doubling the baseline.

The German framework is generally careful to minimise cost-shifting, and explicitly requires that bonus payments must not be financed by those not taking part in the schemes—instead, bonuses should be paid through gain-sharing as a proportion of the win that results from having achieved more efficient services. In contrast to this approach, US policy explicitly permits cost-shifting, and proponents such as Burd clearly see this option as the principal attraction of the policy. His statement regarding the 'responsibility-gap' revealed (see Figure 2.3) that his interest in higher levels of incentives is not driven by, for example, a view that the next stage in making a major step forward in terms of achieving behaviour change and helping people lead healthier lives can only be made if incentive levels are increased from

20% to 30% (or 50%—nor has anyone else ever made such a claim), but that he wants to reduce his company's payments for the obese and smokers. The latter, he feels, unduly burden the community of more responsible and healthy enrollees, and it is hence inappropriate that they should be required to fund care that is required because of avoidable health risk behaviour (in the same way that the community of prudent and careful motorists should not have to pay for the risks that reckless speeding motorists incur because of their actions).

Table 2.4 provides a summary overview of the principal arguments in favour and against personal responsibility in US and German policy. These arguments, as well as the influential views by commentators such as Steve Burd make a number of non-trivial assumptions in relation to concepts such as voluntariness, attributability, causation, praise and blame, and not least the concept of responsibility itself, which will be considered in more detail in the next chapter.

<b>Table 2.4: Principal arguments for and against promoting personal responsibility for health in policy</b>	
Note: Items in “for” and “against” cells are not generally supposed to correspond (i.e., “against” cell is no attempt at countering directly the “for” argument)	
<b>For</b>	<b>Against</b>
<b>Both US and Germany</b>	
<b>Coproduction:</b> Personal action is one of several determinants of health.	<b>Discrimination/victim-blaming:</b> personal responsibility for health (PRH), can become unduly penalising, especially where people are held responsible for factors beyond their control.
<b>Cost-reduction:</b> belief that ‘an ounce of prevention saves a pound in cure’ explicitly underlies German legal requirement that incentive payments may only be made if savings are achieved.	<b>No economic case:</b> Initial German idea for getting more people to cancer screens was not judged cost-effective. In the US, no evidence for claims that influential Safeway wellness programs achieved savings. Overall it’s unclear over what timespan savings should be demonstrated and whether they are realistic, especially longer term.
<b>Voluntary self-harm:</b> Ok to ask users for higher contributions where they require services as result of voluntarily chosen health risks.	
<b>Germany</b>	
<b>Solidarity:</b> Solidarity demands PRH: (1) sickness funds cant ‘make’ people healthy, so cooperation is required (2) more needs can be met if users limit demands, use services only where necessary.	<b>Solidarity:</b> Solidarity can demand limiting PRH: some costs of unhealthy behaviour are absorbed by the solidaristic community, and cost-shifting is not seen as compatible with the principle of risk-sharing.
<b>Compliance/physician-patient relationship:</b> Partly overlapping with cost-reduction argument, but also part of duty to cooperate in treatment/rehabilitation: lacking compliance in cancer prevention can double copay from 1–2%.	<b>Physicians forced in policing role/ physician-patient relationship:</b> checking on compliance and meeting health standards is viewed as incompatible with primary obligation to patients.
<b>Chronically ill have special responsibilities:</b> as they use more services, can be asked to contribute to efficiency.	<b>Arbitrary penalty:</b> not ok to single out group of cancer patients with compliance regulation: other chronically ill are on 1% threshold.
	<b>Coerciveness:</b> Initial idea for getting more people to undergo cancer screens seen as coercive (high false positive/negative rates in breast cancer screens).
<b>US</b>	
<b>Polluter-pays principle/‘Live free or die’ ethos:</b> Individualism demands PRH: people are responsible for their choices, and this includes the consequences of their choices. I am not my neighbour’s keeper, no solidarity. Auto-insurance analogy.	<b>Continuation of medical underwriting/ discrimination:</b> Initial laws on wellness incentives permitted cost-shifting up to 20% of cost of coverage. More was seen as giving rise to discrimination issues. But 2010 reforms increased this to 30% and 50% in exceptional cases, while ending refusal of insurance for pre-existing conditions. Can be seen as medical underwriting by the backdoor.
<b>Fairness/no free-riding in risk-pool:</b> similar to German solidarity conceptualization: because of limited resources it’s ok to penalise those who (allegedly) free-ride. US law explicitly permits cost-shifting from plans to unhealthy, or from healthy to unhealthy.	
<b>Promote productivity, reduce absenteeism:</b> Employers pick up majority of health care cost and feel ‘cheated’ by employees not pulling their weight (eg because of excess weight... )	

## Chapter 3<sup>1</sup>

### **An ethical framework for evaluating health responsibility policies**

#### **3.1 Introduction**

Chapter 2 showed that in policy and practice, proponents of health responsibility draw on different rationales and have different goals in mind, be they to promote the health of people individually and at the population level, to enhance competition between payers of health care, to improve the efficiency or fairness of health care services, to increase productivity in the workplace, or to select for healthy and against unhealthy individuals in the pool of individuals covered by a given company or insurance provider and thereby lower health care expenditure. While it is not necessarily impossible to achieve all of these rationales or goals to a reasonable extent, it also became clear that tensions can arise when they come into conflict. In this chapter I approach the question of how such conflicts might be resolved, and examine more closely the concept of personal responsibility for health from a philosophical perspective. In part, this matches closely with some of the arguments employed in the policy discourse, but in part it also goes beyond this and will hopefully help to explicate some of the confusion and polarization that besets the debate.

---

<sup>1</sup> Text in this chapter draws on, and in part reproduces verbatim, material published in the following previous publications: Schmidt 2007b, 2008a, 2009a, 2009b, 2011a, Schmidt, Halpern, Ash 2012, Schmidt, Voigt, Wikler 2010. Earlier versions of these papers were also presented at several academic meetings: the Priority in Practice Conference, Philosophy Department, University College London, UK, March 2008; the Ninth World Congress on Bioethics, International Association of Bioethics, Rijeka, Croatia, September 2008; the Responsibility & Health Workshop, Hughes Hall Centre for Biomedical Science in Society, Cambridge University, Cambridge, UK, March 2009; the Fourth International Jerusalem Conference on Health Policy, The Israel National Institute for Health Policy and Health Services Research, Jerusalem, Israel, December 2009; at a meeting of the National Institute for Health and Clinical Excellence's Citizen's Council, May 2010, a seminar at the Kollegforschergruppe "Normenbegründung in Medizinethik und Biopolitik" in Münster, January 2011, and a seminar at the The Centre for the Study of Incentives in Health, London, March 2011. I am grateful for valuable discussion with participants at these meetings, and equally for comments from, and discussions with Annette Dufner, Stefanie Ettl, Nir Eyal, Elias Mossialos, Nicole Knaack, Julia Kreis, Kristin Voigt, Alex Voorhoeve, Joy Wang, James Wilson and Dan Wikler, who reviewed drafts of the above papers, or of this chapter, and helped me better understand specific aspects discussed in them.



I begin by contextualising the use of ‘carrots’ and ‘sticks’ for personal responsibility for health in the broader philosophical debate around the scope and limitation of coercion in the health context. This is followed by a conceptual analysis of the core concept of health responsibility: I seek to show that characterizations found in the literature are insufficiently nuanced to be of practical use in planning and evaluating health responsibility policies in contemporary pluralistic societies. In setting out the concept of *health responsibility as co-responsibility* I then present a proposal for a more nuanced concept of personal responsibility that takes up the oftentimes disconnected strands in the philosophical debate, and is nonetheless compatible with the needs of real-world policy.

### **3.2 Personal responsibility and behaviour change in the spectrum of coercion and ‘doing nothing’**

One of the seminal philosophical contributions regarding the question of how to reconcile tensions between state power and individual liberties that is directly applicable to the health care context is John Stuart Mill’s so-called ‘harm principle’, set out in his essay *On Liberty*:

The object of this essay is to assert one very simple principle, as entitled to govern absolutely the dealings of society with the individual in the way of compulsion and control, whether the means used be physical force in the form of legal penalties, or the moral coercion of public opinion. That principle is, that the sole end for which mankind are warranted, individually or collectively in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right. (Mill 1989 [1859] : 13)

It is not uncommon that the citation ends here, and the quote is seen as synonymous with the paragraph’s equally famous conclusion: “Over himself, over his own body and mind, the individual is sovereign” (e.g.: Gillon 1985: 1807; Callahan 1981: 21). However, the section between these two parts is also noteworthy. Thus, before the concluding sentence, immediately following the above section, Mill says:

These are good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him, but not for compelling him, or visiting him with any evil, in case he do otherwise. To justify that, the conduct from which it is desired to deter him must be calculated to produce evil to someone else. The only part of the conduct of any one, for which he is amenable to society, is

that which concerns others. Over himself, over his own body and mind, the individual is sovereign. (Mill *ibid.*)

For the present context, the full quote is relevant as it illustrates that even drawing on a liberal philosopher as Mill one can accept that there are further options between the two poles of hard paternalism<sup>2</sup> on the one hand, and the complete *laissez-faire* of a libertarian minimal state on the other, that we may consider in efforts aimed at behaviour change: for Mill clearly sees it permissible to ‘remonstrate, reason, persuade or entreat’. A recent report by the *Nuffield Council on Bioethics* that was concerned with setting out an ethical framework for broader public health initiatives represented these different options along a so-called ‘intervention ladder’ that ranked interventions by their degree of intrusiveness, see Box 3.1.

**Box 3.1: The intervention ladder (Nuffield Council on Bioethics)**

The range of options available to government and policymakers can be thought of as a ladder of interventions, with progressive steps from individual freedom and responsibility towards state intervention as one moves up the ladder. In considering which ‘rung’ is appropriate for a particular public health goal, the benefits to individuals and society should be weighed against the erosion of individual freedom. Economic costs and benefits would need to be taken into account alongside health and societal benefits. The ladder of possible policy action is as follows:

*Eliminate choice.* Regulate in such a way as to entirely eliminate choice, for example through compulsory isolation of patients with infectious diseases.

*Restrict choice.* Regulate in such a way as to restrict the options available to people with the aim of protecting them, for example removing unhealthy ingredients from foods, or unhealthy foods from shops or restaurants.

*Guide choice through disincentives.* Fiscal and other disincentives can be put in place to influence people not to pursue certain activities, for example through taxes on cigarettes, or by discouraging the use of cars in inner cities through charging schemes or limitations of parking spaces.

*Guide choices through incentives.* Regulations can be offered that guide choices by fiscal and other incentives, for example offering tax-breaks for the purchase of bicycles that are used as a means of travelling to work.

*Guide choices through changing the default policy.* For example, in a restaurant, instead of providing chips as a standard side dish (with healthier options available), menus could be changed to provide a more healthy option as standard (with chips as an option available).

*Enable choice.* Enable individuals to change their behaviours, for example by offering participation in an NHS ‘stop smoking’ programme, building cycle lanes, or providing free fruit in schools.

*Provide information.* Inform and educate the public, for example as part of campaigns to encourage people to walk more or eat five portions of fruit and vegetables per day.

*Do nothing or simply monitoring the current situation.*

---

Full excerpt from: (Nuffield Council 2005: 41–42)

<sup>2</sup> Joel Feinberg defines as follows: “Hard paternalism will accept as a reason for criminal legislation that it is necessary to protect competent adults, against their will, from the harmful consequences even of their fully voluntary choices and undertakings. [...] Since it imposes its own values and judgements on people ‘for their own good’, it seems well named by the label of ‘paternalism’. It is not as clear that ‘soft paternalism’ is ‘paternalistic at all [...] Soft paternalism holds that the state has the right to prevent self-regarding harmful conduct... *when but only when* that conduct is substantially non-voluntary, or when temporary intervention is necessary to establish whether it is voluntary or not” [Emphasis in original] (Feinberg 1986; Dworkin 2010).

The Nuffield Council's mapping can provide a helpful contextualization of health responsibility policies that take the form of 'carrots' or 'sticks', as outlined in Chapter 2. As we saw there, in presentational terms, *prima facie*, most incentive schemes take the form of non-coercive encouragements, suggesting that people are free to take them up or not. That is, on the Nuffield Council's ladder, incentives rank as more intrusive than the options of doing nothing, enabling choice, and guiding choices through changing the default policy. The latter approach has recently attracted some attention under the label of, first, "libertarian paternalism", and then "nudging" (Sunstein 2003; Thaler 2008; Thaler and Sunstein 2003; see also: Marteau, Oliver, and Ashcroft 2008). In contrast to these measures, the hope of those advocating incentives is that they will have more traction, while avoiding the controversies surrounding the more invasive and potentially liberty-infringing measures higher up on the ladder, such as using 'sticks' or disincentives, or restricting or eliminating choice.

Theoretically, the underpinnings of using incentives rest on the same principles of behavioural economics that underpin Cass Sunstein's and Richard Thaler's work, and have also become known as "asymmetric paternalism" (Loewenstein, Brennan, and Volpp 2007; Camerer et al. 2003). The starting point is to challenge the *homo economicus* model of classical economics which holds that people are self-interested rational agents who are generally able to identify the means necessary to achieve goals worthy of pursuit, and to act accordingly. The case of health behaviour, along with other situations, such as retirement planning, complicates this assumption considerably. Even though the desire to lose weight, to drink less or to stop smoking is felt strongly by many, many also fail in acting on it, despite being quite clear about the means that are required, such as eating less, and exercising more.

Researchers in the field of behavioural economics have begun to systematically map these constraints affecting behaviour change. The concepts of "present preference bias" and "quasi hyperbolic discounting" have been coined in response to a series of experiments demonstrating that people generally prefer rewards in the present or near future to ones further in the distance (Laibson 1997; Frederick, Loewenstein, and O'Donoghue 2002; Madden et al. 1997). In other words, the pleasure of an extra helping of cake today will often be preferred over the option of

being healthier later—even if one’s future self would prefer the alternative. The underlying mechanisms that often lead to inertia and procrastination can, however, be turned around by exploiting the very principles that fuel them, for example, by providing immediate feedback and rewards for behaviour change that might not be perceived as rewarding by itself. A further important concept established in observational and experimental research relates to loss aversion (Kahneman and Tversky 1979). What is meant here is that people disproportionately prefer avoiding losses to making gains (of equivalent value). These and further principles of behavioural economics (Loewenstein, Brennan, and Volpp 2007) have led to a range of studies exploring the potential of using losses and gains in motivating behaviour change through fixed sum discounts, cash rewards, lotteries, or deposit contracts in areas such as medication adherence, smoking cessation, weight loss or substance abuse management (Volpp, John et al. 2008; Volpp, Loewenstein et al. 2008; Volpp et al. 2006; Volpp, Troxel et al. 2009; Lussier et al. 2006; Giuffrida and Torgerson 1997; Paul-Ebhohimhen and Avenell 2008; Jeffery 2012; DeFulio 2012; Higgins et al. 2012). Research also suggests that incentives have the potential to benefit low-income groups (Marteau, Ashcroft, and Oliver 2009; Oliver and Brown 2011).

Conceptually, then, it is clear why incentives, in particular, are attractive as tools for promoting health responsibility: while interventions focusing on improving the infrastructure of the environment still depend on people making use of these improvements, incentives seek to focus peoples’ attention more directly on making use of available opportunities, and similar offers that are viewed as having positive value, in a way that seeks to be sensitive to the need to minimise force or coercion.

However, Chapter 2 also illustrated that the extent to which policies can be said to be successful in achieving these goals depends significantly on the way in which incentives are implemented, and I will return below to a framework that seeks to assist in deciding which policies are reasonable, and which ones are not. However, before this step it is also necessary to take a closer look at the very concept of health responsibility. For some of the confusion in the debate arises not only from the fact that people have different rationales in mind when they speak about responsibility, or different ideas about the legitimacy of using policy tools that may entail persuasion or coercion, but also because they have different ideas about the extent to which the

concept of responsibility for health is primarily descriptive or normative, as will be shown next. Clarity on all aspects is required to make progress with the responsibility debate in academia, policy and practice.

### **3.3 Personal responsibility: conceptual analysis**

Why should we talk about personal responsibility in health or other contexts of social policy at all? Alexander Brown suggested that plausible arguments can be made that appeals to responsibility and policies enforcing it rest on five separate rationales. Accordingly, responsibility-sensitive policies may enhance fairness in the distribution of resources, maximise utility of available resources for all, or promote individual self-respect, autonomy, and human flourishing (Brown 2005: 28-33). Some of the policies to realise these rationales may entail penalties, but not all do.

However, for many commentators, ascribing responsibility is intrinsically linked to holding people responsible, and proponents argue that responsibilities without sanctions appear pointless, while opponents caution that imposing sanctions often entails the risk of misunderstanding the fundamental determinants of health, and risks penalising people unduly. Further to the *Black report* (Black 1980) and the *Whitehall studies* (Marmot, Shipley, and Rose 1984; Marmot et al. 1991) which paved the way for much research on the social gradient in health, the 2008 report by the WHO *Commission on Social Determinants of Health* emphasises the importance of the environment in which people live, showing, for example, that a boy growing up in the deprived Glasgow suburb of Calton will live on average 28 years less than a boy born in nearby affluent Lenzie (54 vs. 82 years, see: World Health Organization 2008). This suggests strongly that the role of individual behaviour alone in relation to health outcomes is limited, to say the least (Venkatapuram and Marmot 2009). Further, following Geoffrey Rose's seminal paper "Sick individuals and sick populations" (Rose 1985), the argument can be made that a focus on individual behaviour, consumerist models and "over-dependence on individual choice will not achieve changes in society norms" (Doyle, Furey, and Flowers 2006: 396).

Given the important role of environmental factors, then, many commentators argue that focusing on broader political and public health measures is, first, a more efficient approach for health promotion, and second, more appropriate in fairness

terms, than seeking to get individuals to change their behaviour (Resnik 2007: 445; Minkler 1999: 126-131; Raikka 1996: 360). For, even if all people are held responsible for the same type of health outcomes—say, being overweight and experiencing associated poor health—such policies would raise problematic equity issues and could amount to “victim-blaming”, where people who are already in disadvantaged social positions are held responsible for factors that are largely beyond their control (Minkler 1986: 551; Wikler 2004: 336; Daniels 2007: 67-68, 148). In particular, there are worries that responsibility-emphasising policies may stigmatise people individually or particular socio-economic groups (Roemer 1995: 21; Minkler 2000: 18).

In addition, a branch of feminist theory, traditionally concerned with an analysis of obligations and entitlements against a background of the structure of (power-)relationships between people, has emphasised the “social cooperation model”. On this view, people are embedded in different webs of interconnectedness, which makes it difficult to attribute full responsibility to individual people alone. Susan Sherwin, for example, has argued against appeals to responsibility that penalise people, although she appears to accept that personal responsibility may be a meaningful concept if it has an empowering effect (Sherwin 1998: 165).

Clearly, these lines of argument, research and evidence need to be taken seriously, and it would be naïve to ignore the constraints and opportunities associated with a particular upbringing, geographical location, or income bracket, as opposed to another. However, as noted in Chapter 2, not all responsibilities are concerned with penalising people for bad health outcomes, and it would be short-sighted to reduce the debate about personal responsibility to this one form.

Moreover, as a review of the literature shows, there is a range of different things commentators may mean when they say that “person X is responsible for p”. Sometimes, distinct notions are made explicit, but other times, several meanings may be in use simultaneously, whether explicitly or implicitly. Not all authors seem to be aware of previously published conceptualisations, in part, perhaps, because discussions have taken place in different specialist philosophical, medical and health policy publications, with insufficient interdisciplinary exchange. Irrespective of the

reasons for the sources of the range of different meanings it is crucial to be aware that much confusion arises from not distinguishing clearly between these different senses, or from not being explicit about which sense is intended in endorsements or criticisms of particular responsibility-related policies.

In the following I seek to provide an overview of the different notions that have currency in the literature. In a second step, I will present a synthesis that aims to illustrate that it is not useful to reduce the concept of personal responsibility to just one meaning, as it is, by necessity, multi-faceted, due to the different roles that health behaviour plays in, among other things, medicine, health policy, law, sociology, philosophy, and politics. While a reduction to one single meaning has a certain aesthetical and methodological appeal, I contend that this approach is of little use in making progress with the debate around personal responsibility as it fails to acknowledge that the concept plays a role in more than one discourse.

In terms of concepts that have been set out to be applied in the context of health care (or are otherwise directly applicable), the following influential examples, roughly in chronological order, stress different aspects. Daniel Wikler differentiated between “causal ... responsib[ility versus] responsib[ility] ... [as] being at fault and accountable” (Wikler 1978: 333) and therefore makes a principal distinction between behaviour as having contributed to a health outcome that is judged as negative, and considering sanctions for this outcome. The first may lead to the latter, but there is no intrinsic connection: separate justification is required in moving from one to the other. In a similar vein, Ronald Dworkin distinguished between three senses: one’s purely functional role as a biological organism (“role responsibility”), the effect of personal choices (“causal responsibility”), and claims others may have against people whose health behaviour leads to avoidable cost (“responsibility based on liability”, Dworkin 1981: 28). Thomas Scanlon proposed a split between “substantive responsibility ... [and] moral responsibility” (Scanlon 1998: 21-22, 248, 272, 278), where substantive responsibility is relevant for questions of distributive justice as it demands us to consider centrally the opportunities a person has had, and the choices that she made, in assessing what we owe her, and what she owes the community. Moral responsibility, by contrast, is concerned with the circumstances that determine when it is appropriate to take a person’s attitudes or actions as the basis for moral

appraisal. More recently, Zofia Stemplowska defined “agent responsibility” as meaning that one has responsibility for having brought about a certain state of affairs (thus overlapping with Wikler’s “causal responsibility” and Dworkin’s “role responsibility”), which is separated from “consequential responsibility”, meaning that “the burdens (or benefits) that come with or constitute [an action] are justly one’s to bear (or to enjoy)” (Stemplowska 2008: 241).

The key distinctions made in these four approaches hence differentiate between descriptive and evaluative senses of personal responsibility that may be, but are not necessarily connected: we may say someone is responsible for an outcome simply because her behaviour played a causal role in bringing it about, and we may also find that such behaviour makes her praise or blameworthy, but such assessments can be separate from seeing legal or otherwise retributive sanctions justified.

Distinctions are also made regarding different types of responsibilities that may or may not be deserving of sanctions. Alexander Cappelen and Ole Norheim distinguish two different senses in separating “responsib[ility] for ... choices ... [from] responsib[ility] for the consequences of ... choices” (Cappelen and Norheim 2005: 478). They argue that within a liberal egalitarian approach it can be acceptable to hold people responsible for the choices they have made regarding health risks, but not directly for the actual consequences of their choices. This is so because of complications around determining exactly the degree to which certain health outcomes can be said to be under the control of individuals, both in terms of epidemiological factors and those generally discussed as relevant in the aforementioned literature on the social determinants of health. Therefore, as argued earlier along identical lines by Julian Le Grand, Cappelen and Norheim see tobacco taxation justified on the basis that it is a penalty for taking a health risk, even if no demonstration is required as to the extent to which an individual’s action *de facto* led to cancer (Le Grand 1991: 121). Without suggesting new terminology John Roemer also made a noteworthy contribution on the subject of how to determine the extent of an individuals’ substantive responsibility in the context of considerable variation of social and epidemiological factors. According to Roemer, responsibility assessments should be linked to the degree to which a person’s risk behaviour departs from standards that are typical for the kind of group a person might be assigned to in view



of their health risk propensity (Roemer 1995: 20; see also: Roemer 1994; Roemer 1993).

A separate attempt for conceptual clarity and progress regarding the questions of which kinds of responsibilities can justifiably attract penalties can be found in proposals that suggest that the central distinguishing feature is to do with whether responsibilities relate to past or future actions. Hence Micha Werner, Henk ten Have, Georg Marckmann, Eli Feiring and others separate between “prospective ... [vs.] retrospective responsibility” (Werner 2002: 525; ten Have 1994: 120; Marckmann 2005: 302; Marckmann 2007: 111), or “forward-looking ... responsibility [vs.] backward-looking ... responsibility” (Feiring 2008: 33), generally arguing that retrospective responsibility attributions should be avoided because of their potential to hold people responsible for factors that are in fact beyond their control, and that, instead, the focus should be on (non-penalising) prospective responsibility. Shlomi Segall, by contrast, does not seek to negate that in some cases—including retrospective or backward looking responsibilities—people may reasonably be said to be responsible for a poor health outcome. He has sympathies for a classical luck-egalitarian account, that holds, broadly, that distributional equality does not require equality of outcome, but merely the elimination of inequalities that are beyond people’s individual control (see: Arneson 1989). However, Segall suggests that the value of solidarity may act as a waiver in some cases, requiring us not to abandon people if they are in a position of disadvantage even if this is, to some extent, reasonably attributable to their actions (Segall 2007: 197; see also Buyx 2008: 1515).

These and further characterisations have been set out for different reasons: some sought to address particular policy issues, others evolved from certain problems that arose in theoretical political philosophy, notably the problem of abandonment in (luck)egalitarian frameworks (Anderson 1999; Arneson 2000). Here, the purpose is not to settle deeper theoretical disputes, but to bring the rich and nonetheless to some extent surprisingly disconnected literature on health responsibility to bear on the issue at hand. This is necessary to make progress with actual policy proposals that, in very direct ways, set out responsibilities that often have significant penalties and rewards attached. In addition, I also wish to

demonstrate that for purposes of health promotion and routine interactions of people with health professionals, it is crucially important to understand that talk about personal responsibility is not necessarily connected with the discourse about rewards and penalties.

With several of the above commentators, I hence agree that at the most basic level, it is conceptually important to distinguish whether we are ascribing responsibility in a backward-looking sense (where, for example, we assess someone's past behaviour that is correlated to some health outcome, be it positive or negative) or in a forward looking one (where we may want to specify what people should do in the future). The distinction is central for framing more clearly the issues at hand. However, I have less faith than the cited commentators that a focus on prospective responsibilities alone will avoid thorny fairness issues, or be sufficient to preserve a meaningful concept of personal responsibility for health promotion and health policy. Nor am I persuaded that reference to solidarity alone will be sufficient to attenuate the potentially penalising elements of responsibility ascriptions. As Chapter 2 has shown, solidarity can cut both ways and may undermine, as well as support, understandings of responsibility that can entail penalties.

In both forward and backward looking perspectives, several different things can plausibly be meant when we ascribe responsibility, as has emerged from the above review of different concepts. In a backward-looking sense, the phrase "person X is responsible for p", with "p" standing for a health outcome that is judged to be negative, may mean:

- (1) X's behaviour has played a certain causal role in having brought about p.
- (2) X's behaviour has played a certain causal role in having brought about p, and should recognise this.
- (3) X's behaviour has played a certain causal role in having brought about p, should recognise this, and try to avoid doing so in the future.
- (4) X's behaviour has played a certain causal role in having brought about p, should recognise this, and make good any costs (with or without being blamed) for reasons of distributive justice.

- (5) X's behaviour has played a certain causal role in having brought about p, should recognise this, try to avoid doing so in the future, and make good any costs (with or without being blamed) for reasons of distributive justice.
- (6) X's behaviour has played a certain causal role in having brought about p, should recognise this, try to avoid doing so in the future, make good any costs, and, in cases where X requires treatment, may be given a lower priority than patients whose behaviour played none or a lesser role in contributing to their health care needs (typically with attribution of blame).

As can be gleaned from the brief review above, it is not uncommon for commentators to focus on the last type only, and/or to jump straight from the first to the last type, assuming that having established some degree of causal or role responsibility, a person must also be held responsible (see: Daniels 2007: 67-69; Heath 2008: 787; Cappelen 2005: 477). But this is far from necessary.

There are several reasons why we may find it adequate and useful to draw on some notion of non-penalising retrospective responsibility, whether in abstract policy, specific prevention campaigns, or consultations with health care professionals. For example, in a given case where a person is responsible in one of the first three senses above, there may remain some degree of freedom for personal action and behaviour change even if environmental constraints have played a role, perhaps even a major one. Realising the scope for action in this area is important for avoiding fatalism and resignation, which may have a powerful grip on people struggling to maintain or improve their health. While it is difficult to disagree with the strong emphasis that proponents of the social determinants of health approach put on the general need for improving environmental conditions, an exclusive or overly strong focus on the environment can overlook the degrees of freedom that people have, even in constrained conditions. For people to take action, then, it is necessary for them to realise the extent to which they contributed to, say, a bad health outcome, and, in this merely functional sense, to realise that they are, and can be, responsible for good or bad health. An important qualification is of course Kant's old adage of "ought implies can", which has particular relevance in this context. For it would be pointless, if not outright Kafkaesque or cynical, to specify responsibilities where, due

to strong environmental constraints, it is simply impossible for people to act accordingly.

It is also important to recognise that talk of responsibility in a forward looking sense is in many ways quite different from the more common backward-looking perspective. Hence, what we may mean here when we say that “person X is responsible for p”, and “p” is an action judged to be conducive to good health, may be:

- (1) X should do p as no-one else can, in principle (or will, practically) do p for X (e.g., exercise more, eat less).
- (2) X should do p, as this will be good for the health of X.
- (3) X should do p, as this will be good for the health of others, or the operation of the health care system, even though X won't be penalised if p is not done.
- (4) X should do p, as this will be good for the health of others, or the operation of the health care system, and X knows that a penalty will be imposed if p is not done (or a reward given, if p is done).

Again, it is far from necessary that the first or second type of responsibility, which may be said to relate primarily to prudential obligations, automatically leads to the last type, which, together with the third, also invokes notions of justice. Some health-related behaviours simply require that people individually do them, as no-one else will do them for them, and not even the most optimal environmental conditions will ‘make them do them’, in some sort of mechanistic way. It is in this somewhat banal, but nonetheless crucially important sense, that a range of health-related behaviours are personal responsibilities. Noting them and appealing to them in health promotion activities is relevant since—environmental constraints permitting—in a significant sense it is up to us to decide on whether we wash our hands regularly, brush our teeth, exercise, see our GP when we are sick, are honest about our health-relevant information, take part in public health programmes, and so on. Advocating such responsibilities can result in clear personal benefits and is also likely to complement the social determinants of health approach as it can help identify particular social or other structural constraints that make it difficult for people to live healthily.

Prospective and retrospective responsibilities can, as some of the above commentators suggest, be advocated in an either/or fashion—for example, with Wikler it would clearly be possible to argue that (retrospective) personal responsibility should at best play a peripheral role (Wikler 2004: 134), but that a prospective notion of non-penalising responsibility in the sense of (1)–(3) above may nonetheless be permissible. But it would be more difficult to advocate retrospective responsibilities without prospective ones. For in the case of penalties it seems reasonable to hold people responsible only if they have had knowledge of the penalty that is attached to certain choices or behaviours at the point of time where they acted in ways that were likely to put them at a disadvantage. Yet, this principle seems not heeded universally in policy: as we have seen in Chapter 2, in the case of the German policies on repaying cost of treatment for complications arising from tattoos, piercing or non-medically indicated plastic surgery (Section 2.3.1), the penalty was introduced both for people who may decide to undergo one of these procedures after introduction of the law, and for all those who happened to have done so in the past (not knowing, clearly, that at some point these actions would incur a penalty). By contrast, the policy on higher co-payments for sufferers of breast, colon, or cervical cancer patients—introduced in the same law—did specify an age cut-off, meaning that the policy will not retrospectively punish those who failed to behave in ways that are deemed legally appropriate. Even if commentators are likely to continue to differ about the role that opportunity of choice should have in holding people responsible, it is clear that advocates of personal responsibility policies would have a stronger case where they set out and justify explicitly prospective responsibilities first, before implementing retrospective ones, in particular where significant rewards or punishments are associated with the measures.

### **3.4 Personal responsibility as co-responsibility**

So what would a plausible concept of health responsibility look like? Further to the above analysis, I draw on Werner's general characterisation of responsibility in seeking to answer this question (Werner 2002: 521-527). Werner separates four dimensions of responsibility: *someone* (a subject of responsibility) is responsible for *something* (the object of responsibility) towards *someone* or some *entity* (the judicial authority) in view of particular *normative standards* (the normative background). To

unpack this approach for the present context in somewhat more detail, it then follows that:

- *Subjects of responsibility* are, first, patients, and much of the literature focuses on patient obligations (see, for example: English 2005; Gauthier 2005; Kelley 2005; Resnik 2005). However, as the analysis in Chapter 2 showed, people in other health states may also have obligations. These are the healthy; those who are unwell, but not yet in need of treatment; and those who are recovering from an illness.
- *Objects of responsibility* are typically past or future actions that relate to one's own health, the health of others, or the operation of the health care system. All three dimensions again being commonly found in policy and law.
  - Self-directed responsibilities may concern leading a healthy life as part of a conception of a 'good life', or, in a more instrumental sense, to achieve particular life plans.
  - Responsibilities towards others can take the form of not harming them. They may also relate to caring for health needs of those under one's guardianship, such as children, protecting and promoting their (future) opportunity ranges. Donating blood or organs are further forms of obligations towards others, if arguably much weaker ones than not harming them.
  - Responsibilities towards the health care system concerns contributing to its fair and efficient operation so that it can serve as many people in need as possible. For example, missing appointments, or not cancelling them in time may deprive others of medical attention, and may have considerable financial cost, as may wasting medicines or using services unnecessarily.
- *The judicial authority*: Where health responsibilities are codified explicitly, the issuing authority or its agents (health care providers or medical professionals) may carry out assessments of whether the obligations have

been met, and whether some form of positive or negative response should follow. But obligations may also be set out merely as ideals or aspirations.

Before turning to the remaining question of the normative standards, that determine which responsibilities are acceptable and which ones are not, I will summarise the key elements of the discussion so far in order to set out the core of what I take to be an appropriate concept of responsibility in the health care context: *health responsibility as co-responsibility*.

Health responsibilities concern one's reasonable prospective and retrospective obligations as a healthy person, patient or convalescent, to lead a healthy life, to respect the health of others, and to contribute to an efficient health care system, insofar as available choices and external factors permit this. Prospective responsibilities relate to appeals to act in a certain way in the future, whether in a purely descriptive or functional sense ('no-one else will go jogging for you, brush your teeth', etc) or in a way that appeals to an obligation to realise certain values, which may or may not entail rewards or penalties. Retrospective responsibilities relate to an assessment of past behaviour in relation to obligations that, ideally, have been specified earlier and were known to the addressees. Retrospective assessments can have a purely descriptive explanatory meaning, for example, where a certain degree of causal responsibility for poor health is conveyed, or appeals are made for behaviour change, on the grounds that this is likely to improve a person's health. Retrospective assessments can also have positive consequences in the form of praise, or financial or other bonuses that may have been offered as incentives. Equally, they may have a potentially negative character, for example, using measures such as not-for-payment bills to convey the cost of treatment; requiring (higher) co-payments; or assigning lower priority in treatment.

Since health is affected both by personal behaviour and factors generally beyond immediate individual control (socio-economic status, access to health care, infrastructural arrangements, etc), it is neither an exclusive matter of personal or social responsibility. As the element of personal control admits of degrees, conceptually, substantive responsibility also needs to admit of degrees. By necessity,

health responsibilities are therefore *co-responsibilities*.<sup>3</sup> This is relevant both for the assessment of the causal factors that led to a particular health state, as well as for attributions of praise or blame, and decisions about possible positive or negative sanctions.

In one sense, this concept might seem disappointingly vague and bland. However, it appears to be the most appropriate model to capture the various interwoven dimensions of health responsibility that can reasonably be subsumed under the term, and to do justice to the fact that an overly narrow focus on blame and punishment detracts from preserving an important core of the concept that can be independent of sanctions. I also contend that both conceptually and in practice, attributing responsibility must admit of degrees, despite the fact that holding people accountable must often be a matter of either/or, and I turn next turn to the question of how particular forms of health responsibility might be justified.

### **3.5 Assessing the reasonableness of health responsibility policies**

In one sense, the question of health responsibilities might simply be a matter of choosing “the right” normative framework. Various political perspectives have different ways of explaining which of the above notions of responsibility should be central, and which ones should be more peripheral, as is clear from the background discussions to the policies discussed in Chapter 2. Equally, as already alluded to in the discussion of the core concept of personal responsibility above, there are different accentuations in philosophical contributions, such as luck-egalitarian ones (Arneson 1997; Dworkin 2000; Roemer 1994; Roemer 1995) or communitarian (Callahan 1998), or libertarian accounts (Engelhardt 1981). However, there are two principal problems with resorting to a single foundational normative theory. First, in value pluralistic societies, agreement about what constitutes the right framework remains generally elusive: saying, for example, that people should be held substantively

---

<sup>3</sup> Note that the concept of co-responsibility also features in Article 1 of the German Social Security Code (SGB V), although in a somewhat different and narrower sense, as responsibility for health is there characterised as shared between the health care system and patients. See also Ian Forde’s and Rosalind Raine’s characterisation of health as co-production: “Responsibility for better health should be shared between society and the individual, ... society’s efforts for health improvement should be dovetailed with individuals’ and families’ efforts.” Central to their discussion is that policies are required that “support ... people to engage with decisions about their own health” (Forde and Raine 2008).



responsible for choices that lead to poor health because a certain luck-egalitarian theory says so will appeal to luck-egalitarians—but not to those who do not share its premises. Second, and perhaps even more significantly, even if we suppose that we are able to find a country in which all residents (or just citizens) can agree on a single monolithic theoretical account, whether political or philosophical, such value systems are typically of a very general nature, and do not tell us *ad more geometrico* how to decide in designing and evaluating concrete policies which, as Chapter 2 showed, have a great number of design features each of which can give rise to distinct normative issues.

Of course, this situation is not unique to the health responsibility debate. For example, regarding the controversial question of just resource allocation, which, according to Norman Daniels, is: “How can we meet health needs fairly when we can’t meet them all?” we are equally faced with a range of substantive positions that offer different perspectives. To make progress in practice, Daniels suggested a proceduralist approach in which general principles of justice are supplemented with fair processes for limit-setting, drawing on the framework of Accountability for Reasonableness, that he initially developed with Jim Sabin (Daniels 2000; Daniels 1999, 2007). This approach requires that policies meet four conditions concerning (1) publicity (making publically accessible decisions and their rationales), (2) relevance (providing a reasonable explanation for decisions, see below), (3) revision and appeals (putting in place mechanisms for challenge and dispute resolution), and (4) regulation (voluntary or public regulation to enable the three aforementioned conditions). The relevance condition is specified in its briefest form as follows:

The rationales for limit-setting decisions should aim to provide a reasonable explanation of how the organization seeks to provide “value for money” in meeting the varied health needs of a defined population under reasonable resource constraints. Specifically, a rationale will be “reasonable” if it appeals to evidence, reasons and principles that are accepted as relevant by [fair minded] people who are disposed to finding mutually justifiable terms of cooperation. Where possible, the relevance of reasons should be vetted by stakeholders in these decisions ... (Daniels, 2007: 110)

While the approach has been the subject of some criticism,<sup>4</sup> I adopt the Accountability for Reasonableness approach to make progress with the debate

---

<sup>4</sup> A relevant charge for the present context has been pressed by Richard Ashcroft. Qualifying his attack somewhat by calling it a polemic he is concerned that the widespread interest in policy circles

around personal responsibility for health in a pluralist society. Moreover, I also see this approach as compatible with Thomas Scanlon's contractualism which can provide a particularly fitting normative anchoring.

Contractualism is based on the idea that people seek to justify to each other actions, principles or policies that govern their conduct. It is typically both a theory about the legitimacy of authority, and about the motivation, form and content of particular moral norms (Ashford and Mulgan 2009). Two principal types can be distinguished. A contractarian branch traces its roots to Hobbes' assumption of a so-called 'natural state' in a pre-social society, characterised by the *bellum omnium contra omnes*. With Jan Narveson, the argument is hence that social contracts and morality more broadly would be accepted by self-interested individuals, "first because we are vulnerable to the depredations of others, and second because we can all benefit from cooperation with others" (Narveson 1988: 148).

Contractualism proper differs from contractarianism in motivation and method. Here, the focus is not on self-interested bargaining individuals, but on people who

---

in procedural justice accounts often amounts to "the adoption of a 'presumption of undecidability'", eschewing robust normative defence: "If bioethicists have anything distinctive to contribute to formation of public health policy and planning, it is in our training in the analysis of complex ethical problems and argument about normative conclusions. If we act on a presumption that the ethical problems of resource allocation, pandemic response and so on, are too complex or controversial for us to illuminate other than by listing principles and describing, once again, principles of procedural justice, then we are abrogating our responsibilities—or admitting our redundancy" (Ashcroft 2008: 7). It is indeed regrettable when the rigorous analysis of problems and argument about normative conclusions is abandoned. However, it is not clear that a procedural justice account requires, or by necessity leads to this consequence. Quite the contrary, as is clear from the above-cited quotation setting out the relevance condition, it explicitly invites such activity. The approach is, however, more modest in accepting that a single comprehensive normative position is not likely to be supported by all members of a value pluralist society, and that hence some trade-offs are required, that people who are interested in finding mutually agreeable terms of cooperation would find acceptable. Therefore, while there are, as noted immediately above, comprehensive normative (luck-egalitarian, communitarian, or libertarian) positions that could provide broad guidance on how to conceptualise the concept of personal responsibility for health, it is not clear which of these we should accept as the master theory, superior to a procedural justice account (and Ashcroft provides no guidance on this difficulty). I therefore proceed here on the assumption that setting out a procedural justice-based list of tests or areas in which justification is owed is the most appropriate way forward, since established comprehensive normative frameworks do not command universal support, lack sufficient specificity, and because it appears that a number of relevant ethical issues can be addressed effectively without recourse to first principles (such as the impact of programmes on the doctor patient relationship, or the acceptability of different rationales, for example in relation to cost-saving). This does not mean that normative theory should be ignored, but rather, that it is one input among many, and that those holding particular normative positions need to accept that there can be competing positions, and that a procedural justice account provides a better basis for making progress in policy and practice than the attempt to persuade everyone to become luck-egalitarian, utilitarian, etc.

share a commitment to justify publicly the standards of morality applicable to all (Scanlon 1998: 191-194), offering a far more promising basis for the context of public health ethics than the contractarian approach (Ashcroft 2006: 11).

Historically, contractualism can be traced to Kant's emphasis that moral principles must be such that they can be justified to all, and to Rousseau's notion of the 'general will' that would be adopted by all free and equal citizens. John Rawls' work is widely regarded as influential in resurrecting this form of social contract, and Scanlon has presented a distinct form in his 1998 book *What We Owe to Each Other*.

Scanlon's contracting parties are not placed behind a Rawlsian veil of ignorance, but should be understood as agents who are aware of their circumstances, with distinct "generic reasons" regarding their willingness to accept certain "principles that may affect them. Generic reasons are reasons that people have in virtue of their situation, characterised in general terms, and such things as their aims and capabilities and the conditions in which they are placed" (Scanlon 1998: 204). These reasons feature prominently in the concept of "reasonable rejectability" that is central to Scanlon's approach:

[Contractualism] holds that an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a the basis of informed, unforced general agreement. (Scanlon 1998: 153) [...] In order to decide whether a principle could be reasonably rejected, we need to consider it from a number of standpoints. From the point of view of those who will be its main beneficiaries there may be strong generic reasons to insist on the principle, and to reject anything that offers less. From the point of view of the agents who will be constrained by it, or of those who would be beneficiaries of an alternative principle, there may be reason to reject it in favor of something different or less demanding. (Scanlon 1998: 213)

Scanlon asserts that his approach offers an accurate description of moral motivation, which rests on "the positive value of living with others on terms they could not reasonably reject" (Scanlon 1998: 162). Justification is hence important not merely in formal or strategic terms, as for contractarians (because it enables a person to bring others to do what serves her interests), but *substantively*: justifying actions to others is "to embrace the value of mutual recognition" (Stratton-Lake 2004: 15). It is what makes one a moral agent, and Scanlon emphasises that in this sense his view has "substantive moral content" (Stratton-Lake 2004: 134).

It was said above that the approach pursued here would be not to draw on a particular moral theory in order to justify health responsibility policies, and it may therefore be surprising to see this anchoring in a substantive normative theory. However Scanlon's contractualism offers a formal, rather than a prescriptive structure, and is well suited to underpin approaches such as Daniels' and Sabin's Accountability for reasonableness approach, and other institutions found in democratic arrangements, such as consultations in the case of potentially controversial planning projects, whether these relate to new railways, buildings, or health policies. In all cases the fundamental motivation of such enterprises is to elicit a broad range of relevant reasons, to ensure that the changes in view of reasonable concerns are made, or otherwise, that explicit justification is provided why particular objections were not deemed relevant.

In this process Scanlon draws particular attention to reasons that may be cited to reject a policy, and this is of particular relevance to the present context, as it is possible that those who may be disadvantaged by health responsibility policies are to be found among societal groups that have no strong lobbies to argue their case. Making genuine efforts at understanding what reasons those have who object to certain policies, and agreeing that mutual recognition is a product of justifying to each other why one design should be more acceptable than another, can hence been seen as substantively rich approach, but it is nonetheless one that aligns closely with procedural justice arrangements in place in most democratic societies. However, despite this helpful general guidance, and in common with other normative approaches, Scanlon's contractualism of course does not give us a finer grid that would help us see in which areas exactly we should ascertain whether people might have reasons to reject health responsibility policies.

In order to specify the areas in which justification is owed, I set out below seven "tests" that concern the impact of a policy in the planning, monitoring, or evaluation phase on key normative and structural values and components that are integral to practically all health care systems.<sup>5</sup> These tests concern evidence,

---

<sup>5</sup> This framework was first published in Schmidt, H. 2008. Bonuses as incentives and rewards for health responsibility: a good thing?, *Journal of Medicine and Philosophy* 2008, 33:198–220. Note that by coincidence Julian Le Grand and Divya Srivastava set out a similar approach in a Report for

rationale, and feasibility; intrusiveness; equity; solidarity/risk-pooling; attributability and opportunity of choice; affected third parties; and coherence (see Box 3.2).

**Box 3.2: Seven tests to evaluate the appropriateness of health responsibility policies**

***Evidence, rationale, and feasibility***

What are the policy's principal rationales and goals? Have they been justified in an open and transparent manner, with opportunity for comment by all those affected by the policy? How sure can we be that the policy will achieve its aim(s), in principle and in practice? Are the required efforts and cost proportionate in view of the goals?

***Intrusiveness and coerciveness***

Are there ways in which the goal of the policy could be achieved in less intrusive ways? If not, is the extent of intrusiveness justifiable in view of the expected benefits?

***Equity***

Are there some groups (such as particular socio-economic, ethnic, or regional subgroups) who are likely to experience disproportionate benefits or burdens as a result of the policy? At what point would it be reasonable to reject a policy because of inequitable impact?

***Solidarity/risk-pooling***

Insofar as the health care system has an implicit or explicit principle of solidarity or risk-pooling: how does the policy affect it? If it should undermine solidarity or risk-pooling: are all affected clear about this, and can the effect be justified?

***Attributability/opportunity of choice***

To what extent are penalties or rewards based on actions that can be attributed to people's free and voluntary choices? Where peoples' opportunity of choice is limited: can waivers or alternative standards be implemented? Should rewards be given, even if people have not changed their behaviour, but just happen to satisfy the policy's criteria?

***Affected third parties***

Does the policy have an effect on the relationship people have with, for example, their physicians or employer? Insofar as physicians are involved in assessing whether or not someone has complied with their responsibilities: is their involvement justifiable and accepted by them and their patients? What information should employers have (or not) about people's compliance with responsibilities?

***Coherence***

How does the policy compare with standards of responsibility, attributability, and blame in other areas of social policy and the law? Since tensions can be resolved in more than one way: in which way should they be addressed?

I illustrate their relevance by focusing on financial incentive systems introduced in chapter 2, which are also the subject of the research in the subsequent chapters. To some extent, this illustration is somewhat general, as the framework is not applied to a single specific policy, but instead comments on overarching features of the schemes described there. However, an example of a direct application can be found in Chapter 7, where the framework is used to show how an incentive policy seeking to improve the uptake of colon cancer prevention might be justified. For now, the discussion should help illustrate how the approach can be used in practice in general terms.

---

Health England (Le Grand and Srivastava 2009: 35), scoring incentive schemes using the following criteria: effectiveness, (low) cost, equity, feasibility, local/individual autonomy, which, in different guise, are also included in the framework presented here. However, my framework also goes further in recognising the importance of assessing the impact on solidarity/risk-pooling, affected third parties, and attributability.

### *3.5.1 Evidence, rationale, and feasibility*

In Chapter 2, several different rationales were shown to underlie the policy initiatives in Germany and the USA. The first step in planning personal responsibility measures is hence to justify these publicly, and to enable those affected by them to contribute their views, which includes possible arguments they would make to reasonably reject them, or request modifications.

Insofar as incentive programmes are offered on a voluntary basis, people might not be concerned about the rationale of using them to improve health. However, in practice, the implementation of incentives typically means not just that some people are offered an additional ‘carrot’, but that those not taking part are denied one, and, in effect incur higher health care costs, as the description of the US incentives schemes clearly showed (see Section 2.4.1). The German programmes generally operate on far lower levels, but in principle the same issues are raised. In both cases, what is offered as a ‘carrot’ will seem to many far more like a ‘stick’. In view of this situation it is desirable to justify implementations such as the above explicitly, and it is especially important to provide evidence that the programmes have a reasonable chance of success, both in terms of helping people change their behaviour, and in terms of achieving goals such as cost reduction, that, as noted, also feature prominently.

While much of the focus in the debate around incentives and personal responsibility is focussed on hard questions of distributive justice, the acceptability of programmes may well turn to a significant extent on an empirical analysis of the effectiveness of incentive programmes in practice, and the actual costs associated with particular risk factors and ageing populations, which are often simply assumed.

In this context it needs to be noted that consensus remains elusive regarding the question of whether, overall, prevention (whether achieved through incentive programmes or other means) will curb cost over time, and there is an as yet unresolved dispute about whether increasingly longer life expectancy will in fact lead to higher levels of morbidity and care needs (known as the ‘medicalisation thesis’), or whether longer life will mean that the period in which care is needed is simply condensed over a shorter time than previously (the ‘compression thesis’) with some

arguing that this will not lead to overall increases in health care expenditure and others less convinced about possible savings (Kühn 2005; Metz 2001; see also: Busse et al. 1999).

For example, Pieter van Baal and colleagues used a dynamic population model to analyze health care data from the Netherlands in order to estimate lifetime health care cost conditional on the presence of risk factors. They found that expenditure was highest for the healthy (defined as: non-smoking, BMI between 18.5 and 25: health care cost from age 20 was estimated to be €281,000) and lowest for smokers (€220,000), with obese people in an inter-mediate position (€250,000), largely due to differences in the longer life expectancy of healthy people, and associated cost for care (van Baal et al. 2008). Better health at older ages may condense health care costs over a shorter time, and possibly reduce costs, but longer life at good health may also lead to higher absolute levels of morbidity and need for care (Woolf 2009; Russell 2009; Cohen 2008). If evidence to this effect should become irrefutable, it would seem that from a purely economic perspective that focuses on health care expenditure only, we ought to be more concerned about the lifestyle associated cost-impact of muesli-eating jogging seniors, than about smokers who drink excessively, and are overweight. Of course, the mere fact that prevention programmes may not lead to cost savings in the longer term does not mean that they should not be carried out. But it would seem that other reasons would need to be given in their support.

### *3.5.2 Intrusiveness and coerciveness*

Incentive systems are generally framed as not being particularly intrusive or coercive, as it is commonly suggested that people are free to use them or not. However, high levels of reimbursement, especially where combined with cost-shifting, as illustrated above (see Sections 2.2 and 2.41), can raise doubts about the extent to which people are free not to make use of the offers. Providers also differ in the way they advertise their programmes, and it is not uncommon for insurance holders to receive frequent reminders by mail or other means. Such initiatives may be perceived as “nannying” or forms of “healthism” (Levin 1987; Steinbrook 2006) or perhaps bribery (Ashcroft 2011: 195) and may have a counterproductive effect on health responsibility attitudes and the development of personal autonomy. Programmes that reward participation in pre-symptomatic check-ups, such as cancer

screens, may also be intrusive in the sense that they bring uncertain and unwelcome knowledge about disease susceptibility, possibly leading to anxiety or confusion, even if the degree of intrusiveness can be mitigated by focusing on providing information with appropriate confidence intervals and the option for people to discuss any questions they may have with independent experts. The level of intrusiveness or coerciveness therefore needs to be considered carefully, and is closely linked to the questions regarding evidence and rationale: poor evidence and rationales combined with highly coercive or intrusive measures would make for rather bad policy.

### *3.5.3 Equity*

The fairest way of providing health interventions is often simply to make them available universally for all: this avoids stigmatization and leaves uptake to people who are suitably motivated. On the other hand, such approaches can be prone to problematic self-selection biases. For the present context, not only the penalising effect resulting from cost-shifting that has been noted above needs to be considered (which is likely to disadvantage most those who are generally poorer in health and income), but also the question of whether unequal reaping of benefits should be acceptable.

Certainly, based on data from the Bertelsmann Gesundheitsmonitor, initial evidence from Germany suggests that this is the case: 19%, or almost twice as many people belonging to the fifth (least poor) quintile, used incentive programmes in 2004–2005 as opposed to 11% of the first (poorest) quintile (Braun et al. 2006: 22). An analysis over time, using the same survey data, suggests that overall participation in wellness incentive programmes almost doubled between 2004 and 2008 from 13% to 25% of the insured population. With regard to health status, this doubling could also be observed in the subgroups of those with fair health or better, but usage among those with poor health changed very little over time and remained between 14–17%. Uptake also remained relatively unchanged for the lowest income group (between 10–17%) while all other groups more than, or nearly doubled (Schmidt, Stock, and Doran 2012, see also Figure 8.12). In the US, a recent study examining the health status user profile of new enrollees of Medicare plans that had added gym membership benefits also found that 6% more of those responding to this offer



reported excellent or very good health, compared to the prior situation, in which no such benefits were offered (Cooper and Trivedi 2012), suggesting that selection effects are real.

It is useful, then, to think more clearly about differences between people regarding their existing motivation, actual behaviour change, and factors that may make behaviour change more challenging, when it comes to assessing the fairness of who receives advantages associated with incentives, and who loses out. To approach this question in a structured manner, I suggest here that it is instructive to consider the responses of the following five groups to incentive programmes that are offered universally to all enrollees of a health plan: (1) ‘the lucky ones’; (2) the ‘yes I can’ group; (3) the ‘I’ll do it tomorrow’ group; (4) the ‘unlucky ones’; and (5) the ‘leave me alone’ group.<sup>6</sup> Depending on the exact characteristics of particular schemes, the impact on these groups varies, of course. Nonetheless, a somewhat more abstract consideration is useful for bringing clarity to the question of ‘who benefits’ that is central to equity considerations. The framework also shows that there are significant differences regarding whether or not programmes promote behaviour change among enrollees. What I suggest might be called the ‘five groups problem’ therefore concerns the following question: at what point do inequalities in the capacity to use incentive programmes constitute unfairness, and how should we respond in policy?

### 3.5.3.1 ‘The lucky ones’

For practically any incentive programme there will be people who would qualify for associated reimbursements without any form of behaviour change. Some people simply enjoy eating healthily and exercising regularly, and do so quite effortlessly. Their behaviour is hence compatible with the spirit of an incentive programme, even if the benefit—say, a participation-based reimbursement for going to the gym regularly, or an attainment-based one for meeting certain Body Mass Index (BMI)

---

<sup>6</sup> My analysis is based on a review of the public health literature on incentive use and behaviour change, personal insights resulting from involvement in the evaluation of incentive programmes, and conceptual analysis of the characteristics of incentive users. I began with established concepts and developed a more nuanced framework that can be applied directly to wellness programmes in planning, practice, or evaluation. Differentiation between the five groups is not intended to provide an exhaustive and exclusive model, to suggest that beliefs are never shared across groups, or to imply individuals may not belong to different groups in different stages of their lives. My purpose is to illuminate plausible distinguishing features between different groups of users in the context of incentive programmes.

thresholds—does not lead them to change their actions. Others whose actions may remain unaffected include people whose dispositions are not as well aligned. For example, some people may eat in the most unhealthy ways, never exercise, and still have a favourable BMI values. Despite the dissonance between their motivations and the programme's spirit, they may reap the exact same benefits as their health-conscious counterparts, without any change in behaviour or motivation.

#### 3.5.3.2 The 'yes I can' group

Other people would not normally have performed the benefit-qualifying behaviour, but they may see the incentive as a welcome occasion—though perhaps not the sole reason—for trying to overcome inertia or weakness of the will. The incentive benefit's 'nudge,' coupled with their underlying motivation, provide an effective basis for action. Here, incentives are likely to feel like a deserved reward. They may help initiate behaviour change in the first place, or sustain it, where intrinsic motivation is not yet sufficiently developed. Conceptually and in practical terms, this group is also generally known as the 'group of responders' in the literature—yet, it cannot be assumed that all, or even the vast majority of those offered incentive programmes, are, in fact, responders. It is also plausible to assume that people's responses differ with regard to the mode of incentive. Some may find that a 'soft' participation -incentive that merely requires them to partake in an activity is most effective nudge, feeling overly pressured by a 'hard' attainment-incentive, that requires meeting, for example, BMI targets. For others it may be the other way round, and they may respond better to a more robust challenge that requires them to meet set thresholds. Designing an effective incentive for the 'yes I can' group therefore requires an assessment of what is likely to be the most effective and acceptable approach.

#### 3.5.3.3 The 'I'll do it tomorrow' group

Others, like those in the 'yes I can' group, also have a desire for behaviour change, but, for a range of reasons, often simply cannot bring themselves to act on it. They may simply feel unable to try, or where they try, find that they often fail. The reasons may include unfavourable opportunities in their everyday circumstances, such as poor access to affordable and healthy food, or insufficient time to prepare it. Or they may lack access and time for physical exercise in a safe environment. Such

factors can render attainment-incentive schemes, such as achieving BMI values, significantly more challenging. And people have been brought up differently: some received more encouragement to be self-motivated and -efficacious, and others less. Therefore, even participation-incentives such as lower health care cost in return for gym attendance may be taken up more readily by some than by others. For many in this group, incentives may be extremely tempting, yet they can be as far out of reach as the branches of the fruit-laden trees were for the proverbial Tantalus.

#### 3.6.3.4 The ‘unlucky ones’

For biological, medical or other reasons that are completely external to their volition, some people face such strong constraints that, whatever they would do, they are simply unable to meet the criteria associated with specific attainment or participation-incentives such as BMI targets or gym participation. For example, some people with genetic mutations will always be obese, regardless of how much they exercise or control their energy intake. As with the ‘I’ll do it tomorrow’ group, incentives that are simply out of reach will make little sense for the ‘unlucky ones’.

#### 3.6.3.5 The ‘leave me alone’ group

A last group of people comprises those who would qualify in principle for wellness incentives but may voluntarily decide not to use them. They may already meet targets, or could do so easily, or could effortlessly participate in incentivised activities, but still resist. Reasons may include that they feel patronised or ‘nannied’ by wellness schemes; or concerns that incentives introduce an inappropriate element of competition in health plans that they think ought to be based on a principle of mutuality and fair risk-sharing. Or, on quite practical grounds, they might judge the effort required to register for schemes to be too burdensome.<sup>7</sup>

---

<sup>7</sup> Of course, such reasons may also play a role in the deliberations of the ‘I’ll do it tomorrow’ group and the ‘unlucky ones’. The purpose of setting out the groups in this way was, however, to bring out the clearest distinguishing features between the main groups, rather than to try and provide a wholly exhaustive and exclusive set of categories. For example, a further group (that, to some extent, may also be viewed as a sub-group of the ‘Yes I can’ group) could be called ‘the grumpy ones’: here people are able and successful in securing the incentive premium, but they resent the fact that they did so, for example, where they do not care about the supposed health benefit, but participate because they felt ‘bribed’ by the level of the incentive.

As this conceptualization shows, universally offered wellness incentive programmes can give rise to several general problems, including the following:

- Some people may receive benefits, even if their motivation and behaviour runs counter to the spirit of an incentive programme;
- Behaviour change is not always required, and some people may receive benefits for default behaviour—whether this is the result of deliberate prior choice, or unreflective habit;
- Some people face constraints due to weakness of the will, poorly developed self-efficacy, or strong medical or societal constraints. Meeting targets, or participating in health promotion activities, requires a much greater effort of them, compared to others. Still, where they fail to begin or complete an incentive programme, they must forgo the benefit in the same way as those who had sufficient opportunity of choice, but who voluntarily decided against taking part.

Ethically and in policy terms, this raises the question of how significant such differences are—both quantitatively, regarding the actual number of people in each of the five groups for a given policy, and qualitatively, regarding the question of how possible differences in numbers between groups should be addressed. While the other elements of the overall framework proposed here would clearly also need to be considered, and while the costs of redressing inequities always need to be balanced against other central goals of a health care system (Le Grand 1984: 40; Oliver, Healey, and Le Grand 2002: 566) the principal options would be (1) to continue to offer incentives universally, regardless; (2) to offer them universally but with some modifications that make it easier for people who, for example, face more challenging environments to avail themselves of the options; (3) to offer targeted, instead of universal schemes; and (4) to abandon incentive programmes altogether. Appendix 3A illustrates what these different policy options would mean for the different groups.

#### *3.5.4 Solidarity*

Solidarity, as noted in Chapter 2, is a value that explicitly underlies the German statutory health insurance system. In practice it means that the healthy support the

sick; the young support the old; the employed the unemployed; and the better off the worse off, as insurance contributions are income-tested. More normatively, the concept can be seen as an expression of the recognition of human vulnerability, and it is plausible to describe those brought together in social health insurance systems as being in a solidaristic relationship with others, which has the aim of providing mutual protection against the negative consequences of sickness and disease. To a significant extent, key aspects of the principle can be found in other insurance systems that rely on risk-pooling, even if the value would be more implicit in such cases.

The general concept of solidarity is highly complex. Recent work examining its motivation, function, and role in different European health-care systems has analysed it, among other things, as attitudes of individuals or descriptions of communal arrangements, and has explored notions of fellowship, compassion, charity, altruism, universal or group-specific brotherhood, friendship, interest coalitions, civic duties, or mutual recognition and interdependency of individuals (Prainsack and Buyx 2011; Houtepen and ter Meulen 2000; Ashcroft, Campbell, and Jones 2000). Although there are hence a range of different candidates that might explain why solidarity should matter, it is clear that a functional baseline description of solidarity in public health-care systems would be focused around the notion of achieving collectively a degree of security that could not generally be achieved individually. In this sense, solidarity (and, by extension, the less value-laden concept of risk-sharing) has also been characterised as “the ‘beating heart’ of a social health insurance approach” (Saltman 2004: 29). It is equally clear, then, that any evaluation of appeals to health responsibilities, whether in the form of bonus systems or policies that have a more negative character, needs to consider whether they are likely to enhance this central concept, or will be to its detriment.

### *3.5.5 Attributability and opportunity of choice*

As highlighted above, there are a number of ways in which it can make sense to attribute a good or bad health outcome to a person without linking this assessment to questions of praise or blame, or reward and punishment. Often, causal attributability will only be partial, as a number of other factors, typically arising from the environment within which a person lives or works also need to be considered.

Where negative sanctions are contemplated—whether framed as incentives or disincentives—there needs to be good evidence that the people concerned had a reasonable range of opportunities to avoid what is regarded as a poor health outcome. In this regard the requirement in the US regulations that an alternative standard must be provided for those who feel unable to meet the standards required by particular attainment-incentives programmes are a useful way of acknowledging that peoples' circumstances and opportunities of choice are central and can differ (Le Grand 1984: 46), and that some programmes will simply be incompatible with the choice sets people have in their daily lives. However, the provision is also very narrow in focusing on medical conditions only (or, following the stratification under the five groups problem, the situation of 'the unlucky ones') and hence ignores much of the data that come from the social determinants of health literature, which demonstrates that the socio-economic situation of a person can imply equally powerful, and often directly linked, constraints (i.e. affecting many in the 'I'll do it tomorrow group'). Care is hence required in devising policies that offer fair chances to all.

### *3.5.6 Affected third parties*

Depending on implementation, incentive systems may not involve any third parties; for example, in the case of the German incentives for regular dental check-ups, reception staff may log a visit electronically, and the patient then receives the rebate for any work that needs to be done. However, the determination of other types of incentive-qualifying behaviour may involve health care staff; for example, they are required where rewards are made if key health data, such as blood pressure, stay within a certain range over a year. Those on no-claim plans may wish that health care staff was not involved if they require treatment before completion of the qualifying period, and issues may arise where patients appeal to staff not to record their appointment. Equally, the US requirement that a physician needs to attest that a person is unable to meet a standard for an attainment-incentive can lead to similar situations. Most of these situations are likely to lead to awkward situations, but depending on the size of the incentive at stake, more serious tensions may arise, and health care professionals may not be pleased with being put in an actual, or perceived, policing position, which may have a detrimental effect on the doctor-patient relationship (Bishop 2006: 757) or the trust relationship (Marteau, Ashcroft, and Oliver 2009: 984).

Another relationship that needs to be considered is that between incentive programme participants and their employers, in particular in cases where incentive programmes are offered in the work place. For obvious reasons, employers are likely to be interested in their employees' health status, and while most countries have in place data protection legislation that regulates access, the implementation of wellness programmes provides opportunity to review compliance and adequacy.

### *3.5.7 Coherence*

The coherence test asks how benefits or disadvantages that result from a personal responsibility policy fit in with the wider context of social policy and law. It is probably more relevant for cases where explicitly penalising sanctions for contributions to a bad health outcome are envisaged. For example, the concept of contributory negligence as applied in jurisprudence regarding road traffic accidents offers an approach where similar questions are addressed on a day-to-day basis. However, while coherence across different areas of social policy and the justice system more widely is clearly desirable, possible conflicts can be resolved either by aligning a particular health responsibility policy with the wider context, or, alternatively, it may be that the health context shows the relevance of significant constraints that require us to re-assess the justification of other policies, provided they are similar in all relevant aspects.

The seven tests set out here in the context of a contractualism-based procedural justice account are intended to enable policy makers, practitioners and ordinary citizens to scrutinise and evaluate in a clearly focused way policies seeking to promote personal responsibility for health. Applying the framework can therefore help in agreeing mutually acceptable justifications for particular policies. At the same time, the process may also mean that some constituents of basic evaluative reference points shift, as the analysis may indicate that revisions are necessary. The process of applying one's reference framework to specific cases, and reviewing whether in the case of clashes the framework, or the case at hand needs to be adjusted, is sometimes described as striving to achieve a "reflective equilibrium" which consists:

in working back and forth among our considered judgments (some say our 'intuitions') about particular instances or cases, the principles or rules that we

believe govern them, and the theoretical considerations that we believe bear on accepting these considered judgments, principles, or rules, revising any of these elements wherever necessary in order to achieve an acceptable coherence among them. The method succeeds and we achieve reflective equilibrium when we arrive at an acceptable coherence among these beliefs. An acceptable coherence requires that our beliefs not only be consistent with each other (a weak requirement), but that some of these beliefs provide support or provide a best explanation for others. [...] The key idea underlying this view of justification is that we “test” various parts of our system of beliefs against the other beliefs we hold, looking for ways in which some of these beliefs support others, seeking coherence among the widest set of beliefs, and revising and refining them at all levels when challenges to some arise from others. (Daniels 2011)

Thus, agreeing which policies that seek to promote personal responsibility are reasonable requires us not only to apply our personal normative reference framework to this specific matter but, in doing so, to accept the possibility that some parts of our reference framework may require revision. In this sense willingness to engage in such a process can also help refine ethical theories, in addition to providing a workable vehicle for making progress in a particularly controversial area of health care policy and practice.

### **3.6 Conclusion**

It must be admitted that the approach of *health responsibility as co-responsibility* set out here is somewhat less clear-cut than one of the for-or-against personal responsibility stances often encountered in the literature, and especially in political debates. With a number of different types of forward and backward-looking responsibilities, a procedural justice account supplemented with seven tests to specify the areas in which justification is owed (without a single test whose outcome would necessarily “trump” all others) the situation seems to be messy.

Still, I contend that this situation is far preferable to any of the alternative options, if we want to avoid the victim blaming potential that personal responsibility policies are prone to have, and equally if we want to bypass the potentially fatalistic implications that an exclusive focus on the social determinants of health can have, and instead seek to preserve a meaningful concept of health responsibility that is appropriate in descriptive, epidemiological, and moral terms. Particular policies that seek to implement personal responsibility standards hence depend on a holistic justification in a number of different areas. These areas, as circumscribed by the



seven tests outlined above, concern central values that are integral to the provision of health care, and I believe that much progress can be made in policy and practice if, in a transparent and open process, valid and explicit reasons, and sound evidence to support them, are given in the design and evaluation of personal responsibility policies.<sup>8</sup>

The following chapters will examine attitudes of the public and of health care professionals regarding central normative elements of this framework, to assess the relevance of this analysis, and the importance of central features. While this chapter only set out the framework in general terms, Chapter 7 will apply it to the specific case of colon cancer prevention, which will show more clearly the scope and limitations of the approach.

---

<sup>8</sup> See also footnote 4 in this chapter on the question of why a procedural justice based account is the most promising way forward, despite inevitable shortcomings, as highlighted by Ashcroft 2008.

## Chapter 4<sup>1</sup>

### **Incentivising individual behaviour: review of the survey and interview literature, and development of an instrument for new research**

#### **4.1 Introduction**

The previous chapters have demonstrated that internationally there has been an increasing focus on personal responsibility in health care policy and practice, with interventions ranging from benevolent encouragements to genuine financial incentives and explicit penalties. To date, policies in both countries have been shaped largely by discourse among politicians, policy makers, employers and academic health policy analysts, including ethicists. While physicians have also been active participants in the debate, the voice of those most directly affected by the interventions, i.e. the public, has not been heard in equal measure, although different forms of public or consumer engagement have been used in other health policy and research contexts.<sup>2</sup>

While it appears that there is no systematic work that engages consumers in the design or implementation of incentives, key rationales for public engagement that are found elsewhere, including identifying consumers' preferences to improve the effectiveness of interventions and maximise the acceptability and legitimacy of potentially controversial policies through engagement (Kreis and Schmidt 2013), apply just as strongly in the case of wellness incentives (Daniels 2009b; Schreier and Diederich 2008; Institute of Medicine 2003; Merzel and D'Afflitti 2003). For the present context, knowing more about views of the public and physicians is certainly of relevance in view of the approach set out in the previous chapter, that seeks to adapt

---

<sup>1</sup> Text in this chapter draws on, and in part reproduces verbatim, material drafted by HS and published in Kreis and Schmidt 2013.

<sup>2</sup> Including priorities for health research (Caron-Flinterman et al. 2005; Brown et al. 2006; Gooberman-Hill, Horwood, and Calnan 2008; O'Donnell and Entwistle 2004), designing randomised controlled trials for specific diseases (Koops and Lindley 2002; Langston et al. 2005; Marsden and Bradburn 2004; Helfand 2010), community-based participatory research (Dalal et al. 2009; Gray et al. 2000; Savage et al. 2006), medical device development (Bridgelal Ram, Grocott, and Weir 2008), guideline development (Boivin et al. 2010; Jarrett L. and the Patient Involvement Unit 2004), systematic reviews (Helfand 2010; Braye and Preston-Shoot 2005; Serrano-Aguilar et al. 2009) and health technology assessment (Facey et al. 2010; Gauvin FP et al. 2010; Abelson et al. 2007; Lehoux et al. 2009).

Norman Daniels and Jim Sabin's accountability for reasonableness framework to questions of responsibility.

It was proposed there to assess policies at the planning, implementation or evaluation phase by examining evidence, rationale, and feasibility; intrusiveness; equity; solidarity/risk-pooling; attributability and opportunity of choice; affected third parties; and coherence. It is obvious that the broad set of reasons that will typically be drawn on in justifying policies would be incomplete without any information on how those directly affected by them experience them. This seems certainly plausible regarding questions around the acceptability of particular rationales, and, for example, the degree to which interventions are felt to be intrusive and their expressive or symbolic value, e.g. whether they are demeaning. Moreover, as the conceptual analysis has suggested, some incentive programmes carry the risk that they may disproportionately disadvantage groups of lower socio-economic status, who are often marginalised and without strong lobbies that would ensure that their views are heard. Engagement, or research on the attitudes of this group of people, can therefore help to achieve a more balanced set of reasons that are relevant for the assessment of incentive programmes. In a more pragmatic sense, understanding the public's views is also relevant for an effective use of incentive programmes: for they are likely to have the greatest chance of success where views about the reasonableness that employers or policy makers may have are aligned with similar assessments by the public.

Part of the intention of this thesis is therefore to carry out original research into views of members of the public on salient aspects that determine the acceptability and effectiveness of incentive programmes. Equally the plan is to gain a better understanding of the views of physicians, as, again, the previous chapters have indicated that their active support is required for incentives to be effective (but cannot always be presumed). To ensure that the empirical work carried out here will not duplicate previous work, and instead builds on it, the first part of this chapter presents an overview of themes from the recent survey and interview literature that is relevant for the questions being addressed here. The second part develops an instrument for the empirical work being carried out in the following chapters.

## **4.2 Key themes in relevant survey and interview work so far**

This section describes the structure of the review and presents findings regarding attitudes towards the role of responsibility in the special case of organ transplantation, and then in health care more generally, focussing first on views of the public and patients, and then on those of physicians.

### *4.2.1 Structure of review*

The review of the published literature on survey and interview work follows broadly the useful proposal for structuring systematic reviews in bioethics that Daniel Strech, Matthias Synofzik and Georg Marckmann set out, building on earlier related work (Strech, Synofzik, and Marckmann 2008; see also: McCullough, Coverdale, and Chervenak 2007). Strech et al. reviewed the related, but far more general, field of empirical work on rationing, and were puzzled by the difficulty of establishing a coherent pattern of conclusions from different qualitative research findings, due to factors such as heterogeneous methodologies, sample compositions and angles of research questions. Adapting the approach of systematic reviews as used in medical research, they proposed that progress could be made by adopting the following seven step approach: (1) careful definition of the review question; (2) selection of relevant databases; (3) application of ancillary search strategies (review of bibliographies from relevant references, hand search of journals not or not completely captured by databases used); (4) development of search algorithms (through index mapping, i.e.: identifying database-related search terms by reviewing indexes of articles already known from prior non-systematic literature reviews, and cluster modelling, i.e.: combining database-specific search terms with Boolean operators); (5) relevance assessment of the retrieved references (through blinded reviewers, predetermined inclusion and exclusion criteria); (6) quality assessment of included studies (including a justification and explication of the assessment tools); and (7) data analysis and presentation (including a justification and explication of methods).

The review here does not attempt to present an exhaustive systematic review of qualitative research findings on personal responsibility in the literature, nor does it aim to weigh the relevance of different arguments. Rather it seeks to ensure that the work to be carried out here in relation to discrete issues arising in the context of incentive programmes will not duplicate previous efforts, and instead will build on these, by

being sensitive to themes that emerge from previous related research. It is therefore not necessary strictly to follow the approach outlined above, in particular regarding steps 5–7. Nonetheless, the framework provides a helpful general approach which was adapted as follows.

The general review question was: “what empirical research is there on attitudes of the public or physicians to using incentive systems as tools for promoting personal responsibility for health, and has work similar to that proposed here already been done?” Based on previous research on conceptual issues around personal, and confirmed by informal consultations with colleagues working in the field, it was clear that in terms of databases PubMed, which comprises more than 20 million citations for biomedical literature from Medline, life science journals, and online books, would be an appropriate resource (step 2). 10 relevant papers were already known prior to the search and these papers and the cited literature were reviewed to establish an initial list of search terms. Most papers were included in PubMed, which allowed an identification of the relevant data-base specific search terms, called “medical subject headings” (MeSH) in the case of PubMed (step 3 and 4). Box 4.1 lists the characterizations of 5 MeSH terms that stood out as particularly relevant. While for an exhaustive review it might perhaps be appropriate to screen all results using these MeSH terms individually or in combination, as they are likely to have a high degree of sensitivity, the strategy would be poor in terms of specificity, as it would identify a far too wide range of hits. For example, apart from “Health Care Surveys” and “Public Opinion” each term individually identifies more than 100,000 sources.<sup>3</sup> The MeSH terms resulting from the index mapping exercise were therefore combined with relevant title keywords in five cluster models.

---

<sup>3</sup> Health Surveys: 314,766; Attitude of Health Personnel: 100,287, in combination: 3195.

**Box 4.1: MeSH terms for literature on empirical work relevant for questions around personal responsibility**

Health Surveys: “A systematic collection of factual data pertaining to health and disease in a human population within a given geographic area”

Health Care Surveys: “Statistical measures of utilization and other aspects of the provision of health care services including hospitalization and ambulatory care.”

Questionnaires: “Predetermined sets of questions used to collect data—clinical data, social status, occupational group, etc. The term is often applied to a self-completed survey instrument”

Public Opinion: “The attitude of a significant portion of a population toward any given proposition, based upon a measurable amount of factual evidence, and involving some degree of reflection, analysis, and reasoning”

Attitude of Health Personnel: “Attitudes of personnel toward their patients, other professionals, toward the medical care system, etc.”

Attitude to Health: “Public attitudes toward health, disease, and the medical care system.”

Note: while a MeSH term “social responsibility” exists (“The obligations and accountability assumed in carrying out actions or ideas on behalf of others”), this is the only term identified in the MeSH browser when searching for “responsibility”. There is no MeSH term “personal responsibility”, or “individual responsibility”, see: <http://www.nlm.nih.gov/mesh/mbinfo.html> (accessed Sept 2009)

Based on familiarity with the conceptual literature and informal consultation with colleagues working in the field, it seemed that relevant studies could be identified by using “responsibility”, “incentive”, “priority setting”, “rationing”, and “allocation” as title key words in combination with the MeSH terms. It is plausible to assume that the first two terms would occur in a publication that presented findings from work concerned with personal responsibility and/or incentive use, and equally, since personal responsibility is often discussed as one of several criteria in rationing, that publications on this topic might include findings on views of the public or physicians on personal responsibility or incentives.

The search was conducted from 2–10 December 2010, using the PubMed online search tool in Endnote X3, limiting results to the last 15 years, i.e. 1995–2010. As Table 4.1 shows, the cluster models produced a total of 1,190 hits. After the first screen (titles only) 184 remained, after the second screen (review of abstracts) 47 were identified as potentially relevant. These 47 publications (which include the previously known publications) were retrieved, and after a third screen (full text analysis) 18 remained.<sup>4</sup>

<sup>4</sup> All searches are available as archived Endnote libraries.

<b>Table 4.1: Quantitative findings from literature search on empirical work relevant for questions around personal responsibility</b>				
<b>Title key word/MeSH combination</b>	<b>Hits</b>	<b>1st screen</b>	<b>2nd screen</b>	<b>3rd screen</b>
Title: Responsibility [and MeSH]				
- Health Surveys	46	11		
- Health Care Surveys	6	1		
- Questionnaires	116	7		
- Public Opinion	18	5		
- Attitude of Health Personnel	157	19		
- Attitude to Health	215	22		
Total	558	55		
Title: Incentive [and MeSH]				
- Health Surveys	12	1		
- Health Care Surveys	13	0		
- Questionnaires	40	2		
- Public Opinion	1	1		
- Attitude of Health Personnel	13	0		
- Attitude to Health	41	18		
Total	100	22		
Title: Priority setting [and MeSH]				
- Health Surveys	12	5		
- Health Care Surveys	3	0		
- Questionnaires	22	6		
- Public Opinion	15	13		
- Attitude of Health Personnel	22	5		
- Attitude to Health	19	0		
Total	93	29		
Title: rationing [and MeSH]				
- Health Surveys	14	6		
- Health Care Surveys	5	0		
- Questionnaires	19	8		
- Public Opinion	32	10		
- Attitude of Health Personnel	39	5		
- Attitude to Health	61	17		
Total	170	46		
Title: allocation [and MeSH]				
- Health Surveys	91	4		
- Health Care Surveys	13	2		
- Questionnaires	57	4		
- Public Opinion	20	9		
- Attitude of Health Personnel	31	5		
- Attitude to Health	57	8		
Total	269	32		
Grand total overall	1,190			
Grand total after screening title		184		
Grand total after screening abstract			47	
Grand total after reading paper				18

The relevance assessment (stage 5 in the above approach) for all three screens drew on the following predetermined inclusion and exclusion criteria: publications should:

- (a) provide data through interviews, focus groups, surveys or polls;
- (b) present an analysis of views of physicians, the public or patients, with patients needing to suffer from conditions relevant to those considered here, i.e. cancer and obesity-related diseases, or ones raising similar issues (such as alcohol-associated diseases<sup>5</sup>);
- (c) be conducted in a developed or high-income country;
- (d) have at least one question on either personal responsibility or incentive use, where surveys or interviews are primarily concerned with broader rationing or resource allocation questions; and
- (e) studies concerned with intensive care situations were excluded, because the situation was viewed as too dissimilar to the types of situations considered here.

#### 4.2.2 Findings

In analysing the full text versions of the articles it became clear that only 18 were accessible, relevant, and presented new and original research.<sup>6</sup> None of the remaining studies were concerned with questions that could be seen as overlapping directly with the research planned here, i.e., there was no interview or survey work that would have examined attitudes of the public or physicians towards incentives to promote weight loss, or incentives for promoting participation in cancer screening and compliance in treatment. Four of the 18 studies addressed responsibility in the context of organ transplantation; the remainder considered it as part of broader questions around resource allocation. Six studies focussed on analysing views in Germany, five on the UK, two compared various countries, two examined the US, and one each concentrated on Norway, Finland, and Australia. Thirteen studies related to views of the public, one to those of patients,<sup>7</sup> two to physician attitudes, and one with a

---

<sup>5</sup> As will be explained below, there are however some significant differences regarding normative responsibility assessments in the context of organ transplantation as opposed to obesity policy. Therefore, only a subsection of transplant-related papers that were particularly relevant was included in the review here.

<sup>6</sup> The remainder turning out to be concerned with broader ration issues or specific responsibilities of physicians (14); more narrow aspects arising in the context of organ donation (4); evaluations of survey or engagement methods (3); presenting preliminary findings of research for which final reports were also available (2); not accessible (2); or in commentary form (1).

<sup>7</sup> Note that a several papers that were identified in the first screen (title only) also concerned patient views, and while it was clear that the papers concerned interview or survey work, the diseases in question were viewed as too dissimilar to warrant further inclusion, and were therefore excluded (2 studies on Alzheimer, 2 on HIV, 1 on rheumatoid arthritis, 1 on asthma, and 1 on schizophrenia). Other



combination of aforementioned. In the following, three summaries will be provided of themes emerging from, first studies focussing on organ transplantation, second, on views of members of the public and patients, and third, of physicians. The second section also includes three particularly relevant studies that have been published after completion of the review, one (by Judith Long, Marie Helweg-Larsen and Kevin Volpp) presenting survey findings from patients in waiting rooms in two US primary care clinics; another relating to a representative sample of the US population (by Sarah Gollust and Julie Lynch) and a third comparing attitudes to incentives in a more experimental setting of US and UK residents (Marianne Promberger, Rebecca Brown, Richard Ashcroft and Theresa Marteau), bringing the total to 21.

#### 4.2.2.1 Responsibility and organ transplantation

The case of organ transplantation differs in significant ways from the conditions that form the focus of the research intended here, but also has some relevant parallels, which is why it seemed appropriate to discuss this area separately. The salient difference between discussions about personal responsibility in the case of organ donation, on the one hand, and responsibility for body weight or participation in screening and complying in treatment, on the other hand, is that in the case of organs, the resource that is to be distributed, viz., transplants, are scarce to the extent that many people on waiting lists will die while waiting for an organ. Some people require an organ because of natural pathologies that are independent of their behaviour, whereas in the case of others, behaviour may have played a significant role. Thus, there are some people on waiting lists who had some opportunity to change their behaviour, but failed to do so. Had they behaved differently, they might not have required a donor organ. Moreover, they would not have come in a situation where they might deprive a person who equally needs a donor organ of receiving it, in cases where they are given preference over people who had no way whatsoever of avoiding their need. While in some case this merely means that they have a longer wait because

---

papers excluded at that stage were 6 that focused on developing country, or specific minority group settings; 22 that were on topic, but not surveys or interview based; 41 that fitted methodology-wise but were clearly not concerned with personal responsibility or incentives (mainly general rationing issues, or using rationing as an example to study certain methodological issues); 4 that concerned the use of incentives at the physician or institutional, but not individual level; 14 that assessed the clinical or otherwise scientific (in the narrow sense) effectiveness of particular incentive programmes; and 39 that were either off-topic, inaccessible, clearly irrelevant, or in a language other than English, French or German.

others failed to behave in ways that are regarded as responsible at an earlier stage, in other cases they may die. Such consequences generally do not materialise in the same way when it comes to failing to maintain a healthy body weight, participating in screenings, or complying with treatment: here, the costs are primarily financial. Even though opportunity costs may result due to attention of health care staff being unavailable for people who require treatment because of events that are wholly beyond their control, this will hardly ever lead to death.

This preface is necessary to understand that although findings on normative assessments of responsibility for smoking or alcohol consumption leading to the need for an organ transplant share some similarities with responsibility assessments regarding excessive food intake (or insufficient energy expenditure), people may feel more strongly about responsibility for transplant needs because of the more immediate life or death consequences for people who had no opportunity whatsoever to influence their need. It is not surprising, then, that Peter Ubel and colleagues found in a questionnaire-based study of 407 prospective jurors in the US that subjects allocated significantly fewer than half of the organs to those with unhealthy behaviours and worse prognoses, and that significantly fewer organs were allocated to patients viewed as responsible for causing their diseases<sup>8</sup> (with subjects who never smoked being the most likely to follow this pattern, Ubel et al. 2001). Equally, Georg Schomerus and colleagues found in a telephone survey involving the adult German population (randomised sample, weighed results,  $n=1,012$ ) that out of nine conditions alcoholism was considered to be particularly “self-inflicted”, evoking a high desire for social distance, and being clearly associated negatively with resource allocation decisions<sup>9</sup> (Schomerus, Matschinger, and Angermeyer 2006). Eve Wittenberg and colleagues,

---

<sup>8</sup> Subjects were presented with four scenarios and asked to distribute 100 transplantable organs among two groups of 100 patients each. In each scenario, one group of patients, but not the other, was described as having a history of unhealthy behaviour (alcohol or cigarette use) associated with a poorer prognosis. In some scenarios, alcohol or cigarette use was said to cause the organ failure. In others, it only contributed to the patients’ transplant prognosis.

<sup>9</sup> Participants were asked to name three out of nine conditions (Cancer, myocardial infarction, AIDS, diabetes, Alzheimer’s disease, rheumatism, depression, alcoholism, schizophrenia) for which they would prefer resources not to be cut should general cutbacks within the health care budget be necessary. For all conditions respondents were asked about personal attitudes and illness beliefs. The authors note that the “perceived personal responsibility was highest with alcoholism compared to all other diseases. More than 8 out of 10 respondents rated alcoholism with ‘4’ or ‘5’ on a five-point Likert scale with the extreme ‘depends very much on oneself’, followed by AIDS, where in a similar manner seven out of 10 respondents held sufferers responsible for their illness.” (Schomerus, Matschinger, and Angermeyer 2006: 207)

who administered a questionnaire with two scenarios relating to liver disease and asthma to a cross-sectional sample of 800 randomly selected US residents (response rate 43%) also found that

respondents who believed that alcohol-induced liver disease and asthma caused by in-home air pollution were the result of personal responsibility allocated fewer treatment resources to these patients than to the patients with inherited liver disease and those with asthma related to outdoor air pollution. (Wittenberg et al. 2003: 197)

The two principal reasons for allocating livers to patients with inherited disease were “alcohol consumption is voluntary” (17%) and “alcoholics are responsible for their illness” (15%, *ibid.*: 200–201).

However, a recent systematic review of studies that explored community preferences for solid organ (heart, lung, liver, and kidney) allocation by Allison Tong and colleagues was less certain that assessments of personal responsibility outweighed other considerations in allocation decisions. The research team analysed fifteen studies involving more than 5,563 respondents (including Ubel et al. 2001, but not including Wittenberg et al. 2003, or Schomerus et al. 2006<sup>10</sup>), and identified seven themes that could describe community preferences for organ allocation. These were maximum benefit; survival and quality of life; social valuation; moral deservingness; “the ‘worthiness’ of recipients based on their social standing and lifestyle decisions”; prejudice; ‘fair innings’; ‘first come, first served’; and medical urgency. The authors conclude that “[d]espite the emergence of these general themes, there was no evidence of an emerging consensus with evidence rather suggesting conflicting preferences between individuals” (Tong et al. 2010: 803), although it might be said that the discussion of the relationship between the themes could have been made more explicit. The authors also observe—similar to Strech et al.’s above-noted commentary—that generalizations across studies may be complicated by differences in underlying methodologies and theoretical frameworks (*ibid.*).

For the purposes here it is not necessary to adjudicate the question of the salience of the criterion of personal responsibility in the case of organ transplants over other criteria. But it can be noted that it certainly is one that has particular relevance at the sharp end of rationing decisions, drawing attention to the importance of assessing the

---

<sup>10</sup> Regarding the reason for why the remaining papers were not considered here, see footnote 4.

extent to which a need can in fact be said to have been avoidable, which requires an assessment of the opportunities of choice that people concerned have had. Equally, the case of organ transplants highlights that the magnitude of the ‘cost’ of failing to behave in ways that are regarded as responsible requires scrutiny. For while it might be argued that the consequence of people dying gives strong reasons to appeal to personal responsibility (in whatever way: verbally, or through some sort of sanction), there might be less reason to do so if the consequences are less significant. The next section will shed some light on the relationship between responsibility and other goals.

#### 4.2.2.2 Responsibility and views of members of the public and patients

As attitudes around responsibility are often influenced by developments in politics and policy, and since recent years have seen an increased focus on responsibility, the summary of studies concerned with views of the public and patients will proceed backwards in time, beginning with the most recent work, citing other work where fitting. This approach also has the advantage that it begins with some of the most relevant studies for the questions that are of interest here.

Marianne Promberger, Rebecca Brown, Richard Ashcroft and Theresa Marteau focused on the acceptability of financial incentives and examined attitudes through an analysis of data from two non-probability-based convenience samples, with respondents being members of online access panels in the UK and US, n=88, 100, (Promberger et al. 2011). The team asked respondents to evaluate the acceptability and fairness of using medical interventions (pills and injections) versus financial rewards and penalties (levels of rewards or penalties were not specified). Respondents were told to assume that medical and financial interventions were equally effective, and asked to state which type of intervention should be funded by a health care system in relation to five health-related behaviours (weight loss, smoking cessation, and adherence to treatment programmes regarding drug addiction, serious mental illness, and post-operative physiotherapy). While only the case of weight loss is directly relevant for the research planned here, it is nonetheless noteworthy that the authors conclude from their finding that across all five conditions, and in both the US and the UK samples, “financial incentives, whether rewards or penalties, are judged as less acceptable than medical interventions” (Promberger et al. 2011: 4). They hypothesise that the reasons may be that financial incentives constitute a violation of a cultural

norm, or that a sense of injustice may lead people to find it inappropriate to offer financial advantages to people who could have avoided poor health through behaviour. For weight loss, participants in both samples overall favoured pills and injections over rewards and penalties, but were most opposed to the latter, preferring reward schemes. Respondents in both samples also strongly agreed that smokers were responsible for the consequences of their behaviour, whereas assessments were less clear-cut with regard to overweight.

A study by Sarah Gollust and Julie Lynch concentrated on the US and examined the extent to which cues about a person's race, income, gender or responsibility for poor health mattered in relation to the perceived deservingness of care (Gollust and Lynch 2011). The authors also used an online access panel, albeit a larger one, with the sample being representative of the US population (two waves with a total of three vignettes were fielded with  $n=1,334$  completing both surveys). With regard to health, the authors note that respondents were more likely to blame diabetes patients for their condition if respondents were provided cues that patients had poor diet and exercise habits than when they were said to have had a family history of the disease. However, less blame was attributed when patients were characterised as working class, compared to middle class, with no difference in this pattern between higher income and lower income respondents. Equally, both groups agreed that poorer patients who suffered diabetes should shoulder a smaller burden of health care cost.

Judith Long, Marie Helweg-Larsen and Kevin Volpp fielded self-administered surveys to a convenience sample of 515 patients in waiting rooms of two university-based primary care centres in Philadelphia, USA, focussing on incentives to quit smoking, lose weight, and control blood pressure, diabetes and cholesterol. Overall, acceptance of using incentives was found to be equivocal, with 36–42% agreeing that incentives were a good/excellent idea, and 41–44% finding that it was a bad/very bad idea. 'Carrots' were preferred over 'sticks' for smoking and weight loss (smoking 67% vs. 54%, weight loss: 43% vs. 32%).<sup>11</sup> With regard to blame attributions, 76–79% of people disagreed with the statement that people with hypertension or diabetes had themselves to blame for poor health, with markedly less in the case of obesity

---

<sup>11</sup> No data on effectiveness was given for the other conditions.

(55%), and smoking (24%). Thirty per cent agreed that paying smokers to quit maybe one of the only effective ways to achieve behaviour change (smokers showing statistically significant higher odds) with 47% disagreeing.<sup>12</sup> Sixty-two per cent found it acceptable if companies offered incentives for health improvement, but 53% agreed that people should not be paid for behaviour they should display anyway. Regarding appropriate levels of incentives, respondents were presented with three brackets, \$0, \$50–500 and >\$1,000 for each of the four conditions. There was little variation across these, and 51–53% opted for \$0, 35–40% for the middle bracket and 10–13% for the highest; the median was \$0 for all four conditions. Acceptable levels for incentives in the form of ‘sticks’ were not assessed (Long, Helweg-Larsen, and Volpp 2008).

Adele Diederich, Margrit Schreier and colleagues carried out a number of different interviews and survey studies as part of a larger interdisciplinary research collaboration on priority setting in health care, funded by the German Research Foundation.<sup>13</sup> In a nationwide representative survey of 2,032 Germans, 34 questions in ten clusters that concerned different aspects of rationing-related issues were posed. One of the clusters addressed health behaviour and included three questions on higher co-payments where skin cancer may be correlated with prior solarium use; where other potentially risky behaviours are given, or where there is lack in compliance in treatment, see Table 4.2.

---

<sup>12</sup> No data on effectiveness was given for the other conditions.

<sup>13</sup> Priorisierung in der Medizin – DFG-Forschergruppe FOR655, see: <http://www.priorisierung-in-der-medizin.de/> (accessed 16 Dec 2010)

**Table 4.2: Health behaviour attitudes/Germany: findings from a representative survey (Diederich/Schreier 2010)\******Solarium use***

To what extent do you agree with the following statement: ‘people who frequently use a solarium and subsequently suffer from skin cancer should pay a proportion of the cost of treatment’

Agree completely	Tend to agree	Tend not to agree	Disagree completely	Don't know	Answer refused
45.7	26.2	13.2	12.9	1.4	0.6

***Importance of potentially risky behaviours***

There are many behaviours that have a negative effect on health and can increase the risk of illness. In your view, in which cases should patients incur higher co-payments?

	Yes	No	Don't know	Answer refused
Consuming unhealthy foods	45.9	45.9	7.2	1.0
High use of alcohol	70.9	25.6	3.1	0.4
Smoking	67.8	28.7	3.1	0.4
Adventure sports	74.2	23.3	2.3	0.1
Sunbathing/solarium use	65.0	31.1	3.5	0.3
Use of illicit drugs (e.g. heroin)	76.4	21.1	2.0	0.5
Lack of exercise	38.1	54.0	7.2	0.7
Other	5.3	11.2	74.9	8.6

***Compliance***

Patients don't always comply with treatment. However, if, for example, antibiotics are not taken as prescribed they are likely not to be effective, and treatment may be unsuccessful. Do you think that patient who are known not to comply should pay some of the cost of treatment themselves?

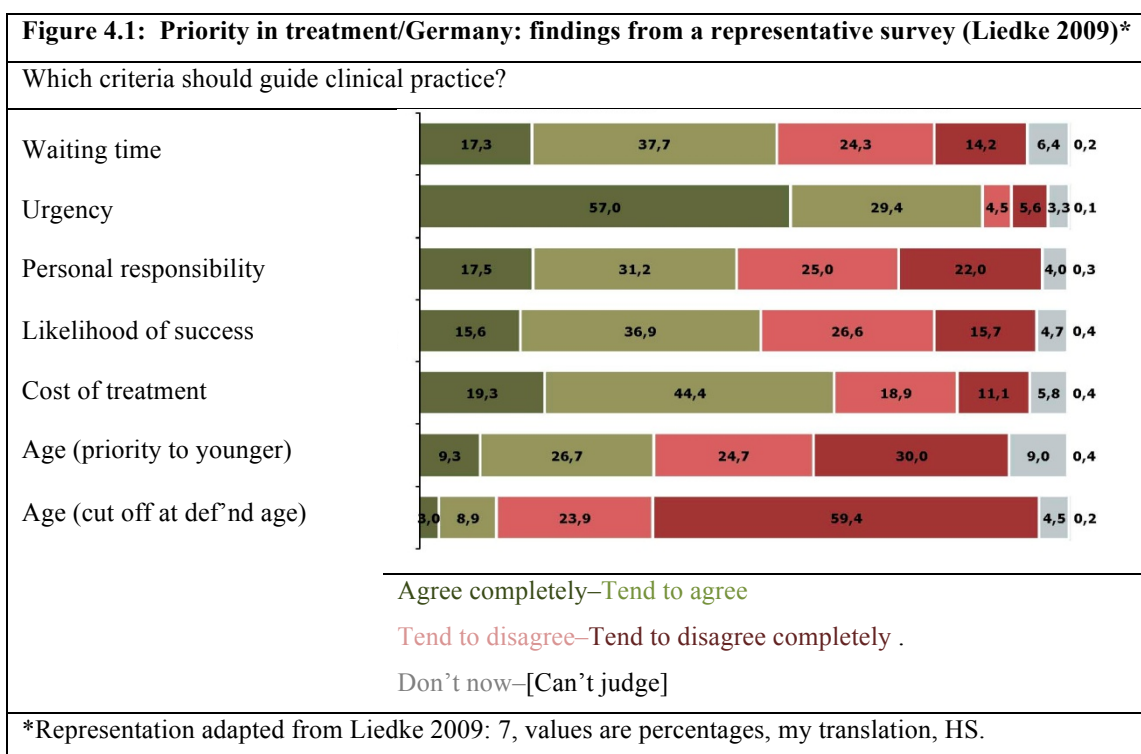
Yes	No	Don't know	Answer refused
72.4	23.0	4.2	0.3

\*Representation adapted from Diederich and Schreier 2010: 22–23, values are percentages, my translation, HS.

As Diederich and Schreier note in the discussion of the findings, there appears to be broad support for implementing some form of cost-shifting for a number of behaviours that may lead to health care needs, with the exception of ‘lack of exercise’, which 54% reject as a criterion, and ‘consuming unhealthy foods’. Regarding the latter, 45.9% are in favour, while the exact same number of people is opposed, with the percentage of people responding ‘don't know’ or ‘refuse to answer’ the highest, in comparison to the other items on the list (Diederich and Schreier 2010: 23-24). Non-compliance in treatment as a criterion for higher health care cost also receives much endorsement, scoring the third highest approval ranking at 72.2% (with extreme sports at 74.2%, and use of illicit drugs at 76.4%). Note, however, that respondents were not provided with figures for the level of co-payments, which, it might be argued, would be likely to affect their responses. It is also noteworthy that in earlier qualitative work using semi-structured interviews that were carried out in preparation of the survey, 9 of

45 respondents pointed out that the acceptability of higher co-payment depended on their level, even though 21 respondents agreed to the general principle that smokers should incur a higher burden, as they has caused their health care needs (Diederich and Schreier 2010: 899).

Posteriorisation, or assigning lower priority, can take different forms. One approach, as envisaged by Diederich and colleagues, focuses on financial disadvantage. Another question is what role personal responsibility might play in assigning priority in clinical practice, which was explored almost concurrently in research commissioned by the Allianz Deutschland AG in 2009 in the form of a telephone survey among randomly selected German adults ( $n=1,039$ ). In the survey, respondents were asked to rank seven criteria that might determine priority in treatment, as represented in Figure 4.1.



Given that 48.7% of respondents are sympathetic to giving people who bear some responsibility for their actions lower priority in treatment, but almost as many (47%) are opposed, Annika Liedke comments in her presentation of the findings that posteriorisation would not be straightforward, speculating that those opposed to using the criterion based their concerns on fairness grounds (Liedtke 2009: 7).



In a questionnaire based survey conducted in Australia in 2005, Erik Nord and colleagues sought to assess the extent to which there is support for drawing on the Quality Adjusted Life Year (QALY) approach in allocating health care resources. 2,000 questionnaires were distributed of which 551 were returned. One question concerned the role of smoking and asked:

Which of the following do you agree with?

1. If there is not enough money to treat everybody with heart disease or lung cancer, non-smokers should have some priority over smokers.
2. Smokers should have the same priority with respect to treatment for heart disease and lung cancer as everybody else.

In response, 59.5% believed that non-smokers should be prioritised, with the remainder opting against. The researchers point out that the bias was not explainable “because [respondents] thought outcomes tend to be better in non-smokers than in smokers [but instead] the great majority blamed smokers for self inflicted conditions” (Nord 2006: 434).

In the UK, Ann Bowling analysed a question on the role of personal responsibility in a survey on general rationing issues conducted in 2000 on a representative sample of UK residents, who participated in 2,005 face to face interviews. In response to the statement “People who contribute to their own illness, for example, through smoking, obesity, or excessive drinking, should have lower priority for their healthcare than others” the following responses were received: Strongly disagree (10%); disagree (33%), neither disagree nor agree (15%); agree (33%) strongly agree (9%); (Bowling 1996: 12). Here 42% were in favour of responsibility as a criterion for lower priority, while 43% were opposed, a result broadly similar to the findings Liedke reported for Germany, as noted above.

In a similar vein, earlier research by Herbert Matschinger and Matthias Angermeyer suggested that factors such as the urgency or severity of disease could rank higher than personal responsibility when it comes to funding priorities. The researchers analysed 5,025 interviews of respondents constituting a representative sample of the German population and asked them to select three out of nine conditions for which available resources should on no account be shortened. The vast majority (89%) selected cancer, almost half of the respondents selected HIV/AIDS (51%) and cardiovascular diseases (49%). This was followed by diabetes (33%) and Alzheimer’s

disease (28%); next was rheumatism (19%), schizophrenia (10%), depression (7%), and alcoholism (6%). While the study's clearest finding is that respondents differentiate between priority for the medically (in the narrow sense) ill as opposed to the mentally ill, the extent to which personal responsibility might play a role in health care needs is not as well aligned (compare, for example, the melanoma scenario explored by Diederich and Schreier above, and equally responsibility is attributed by many with regard to HIV/AIDS and cardiovascular diseases, making up the top three priorities in which no cuts should be made (Matschinger and Angermeyer 2004).

Three studies analysed on data from the 1998 *Eurobarometer*. Derek King and Alan Maynard drew on the UK segment, which was based on a representative sample of 1,055 respondents. Personal responsibility issues were explored in one of eight questions with relevance for rationing, and similar to the examples considered above, respondents were asked to rank "lifestyle" in relation to other factors regarding priority in treatment. King and Maynard presented the findings, but do not discuss them further, see Figure 4.5 (King and Maynard 1999). Reinhard Busse examined the *Eurobarometer* findings for Germany, which were in this case based on a representative sample of 2,000 respondents, see Table 4.3.

<b>Table 4.3: "Lifestyle" in relation to other factors regarding priority in treatment: findings from the 1998 Eurobarometer survey*</b>						
"Which factor that a doctor or other health professional might take into account when deciding which patients should be given priority for treatment do you consider to be the most important?"						
	Germany	UK	France	Italy	Netherlands	Sweden
Waiting time	32.4	36.4	21.3	32.3	39.7	45.0
Each case should be treated individually**	18.8	17.1	40.3	38.6	14.8	18.5
Family commitments	13.6	14.1	11.4	4.8	9.5	12.2
Treatment outcomes	12.7	14.0	9.4	11.9	16.0	9.0
Lifestyle	8.3	9.1	5.4	2.2	11.7	3.5
Age	4.1	6.9	8.0	5.5	1.9	6.4
Don't know	10.0	2.4	4.2	4.8	6.4	5.4
*Representation adapted from: (Busse 1999: 86; King and Maynard 1999: 46; Mossialos and King 1999: 100), values are percentages.						
** Spontaneous statement by respondents, not on initial list of options.						

Although, due to the different methods of ranking and different wording and substantive content of items, a comparison to the survey findings presented by Liedke above are not straightforward, there appears to be far less support for considering personal responsibility related factors in this (earlier) study. Contextualising these

findings in prior national survey work, and giving more meaning to what might meant by the notion of ‘giving priority’, Busse notes, regarding coverage by sickness funds:

In a 1993 Forsa survey, 55% of the 1,005 persons interviewed were of the opinion that sickness funds should pay for ‘everything’ while 41% thought that they should not cover certain diseases: smoking-related diseases 32%; alcohol-related diseases 28%; injuries through risky sports 26%; drug abuse 23%; abortion 11%; stress-induced diseases 3%; pregnancy 1%. Similarly, Polis’s 1995 survey in Northrhine-Westfalia showed that 23% thought that self-induced diseases should not be paid by the sickness funds. Seventy-five per cent said that raising sickness fund contributions could be avoided through increased efficiency in the health care system. In Forsa’s 1995 survey, 41% favoured the inclusion of health risks in the calculation of sickness fund benefits, mainly through bonuses for healthy lifestyle (29%) and less frequently through extra contributions for persons with risky behaviours (7%). (Busse 1999: 83)

It is noteworthy that despite a significant spread, these surveys indicate stronger support for implementing responsibility in practice than the *Eurobarometer* findings, with the 1995 Forsa study suggesting four times as much support for using ‘carrot’ approaches, as opposed to ‘sticks’. Elias Mossialos and Derek King examined the *Eurobarometer* findings across the six countries represented in Figure 4.4., but, as in King’s publication with Maynard, did not consider the “lifestyle” question in more detail in the description of the findings, nor in the discussion section, focussing instead on the remaining criteria. In part, this was a consequence of the paper’s main focus, in part it may be to do with an earlier discussion in the paper regarding the relevance of the “lifestyle” criterion. The authors note that in “the countries where attempts at priority setting have been made by government committees, however, lifestyle has generally been ruled out as a consideration determining access to services, though it may be taken into account in assessing outcomes” (Mossialos and King 1999: 90). As Chapter 2 here has shown, this practice has certainly changed since 1998. After a brief review of some of the survey literature in different European countries the authors also concluded that: “[o]verall, there appears to be little consensus on the relevance of lifestyle to priority setting decisions. In fact, most of the evidence suggests a sharp split in opinion between those who support the inclusion of lifestyle as a priority setting criterion and those who do not.” (Mossialos and King 1999: 92)

Such sharp differences of opinion may partly be the result of deeper implicit or explicit normative positions; in part, they may also reflect that the issues raised are highly complex, and that attitudes may change, after a period of deliberation. Paul

Dolan and colleagues sought to assess this hypothesis through research in which groups of people would discuss relevant issues, and any changes in the attitudes over time would be monitored. A total of 60 people met in ten groups of five to seven people on two occasions, separated by a fortnight. Participants were randomly chosen patients, recruited via two general practice lists in York. All discussions were moderated by the same two researchers, who mainly sought to ensure that everyone had an opportunity to be heard. A questionnaire was filled in by all participants in the first meeting, to record initial assessments, and another questionnaire was filled in at the second meeting, to record changes that were likely to be the results of discussion with others, or responses to a set of trade-off scenarios and other debate or information materials which the researchers presented. Table 4.4 shows the changes in attitudes.

Group features	Initial response assessment		Final response	
	Lower priority	Higher priority	Lower priority	Higher priority
Children	70 (42)	2(1)	2(1)	62 (37)
Illegal drug users	57 (34)	8(5)	43 (26)**	3(2)
Smokers	57 (34)	0	32(19)**	0
Heavy drinkers	55 (33)	3(2)	37 (19)**	2(1)
Disabled	0	48(29)	0	35(21)
Elderly	7(4)	47 (28)	5(3)*	27 (16)
Private health insurance	30 (18)	2(1)	33 (20)	2(1)
With children	2(1)	30 (18)	3(2)	20 (12)
Unhealthy diet	23 (14)	7(4)	17 (10)	5(3)
Rich	20 (12)	0	23(14)	0
Rarely exercise	17 (10)	0	20(12)	0
Poor	0	17(10)	0	10(6)
Low education	0	15(9)	2(1)	8(5)
Homosexual	13 (8)	0	10(6)	0
Important	13 (8)	0	5(3)	0
Women	0	7(4)	0	3(2)
Married	0	5(3)	2(1)	0
Contributed a lot	2(1)	3(2)	2(1)	2(1)
Unemployed	0	3(2)	0	3(2)
White	0	0	0	0
Men	0	0	0	3(2)

\*Representation adapted from: (Dolan, Cookson, and Ferguson 1999: 918. Values are percentages (numbers) of 60 respondents; responses of remaining respondents were that the same priority should be given to that group.

\*\* P<0.01 compared with initial responses.

For the issues considered here, it is noteworthy that there is a clear shift in attitudes regarding the assessment of smokers, heavy drinkers, and illicit drug users. While on the initial assessment more than 55% wished to assign them lower priority,

on the second assessment this dropped to between 32 and 47%. Dolan and colleagues conclude from this that

the public's views about setting priorities in health care are systematically different when they have been given an opportunity to discuss the issues. If the considered opinions of the general public are required, surveys that do not allow respondents time or opportunity for reflection may be of doubtful value (Dolan, Cookson, and Ferguson 1999: 916).

The more fundamental critique of using survey data that the authors present will be considered again in the discussion section below. For now, it can be summarised that, although as already noted above, it is far from straightforward to aggregate findings from survey research across studies in a robust manner, it is clear that a range of different surveys carried out over more than a decade demonstrated considerable support for considering personal responsibility, in particular in relation to the somewhat general, and to some extent unspecific, question of whether people who drink excessively, exercise insufficiently, and so on, should be given 'lower priority' in the context of clinical practice, treatment, or cost-sharing. In all studies the concepts of 'causality'/'causal responsibility', 'voluntariness', 'opportunity of choice' arise as relevant in seeking explanations for the findings, although the data are generally insufficient to surmise what assumptions respondents made in their assessments.

#### 4.2.2.3 Responsibility and views of physicians

The attitudes of physicians towards patient responsibility have been studied far less extensively than those of the public and patients, and only three surveys were identified here. Piija Jallinoja and colleagues focussed less on health policy and fairness issues, but sought to address physicians' and nurses' views on patient and professional roles in the management of "lifestyle-related" diseases such as adult obesity, dyslipidemia, high blood pressure, type 2 diabetes, and smoking. Using a structured written questionnaire, the researchers surveyed 220 Finnish professionals in primary care (Jallinoja et al. 2007). Among other things they found that:

[a] majority of physicians (88%) and nurses (95%) agreed that patients themselves must accept the responsibility for lifestyle-related decisions. In respect of all the conditions presented, a majority of physicians and nurses considered that patients' unwillingness to change is always or nearly always a key barrier to treatment. Patients' insufficient knowledge of the risk of the condition was much more seldom regarded as a barrier; an opinion most pronounced in respect of adult obesity and smoking (Jallinoja et al. 2007; 246).

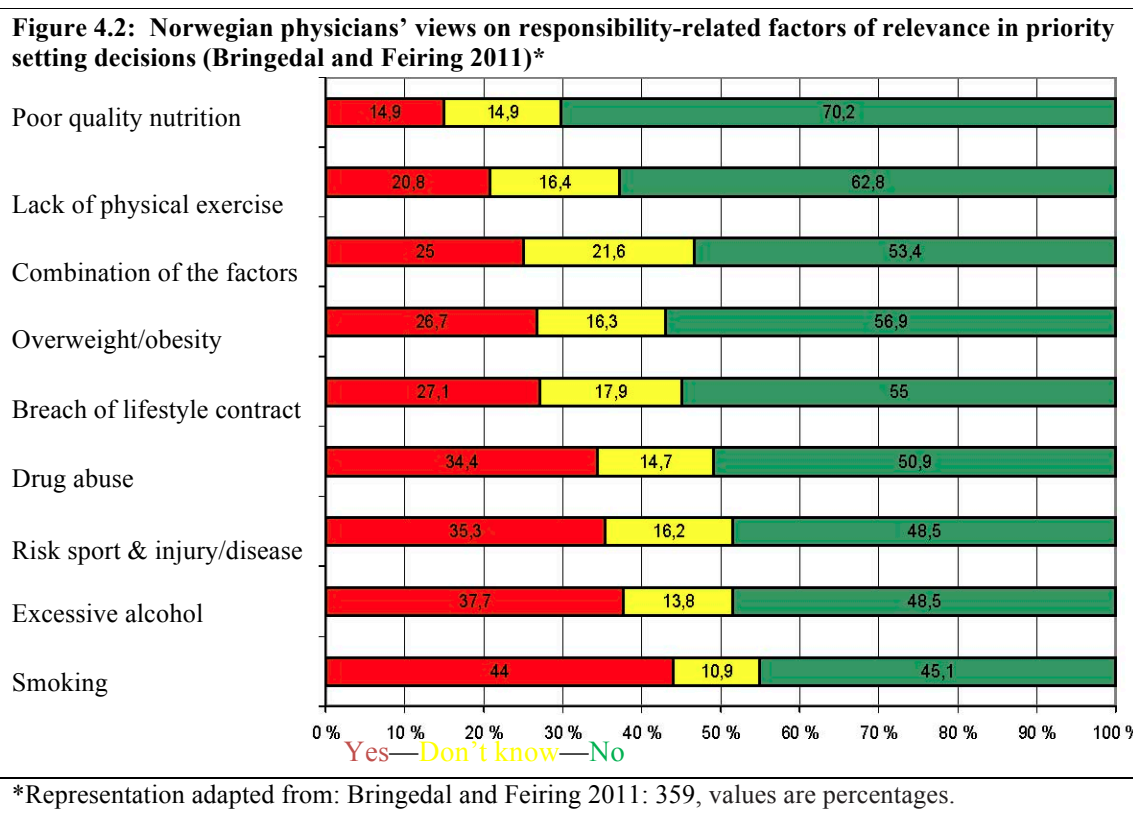
Later the authors noted that both physicians and nurses were faced with a dilemma regarding the role of patients' behaviours in the treatment of diseases such as obesity and smoking, as the patient's cooperation was regarded as central in disease management but at the same time often seen as a major potential barrier to treatment:

The primary care physicians and nurses studied here hold a view that while patients are responsible for their lifestyle change, for the majority they are not able to act in their best interests. The dilemma of this situation might be frustrating for professionals, and it might also lead some to neglect lifestyle counselling as useless and rely predominantly on pharmacotherapy in the treatment of lifestyle-related conditions (Jallinoja et al. 2007: 248).

Berit Bringedal and Eli Feiring focussed centrally on rationing-related issues.

Drawing on the 2008 Norwegian Medical Doctor Survey, which is composed of a representative sample of 1,650 Norwegian practicing medical doctors, data were collected using a cross-sectional questionnaire, with 1,072 responses being received. The authors sought to address three principal questions: (1) Do Norwegian medical doctors find personal responsibility relevant to priority setting decisions?; (2) Which dimensions of responsibility do the respondents find most relevant?; (3) Can different views between subgroups of doctors be traced to priority setting experience? Respondents were asked to assess a set of general propositions regarding the role of personal responsibility in priority setting, and to evaluate the importance of different responsibility-related factors that might be of relevance in priority setting decisions, see Table 4.5 and Figure 4.2.

	Partly/compl disagree	Neutral	Partly/compl agree
Healthcare priority should depend on the patient's personal responsibility for the disease	57,0 (602)	25,9 (273)	17,1 (181)
Access to expensive treatment should depend on the patient's personal responsibility for disease	55,6 (586)	25,3 (267)	19,1 (201)
Access to scarce organ transplants should depend on the patient's responsibility for the disease	49,3 (520)	23,8 (251)	26,9 (283)
Lower priority should be allotted to patients who violate a contract of changes in lifestyle	49,7 (523)	26,3 (277)	24,0 (252)
A patient who is responsible for the disease should pay additional co-payments	73,9 (780)	18,6 (196)	7,6 (80)
*Representation adapted from Bringedal and Feiring 2011, values are percentages, n in parentheses.			



In relation to the statements reproduced in Figure 4.2, Bringedal and Feiring highlighted that 43% choose the alternatives ‘partly/ completely agree/indifferent’ in response to the statement “Healthcare priority should depend on the patient’s personal responsibility for the disease”, while 44.4 % chose these alternatives to the statement “Access to expensive treatment should depend on the patient’s personal responsibility for the disease”. They emphasised that these assessments are in conflict with Norwegian legal provisions which state explicitly that only severity of disease, likelihood to benefit, and cost are to be considered. They also found that responsibility ascriptions vary with age and gender, with older doctors and females more reluctant to consider responsibility. There was no significant variation between GPs and other physicians (Bringedal and Feiring 2011).

That doctors have sympathy for considering personal responsibility, even if professional guidance suggests they should not, also became apparent in a UK survey on rationing issues carried out online by the Medical Hospital Doctor Magazine in 2008. 873 physicians participated in total, and 686 responded to a question that sought to assess whether there were conditions in which patients should be refused treatment (Medical Hospital Doctor Magazine 2008), see Table 4.6.

<b>Table 4.6: Attitudes of UK doctors towards refusing treatment (Medical Hospital Doctor Magazine 2008)*</b>		
“In which, if any, of the following scenarios do you believe the patient should be refused clinically indicated treatment on the NHS?”*		
	Favourable Responses (%)	Favourable Responses (count)
An intractable smoker who requires a bypass	48.0	329
An obese patient who requires a hip replacement	25.8	177
An alcoholic who refuses to stop drinking but who requires a liver transplant	93.7	643
A heroin addict who requires maintenance therapy	17.8	122
An injury resulting from a high-risk sport, such as boxing	10.5	72
An elderly patient who is likely to gain benefit from treatment but only for a short period (e.g. a statin for a 95-year-old after an MI)	27.8	191
	Answered question:	686
	Skipped question:	187

\* \*Representation adapted from: Medical Hospital Doctor Magazine 2008, multiple selections possible.

Unsurprisingly, these findings caused considerable media interest, prompting the British Medical Association (BMA) to publish a press release, in which the Chairman of the BMA’s Medical Ethics Committee, Dr Tony Calland stated: “Someone’s age or lifestyle choices should not impact on their eligibility for NHS treatment and the BMA would be against any such moves. The BMA is against blanket bans based on age or other arbitrary factors” (British Medical Association 2008).

While survey or interview work on physicians’ attitudes towards personal responsibility is in short supply, and does not concern the countries that are intended to be studied here, the data that do exist appear to suggest that physician’s views are not that dissimilar from those of the general population, but rather stand in a continuum. The example of the Finnish physicians shows that the concept of (non-normative) causal responsibility is of importance in every day clinical interactions, while the Norwegian and British surveys showed physicians do not necessarily shy away from making value judgements. As in the case of the surveys of the public and patients, this raises deeper questions about their assumptions regarding, in particular, the concepts of ‘voluntariness’, and ‘opportunity of choice’.

### **4.3 Discussion of findings and development of a new survey instrument**

The studies summarised in outline here provide a range of interesting insights. At the same time, they are also limited in significant ways, both for the purposes pursued in this thesis, as well as independently of it.



First, as noted, much research asks about conditions or behaviours that may justify giving people “lower priority” or, for example, demanding “higher co-payments”. Yet, none of the studies set out what level of co-payments exactly is envisaged. Apart from the Hospital Doctor Magazine’s survey which asks physicians to state under what conditions patients should be “refused” treatment, studies considering other forms of posteriorisation are opaque about how this would be achieved. Both readers and survey respondents must therefore be unclear about what precisely it should mean to receive less priority in organ allocation or clinical practice. Long et al. asked respondents to determine what levels would be appropriate for ‘carrots’, but the categories were very broad; the incentive programme used in the scenario was not framed in the way they are typically used in practice,<sup>14</sup> and the sample was a relatively small convenience sample. Lack of evidence about the acceptable magnitude of, particularly, penalties matters, as arguably, the results of, for example Diederich’s and Schreier’s interesting survey might look very different if numbers had been attached to the loose descriptor of “higher co-payments”: did the 65% of respondents who thought that solarium users should have higher co-payments have in mind one Euro, ten, 100 or 1,000? We don’t know. Yet, the discussions about responsibility appeals being coercive or unfair hinge to a very significant extent on the magnitude of the financial (or other) burden that people may be asked to bear. It therefore seems desirable to be precise about this aspect in the work developed here.

A further problem has to do with the fact that the focus of the research reviewed here is about examining the conditions under which people might be disadvantaged or penalised, rather than incentivised. As the previous chapters here have shown, to some extent, it is not always easy to discern the difference between ‘carrots’ and ‘sticks’, and often policies that are offered as ‘carrots’ turn out to be ‘sticks’. Yet, with the exception of Long et al. and Promberger et al., most of the work frames choices in such a way that respondents are asked to assess under what conditions ‘sticks’ might be appropriate. But, as the example of the BMA’s response to the Hospital Doctor Magazine’s survey showed, such policies are generally unlikely to be implemented,

---

<sup>14</sup> The authors note: “The survey questions focused on rewards for change in behavior (losing weight), though most existing P4P programs reward absolute achievement (getting weight below some pre-set threshold). We did this because we feel that tying incentives to changes in behavior is more likely to increase the rate of healthy behaviors than if such programs focus on absolute achievement.” (Long, Helweg-Larsen, and Volpp 2008: 1651)

due to political and professional reservations. A similar reaction could be registered when Diederich and Schreier published the initial survey findings in the summer of 2010 in Germany. On the one hand, Marco Wanderwitz, MP, a backbencher of the conservative party, welcomed the findings and argued that they rightly underlined that people who voluntarily live in unhealthy ways would need to shoulder the financial consequences. At the same time, Prof Karl Lauterbach, MP, a health policy expert and ministerial adviser, simply chided the proposal of imposing higher cost as “bonkers”, with several newspapers suggesting that this was a typical proposal for the summer’s ‘silly season’ (Handelsblatt 2010).

Indeed, it would not be imaginable to implement higher co-payments (of whatever level) for the range of activities explored in the survey at this point in Germany. Yet, the aforementioned bonus programmes can easily have a similar penalising effect. Assessing attitudes on the acceptability of these initiatives, which are already in practice, would have been a useful complement to studies asking about the conditions under which of higher co-payments would be acceptable. Promberger et al.’s comparative examination of attitudes towards ‘carrots’ and ‘sticks’, and the finding that respondents in the US and UK samples prefer rewards over penalties is noteworthy, but it would also be interesting to know whether such attitudes vary with the size of the incentive—as noted, one major development in recent US reforms related to a significant increase in the levels of incentives. It therefore seems that the work proposed in this thesis can extend some of the general work already done and make it more specifically relevant for the actual policy context, exploring, in particular, some of the nuances in relation to the framing of ‘carrots’ and ‘sticks’.

This approach should also go some way towards addressing the concerns Dolan and colleagues have regarding the use of survey results. Of course, the conventional form of surveys means that the demand to “allow respondents time or opportunity for reflection” in order to ensure that results are not of “doubtful value” is likely to be overly demanding (Dolan, Cookson, and Ferguson 1999: 916). But in a constructive way the comment can be understood to emphasise that questions be phrased in a way that makes them as concrete as possible, avoiding highly abstract and general scenarios.

In line with this, a first draft instrument was developed, which was then refined further in a series of pre-tests for use in semi-structured interviews with primary care physicians and oncologists (Chapter 5), and in population level surveys (Chapter 6). The initial draft instrument is at Appendix 4A.<sup>15</sup> Below, a brief overview of its structure and the rationales for individual items is provided.

#### *The role of personal responsibility and other factors*

To begin with, a question is asked to assess how important respondents find the role of individual behaviour in relation to other relevant factors that can influence one's health, such as the environment, genetics, access to health care and so on. This has several purposes. First, to alert respondents to the possibility that the scope of individual behaviour may have limitations, and also to see where respondents 'are coming from', as views about the acceptability of rewards and penalties are likely to differ depending on one's general view of the role of individual behaviour. The question also helps to assess empirically whether the conceptual notion of health as co-production (Section 2.3.1, 3.4), that was highlighted in the previous Chapters 2 and 3, has currency among respondents. Since the extent of control that people have over their health was shown to be an important factor in evaluating the fairness of programmes, it will also be instructive to examine the extent to which there are differences among groups with differing income or health status regarding the role of personal behaviour.

#### *The car insurance analogy*

The analogy that people who are unhealthy behave like risky drivers was influential in the recent changes in US policy, but it is not clear to what extent it meets with approval by physicians or the public. The scenario is highly relevant for questions around solidarity, voluntariness, intrusion and opportunity of choice, and therefore central to several elements of the procedural justice framework set out in the previous chapter. The analogy also provides a crisp scenario that can help understand whether there are differences among German and American respondents.

---

<sup>15</sup> Note that the instrument has changed considerably after pre-testing (as was to be expected). It might therefore be more instructive to review the instruments as used for the physician surveys (Appendix 5D and E) and population level surveys (as reproduced in Chapter 6).

*Weigh loss incentive*

A commonly used attainment-incentive is to provide a ‘carrot’ or ‘stick’ for meeting (or not meeting) specific Body Mass Index targets. Using a highly naturalistic scenario and realistic monetary levels of incentives, rather than exploring acceptability in the abstract, can further a better understanding of the acceptability of real-world applications of responsibility policies.

*Appropriate levels of ‘carrots’ and ‘sticks’ for weight loss*

To explore what levels would be acceptable in relation to a typical incentive scenario, respondents are invited to suggest amounts within the scope of the newly increased levels established under US health reform, for both a ‘carrot’ and (a) version(s)<sup>16</sup> of a ‘stick’ scenario.

*Motivation, behaviour change and deservingness in relation to weight loss*

In a range of different scenarios that partly mirror the ‘five groups problem’ (Section 3.5.3) it will be explored to what extent respondents find policies fair in which people: (1) obtain a reward without behaviour change, just because they happen to meet required targets, (a) with their personal motivation and lifestyle aligned with the programme’s goal, (b) with their personal motivation and lifestyle running counter to the programme’s goal; (2) obtain a reward for successful behaviour change, as a result of being motivated by the programme; (3) lose out on a reward as they genuinely tried, but failed to meet targets; (4) lose out on a reward because they object to the programme’s goal. These scenarios again relate to central issues around opportunity of choice and desert, and build, among other things, on the above-cited prior work by Long et al. and Gollust and Lynch.

*Economic status and weigh loss incentives*

The scenario above chiefly relates deservingness and acceptability of incentive rewards to personal effort. However the previous chapter also referred to survey data that suggests that healthier and economically better off people are more likely to benefit from incentive programmes (Section 3.5.3). The economic status of the primary beneficiaries can also matter for the acceptability of programmes and will

---

<sup>16</sup> It was initially envisaged to use only two scenarios, but the population level surveys eventually included three different ones (see Section 6.2.2).

hence be explored in a separate scenario in which it is suggested that twice as many of the better off receive incentives, compared to the worst off.

#### *Colon cancer prevention policy*

The German cancer prevention policy is remarkable on a number of grounds and attracted much criticism due to its ambiguous framing that combines ‘carrot’ and ‘stick’ elements, and the potential to be coercive and penalising, as described in detail in Chapter 2. Yet, there is no data on population-level views. The instrument therefore includes an abstracted scenario, focussing on colon cancer prevention. Since the policy combines two elements (attending counselling and compliance in treatment), individual elements are then explored in separate scenarios. The first of these uses the initially envisaged plan to require undergoing screening (and not just counselling), to avoid a doubling of the co-payment threshold. The second explores whether the newly established, and largely unique, obligation to think about one’s health (in the context of incentivised counselling sessions) meets with approval or opposition.

#### *Framing of incentives*

The discussion in Chapters 2 and 3, and the prior work by Promberger et al. and Long et al. showed that attitudes towards ‘carrots’ and ‘sticks’ can vary, and that it cannot be taken for granted that ‘carrots’ will always be perceived as purely benign offers of assistance. For both the weight loss and the colon cancer prevention policy respondents will therefore be asked to state in what way they view the programmes (as rewards, penalties, or both at the same time).

#### *Impact on the doctor-patient relationship*

Since incentives are not merely transactions between individuals and a faceless health care system, but often mediated by health professionals with the potential to cause tension in the doctor-patient relationship, the instrument includes two scenarios relating to weight loss and colon cancer prevention that explore how likely patients are to request physicians to record an outcome in their favour so that they can secure an incentive; how likely physicians are to comply with such requests; and how complying physicians are viewed in terms of acting professionally.

#### **4.4 Conclusion**

While there is a substantial amount of survey literature on the role of personal responsibility, much of this is of a general or abstract nature, and only two of 21 relevant studies have examined the specific context of incentives. Themes that are of central relevance in the survey literature map closely onto the issues arising from Chapters 2 and 3 and concern, among other things, attitudes towards coercion, deservingness, effectiveness, and acceptability. While the principle of disadvantaging those who are viewed as having behaved in ways that lead to avoidable health risks is documented relatively extensively, nothing is known about the normatively relevant acceptable magnitude of penalties, and very little is known about the magnitude of acceptable rewards. The instrument developed here seeks to build on the existing literature and to help clarify central elements that are salient in the ethical discussion. By itself, this work does little to settle the question of how to proceed in implementing incentive policies. However, as one of several inputs, it can help establish a broader set of relevant reasons that feed into a procedural justice account (see Section 3.5).

## Chapter 5<sup>1</sup>

### Health responsibility: physicians' views

#### 5.1 Introduction and Aims

The approach of combining Norman Daniels' and James Sabin's Accountability for Reasonableness framework with Thomas Scanlon's Contractualism (Section 3.5), places great emphasis on the identification and due consideration of relevant reasons which different stakeholders may have, in particular, if they are cited in arguments seeking to reject a policy. One particularly important group of stakeholders is physicians. As Chapter 2 showed, physicians play a central role in implementing incentive programmes, typically have a keen interest in identifying measures that help patients to stay healthy, and are also likely to hold implicit or explicit views about what constitutes reasonable resource use within a health care system.

The consideration of physicians' views is then relevant not only on normative grounds, but also on practical ones, as physicians who object to particular elements of a policy may undermine its implementation. Conversely, proper alignment of policy goals and physician views can maximise benefits. The discussion in Chapter 2 of the responses of the professional interest groups of physicians in Germany and the US has shown that such alignment cannot be taken for granted, as there was strong opposition to several elements including the objectives of incentive programmes and a perceived policing role for physicians (Section 2.3.3 and 2.4.2).

---

<sup>1</sup> I am very grateful to the following friends and colleagues for invaluable advice and assistance in developing the instrument and recruiting interviewees: Johann Ach, David Ash, Ronal Barg, Gene Bishop, Jochen Breinlinger-O'Reilly, Berit Bringedal, James Brooke-Turner, Alena Buyx, Shawneequa Callier, Hanno Charisius, Carmen Guerra, Rizwan Haq, Ingo Härtel, Ingo Höhr, Peter Hasselblatt, Paul Henning, Nilophar Lafrai, Tim Kennerly, Sabine Kies, Jan Köser, Julia Kreis, Srinivas Kuruganti, Stefan Lenz, Wolf-Dieter Ludwig, Georg Marckmann, Claus-Dieter Middel, Mathias Mrotzek, Jens Niehoff, Charles Orellana, Engin Osmanoglou, Birte Pantenburg, Elfteria Panagiotou, Lutz Rabe, Annette Rid, Markus Rudolphi, Bettina Schöne-Seiffert, Dorothea Schmidt, Lea Schmidt, Diana Seemann, Julian Strauss, Simon Tönsmeier, Sridhar Venkatapuram, Dan Wang, Joy Wang, and Andreas Vieth. Findings were also presented at the health service research Work in Progress seminar at the Perelman School of Medicine University of Pennsylvania, on 7 December 2011, and the discussion with participants has been most helpful for reviewing the analysis here. I am especially grateful to the physicians who generously agreed to give of their time, and provided most valuable views, but, for methodological reasons, must be unnamed here.

On a more conceptual level, Chapter 3 noted that it would be short-sighted to assume that analyses of incentive programmes only needed to focus on what they mean for users of a health care system: the impact on third parties, such as physicians, and in particular on the doctor-patient relationship, also needed to be considered. However, in both chapters, the level of analysis was by necessity of a somewhat general nature, either because it took the form of conceptual analysis, or of analysis of position statements that focused on the most important elements of the reforms. But the nuances of incentive programmes can be as important as their broad outlines, or overall goals. Moreover, as the previous chapter has shown, prior survey and interview work tells us little about a number of specific aspects that are raised by incentives, such as what amount of reward or penalty should be acceptable; whether 'carrots' or 'sticks' should be seen as more acceptable; and whether those involved in incentive programmes, including physicians, always 'see through' the implications of different ways of framing programmes.

This chapter seeks to complement current research by presenting findings from interviews with primary care physicians and oncologists in Germany and the US. There are, of course, several methods that can be used to provide a physician perspective, including nationally representative surveys of all types of physicians, surveys of specialised subgroups, carried out in a structured, semi-structured or unstructured way, or perhaps focus groups. These and further methods clearly all have different advantages and disadvantages. A qualitative research approach that allows an exploration of “the meanings, values and experiences of purposefully sampled individuals and groups in their 'natural' context” (Kitto, Chesters, and Grbich 2008: 243) by drawing on semi-structured interviews seemed both feasible and appropriate in this context. While this method will not enable measuring in quantitative terms the extent of agreement and disagreement regarding physicians' views on incentive policies, it can help shed more light on the reasons that may underlie some of the reported resistance and prepare the ground for further quantitative research. It is hence in this sense that the chapter intends to respond to the thesis' sub-research question 3:

In the views of members of the public and physicians in the US and Germany: which aspects of cancer care and obesity policies focussing on individual behaviour and contribution to efficient services are reasonable, and which ones are not?



After an outline of the methods, the interview findings are presented, and subsequently discussed, along with relevant limitations.

## **5.2 Methods**

This section describes the interview method; approach to sampling and recruitment; testing and development of the instrument; ethics review; data and methods of analysis, including the development of the code list, as well as use of computer-assisted qualitative data analysis software (NVIVO Version 8.03).

### *5.2.1 Interview method*

Data were gathered in 20 semi-structured interviews of approximately 30 minutes with five primary care physicians each in Berlin, Germany, and Philadelphia, USA, and the same number of oncologists in both cities. Interviews were conducted between March and September 2011. Interviewees were given the option of both a face-to-face and a telephone interview. Mixing these modes seemed acceptable as neither the topic nor individual questions raised significant mode-dependent issues that would have biased responses. All interviews were carried out by me.

### *5.2.2 Sample frame and actual sample*

The sampling was non-random and purposive. As Kelly Devers and Richard Frankel observe, purposive sampling is “designed to enhance understandings of selected individuals or groups’ experience(s) or for developing theories and concepts. Typically this goal is sought to be accomplished by selecting ‘information rich’ cases, that is individuals, groups, organizations, or behaviours that provide the greatest insight into the research question.” (Devers and Frankel 2000: 264). The authors also draw on Matthew Miles and Michael Huberman in observing that this strategy is particularly successful for the three groups of typical cases, deviant or extreme cases, and negative or disconfirming cases (“exceptions to the rule”, Miles and Huberman 1994: 34). While the latter two cases were not of interest to the study here, efforts were made to ensure that the sample approximated typical cases.

Accordingly, the sample frame inclusion criteria required that physicians were between 35 and 65 years old; had at least 5 years of qualified professional experience in primary care or oncology; and worked in a single or group practice or hospital in

metropolitan areas of Berlin and Philadelphia, cities with broadly comparable characteristics. The term ‘oncologists’ was understood to refer to physicians diagnosing and treating cancer, not merely carrying out the screening (which, of course, can often involve treatment too).<sup>2</sup>

I also aimed for roughly equal numbers of male and female interviewees. For each group of five physicians, I sought to recruit at least two who care for patients with predominantly low socio-economic status, two with high socio-economic status, and one with patients who had a medium status. Exclusion criteria were: first, physicians practicing wholly or partly alternative medicine, homeopathy or other forms of medicine that stand in some tension to the paradigm of evidence based medicine; second, individuals who had already formed an explicit opinion and were known for their views in the press or academic literature. The total target number of 20 may appear low, but is not uncommon for this type of research.<sup>3</sup> The first set of interviews was carried out with German primary care physicians, and the mode of recruitment was refined in light of the experiences, as set out at Appendix 5A which describes the method of recruiting physicians via professional organisations, snowball sampling, and profiling the socio-economic profile of residents in the respective practice areas. Table 5.1 provides more detail on the characteristics of interviewees who participated in the research.

---

<sup>2</sup> The rationale was twofold. First, it was to ensure that physicians would be as familiar as possible with the epidemiology and treatment options of colon cancer, as well as the cost implications. Second, a controversial feature of the German policy relates to the role of physicians in reporting patients who refuse treatment, and it was intended to hear from oncologists who would face a realistic chance of finding themselves in such a situation.

<sup>3</sup> See, for example: Aira, Marja, Jussi Kauhanen, Pekka Larivaara, and Pertti Rautio. 2003. Factors influencing inquiry about patients’ alcohol consumption by primary health care physicians: qualitative semi-structured interview study. *Family Practice* 20 (3):270–275, (35 physicians); Fröjd, Camilla, Claudia Lampic, Gunnar Larsson, Gunnar Birgegård, and Louise von Essen. 2007. Patient attitudes, behaviours, and other factors considered by doctors when estimating cancer patients’ anxiety and desire for information. *Scandinavian Journal of Caring Sciences* 21 (4):523–529, (19 physicians); Guerra, Carmen, Samantha Jacobs, John Holmes, and Judy Shea. 2007. Are Physicians Discussing Prostate Cancer Screening with Their Patients and Why or Why Not? A Pilot Study. *Journal of General Internal Medicine* 22 (7):901–907, (18 physicians).

Table 5.1: code list of interviewed physicians and key attributes								
	Physician code*	Interview date	Principal Job title#	SES profile of patients^	Gender	Practice size	Phone or face-to-face	Sequence in interview
<i>Primary care physicians, Berlin</i>								
1	G-1-PC-H	10 Mar	Primary care physician	high	M	Single practice	Face	3
2	G-2-PC-L	11 Mar	Primary care physician	low	F	Group practice (3)	Face	4
3	G-3-PC-L	11 Mar	Primary care physician	low	M	Single practice	Face	5
4	G-4-PC-M	18 Mar	Primary care physician	mid	F	Single practice	Face	6
5	G-5-PC-H	18 Mar	Primary care physician	high	M	Group practice (2)	Face	7
<i>Primary care physicians, Philadelphia</i>								
6	US-1-PC-L	6 June	Primary care physician	low	F	Group practice (4)	Phone	9
7	US-2-PC-H	7 June	Primary care physician	high	F	Group practice (3)	Phone	10
8	US-3-PC-M	1 July	Primary care physician	mid	F	Group practice (6)	Face	11
9	US-4-PC-H	15 Aug	Primary care physician/general internist	high	F	Group practice (6)	Phone	15
10	US-5-PC-L	13 Sept	Primary care physician/general internist	low	M	Hospital	Face	18
<i>Oncologists, Berlin</i>								
11	G-1-ONC-M	9 Mar	Gynaecologist/oncologist	mixed	F	Hospital	Face	1
12	G-2-ONC-L	10 Mar	Hematologist/internal oncologist	low	M	Group practice (2)	Face	2
13	G-3-ONC-M	19 Apr	Hematologist/gastrointestinal oncologist	mixed	M	Hospital	Face	8
14	G-4-ONC-M	23 Aug	Internal oncologist	mixed	F	Hospital	Face	16
15	G-5-ONC-M	26 Sept	Haematologist/oncologist	mixed	M	Hospital	Phone	20
<i>Oncologists, Philadelphia</i>								
16	US-1-ONC-M	18 July	Haematologist/oncologist	mixed	F	Hospital	Face	12
17	US-2-ONC-M	19 July	Gastrointestinal oncologist	mixed	F	Hospital	Face	13
18	US-3-ONC-M	10 Aug	Haematologist/oncologist	mixed	M	Group practice (7)	Phone	14
19	US-4-ONC-M	6 Sept	Gastrointestinal oncologist	mixed	M	Hospital	Face	17
20	US-5-ONC-M	15 Sept	Gastrointestinal oncologist	mixed	M	Hospital	Phone	19

\* Note: The first section of the code indicates the country (Germany =G, United States =US); the second the sequence of interview within the specialism; the third identifies whether the interviewee was a primary care physician (PC) or oncologist (ONC); and the last indicates whether the physician's patient population had primarily low (L), medium (M), or high (H) socio-economic status, with M representing both "medium" and "mixed". The latter was assigned for all hospital doctors, who typically saw a wide range of patients.

### 5.2.3 *Instrument*

The instrument developed in Chapter 4 was pilot tested in 16 iterative interviews in German and English versions with German and English native speakers.

Translations were made by me. Eight German and eight English interviews were conducted. The core questions for physicians and those to be used in the survey (Chapter 6) were identical, but the scenario in two questions was adjusted to be tailored to the physician situation (questions 6a and b). The physician version was tested in four interviews with an oncologist, a cardiologist, a primary care physician and a medical student, with the latter two conducted with German native speakers in German, the former two with English native speakers in English. Likert scales were removed in the guide for physicians, and replaced with questions that sought to elicit open responses.

Interviews were timed, and volunteers were instructed to behave as if they were in a real interview, and to ask questions instantly if they felt that they were unable to provide an answer. At the end of each interview volunteers were asked which general observations they had on the experience, which questions they found most difficult, whether the scales (used in the survey version only) were comprehensible, whether they thought that questions were sufficiently neutral or leading in any way, and whether they felt uncomfortable providing answers in specific cases.

Substantial revisions to the form and content were made in the process, and the most important ones include the following:

- As noted above, the interviews were intended to explore the views of physicians on policies that may put them in a position where a patient asks them to lie (so that the patient could obtain an advantage associated with an incentive). This endeavour creates a potential dilemma. One could either ask this question directly, as envisaged in the initial draft of the questions (“Only you and the patient are in the room, when the patient asks you not to tell the health plan... Would you comply?”). However, this would carry the risk that respondents in the interview situation would not answer what they honestly believed in order to ‘please’ the interviewer. The alternative is to phrase the question more indirectly: “Only you and the patient are in the room.... Out of ten doctors, how

many do you think would comply with the patients' request?" The latter option allows the physician to 'hide' their view behind a statement about what 'most people would do', but, of course also asks about that assessment, and not his or her personal view. Both options have their merits and drawbacks, and in the end I decided to adopt the latter version.

- While the initial instrument had no country identifier for particular policies, some context needed to be given, as each of the incentive policies in questions 3 and 8 were realistic in either Germany or the US, but not in both. In the first trial version the incentive policy was introduced in the English version as "In Germany, a sickness fund plans to" and in the German version as "An American health plan intends to"; however, it became clear that while such labelling might provide ready orientation, it could also lead to a mixing of attitudes about the policy with attitudes about national stereotypes, which can be strong in both directions. The policies were therefore introduced as "In another country..." where applicable.
- For similar reasons, first names in the scenario where respondents are asked to comment on the impact of the weight loss policy on five different people were anonymised (questions 14a–e: "Person A, B, C" etc), to avoid gender or name connotation bias.
- Regarding levels of incentive amounts, references to percentages of gross insurance contributions or income were replaced with absolute numbers, which are easier to understand (questions 5a, b, and 8).
- Questions were also reordered to position more difficult questions after respondents were 'warmed up', and a number of revisions were made to make questions as neutral as possible.

The final version of the guide for interviews with physicians is reproduced at Appendix 5G (English version) and Appendix 5H (German version).

#### *5.2.4 Ethics Review*

The LSE's policy on ethical review was followed strictly. In accordance with the procedures the research was determined to be exempt from review by the thesis supervisor Prof. Julian Le Grand on 16 Dec 2010. All data have been anonymised, and statements are not attributed to individuals, their practices or hospitals in either

the thesis or ensuing publications. All audio recordings and transcripts are stored on password protected hard drives. Hard copy records carry initials and dates as identifiers only. Moreover, the nature of questions is such that they are unlikely to jeopardise employability or pose any other significant risk to interviewees.

#### *5.2.5 Data and data analysis*

Interviews were recorded and transcribed. Handwritten notes made during the interview were also used in the analysis, which was carried out using qualitative data analysis software. More detail on each of these steps is provided next.

#### Data

All interviewees consented to recording and the storage of data as outlined immediately above. Interviews lasted between 25 and 40 minutes, with the vast majority taking around 30 minutes. Each interview was recorded and transcribed. Handwritten notes of key points were made on a hard copy of the guide. Interviewees were reminded at the beginning that the main purpose was to hear their reasons in favour or against particular policies, rather than yes or no answers. No interviewee refused to answer a particular question or felt noticeably uncomfortable. Some questions were nonetheless skipped in the flow of the discussion, and this is indicated in the finding sections accordingly. In two face-to-face interviews, respondents wished to have the questions in front of them from the outset as we spoke (G-2-PC-L and G-4-PC-M). To avoid allowing these interviewees to consider earlier questions in the context of later ones, interviewees were asked to cover up questions that were still to be asked, to establish a similar situation to the other interviewees who only heard the questions.

One interviewee wished to read on the topic before the interview, but enquired by email whether this would be a problem. We agreed that it would be preferable if he read more after the interview, and that I would send some introductory papers. Interviewees were given the option of providing further comment or clarifications after the interview by email, but none were received. In one interview the recording device ran out of battery, and responses to questions 7–12 were not recorded (G-4-PC-M). Summaries were filled in based on the handwritten notes.

Interviews conducted in German were translated by me in the process of transcribing them, and exist in written English only, mainly to facilitate coding as described further below. For all interviews, transcriptions were made of responses to the individual survey questions only, and exclude the introduction and exit conversation. Long pauses, laughter, sighs and other semantically relevant utterances have been indicated, but similar minor pauses have only been recorded as [...]. On some occasions respondents repeated the question before answering it, and while this may be indicative of some difficulty in comprehending the question that was asked, such and similar redundancies have also not been transcribed in the interest of ease of use of the data. Equally, on occasion I repeated a key point that an interviewee made, but that seemed implicit, to ensure that my understanding was adequate. Instead of reflecting such interaction as a verbatim back and forth exchange, the statement has been inserted in the transcripts between square brackets in the section in which the respondents' comments have been noted, followed by what the respondent said in addition to my statement.

Clearly, such editing entails some judgement calls, and on a purist view it would be preferable to transcribe, code and analyse exhaustively not only all verbal, but also non-verbal aspects of the encounter, from the first to last contact in relation to the interview. For example, while most interviews were conducted in workplace settings, one took place in a café (G-3-PC-L) and one, whilst the interviewee was driving (US-3-ONC-M). Rapport was established in different ways, with some interviewees seeming to prefer to hear more about my background and the overall context of the research, and others less. These and further factors can, of course, matter as data, and may warrant consideration in analysis and interpretation of findings.<sup>1</sup>

However, such extensive transcription appeared unnecessary in view of the goals of the research here. While the methodological approach shares some

---

<sup>1</sup> For example, Jennifer Mason outlines three basic approaches to analysing qualitative data, distinguishing “literal”, “interpretive” and “reflexive” modes (Mason 1996: 54). While the first focuses on a narrow analysis of language and grammatical structure, the last seeks to consider also the effect which interactions between the interviewer and the interviewee have. The second, by contrast, is aimed more directly at making sense of the respondents' accounts, and is pursued here, cognisant of the interconnectedness of the different approaches, which, depending on research angle and question might even be seen as strands of the same approach, rather than distinct instantiations.

similarities with those used in grounded theory research, the aims here are more modest. As in the case of grounded theory (or rather: one branch thereof, as the field itself is far from monolithic in its methods, with many rejecting overly structured approaches see, e.g., Chamaz 2006), it is intended to progress from codes to categories to themes (Saldana 2011: 12). However, there is no intention here to explore in more depth the relationships between categories with an aim to develop a new theory. Moreover, the interview guide is sufficiently structured to elicit responses on a narrow range of issues, and while behaviours such as throat clearing, changes in posture, or a shift of gaze may nonetheless all be of semantic relevance in interpreting an interviewee's responses, all verbal statements were sufficiently unambiguous that I felt confident that the edited transcript captured attitudes that were salient and relevant for the overall goal of the research as described above.

### Analysis

The principal sources of analysis were the transcribed responses of interviewees. The data were organised using a computer-assisted qualitative data analysis software (CAQDAS), NVIVO, Version 8.03, see Appendix 5B. Transcriptions were imported as Microsoft word documents, in the form described above, and associated in NVIVO with the identifier reproduced in column two of Table 5.1, and the four attributes of gender, location of work, professional specialism and socio-economic status of patients.

### **5.3 Findings**

This section summarises the findings, using the codes reproduced in column two of Table 5.1 to denote key characteristics of individual physicians. Each section is preceded by the English version of the questions, as asked in the interviews. In most cases responses are first presented in tabular format, in addition to a qualitative analysis that follows. This is not to suggest that particular quantitative inferences might be made, but to provide an initial orientation of the spectrum of responses, which may also be of use in putting the qualitative analysis into context.

With few exceptions, the responses cited relate directly to the question being asked. In a small number of cases respondents made comments at other stages of the interview that were of little or no relevance to the question being asked there, but



matched with their responses to other questions. In such instances some quotes were reproduced in these, better fitting places. The symbol [\*] after the quote indicates that the quote was initially provided in relation to another question. Where striking differences between professional groups or German and American physicians were noticed these were highlighted. However, in view of the sample size, these differences need to be interpreted with some caution.<sup>2</sup>

### 5.3.1 The relationship of personal responsibility and other determinants of health<sup>3</sup>

<b>Question 1</b>																					
Overall, when it comes to our health, how important are personal choices in relation to other factors: such as environmental conditions (clean air and water), income, education, access to the health care system or inherited genetic factors? Please choose one of the following options (and expand on why, if you wish):																					
1. Personal choices are really the most important factor that determines health: we are largely responsible for how healthy we are.																					
2. The role of personal choices is an important factor, and we always bear some responsibility for how healthy we are. But sometimes other factors can be more important: it's therefore not possible to say which factor is the most important one.																					
3. The importance of personal choices is overstated: many of the factors above such as the environment, how much money we have, and what genes we inherited influence quite strongly what kind of choice we can make, and they are practically always far more important.																					
4. Don't know/none of the above.																					
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists					
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M	
1																					
2	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x					x
3									x	x							x	x	x		
4																					

Note: Two vertical "x"s indicate that the respondent assigned an in-between value.

There is relatively little spread in the responses given to question one. Most respondents agree that the second option—emphasising that personal responsibility is one of several factors that needs to be considered alongside these—is an adequate characterisation, although, as three respondent noted, different answers would be given for different diseases or disease groups.

<sup>2</sup> In general, the sample size did not seem to merit a more structured approach in which, for example, first all responses of German primary care doctors were compared and contrasted with those of their American counterparts, followed by a similar approach for oncologists.

<sup>3</sup> Note: no titles or headings were provided for the individual questions in the interview, these have only been added retrospectively.

US-5-PC-L observed that cultural settings could have a strong grip on health behaviour, as for example the undesirability of thinness among some African Americans and people with Caribbean roots had a negative effect on healthy-weight campaigns. US-2-ONC-M, opting for the third option, highlighted socio-economic constraints and access to health care as the most significant factor: “Working poor have pretty much no access to care.” US-3-ONC-M, also settling for this option, agreed that all factors played a role but ultimately concluded: “Certain things like smoking and obesity you can control, but I’m a bigger believer in fate and random DNA mutations than I am in personal responsibility”, with US-4-PC-H also highlighting genetics and the environment. US-4-ONC-M noted:

Personal health is an amalgam of genetic factors and choices that we make, and... it’s a bit like the illusion of free will that we think we have free will in a lot of aspects of our lives and we think that we have choices that are free, but there’s an enormous tilt on the table upon which those choices are made. The background, and the environment in which we find ourselves—whether it’s economic or political—is a huge determinant of how we’re going to choose.

G-3-PC-L chose the second option, but emphasised that from a primary care physician’s perspective responsibility nonetheless played a special role. According to G-3-PC-L, doctors had little or no control over the other determinants of health of patients, but

without personal engagement we couldn’t make any progress in our everyday work... the thing is, personal responsibility is the area where you can actually do something, whereas all the other things are pretty hard to influence [for us].

While, overall, there was little difference in reasons given between primary care physicians and oncologists, with regard to German and American doctors, practically all of the latter emphasised that access to health care could not be taken for granted and deserved special attention, with frequent reference to ‘food deserts’ and problems with access to healthy food.

### 5.3.2 The car insurance analogy

<b>Question 2</b>																				
In auto insurance, risky drivers have higher premiums. Some say that the same model should be used in health insurance. They argue that, for example, obese people should pay higher premiums than people with normal weight. What do you think about this proposal?																				
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
Y	(x)													(x)				(x)		(x)
N		x	x	x	x	x	x	x	x	(x)	x	(x)	x		x	(x)	x		(x)	

Note: "Y" indicates approval, "N" indicates disapproval. "X" indicates approval/disapproval without reservations, "(X)" indicates approval/disapproval with reservations.

This question invited views on the common analogy which proponents of more drastic approaches to personal responsibility make (see Section 2.4.2). While no physician wholeheartedly sympathised with the view, US-3-ONC-M opposed it for colon cancer where personal control was viewed as highly limited, but accepted it for cases like obesity and smoking, although the decision was "tough". US-1-ONC-M rejected the policy, but had some sympathy in principle:

If you're educated on your choices, and you make choices that are deleterious to your health that's fine, ... this is the United States of America, you're free to make those choices, but why should I pay for it.

Equally G-1-PC-H was in favour and optimistic about the effectiveness of the measure, believing that it could lead people to lose weight, and improve their health. But he nonetheless had equity concerns: "From a medical and motivational point of view I think this is good. Ethically, I find it questionable". A 'carrots' approach hence seemed preferable to G-1-PC-H, but if that was not feasible this could constitute a legitimate second best.

Overall, the proposal attracted some strong, mostly negative views. The following points were made in opposition. First, it was noted that it was problematic to single out the obese. If one followed the argument, there appeared to be other comparable cases. As G-2-PC-L said (and many others echoed):

I don't think you can assess every risk factor...one person goes skiing, another goes bungee jumping, another eats too much, one smokes,... how do you want to do this in a fair way? ... your metabolism plays a role, and genetic factors.

Second, it was highlighted that obesity rates correlated with income and that such approaches would constitute a ‘double whammy’, affecting the poorer disproportionately, “possibly pricing them out of health insurance” (US-4-PC-H). Third, and relatedly, American physicians flagged that although risk-adjusted premiums might be plausible on a prima facie economic level, one needed to look further, as higher premiums could become a barrier to care, and in this way “could come back to really hurt people in society as a whole” (US-3-PC-M, also US-5-PC-L). Two American physicians also highlighted discrimination legislation and potential for civil litigation as standing in the way of such policies. Lastly, G-4-ONC-M who had general reservations, also cautioned, on a practical note, that it would be “really difficult to implement objectively and fairly.”

### 5.3.3 Money for not being overweight

<b>Question 3</b>																				
Consider the following real policy. A large employer wants to create incentives for people to be healthy. Employees who show that over a year they are not overweight (in medical terms), would get a cash reward of \$150. Participation is voluntary, that is, no one is forced to be weighed. Do you think the health plan should use the policy?																				
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
Y	x		x				x		x	x		x	(x)		x	x	(x)	x	(x)	x
N		x		x	(x)	(x)		(x)			(x)			x						

Note: “Y” indicates approval, “N” indicates disapproval. “X” indicates approval/disapproval without reservations, “(X)” indicates approval/disapproval with reservations.

The question was modelled on a common US incentive. The amount at stake was relatively high for the German context, but average for the US. Views among respondents differed considerably. G-2-ONC-L approved in principle and viewed money as a powerful motivating force, cautioning also, however, that it might be pursued with different goals in mind, and that thorough evaluation was required:

I like this. You know I think that’s the biggest incentive, ... it’s better than something idealistic. If you show people money they’ll do something. And if it promotes health, then sure, I’m ok with it. So you could try this and then after a year or two you check what are the pros and cons, but generally as a project I think that’s a good thing. So, of course, it depends on what your goals are.... sickness funds obviously want to reduce expenditure and increase

profit... they're basically just companies, and the patient wants to make money and yeah, put those two things together and you might get health out of it.

G-3-ONC-M similarly had some sympathy, but pointed out that the policy was not necessarily an incentive for behaviour change:

I wouldn't want to reward normal weight because people have intrinsic motivation to keep that, but I think it's ok to give a reward for someone who's overweight when they lose weight. And for the others, this just looks like a penalty program. And that's why I would start with rewarding the desired behavior of someone who is departing from the norm.

Concerns about who benefits and what impact money had on motivation also played an important role in the views of those who opposed the policy:

...in terms of effectiveness, I think there are limitations. I'm sure there are some patients for whom this works, but generally I think that would work for 10–20 per cent. (G-2-PC-L)

US-1-ONC-M estimated that for 25% “it could help. I don't think it would get everyone”. G-4-ONC-M noted: “If someone is difficult to motivate I don't think that sum will make a difference. [And] it won't really achieve a change in attitude.”

Several physicians observed on the occasion of this policy, and regarding other interview scenarios, that they were not happy with the implications of simply rewarding normal weight and that a more dynamic approach would be preferable:

...so the key thing... don't just take absolute value for BMI or cholesterol, but you set them in relation to where people started out from. (G-1-PC-H)

5.3.4. Money for not being overweight: Carrots? Sticks?

<b>Question 4</b>																				
Two people are trying to describe the key feature of the policy just described. Which description do you find most adequate?																				
“This policy works mainly like a ‘carrot’: it rewards those with normal weight”																				
“This policy works mainly like a ‘stick’: it penalizes those who are overweight”																				
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
‘carrot’	x		x			(x)	x	x	x		x			x	x	x	x			x
‘stick’					x						x			x						
Both		x		x				x					x						x	
Neither																				

Note: “X” indicates assessments without reservations, “(X)” indicates assessments with reservations.

As noted in earlier chapters, the framing of incentives matters considerably regarding actual and perceived fairness. With regard to the above-introduced approach, some respondents were quick and unambiguous in their reply: “That’s clearly an incentive programme. Psychologically that’s how I see it. This is how it works in practice” (G-1-PC-H). In a similar vein, US-3-ONC-M said: “Well they’re not taking something you have, they’re giving you something you don’t have”, and equally US-5-ONC-M: “You’re not taking money away from the people who are overweight, you’re giving money to the people who are controlling their weight.” However, others were less sure, and seemed to think that not being able to receive a reward could be a form of penalty. Switching from their own perspective to that of people affected by the policy, US-3-PC-M notes:

I think it really depends on the outlook and the point of view of the employees... you could see it either way. You know, patients who are normal weight may view it like a carrot, and people who are overweight, or smoke, or whatever the reward is for, may view it as punitive. So it’s really dependent on the point of view of the person.

Similarly G-2-PC-L focused on how programmes were described:

Well I think this really depends on how it’s sold to the patient. I guess generally, this will be sold as an incentive programme...I think it’s both at the same time [laughs]...it’s just not treating people equally and you can see it this way, or you can see it that way. (G-2-PC-L)

G-4-PC-M agreed that both descriptions had validity, and offered an explanation for why sickness funds might have an interest in framing them as carrots, also expressing frustration regarding the effectiveness of programmes, and their impact on physician time and budgets:

These programmes already exist, of course...many sickness funds offer these, you know.. and I don’t know... this is really to lure customers because if you think about what they are paying for this and what a doctor gets for it—doing the examinations that are required to do—and if you think that the patients get more than double of what the doctor gets, then there’s something wrong, you know.

5.3.5. Money for not being overweight: how much, if it reduces insurance costs?

How much, if it results in a surcharge?

**Question 5a**

A question about how high incentives should be.

An employer’s health plan wants to introduce an incentive program like the one above, where employees who show that over a year they are not overweight (in medical terms), would get a cash reward of a certain amount.

The plan has to decide what the maximum level of incentives should be, for employees who are on average national income. Please select of the following options the one that seems most fair to you.

**Question 5b**

Another health plan also wants to offer incentives to people whose weight is normal. However, this plan wants to use a different approach. The employers’ idea is this: they increase insurance contributions for all plan members by a certain amount. This amount is then offered to employees as an incentive. That means, employees whose weight is normal get this amount reimbursed and avoid the higher premiums.

So the question is, by how much should insurers be allowed to increase insurance contributions for employees who are on average income? Again, please select from the following options the one that seems most fair to you:

	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
<b>Reduce</b>																				
Against				x							x		x	x						
\$100	X	(x)	x					x	x						x					
\$250																	x	x		x
\$500					(x)											(x)				
\$1,000																				
\$1,500																				
\$2,500																				
Open						x				(x)										
<b>Increase</b>					-		-													-
Against	X	x	x	(x)		x	x	x			x	x	x	x	x	x	x			
\$100																				
\$250																		x		x
\$500																				
\$1,000										(x)										
\$1,500																				
\$2,500																				
Open																				

Note: “X” indicates firm approval, (X) indicates that respondent has some reservations, but thought that this would be the right amount if such programs were to be pursued (despite reservations, the respondents did not think that such programs should not be used). Where respondents indicated an amount that fell between two brackets the one that was closer was selected. “-“ indicates no data.

As noted in the previous chapter, while a considerable amount of survey research has explored the question of whether those not behaving responsibly should be given

lower priority, the acceptable magnitude of such disadvantage is underexplored. Questions 5a and 5b sought to address this point in two ways, first, by exploring differential treatment by means of a discount versus a surcharge model, and second, by asking about the size of the incentive that might be permissible in each of these approaches.

The increments were chosen in steps that would allow alignment with the current US wellness incentive policy, and, in ascending steps, for employees on average income, represented approximately 2%, 5%, 10%, 20%, 30%, and 50% of the cost of coverage. As noted in Chapter 2, 2010 US health reform has increased the permissible levels of wellness incentives from previously 20% to 30, with the option of 50% in exceptional cases. In Germany, the highest levels that insured are typically able to claim annually are between \$100 and \$250 (around the nominal amount assumed here for the first two increments). However, this sum would usually involve completion of not just one, but several programmes or activities (see Section 2.3.1).

Overall, it is striking that while there is much variation regarding acceptable levels in the discount model, there is considerable consistency when it comes to the surcharge approach, which is widely rejected by respondents. Regarding the discount model, the German physicians who approve of it largely view current customary levels as adequate. G5-ONC-M was also concerned that higher levels would indirectly lead to different cost of health care, opting for the lowest level and opposing the surcharge model, as the issue became more pronounced here.

Approving US physicians set a similar, if somewhat higher ballpark range. US-1-ONC-M commented: "I think 500 is very generous... even for \$150 people will do it. Especially if it's seen as a windfall not part of a pay-check." G-5-PC-H would set a similarly low value if the programme was offered universally, but would set it as high as \$500 if it would be targeted to the overweight. US-1-PC-H would leave the decision to health plans, but thought it would "have to be a large amount of money to be effective." US-5-PC-L was also content to leave the determination of the level to the health plan, arguing as follows:



If the motives of the company or insurer are primarily financial, then whatever incentive they deem appropriate to essentially kind of get their money back is fine by me... if I were to disagree with the policy it's not based on the dollar amount, it's based on the principle. The dollar amount doesn't trouble me, I think the only concern I might have is that if it's kind of an dichotomous, you make it or you don't...in practical terms...it's probably better to have tiers so that achievement of a certain weight results in a certain reward. An even greater achievement in terms of weight would result in a greater reward.

Further salient comments that were made regarding the discount version also related to the nexus of amount of money and impact on behaviour, if differing in angle:

So if you look at this from a purely economical perspective you could go all the way up to € 2,000, but just to motivate people, you don't need that much. € 40 a year is sufficient... there are people I see here who easily make € 5,000-6,000 net a month, and they see me for these € 10 they can get per quarter, and wouldn't come otherwise but what they want is to have € 40 back at the end of the year\* (G-1-PC-H)

With regard to the surcharge model, G-5-PC-H had reservations in principle, but found that if a programme had been shown to be effective in robust evaluation, it should nonetheless be considered (although no threshold was given). US-3-ONC-M thought that in both the discount and the surcharge version the amount should not exceed 10% of the insurance premium (which, in the depiction above, has been reflected as 5% of the cost of coverage, assuming that cost might be shared between employee and employer).

Many felt that the surcharge model was penalising and not motivating:

That's almost a punitive...increase in insurance premiums, or you know, salary reduction, however you want to couch it. But I think that's sending a bad...that's not an encouraging message, that's really a stick, that's not a carrot, that's really a stick. So I would not be in favor of that, that's not encouraging, that's depressing. (US-3-PC-M)

I'm against it because, as I said, carrots work better than sticks...it's not the money, because in the end the money is the same.. I think human nature being what it is, you will get people more excited about positive reinforcement as opposed to negative penalties. (US-1-ONC-M)

There were also worries that where people were not able to comply with the incentive conditions, the increase could become a barrier to access to health care:

People's ability to obtain health care should be a right, I'm distressed by the economic barriers that a lot people face in getting health care, so I do think that it is improper for health insurance companies to place barriers. (US-4-PC-H)

In a more indirect way, it was also suggested that such policies could have a negative impact on social cohesion:

I just think you shouldn't do this. That's really unfair. That's not treating people equally. You know, that creates tensions in society, and one group is angry with the other because their contributions are higher because the others are too fat and so on. (G-2-PC-L)

In a similar vein, G-4-ONC-M feared:

If you increase this for all then you basically have a hunt, the hunt is on for the obese, because the sickness fund would justify this way in a letter to everyone with something like "we have to increase the premiums because so many people are overweight" and then somewhere in an aside they would say if you're slim you pay €100 less or whatever.

The policies also led to consideration of alternative, better approaches. US-3-PC-M thought money should play mainly the role of "a token, or a signal from the administration that we believe in health, which is why I think paying for gym membership is a great thing". G-4-ONC-M favoured non-financial incentives over financial ones, and advocated free or subsidised wellness holidays. US-1-ONC-M equally thought:

Honestly, I think there are other ways to do this. You could put a gym at work and give people time to go use it. You could make better food choices available. You could have a contest or weight watchers at work, you know, so that money could be spent less on the individual and more on the group to help...

Two comments were also made regarding the form in which incentives might be provided with a suggestion that a "nominal reward, or relatively nominal reward, from the employer is qualitatively different.... than tampering with the health insurer" (US-4-PC-H), and that, as noted above, a windfall will be more effective than a paycheck reduction (US-1-ONC-M).

5.3.6 Money for not being overweight: how many doctors would lie to help a patient, and how would doctors view a colleague who lies?

**Question 6a**  
 Assume you are seeing a patient who is enrolled in a plan as described above, where you can get \$150 if you are either not overweight, or succeed in not being overweight (in medical terms), over a period of one year.  
 A doctor needs to confirm the patients weight to the health plan.  
 The patient who has come to see you has lost 10 pounds over a year, but on the day he is weighed to record his weight for the health plan, he is 2 pounds over the specified weight target. This is the last day of the qualifying period, and the patient cannot come back another day. Only you and the patient are in the room, when the patient asks you to state that he has in fact achieved the target.  
 Out of ten doctors, how many do you think would comply with the patients' request to write down 2 pounds less?

	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
0																-				
1																				x
2												x								
3										x		x						x		
4																			x	
5			x			x	x		x				x							
6		x																		
7		x						x												
8						x		x			x			x			x			
9	x				x										x					
10				x																

Note: Two "x" indicate that the respondent assigned an in-between value. "-" indicates no data.

**Question 6b**

Let's suppose there is a doctor who accepts the patient's requests to record that he has met the target. How would most of your peers judge the behavior of their colleague who puts the patient's interest above complying with the policy's rules?

The involvement of physicians in attesting the achievement of a standard required by an incentive programme is not always necessary. However, in Germany it is common, and several primary care physicians referred to their prior involvement in such activities. The interview included two variations of the theme, with the first one entailing a relatively minor request for a cover-up, and the second one regarding cancer-care a major one, with more money at stake for patients and a far higher risk of being found guilty of professional misconduct for physician.

The depiction above indicates that most respondents would think that significantly more than 50% of their colleagues would lie, with German primary care

physicians giving the highest numbers, and US oncologists trailing the field somewhat. The second question asked how most physicians would view a lying colleague, and as noted, was intended to allow respondents a ‘protection zone’ so that if necessary, they could couch their own views in those attributed to colleagues. In many cases, however, respondents were not shy to say what they thought of lying colleagues, and two clusters could be distinguished, ‘lying for good reasons’, and ‘lying for bad reasons’. G-4-ONC-M showed understanding for the former case:

Doctors feel that they have an ethical obligation to their patients, and they don’t want to punish. And I would just really not want to do this. There would be all sorts of things surfacing here, like: ‘I just really need the money’, you know, you might have someone who’s unemployed and they try really hard... [so, complying with the patient’s request is] not necessarily a sign of character fault, it’s just in favor of the patient...showing compassion, showing empathy.

G-4-ONC-M commented on both good and bad reasons:

If that helps to motivate the patient, I would do that too [lying], I think that’s ok. Ethically and morally justified. But there are colleagues who simply don’t want to discuss things, they just fill in everything the patient wants so that they don’t have a lot of work with them, without actually doing anything to help them. That I find completely wrong. You know, psychology is the most important thing in our job and these [patients] have, after all, made an effort.\*

G-2-PC-L framed things in a similar way:

Depends on why he does it [lying]. If it’s just to keep a patient so he comes back the next quarter... that’s different than if you’re thinking the patient really tried, but still doesn’t succeed—then most people would think that’s ok.

The reference to the patient returning the next quarter and physicians just filling in more or less whatever patients present them relates to the circumstance that German patients are free to choose primary care physicians, and can ‘shop around’. A significant part of a primary care physician’s income consists in capitated payments, which means that they have an interest (one might argue: an incentive) to be accommodating of patients’ requests.

On a more technical level, US-4-PC-H observed that “Two pounds one way or the other is not a big deal”, given the patient had taken action, and made some progress. G-5-ONC-M and G-2-ONC-L also acknowledged such motives, but the latter was concerned that such a view is overly narrow in its focus on the patient perspective:

That’s understandable from a human point of view [that doctors would be lying], and at the same time it’s also compromising data. You know, I think in a professional life you have to be disciplined and stick to the truth. If you start manipulating diagnoses that’s the beginning of the end.

One way in which such interferences might become a problem concerns the evaluation of incentive programmes. US-1-ONC-M also accepts that doctors may wish to give in to patients' requests, but refers to a conception of professional and personal ethos that she sees at stake, pointing also to ways of avoiding or mitigating risks:

I will not do it. It's an absolute policy for me as a physician because honestly it comes up a lot. Patients want you to lie for them frequently. And you know, it's my license, it's my ethics. I can't do that. Even if I'm put into a position where the patient, you know... would reap some benefit for it. I don't know. Honestly I don't know. I think maybe outwardly docs would say no, no, that's wrong, but maybe in their office with a closed door they might do it. So I guess maybe I don't have a very good view of human nature and I think that people would feel pressured, and you know, there is a relationship between the doctor and the patient and I think a lot of doctors have trouble being the boss, saying I'm the one with the license and I'm the doctor. So I don't think they would do it so much because they want to cheat, but that they feel pressure, the patient's, you know, pushing them. And that happens a lot. That happens a lot in medicine, .... there's ways around that. Patients could weigh in at work, they don't need to weigh in with the doctor.

That said, not all doctors abhor their possible policing function, or are too concerned about losing patients. G-1-PC-H is not shy to state:

You have no idea how many people want me to seal their bonus cards for €10 without having actually done anything. But in those cases, I am really tough and I say 'no'. So if someone has lost 20 kilograms and is down from 100 and we're talking about 1 kilogram...whatever. But for all other cases: no!

### 5.3.7. Money for not being overweight: the social gradient

<b>Question 7</b>																				
Poorer people are often less healthy and more overweight than those with more money. Think again about the policy above where you can get \$150 if you are not overweight.																				
Suppose that research shows the following: although the incentive program is offered to all employees of a company, it turns out that almost twice as many of the high-income earners benefit from the program than the low-income earners.																				
Does it concern you in any way that the poor are less likely to benefit?																				
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
Y	-			-	x	x	x	x	x	x	x		(x)	-			x		(x)	
N	-	(x)	(x)	-								(x)			x	(x)		x		(x)
Note: "Y" indicates concern, "N" indicates no concern. "X" indicates approval/disapproval without reservations, "(X)" indicates approval/disapproval with reservations. "-" indicates missing data.																				

This scenario sought to introduce the research findings noted in Chapter 2 which suggested that across different socio-economic groups uptake of universally offered incentives programme was often such, that those better off in terms of wealth and health used them more. As US-4-PC-H, notes, in the case of policies such as the one considered here, this raises the concern that: “we’re benefiting people who do not actually need the benefit” (similarly US-5-PC-L).

Several respondents took a somewhat stoic approach with regard to this situation: “I think that’s inevitable with a policy like this. I don’t think you need to change something here, even if in practical terms there is some unfairness” (G-5-ONC-M). Others thought that unequal benefit was indicative of the fact that “the basic starting point is the wrong one”, calling for alternatives. For some, targeting incentives at those in need of health improvement seemed a promising option, although it was cautioned that this could be administratively more burdensome than offering a universal programme. For example, eligibility criteria would need to be defined and implemented, and adherence subsequently ascertained. In view of such trade-offs G-2-PC-L came down on the side of accepting less-than-perfect policy:

Well I guess there’s no point in making things even more complicated by introducing quotas or to open things to some income brackets...you know, just practically. I think it needs to be a policy that is clear to everyone. And having a clear policy can be more desirable than one that is perfectly just.

US-3-PC-M was also discontent about an insufficiently level playing field and suggested that:

You probably can’t offer a program like that without offering a full educational program that goes along with it, because then it would sort of be a bit unfair for patients who have no idea how to get started to take advantage of that program, to be competing in the same system with people who are very savvy and well educated.

Another related option to achieve more equality was to abandon the idea of equal outcome, and focus on improvements people make that also surfaced elsewhere:

I guess I would really rather see it...as not what is your weight today, but what did you do to make yourself healthier. If you made some changes over the year to make yourself healthier, I think rewarding that would be fairer, because it’s a level playing field. You know if you’ve got someone who’s starting at 350 pounds and they’ve have to make it to 150, that’s not going to happen. So I think that, ...to make it more fair, you would want each employee to do something to improve their health, not just be who they are. (US-1-ONC-M)

5.3.8. Money for not being overweight: fairness for five different people

**Question 14**

Think again about the plan above (you get \$150 if you are not overweight). Below, five different people are described, who are covered by a plan that offers such a scheme. All of them have average incomes and are around 45 years old. Because of their different circumstances and preferences, they differ in how they are affected by the policy. Please give us your ratings of the fairness of this policy for each of the below individuals:

*Person A:*

Is not overweight. He/she is health-conscious by habit, always eats healthily, exercises regularly, has had a normal body weight for years.

In line with the policy, Person A gets \$150. Is this fair?

*Person B:*

Is not overweight. He/she is not health-conscious by habit, eats mostly unhealthily, never exercises, but nonetheless has had normal body weight for years.

In line with the policy, Person B gets \$150. Is this fair?

*Person C:*

Likes to eat, dislikes exercise, and is overweight. However, person C would like to be slimmer and views the program as an opportunity to go on a diet and exercise twice a week. As a result, person C loses weight and is successful in maintaining normal body weight.

In line with the policy, Person C gets \$150. Is this fair?

*Person D:*

Likes to eat, dislikes exercise, and is overweight. However, person D would like to be slimmer and takes the program as an opportunity to go on a diet and exercise twice a week. Person D ends up trying several diets, but can't stick to one. Some weeks there is time for the exercise programs, but other times there isn't. As a result D loses some weight, but is not successful in losing weight and keeping it off and remains overweight.

In line with the policy, Person D does not get \$150. Is this fair?

*Person E:*

Likes to eat, dislikes exercise, and is overweight. Person E eats a lot, and hardly ever exercises, but does not mind the extra pounds.

In line with the policy, Person E does not get \$150. Is this fair?

	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
A	f	(f)	f	U	u	(f)	f	f	(f)	f	(f)	(f)	u	f	f	(f)	f	f	f	f
B	f	(f)	f	U	u	(f)	(f)	(f)	(f)	(u)	U	f	u	f	(f)	(f)	u	f	f	f
C	f	(f)	f	U	f	f	f	f	f	f	F	f	f	f	f	f	f	f	f	f
D	(f)	(u)	f	U	u	(u)	(f)	(u)	(f)	(u)	U	f	(u)	f	(f)	(f)	u	f	f	f
E	f	(u)	f	U	f	(u)	f	(f)	f	f	F	f	f	f	f	f	f	f	f	f

Note: “F” indicates: fair, without qualifications; “u”: unfair, without qualifications. “(f)” indicates: fair, with qualifications; “(u)””: unfair, with qualifications. “–” indicates missing data.

Issues around the impact of a single policy on different types of people were examined further in the above scenario.<sup>4</sup> As the depiction shows, the greatest extent of agreement was given for person C, who, in some sense, constitutes the paradigm

<sup>4</sup> Note: in the interviews, this question was asked at the very end, as a 'sacrificable' question, in case an interview would have to be stopped for shortage of time or other reasons.

case of an incentive user, in terms of motivational structure and effectiveness of the programme. C wants to change her behaviour, and the incentive is effective in helping achieve behaviour change: 18 physicians find the situation fair, without any reservations.<sup>5</sup>

Regarding Person A and B, who differ in that both meet the target without extra effort, but B lives unhealthily, G-1-ONC-M, US-2-ONC-M and US-5-PC-L find that B is less deserving, even if awarding the incentive is, of course, in line with the policy. Those who find the payment to Person A in order, even though there is no behaviour change, argue that there is something like ‘active-avoidance-of-needing-behaviour-change’ that should be rewarded: “they’ve maintained their healthy lifestyle...they could have started eating cheese-steaks and sitting on the couch” (US-4-PC-H). G-4-ONC-M comments:

See, if we were to separate between what’s habit and where people made an effort, there it gets really philosophical...you know, where do you draw the line there? Just because something is easy for you, I don’t give you a reward? That would be a bit like not recognizing the performance of a top athlete, but just saying that’s just because they have a good physique. I think that’s getting too judgmental.

Others reflect less on fairness, and more on what the purpose of the programme is, or should be:

You know, this person does it anyway, they have an inner drive so you can’t really say they are rewarded for it. So is this fair...it’s not unfair...but the question is really more does it make any sense? Is it justified? That seems more relevant than is it fair. (G-3-PC-L)

In this vein, US-3-PC-M notes that the policy still has a certain signalling effect, that might be more important than the side-effect of people reaping benefits, even though they behave as they always have: “Leadership of the company is interested in the fact that you’re maintaining your weight. It sends a message” (US-3-PC-M). According to US-1-ONC-M, who agrees, the message could be that “weight is the best marker of health”.

With regard to C and D, (both of whom made the same effort, but only C is successful), those defending the situation accept a degree of arbitrariness, but also

---

<sup>5</sup> Including reservations, 17 of the 20 physicians find situation A fair (excluding: 11); 14 find situation B fair (7, excluding reservations); 19 find situation C fair (18 excluding reservations); 11 find situation D fair (6 excluding reservations); and 16 find situation E fair (15 excluding reservations).



argue that D could have acted otherwise in the past: “That’s life... she lived less healthily beforehand.” (G-1-PC-H) Regarding person D taken by herself, there is widespread recognition that the effort should somehow count, but equally, that the problem is not so significant that the policy should be abandoned. Several respondents, however, felt that there is something unfortunate about the situation:

You could have [person D] even as someone who is doing exactly what [person C] is doing and still not reaching their goal..... I’ll see folks who seem to be doing similar things and having different results... sometimes it’s a lot easier for one person over the other. Sometimes you’ll have someone who is says yeah I’m exercising two times a week and I cut out soda and they go ahead and lose 50 pounds. Whereas this other one is going, like, four times a week, and you know, working really hard at it, and loses 10. (US-1-ONC-M)

Well it’s not fair if you are thinking about effort, but it’s fair in view of the policy. (G-3-PC-L)

[whether it’s fair nor not] depends on the specific circumstances that prevented that individual from being less successful than the prior individual (US-5-PC-L)

US-3-PC-M is prompted to call for support mechanisms:

You can’t just have a program and not offer any sort of information or education regarding the program, that would not be fair.

With regard to person E, most who accept the policy in principle find that because E voluntarily passes up the opportunity, E is owed nothing. However, looking at the larger picture G-2-ONC-L still perceived a fairness issue: “So in a way you would think this person should get less money than the one who at least tried and didn’t succeed.” There were no clear patterns that seem to have emerged between either German and US physicians, or oncologists and primary care physicians.

### 5.3.9 Money for taking part in colon cancer prevention

#### Question 8

Now some questions about another real example of an incentive policy. My apologies, but the policy is a bit complicated. Please let me know if there is something that should not be clear.

A health plan in another country wants to improve colon cancer care. The plan offers an incentive program that works as follows: Chronically ill on average income never have to pay more than \$550\* per year in co-payments to all health care (co-payments are generally low in this country). However, for colon cancer patients on average income this threshold is increased to \$1,100 per year, unless they comply with two conditions. If they comply, they too, have the lower co-payment-cap of \$550.

The two conditions are as follows:

1. Plan members who are older than 50 years need to attend a counseling session on the advantages and disadvantages of colon cancer screening.
2. If they should require treatment for colon cancer, they should never refuse medically recommended interventions.

So, the level of co-payments is doubled for colon cancer patients, unless they comply with these two conditions

The health plan says that the policy is justified because colon cancer patients have special personal

responsibilities to contribute to cost control.  
Is there anything that you like or dislike about this policy?

	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists					
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M	
Y	(x)																			(x)	(x)
N		x	x	x	x	x	x	x	x	x	x	(x)	x	(x)	x	x	x	(x)			

Note: "Y" indicates approval, "N" indicates disapproval. "X" indicates approval/disapproval without reservations, "(X)" indicates approval/disapproval with reservations.

This policy represents the current German law regarding colon, breast and cervical cancer. There was no physician who wholeheartedly welcomed the approach. G-1-PC-H exclaimed: "Des Menschen Wille ist sein Himmelreich", roughly, "you should always respect what people want, if that's what they desire", expanding:

Everybody should be able to say no when they want to. On the other hand, the sickness funds are a community based on solidarity, and those who aren't behaving in a solidaristic way have to pay more. That's quite clear. [sigh] From an economic perspective, this policy makes sense. Psychologically I really prefer bonus instead of malus systems.\*

US-4-ONC-M and US-5-ONC-M equally emphasised the economic case, as well as the patients' interest: "You do save lives" (US-4-ONC-M). Both thought the policy could be appropriate, as long as there was reasonable access to services, and the clinical case was monitored, to assess more precisely the magnitude of associated benefits. Those opposed did not generally dispute the scientific evidence behind colon cancer screening. Still, G-2-ONC-L also noted

You only have the incidents, the cases per year that arise, you have the people who come to the screening and then it goes down... it's really difficult to do evidence based medicine here,

This was echoed in similar ways by G-3-ONC-M. Both respondents made the point that strong scientific evidence was required to justify any policy that went beyond leaving screening uptake to the voluntary and autonomous decisions of individual health care users.

Still, strong evidence was only a necessary, and not a sufficient condition. Those opposed were principally concerned about the compliance requirement, and had, in general, less reservations about the informational requirement, which will be taken up again in the separate questions 11 and 12 below. One major line of

opposition was based on the view that the policy constituted a ‘double whammy’ for people who are in a highly distressing situation to start with. Respondents generally assumed that no patient would refuse surgery, but that the scenario would revolve around post-operative adjuvant therapy, typically chemo-therapy. US-1-ONC-M highlighted that physician and patient interests would not necessarily align in all cases, and that the option to refuse treatment without additional concerns about financial disadvantages would be especially valuable in the case of physicians who made prolonging life paramount:

You should be able to say ‘I’ve had enough’, ... some physicians are more aggressive in recommending treatment that is kind of diminishing return so you get much toxicity and very little benefit. I know one of my [practice] partners... if there’s no data to use whatever in the 25th line setting, he’ll say: ‘No, no, I’m not giving up, I’m fighting for my patient!’ And if the patient says: ‘Look, the likelihood of having any benefit is very low, and the toxicity is significant, I really don’t want to...’, and all of the sudden he has to pay twice as much,...

US-2-ONC-M was also unequivocal:

Well I think that takes away free choice about your own health, what you want and what you don’t want for you own life for health goals. So if it’s metastatic colon cancer and they say you must have chemo, I don’t think that’s really a fair policy. I don’t think any patient should be forced to do aggressive care if that’s not what they believe is right for them. I mean this health care system certainly allows patients to make any decision that they feel comfortable with. I really don’t think getting away from that... is fair.

US-4-PC-H expressed some of the strongest opposition, commenting on the exceptionalist status which had been created for colon cancer patients, and the way such a policy would be experienced by them:

What I find really offensive about this is treating the colon cancer patients as a separate um, population, and there’s a certain...we’re talking about real people who are real patients who have cancer, now you’re sort of saying to them: ‘Well, ... too bad, you did not do the proper health behavior!’ ...and I feel that, that’s like, that’s a noxious thing to say to somebody. And in terms of that individual, who’s a cancer patient, whose health we’re trying to promote, really .. the last thing I want to .. do in that situation is sort of add to the regret or the financial stress or worry of that person;...As a clinician, if I was in the room trying to take care of this patient, I certainly wouldn’t be trying to add stress or add a feeling of regret or guilt to this person who has just been diagnosed with colon cancer, I would be trying to comfort them and help them be healthy and feel better. So, as a clinician, I sort of recoil from this whole concept.

G-4-ONC-M also noted that she found it “really extreme, if you connect these two conditions [counseling and compliance]”. While the Americans saw as the only acceptable way to proceed here to abandon the policy, G-3-ONC-M focussed on

ways of making the policy less objectionable, highlighting, in particular, the point of time at which the disadvantage is conveyed. He argued that it should be brought forward and that it should be seen as a form of risk-tax, rather than as an amount that sought to correspond in some way to actually materialising health care costs associated with having cancer:

Either you tell everyone who hasn't come to counseling to have an increase of so and so much in contributions...I could live with that I guess, but to also punish someone who actually becomes sick ... they're penalized for not having done this or that. No.

G-5-ONC-M, was opposed to the policy, but also drew on a temporal aspect in his reasoning, arguing that the penalty that was far in the distant was not meaningful because:

I just don't think patients think that far ahead. I don't think they can assess the risks adequately here, ... before they get cancer. There are a lot of psychological factors at stake, too. You know, people just don't want to think about it and so don't go to screenings for that reason. And that's why I don't think that the penalty is the right thing in this case.

5.3.10 Money for taking part in colon cancer prevention: Carrots? Sticks?

<b>Question 9</b>																					
Two people are trying to describe the key feature of the policy just described. Which description do you find most adequate?																					
"This policy works mainly like a 'carrot', because it rewards responsible behaviour."																					
"This policy works mainly like a 'stick', because colon cancer patients who do not satisfy the conditions are faced with double the co-payment rate than other chronically ill."																					
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists					
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M	
'carrot'			x							x				x	x						x
'stick'	x				x	x	x	x	x			x	x			x	x	x	x		
Both		x		x																	
Neither											x										

There was little disagreement when it came to describing whether the policy should be seen as a 'carrot' or 'stick', with most opting for the latter. By contrast, G-3-PC-L, who had also seen the obesity incentive in question 3 as an incentive said: "If you do it, it is an incentive program to take part in prevention", seeing a penalty as a reason for action in the same way as a reward might mobilise people. G-2-PC-L and G-4-

PC-M had described the scenario in question three as both ‘carrot’ and ‘stick’, and saw the same applicable here.

### 5.3.11 Colon cancer prevention: 3 policy options

#### Question 10

Another health plan thinks about implementing the above policy, but has received some letters from colon cancer patient groups who find it unfair. The patient groups make two alternative proposals. The health plan does not want to be accused of being unfair. Executives therefore consider the policy alongside the two alternative proposals, that are sketched out below. Which option would you find most fair?

*Policy 1: Higher copay cap for colon cancer patients.*

As above, all chronically ill have a maximum of \$550 in co-payments, except colon cancer patients who have higher cap of \$1,100.

Colon cancer patients can only reduce their co-payment cap to \$550 if they comply with the two conditions (i.e., they come to counseling and don’t refuse treatment).

*Policy 2: Lower copay cap for colon cancer patients.*

On this policy, all chronically ill, including colon cancer patients, have a maximum of \$550 in co-payments.

However, colon cancer patients have the option to pay less.

If they comply with the two conditions (i.e., they come to counseling and don’t refuse treatment) they have a cap of \$275.

*Policy 3: No difference between colon cancer patients and other chronically ill.*

On this policy, all chronically ill, including colon cancer patients, have the same cap of \$550 in co-payments. But there is no difference between colon cancer patients and other chronically ill: all are treated the same.

	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists					
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M	
1																					(x)
2	x		x		x					x											
3	x	x	x	x		x	x	x	x		x	x	x	x	x	x	x	x	-		

Note: “X” indicates approval without reservations, “(X)” indicates approval with reservations. Two “X” indicate that the respondent assigned an in-between value. “-“ indicates missing data.

Question 10 sought further to explore views on framing, and to assess whether views on the relationship of benefits and risks that are associated with taking part in the current German policy on colon cancer might change, if the incentive structure was varied. As would be expected from the response to question 8, the vast majority was opposed to using incentives in any way. Of those opting for version 2—which turned the mechanism from a surcharge into a discount approach—two German primary care physicians, although none of the oncologists (and overall, none of the

American physicians<sup>6</sup>) changed their mind. G-1-PC-H had already expressed a preference for a carrot approach which was reiterated here, with further reference to the role of solidarity. G-3-PC-L explained:

I would do the middle option but there would be other ways in which you could improve prevention. So now sickness funds have also introduced more prevention programmes and they write to people. They should be the same as all other chronically ill but of course you can still provide incentives. If you really want the colon cancer people, then you can give them some candy... so they could get a reduction [of some kind, perhaps a] € 10 copay reductions, or whatever. But the main point is to connect health conscious behaviour with some kind of reward.

In favouring option 2, G-5-PC-H acknowledged that this would amount to treating colon cancer patients differently from others, but was swayed by the empirical evidence:

You know, we discriminate against those people who have that disease in comparison to others. At the same time, the reasoning is here of course that it's true that this is an effective form of prevention. For most other things, there aren't really preventative measures that are as good. And I think we have to be honest and accept that this is the first really successful programme that we have in Germany, so actually I think option 2 is ok here.

Those opposed were, clearly, not persuaded by such reasoning. G-2-PC-L thought it would be unnecessary bureaucracy and questioned economic rationales and clinical utility:

I think that if it's just about colon cancer screening, I wouldn't make such a big difference. I see this on the one hand from a practical point of view, there are just so many specifications and rules, and way too many exceptions and if you start here, then you can also add skin cancer and gynaecology and then we have this huge apparatus and no one understands it anymore. You know, I think this is about cost-benefits. I'm all up for colon cancer screening, but of course screenings generally are controversial. And the question really is about numbers needed to treat, and you know, I find this is a lot of effort for relatively little gain.

G-3-ONC-M illustrated that the solidarity, which G-1-PC-H invoked in support of option 2, and possibly also option 1, can cut both ways:

I think all should have the same contributions. That would mean that we healthy also have to pay higher contributions. You know, *that's* the solidarity principle, but not to penalise chronically ill.

---

<sup>6</sup> Note, however, that oncologist US-2-ONC-M noted that she might reconsider if the policy in question was about screening only.

5.3.12 Colon cancer screening, the 'right not to know', and the 'right not to think'

<b>Question 11</b>																				
Suppose a health plan wants plan members to do more than just attend a counseling session. Instead, the plan intends to offer the lower co-payment cap only for those who have undergone recommended colon cancer screening. So, in order to reduce their copays they need to come to the screening. However, the proposal has some opponents. What do you think about the following statement? “There should be a right not to know about whether one has a serious disease such as cancer, or will suffer from it in the future. This means that if someone does not want to take a test, there should be no financial disadvantage or penalty.”																				
<b>Question 12</b>																				
Another health plan defends the alternative policy introduced above as the first option, which, as you know, does not require undergoing screening, but attending counseling on the advantages and disadvantages of screenings in order to access the lower co-payment threshold. The plan therefore accepts that there is a right not to know about whether one will suffer from serious diseases such as cancer now or in the future. But a spokesperson also emphasizes the following point: “We do not accept that there is a right not to think about whether one should take part in a colon cancer screening. This is why a financial disadvantage is justified for those who refuse to attend a counseling session in which they can think about the advantages and disadvantages of colon cancer screens.”																				
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	G-1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
<i>Right not to know</i>																				
Y		X		x	x	x	x	x		x	x	(x)	X		x	x		x		
N	x		(x)						(x)					x			(x)		(x)	x
<i>Right not to think</i>																				
Y		(x)		x		x	x		x		x	x			(x)			x		
N	x		x		x			(x)		(x)			(x)	x		(x)	x		x	x
<i>[For context: responses to the Question 8, regarding the current German cancer policy as a whole]</i>																				
Y	(x)																		(x)	(x)
N		X	x	x	x	x	x	x	x	x	x	(x)	X	(x)	x	x	x	(x)		
Note: “Y” indicates that the respondent affirms the right; “X” that the right is disputed.” X” indicates that the assessment is made without any reservations, (X) that it is made with some reservations.																				

Questions 11 and 12 aimed to explore reasons that might be supported in going further than the current German policy, by demanding undergoing screening, as opposed to merely taking part in an information session (as noted: this was the scenario the German Government had favoured when it initially introduced the Bill to Parliament), and to ascertain what physicians made of the 'right not to think' about one's cancer care options, that the *de facto* German law had denied, in imposing a surcharge for those not attending a counselling session.

The overall picture of approval and disapproval is relatively clear regarding taking the law further, with most opposed. But it is quite unclear with regard to the newly established 'right not to think'.

### 5.3.12.1 The 'right not to know'

Regarding the right not to know one might expect that those who objected to the current policy would also be likely to object to a more severe version. This seems to be the case, although G-3-PC-L, US-1-ONC-M, and US-2-ONC-M are opposed to the current policy, and would not support higher charges for those refusing to attend screening, but would be content to support unequal treatment by providing 'carrots' for undergoing screening (that is, they accept the second scenario of question 10, rather than the first, which represents actual policy). G-3-PC-L refers to contractual arrangements, existing reporting mechanisms regarding uncooperative patients, and the preamble of the social security schedule:

I reject that [the right not to know]. So I would say that's ok, ... we force you to take part in these programmes but if you do that, you have a reduction. The theme here is really that you encourage patients to be cooperative, and you know, the professional associations ... we have to report to them already, and if people aren't cooperative then they even cut their sick pay [which is administered by the sickness funds]. There are real sanctions if a patient refuses to cooperate, and the sickness funds also have that as a preamble, that all sick people have to be cooperative in overcoming their illness, and they can charge higher copays if a patient deliberately violates instructions."

US-5-ONC-M's view on the 'right not to know' is:

I totally disagree, I think that's ridiculous. You have health insurance to protect your health so it's not like you're being socially penalized, you're just being told that if you want health insurance, you have to actually do things that will think will keep you healthy.

Similarly, G-4-ONC-M finds that "when we're talking about colon cancer, it's OK to coerce people into their good fortune ('sie zum Glück zwingen'). US-4-ONC-M notes:

I understand the position that people take on [the right not to know]. I'm partly sympathetic to it, but I've seen the victims of this on the other end, and they say: 'oh I should have done this'. So, I don't feel strongly about that objection, and I don't feel strongly about overruling that objection either if it's clearly to the benefit of the population at large.

For US-4-PC-L the clinical benefits also feature centrally:

You know it's funny, I'm a little bit torn about this because I would love to increase people's utilization of colon cancer treatment; that's a positive value to me, to do something that might incentivize people to actually have the colon cancer screening... You know, probably in general, I have a discomfort level with the idea of the financial incentive, but in terms of really trying to persuade or increase utilization of colon cancer screening, is something I'm really in favor of...so yeah.. I have mixed feelings about that one.



With US-2-ONC-M another American join the ranks of those with some sympathy for constrained respect for patient autonomy:

Um, it sounds like someone who belongs to the ACLU [American Civil Liberties Union] or something like that. I think if it's going to have impact on the costs of running a health plan I don't think that that's really right. ... I could see rewarding patients for doing a recommended screening.

GP-1-PC-H, who already expressed sympathy for the current German cancer policy, expectedly, is also against a wide-ranging right to know, emphasising the principle of mutuality and solidarity, and echoing G-3-PC-L:

I think we can ask people that. I think it's ok that people who do more for themselves get a benefit out of that. You know the sickness fund works on the principle of solidarity; we all pay into it, and should help those who are weak and poor but it can't be that some say, you know I'll just free-ride as much as I can. And this helps to avoid such things.

Equally unsurprising, those who were opposed to the current German policy argued for full respect for the right not to know, drawing on notions of free-will, patient autonomy, consent, and the need to understand the difficulties of judging who is able to cope with diagnostic or predictive health information at specific stages of their lives. As US-1-PC-L notes:

There's extenuating circumstances where people just don't want to know and I respect that as long as they understand that that could potentially kill them, and alter their quality of life.

### 5.3.12.2 The right not to think

Those disputing the right not to know typically also dispute the right not to think.

Accordingly, G-1-PC-H states: "...definitely.. I think it's ok to have a penalty here, because they cost a lot more down the line." Only half in jest, US-4-ONC-M comments:

All they're doing, all they're being forced to do, is ponder their mortality, which is not so awfully coercive—they should have been doing it years ago! (laughs) So... I just don't think it meets that standard of interference with people's lives—I think it's ok.

In a similar vein, US-1-ONC-M presents an American version of the solidarity argument:

We're talking about health plan, we're talking about a finite number of dollars, and trying to save money and trying to keep a population healthier...you know, I don't disagree with the statement.

However, not all who uphold the right not to know also uphold the right not to think. Physicians G-5-PC-H, US-3-PC-M and G-3-ONC-M all support the right not to know, but can envision reasons to reject a ‘right not to think’. While G-3-ONC-M gave no particular reason, US-3-PC-M argues “you have to be educated in your choices before you make those choices”. G-5-PC-H explains:

Ok here I agree with the sickness fund. You know if it’s really only about the counseling then I think it’s ok to penalize, not for not doing the screening, but for counseling that’s ok. No I really don’t think there is a right not to think... because, if we are in a system that’s based on solidarity like the statutory sickness funds, then you can’t get out of it, you can’t say I don’t want this. I think some obligations... and this is at such a low level, if it’s just about coming to counseling then it’s not expecting too much. I think you exaggerate liberty if you say you can’t do that.

But some are fundamentally opposed to the approach. G-2-ONC-L says: “...even if it’s just about thinking, I don’t think you can make that a duty. It’s just somehow my basic ethical instincts.” Others are more specific in their objections, and view the approach as an aberration or undesirable singularity within the health care system and social policy more widely. G-4-PC-M notes:

Well how do you want to prove that someone has thought of this? You have to then make a distinction between freedom to use information and freedom not to think. That seems to be the same. So what this means is that there’s no right not to think, but there is a right not to use all information. You know, not thinking, and not wanting to be informed... I don’t think you can measure that. I can’t really prove that just because someone didn’t come to a counseling session, they haven’t thought about it... You know, he could have used other media to inform himself. Or maybe someone who says well there was somebody in my family who had colon cancer, and I really just don’t want to know if I have it, and if I have it I just want it to be over quickly. [...] So you can’t say they haven’t thought about it. You know you also see that in the rationales behind these policies we are discussing here. This is just about structure and shifting money flowing from one place to another but it’s not really about treating diseases differently or earlier.

G-2-PC-L accepts the principle behind the approach, but finds a form of implementation that focuses on aspirations, as opposed to sanctions, more adequate:

I would formulate this as should, not as must... duty really means must and I think there are certain topics that every patient should think about, you know, for example, transplant medicine, and I think there should always be the option that you have the option not to think about it. So 'should': yes! I’m really in favour of that. But I think you have to stop somewhere.

US-4-PC-H also agrees that the outcome of patients having reflected on the advantaged and disadvantages is desirable, but that this could be done in simpler

ways, namely by patients talking with their physicians. Moreover, it was thought that providing incentives might complicate such discussions:

[I]t really should be the role of the physician to talk to the patient about [this]...it also qualitatively feels different to me, to say that I as a physician will discuss the pros and cons of this testing and try to persuade, which I really do try to persuade people to go to this testing...versus I'm going to compel somebody to go to a class and talk about the pros and cons, it kind of sounds almost like a DUI [driving under the influence of alcohol] class or something: everyone needs to take time out and go to the special class. I guess it's because of how I regard the patient physician relationship, versus the insurance company compelling people to do something and/or incentivizing or disincentivizing them.

*5.3.13 Money for not refusing colon cancer treatment: how many doctors would lie to help a patient, and how would doctors view a colleague who lies?*

<b>Question 13a</b>																				
Suppose a person has colon cancer. The physician recommends post-operative treatment that is accepted as the most effective by the professional bodies. However, the patient thinks that a natural remedy would be better, that has been recommended by a friend who read a lot about alternative medicine. She therefore plans to refuse the treatment that the physician recommends. The patient knows that refusing treatment means that his overall co-payments for all health care cost are \$1,100 instead of \$550. She also knows that doctors are required to inform health plans about patients who refuse treatment. Only you and the patient are in the room, when the patient asks you not to tell the health plan about her refusal. Out of ten doctors, how many do you think would comply with the patients' request not to notify the health plan?																				
	Ger. Prim. Care					US Prim. Care					Ger. Oncologists					US Oncologists				
	Gf1-PC-H	G-2-PC-L	G-3-PC-L	G-4-PC-M	G-5-PC-H	US-1-PC-L	US-2-PC-H	US-3-PC-M	US-4-PC-H	US-5-PC-L	G-1-ONC-M	G-2-ONC-L	G-3-ONC-M	G-4-ONC-M	G-5-ONC-M	US-1-ONC-M	US-2-ONC-M	US-3-ONC-M	US-4-ONC-M	US-5-ONC-M
0			X	x	x		x	x		x	x	x		x		X	-	x	x	
1	x	x			x	X				x	x					x				
2		x																		
3														x						
4																				
5									x											
6																				
7																				
8																				
9																				
10																				
Note: Two "X" indicate that the respondent assigned an in-between value. "-" indicates missing data.																				
<b>Question 13b</b>																				
Let's suppose that there is a doctor who accepts the patient's requests not to tell the health plan about her refusal. How would most of your peers judge the behavior of their colleague who puts the patient's interest above complying with the policy's rules?																				

The second scenario around physicians needing to confirm a patient's decision to a health plan or sickness fund was set up in a more drastic manner, and several physicians observed in the interview that it was unrealistic, as health plans or sickness funds might already have the information which doctors are supposed to provide, for example in the form of bills for drugs, or documenting more comprehensive chemotherapy procedures. In many cases this may be true. However, it still seemed reasonable to ask about physicians' attitudes, as first, patients will generally not be aware of the billing mechanism, and, in a scenario as the above, may well ask their doctor regardless. It could then be interesting to see whether doctors would simply view the situation as neutral (and inform patients that the payment systems are set up in a way that would make complying with their request impossible) or whether they would nonetheless feel uncomfortable (by having to refuse a patient a wish, if the patient perceives they have some scope for action). Second, part of the point of the scenario was to test how far doctors would be willing to go in such cases. For, in the most extreme case, it could at least in principle be possible that a physician would try and manipulate records to assist a patient.

It is striking that the depiction of responses above indicates that at most one or two physicians—though not none—were viewed as likely to comply with a patient's request, despite the above qualifications. US-4-PC-H did emphasise that it would be extremely difficult for a physician to find a way around this, but nonetheless said "I'm going to say 50/50 again, I'm not really sure but I would say 50[%]". Most physicians noted that if the patient was about to make a severe mistake, that their obligation was to talk them out of their decision, and that, if the patient should nonetheless pursue it, this could still be explicable, if only as a form of refusing aggressive chemo-therapy. For example G-1-PC-H stated:

Of course I would understand—I wouldn't feel like chemotherapy—and we have to admit, many chemotherapies don't lead to a cure but are really only done palliatively. So really quite questionable. So I do think everybody should have the right to say no to palliative care.

Equally, several physicians commented again on the 'double-whammy' this situation would constitute for the patient, who suffered both physically and financially. Personal implication in this process, even if only as the messenger, made several physicians uncomfortable, and led G-2-ONC-L to think about what such scenarios means for a doctor's role, in negotiating needing to serve both the patients' and the

sickness funds interests: “you know, so this is about health but it’s also about profits for the sickness funds, and to compromise my confidentiality for that makes my stomach turn.”

## **5.4 Discussion**

The discussion section concentrates on differences and similarities between attitudes of American and German physicians, and between the two professional specialisms, insofar as these are relevant to the research question underlying this chapter. That is, the focus is on which aspects of policies seeking to promote personal responsibility, cancer care and weight-control are reasonable, and which ones are not, from physicians’ perspectives.

### *5.4.1 Differences and similarities between American and German physicians*

Again, this section need to be prefaced by saying that no attempts are being made here at extrapolating from the sample that was interviewed to wider physician populations, nor is it intended to suggest that any differences that were noted are by necessity to be found among all physicians of the respective disciplines in the two countries: the comparison is merely between the physicians interviewed as part of this research. The interviews were carried out to identify patterns and themes that emerged and that may be of relevance for assessing the salience of key elements of the ethical framework presented in chapter 3, and for contextualising the findings from the surveys in the following chapters.

Question one, concerning the factors that influence health, brought up one of the clearest differences. Most American physicians were unable to talk about the role of personal responsibility without also emphasising the fact that not everyone had ready access to comprehensive and affordable care. Equally, it was emphasised repeatedly that the conceptual discussion about rewards and penalties was intimately linked to significant questions of affordability and access of care. Both in response to the first two questions, as well as in later questions, it was emphasised that incentives, of whatever kind, and for whatever health-related behaviour, would need to be designed in a way that did not deter people from seeking care, or made care less affordable. For obvious reasons, such emphasis was not found in the German responses: here the concern was more about cream-skimming of the better off.

While the extent of concern differed among American physicians, the risk that incentives may worsen population health, rather than better it, was perceived as real, and marked a first central aspect that determined the acceptability of incentives: namely, that they must be designed in a way that does not undermine affordability and access to care. This was not a concern of German physicians.

A related similarity between physicians in both countries was that the features of programmes presented in the interviews—all of which were either abstractions of real policies, real policies, or policy proposals with strong support from prominent quarters—caused some very strong reactions (although the overall tone of conversation was calm and friendly throughout). Both German and American physicians voiced concern about what the policies meant for them and their understanding of the doctor patient relationship. The German cancer policy was deemed “offensive”, “extreme”, “really horrible”, with its implications making one’s “stomach turn”, and leading in some cases to nothing less but “the beginning of the end”. Particularly strong reactions were noted with regard to being put in situations where patients requested physicians to lie. Even when the amount was not substantial, requests for cover-ups were capable of generating significant unease, not least as physicians were then put in the situation of having to decide where to draw the line. Some resolved this by asserting that they categorically never lied, others appeared more pragmatic, but nonetheless not content with the situation.

Concern was also expressed about the wider social consequences that may arise from obesity policies, which were viewed by some as declaring a “hunt on the obese”. Through their role, physicians are often complicit in enabling such policies, and clearly, programmes that cause much friction will be less successful than those where interests are better aligned. The examples discussed here cover only a small set of situations, but it is clear that physicians will not judge as reasonable policies that put them in uncomfortable positions, in particular where these are viewed as incompatible with their professional ethos; recall also the different statements physicians made about the proper attitude towards their patients, in which sympathy, empathy, compassion, support and assistance should be guiding, rather than administering penalties on behalf of health systems.

Still, with regard to accommodating physicians' concerns, judgements will need to be made. Not every objection to an incentive programme should be viewed as a veto. For example, in the case of what one might call 'incentive envy', that surfaced in the statements of three German physicians (but also one American) they noted that they were not happy that patients received incentives for certain outcomes, but they as physicians did not, or they felt that, in any case, better outcomes would be achieved if both sides were incentivised. In some cases this view reflected concerns about demands on physicians' time that resulted from administering incentives on behalf of sickness funds, but there may also be deeper psychological element that may need to be considered here, having an influence on how well (or not) incentive programmes are accepted by physicians. In terms of approaching the question of what reasonable incentive programmes should look like, this dimension would also need to be considered, if the four physicians should not be alone with their views.

A number of further minor differences seemed to emerge from the responses to the questions, for example, German physicians seemed to aim somewhat lower in setting acceptable thresholds for incentives in the discount-mode; with regard to creating incentives through the surcharge mode, the only three that were in favour of this approach were Americans (see tabular overview to question 5a and 5b, Section 5.3.5); and when it came to estimating how many physicians would lie when requested by their patients, German primary care physicians gave slightly higher numbers than their colleagues (see tabular overview to question 6a, Section 5.3.6 and question 13a, Section 5.3.13). However, in view of the sample size there seems little point of seeking to analyse these differences in more depth, other than noting them in passing.

By contrast, a final, more major and noteworthy, similarity between American and German physicians concerns the implicit and explicit reference to the concept of solidarity. While one would not have been surprised to find German doctors invoke this principle, it was interesting to note that the American interviewees presented arguments that progressed along very similar lines—both with regard to asserting that individuals have claims against the community, as “being in this together” means that the poor and vulnerable deserve support, but also with regard to recognising that the community may have claims against individuals. This became particularly

apparent in the discussion around cancer, perhaps because the element of personal causal responsibility regarding getting cancer is far less significant than in the case of obesity (which, of course, was one of the reasons why these two conditions were chosen). I will return to these arguments in Chapter 7, which will seek to illustrate the utility of the framework set out in Chapter 3 by applying it to the German cancer policy, and seeking to defend a modified version of it. For now it can be noted that perhaps unexpectedly, a significant number of both German and American physician were willing to accept the novel and somewhat unique denial of the right not to think (about colon cancer prevention). They felt that the obligation to think can be defended in the interest of both maximising resources among a community of solidaristic risk-sharers (such as insurance holders), and, in openly paternalistic ways, to promote health of people individually.

#### *5.4.2 Differences and similarities between primary care physicians and oncologists*

The differences between primary care physicians and oncologists that arose in the interviews were mostly on a very small scale, if talk about a scale is meaningful in this context. Oncologists—and US oncologists more so than German ones—were more likely to see some good in the car insurance analogy (see tabular overview to question 2, Section 5.3.2). When it came to physicians lying on request of their patients in the weight-incentive scenario, US oncologists, with one exception, assigned the lowest numbers (see tabular overview to question 6a, Section 5.3.6). They also seemed to be somewhat less concerned if the poorer benefit less from weight-control incentives (see tabular overview to question 7, Section 5.3.7), and US oncologists were the most willing to view as fair the descriptions of five different people covered by a plan that offered a weight-control incentive of the above type (see tabular overview to question 14, Section 5.3.8). However, again, the sample size does not suggest that the magnitude of these differences can be explored at this stage in a meaningful way.

Instead, it seems useful to highlight a last point that surfaced in several places of the interviews, concerning the usefulness of rigid outcome measures—such as fixed BMI values—as opposed to more flexible participation-incentives. Both oncologists and primary care physicians in the US as well as Germany voiced unease about requiring normal body weight. Concerns were based on fairness as well as



effectiveness considerations, and a number of alternative suggestions were made, such as providing gym access instead, or wellness holidays, or access to healthy food. However, none of these arguments were supported by any evidence as to the impact on either fairness or effectiveness. While the former is, however, relatively easy to supplement with conceptual analysis, the latter is far less straightforward. Thus, a problem could arise if it should emerge that attainment-incentives are, in fact, far more effective than process incentives (Schmidt, Halpern, and Ash 2012). This question cannot be resolved from the armchair and required robust empirical trials and evaluations. However, it seems clear from the feedback provided from respondents that even if such programmes were more effective, they would be less welcomed by physicians. Again, this situation would require good dialogue to ensure that even though physicians may be *prima facie* disinclined to find attainment-incentives an appropriate tool, they consider them on the basis of the evidence (on effectiveness and fairness) alone.

### **5.5 Limitations**

In terms of methods, as described, the interviews were coded by myself only, and there was hence no process to ascertain the reliability. At the same time, one coder only also means that there is no need to worry about inter-coder variability, and since coding was done using NVIVO, in principle it is possible for others to ascertain the appropriateness of the overall code map, as well as the coding of individual sections. On the whole, however, it would appear that since the questions were generally rather specific, and responses usually short and to the point, the risk of significant ambiguities was minimised.

The aim here was not to develop a new theory, and although a number of statements referred to quantitative aspects of the distribution of responses, no attempts were made here at making any broader generalisations or inferences from the data. Different physicians, different methods of recruitment, and slight nuances in the wording of questions would, no doubt, have brought about different results. At best, the data here have generated some hypotheses that can be ascertained in more detail in further survey research, that can seek to establish representative findings.

## 5.6 Conclusion

This chapter described the methodology and execution of 20 interviews with oncologists and primary care physicians in Berlin and Philadelphia. Survey items included concepts that were identified as salient in the prior conceptual analysis and review of the survey literature. The chapter addressed sub-research question 3 and sought to shed more light on which aspects of cancer care and obesity policies that focus on individual behaviour change physicians viewed as reasonable, and which ones as less so. With the exception of health-system related background conditions, there were no significant major differences between the four groups of physicians, although considerable variation of reasons could be noted in support and against normative assumptions inbuilt in the survey questions and scenarios. Key factors of relevance in physicians' assessments of the reasonableness of incentive programmes centre around their impact on affordability and access to care; their impact on the doctor-patient relationship and their personal and professional comfort with regard to being put in policing situations or being requested to lie; the extent to which programmes can maximise efficiencies and engage patients (and the public) in health literacy and health promotion activities; and, more indirectly, the extent to which they are capable of supporting the principle of solidarity (or, less normatively laden: risk-pooling) among insurance holders. Most of the interviewed physicians were concerned about inflexible one-size-fits-all incentive programmes and deemed tailored programmes that focus more on process than attainment-incentives more acceptable.

The next chapter will use a modified version of the instrument used here for a survey with the general population, to approach the question of how particular programme features are viewed by another important stakeholder group. The following chapter will draw both perspectives together in applying the framework set out in Chapter 3 to the specific case of colon cancer prevention.

## Chapter 6<sup>1</sup>

### Health responsibility: views of the public

#### 6.1 Introduction and Aims

As noted in Chapters 2 and 3, taking into consideration views of members of the public is of central relevance both for the legitimacy of personal responsibility policies, as well as for considerations of fairness, equity, and more practically, for the feasibility of implementing policies. Chapter 4 showed that a considerable amount of the existing survey literature explores the extent to which health related behaviours that are associated with poor health may justify giving lower priority to patients (or, demanding higher cost sharing). However, there is no reliable data that would allow one to gauge the magnitude of such disadvantage, nor the extent to which realistic policies, as implemented and envisaged by policy makers, meet with public approval or not.

This chapter seeks to build on prior research, and make progress in both areas, by presenting and discussing the findings from surveys at the population level in the US and Germany. In doing so it also seeks to address sub-research question 3:

In the views of members of the public and physicians in Germany and the USA: which aspects of cancer care and obesity policies focussing on individual behaviour change are reasonable, and which ones are not?

The discussion section will also relate the survey findings to themes that emerged from the interviews with physicians. The surveys were based on the instrument used in the physician interviews, but included several significant revisions, as will be described below. A considerable amount of data was generated, giving rise to

---

<sup>1</sup> I am very grateful to the following friends and colleagues for invaluable advice and assistance in developing the instrument, fielding the surveys and analyzing the results: Johann Ach, Sandra Applebaum, David Ash, Jorge Ayala, James Brooke-Turner, Johannes Bruns, Alena Buyx, Anna Bushan, Shawneequa Callier, Joseph Capella, Hanno Charisius, Tim Doran, Ian Forde, Markus Grabka, Scott Halpern, Rizwan Haq, Ingo Höhr, Paul Henning, Tim Kennerly, Amit Kavnekar, Martin Kaiser, Lore Korbei, Jan Köser, Julia Kreis, Srinivas Kuruganti, Karin Lange, Wendy Mansfield, John McMillan, Stephanie Morain, Mathias Mrotzek, Axel Mühlbacher, Engin Osmanoglu, Ursula Poch, Veronika Poch, Marianne Promberger, Ian Riner, Lutz Rabe, Tim Rhode, Andrew Ryan, Bettina Schöne-Seiffert, Dorothea Schmidt, Lea Schmidt, Stefan Subias, Humphrey Taylor, Simon Tönsmeier, Andrea Troxel, Sridhar Venkatapuram, Dan Wang, Joy Wang, Julia Weinmann, Andreas Vieth, Kevin Volpp, and Jingsan Zhu. I am also grateful to the Penn CMU Roybal P30 Center in Behavioral Economics and Health for enabling funding through the Division of Behavioral and Social Research, National Institute on Ageing, US National Institutes of Health.

numerous options for interrogating the data and testing emerging hypotheses. Throughout the chapter the focus is largely on descriptive statistics, with basic inferential statistics employed better to understand the significance of differences between assessments of, for example, age, health, or income groups. Each subsection of the findings is preceded by a summary chart and summary description, followed by a more detailed overview of statistically significant findings. To render the overall research question more concrete, and to provide structure to the presentation of the data, the findings section is organised around the following five questions:

- To what extent do respondents use incentive programmes?
- To what extent do respondents agree with rewards and penalties?
- To what extent do respondents value opportunity of choice, motivation and effort?
- What levels of incentives do respondents regard as acceptable for particular incentive programmes?
- Do respondents believe incentive programmes interfere with the doctor patient relationship?

The cross-cutting discussion section is structured around the following three headings:

- Similarities and differences between the American and German samples.
- Similarities and differences between survey respondents and physicians.
- Winners and losers of incentive programmes.

## **6.2 Methods**

This section describes the types of surveys that were carried out; the approach to sampling; revising and testing the instrument; the ethics review; and types of data as well as methods of analysis, including the use of statistical software.

### *6.2.1 Survey types*

Regarding the survey mode, pilot testing revealed that several questions were overly complex to be used in a telephone survey, and that a written version would be preferable with regard to maximising response rates and validity of responses. Web-based surveys are a relatively new method of survey research, and there is some

debate about how to apply established assessment metrics regarding the robustness of findings from telephone and paper-based written surveys to the online version (Wright 2005; Muntz 2011). Compared to the options of telephone or in-person surveys, online surveys share significant advantages of written surveys, in that they enable people to complete questions in a pace that suits them, and without feeling pressured to conform to perceived expectations of the interviewer, or without becoming nervous about sensitive subjects. Given that some questions concerned lying, body weight and assessments of the fairness of policies that could penalise groups who differed in relevant aspects from the respondents, the anonymity of the written format seemed particularly appropriate. The option of being able to randomise questions within a survey is a major advantage of online surveys, and was relevant for an experiment that formed part of the survey (to exclude ordering bias with regard to sets of questions where this was likely to pose problems, as will be described below).

Three web-based cross sectional surveys were fielded, two in the US, and one in Germany. While, ideally, surveys seeking to compare attitudes in the US and Germany would both have been conducted in probability-based samples that are capable of being representative of the general population, the cost for an instrument of the envisaged length would be around £20,000 per survey, compared to around one tenth for a non-random sample. In view of available funding, the fact that in several ways the planned research explored new ground for the first time (which can justify a more cost-effective approach), as well as the more experimental focus on a set of questions on the adequate levels of incentives, it therefore seemed appropriate to take a pilot approach with two non-random samples, and use a smaller subset of questions in a random sample in one country only. In the following I will refer to the non-random sample in the US as “US1”, and to the random sample as “US2”. The German non-probability-based survey will be referred to as the German survey, or “Ger”, typically in figures. Only findings from US2 enable robust extrapolations to the wider population (Yeager et al. 2011).

### *6.2.2 Instrument*

In reviewing the instrument used for the physician surveys for use in a population level survey, several revisions to the form and content were made, as detailed at

Appendix 6A. In several ways the instrument used in the physician interviews therefore differed from the one used in the survey. However, it was never intended to carry out a direct quantitative comparison between findings in both groups, and it seemed preferable to make the changes to improve the utility of the survey instrument, rather than preserve a higher degree of concordance between both instruments.

Further to pilot testing the initial instrument used in the physician interviews, the English and German versions of the final web-based instrument were pilot tested again with six native speakers each. In terms of structure, since the survey was not short (15–20 minutes), an overall format was chosen that sought to strike a balance between rigour in terms of validity of data generation (which might demand, for example, that the order of all questions be randomised) and ease of use and maximising completion rates (organising questions in a logical manner may guide respondents more easily, but is also likely to introduce bias due to the ordering of questions). Specifically:

- The first question from the physician interviews remained in front to ensure that all respondents had considered the fact that health can be influenced by several factors. Here, ordering was necessary.
- To avoid order effects, question randomisation was used within three scenarios in which respondents were asked to determine amounts that should be provided for a ‘carrot’ incentive (leading to a net reduction in health care cost), a ‘stick’ option (where the incentive is to avoid a surcharge) and a ‘false carrot’ scenario (here, insurance contributions are increased, and the incentive amount reduces, or neutralises, the increase); within the five vignettes in which respondents are asked to rate the fairness of obtaining or not obtaining a weight-control incentive for five groups of people who differ in their behaviour and motivations; and for two scenarios on financial penalties for not attending a colon cancer counselling session, or failing to undergo a colonoscopy.
- Where possible, questions were fitted on to one screen only to reduce the risk of unintentional non-response where questions that require scrolling down might be overlooked. Demographics and other non-evaluative questions were placed at the end. Demographic information relating to age, gender,

race and educational attainment were not collected in the survey, but provided by the sample providers (who hold these data on all panel members).

All policies were presented in a realistic fashion, similar to the way enrollees might learn of a programme through their employer or insurance company. Balanced Likert scales with neutral mid-points were used for all questions concerning agreement with a proposition, or judging the fairness of a programme (Brace 2004: 86). Options were structures such that the top (or left hand side) began with the “strongly disagree”/”completely unfair” choice. Except for the consent page, there were no forced-choice questions, as skip rates in access panels tend to be low, and providing the option of skipping tends to increase overall completion rates and robustness of the collected data (Iarossi 2006: 61).

The survey versions used in US1 and Ger comprised 21 individual substantive questions that concerned respondents’ attitudes, with the wording as reproduced in the text boxes below. Fourteen questions related to demographics and other respondent characteristics. The shorter version used in US2 comprised five substantive questions with identical wording to the questions used in the longer surveys, and three regarding other respondent characteristics (health, insurance status, political affiliation). Surveys US1 and Ger were fielded using the survey platform *SurveyMonkey*, while US2 was administered through a proprietary platform of the company the Knowledge Network (more detail on the companies’ respective approaches is provided below).

### 6.2.3 Ethics Review

Since the research received funding from the Division of Behavioral and Social Research, National Institute on Ageing, US National Institutes of Health (P30AG034546-03, pilot project 9), the protocol was submitted to the University of Pennsylvania’s Office for Regulatory Affairs with a request for exemption from review by the Institutional Review Board, granted on 15 September 2011 (Protocol number 814339, “Survey on Attitudes on Financial Incentives”). Datasets for all three surveys contain a respondent identification number that does not enable anyone outside of *Usamp* or *Knowledge Networks* to identify respondents. All data were analysed and are reported in the aggregate only.

#### 6.2.4 Data and data analysis

This section describes the type of data that were generated and the methods of analysis.

##### 6.2.4.1 Data

Data from US2 were preferences of sets of ordinal or nominal choices. US1 and Ger also provided two scenarios requesting continuous data: first, regarding the number of patients who would ask a physician to lie to obtain an incentive (on a scale of 0–10), and second, regarding the amount of money that should be provided in the three different incentive scenarios (with amounts to be specified between \$0–\$2,500). Ger and US1 furthermore included a text box at the end of the survey for open responses, inviting general comments and feedback. Total number of respondents completing the survey were US2: n=1,114, US1: n=1,000, Ger: n=1,000.

Data were cleaned, and values that suggested that respondents were not taking the survey seriously were set to ‘missing’. For the question of how many patients would request a lie, this was necessary in eight cases in US1 and in 17 in Ger. In the ‘carrot’ scenario, US1 included two invalid amounts, and Ger nine; the ‘false carrot’ three in US1 and one in Ger, and the ‘stick’ one in each. For weight, five entries were set to status ‘missing’ in US1, and four in Ger, where the same was done for height in five cases. 14 respondents in US1, and 31 in Ger skipped the income question: in these cases income data provided by the *Usamp* was used.

##### 6.4.2.2 Analysis

All distributions and statistical analyses were generated using the statistical software JMP Pro Version 9.0.0. To test the distribution of the dependent variables across each level of the categorical or nominal independent variables, mosaic plots were generated and Pearson Chi-square tests performed. Tests compared differences in:

- age (three levels: 20–35, 36–49, 50–65);
- gender (female/male);
- income (three levels: low, middle and high-income, see below);
- education (four levels: less than high school, high school graduate, some college or university, BA or higher degree);



- political orientation (in the case of US1 three levels: democrat, republican and independent/other/no preference, in the case of Ger seven levels: Left Party, Green Party, Social Democrats, Liberal/Freedom Party, Christian/Conservative Party, Right wing, and other/no preference);
- insurance status (in US1: two levels: covered or not covered, in Ger three levels: sickness fund, private health insurance, and other);
- health status (five levels: poor, fair, good, very good, excellent);
- Body Mass Index (three levels, underweight: BMI <18.5, normal: 18.5–25, overweight: 25–30, and obese: >30);
- cancer status (three levels: never had cancer diagnosis, cancer diagnosis although no colon cancer, colon cancer diagnosis); and
- use of colon cancer screening (three levels: no colon cancer screening, colon cancer screening although no colonoscopy, colonoscopy).

Strictly speaking the Chi-square test can determine the significance of differences between groups only in random samples. As noted, it was not feasible for all surveys to be fielded in such samples, and the results for samples US1 and Ger therefore need to be treated with caution. Findings regarding distributions are reported in percentages, and Pearson Chi-square p-value levels indicate statistical significance in between-group comparisons, with alpha set to  $p < .05$ , and all p-values reported exactly (although in figures, for brevity, the convention of [\*] to indicate  $p < .05$ ; [\*\*] for  $p < .01$ , and [\*\*\*] for  $p < .001$  is used). Relatively rare suspect chi-squares (resulting where more than 20% of cells have an expected count of 5 or less) are not generally reported, but the occurrence is stated in the overview Table of statistically significant differences at Appendix 6B, and in the detailed reports in Volume B.

Continuous data regarding the number of people who would request a physician to lie were distributed almost normally, and analysis of variance (ANOVA) was done to determine statistical significance in comparing means. Results are reported by stating the mean (M) and standard deviation of the mean (SD), with significance levels and mode of reporting as described above. Data with regard to the monetary levels that should be provided for particular incentives types were distributed non-normally, with the vast majority at the lower end of the spectrum. While analysing means can nonetheless be interesting and has been done

as part of the background analysis, the focus here is on medians, with means only provided in a general overview, to indicate the spread of data points. Findings are reported stating mean, median and the inter-quartile range (i.e. 25% above and below the median). Significance of differences between distributions within groups was determined using the Wilcoxon/Kruskal-Wallis tests, which merge data points of two or more sets, rank the whole set and then re-sort the ranked values into the initial groups, to measure differences. Significance levels and mode of reporting are provided in the same way described above.

### 6.2.5 Samples

This section describes the sample frame, the way samples were obtained, and key demographic characteristics of respondents in comparison with the general population.

#### 6.2.5.1 Sample Frame

The sample frame for all three surveys comprised adults of working age (20–65), as most US incentive programmes are accessed through employment-based insurance. The upper age range still included people for whom the colon cancer policy would be personally relevant. For US1 and Ger, census-adjusted proportional quota sampling with regard to income was done, to achieve an income distribution that would be similar to the general population. Three main income groups that are commonly used in poverty research were constructed: respondents were grouped in the strata of up to 60% of median gross household income (low-income), those with 60–200% (middle income), and those with above 200% (high-income). To achieve broadly proportional representation within each of the three income groups, the survey company was provided with frequencies for income groups in steps of approximately \$10,000.<sup>2</sup>

---

<sup>2</sup> This approach was straightforward for the US sample, as the Census Office provides detailed income data expressed as median household income. By contrast, the German *Statistische Bundesamt* and other authoritative bodies gathering and analysing such data draw on annual median equivalised disposable household income, and some conversions, using tax data were required. For the US, the median gross HHI in 2010 was \$49,445 [≈£31,970], and the low-income group with <60% of median HHI comprised all those up to \$29,667 (rounded to: less than \$29,999); the middle income group: 60–200% HHI equalled between \$29,668–98,890 (rounded to: between \$30,000 and \$99,999); and the high-income group >200% included all above \$98,891 (rounded to more than \$100,000). For Germany, Median gross (HHI) in 2006 was 32,048 € [≈£27.786], and the low-income group with <60% of median HHI comprised all those up to 19,228€ (rounded to: less than 20,000€); the middle income group: 60–200% equalled between 19,229€–64,096 € (rounded to: between 20,000 € and 64,999€), and the high-income group >200% included all above 64,098€ (rounded to more than

The sample size of 1,000 seemed appropriate in view of the expected variation regarding the question items, and also in view of informal power calculations regarding the experiment on acceptable amounts for the ‘carrot’, ‘false carrot’ and ‘stick’ scenarios, which suggested that, based on the pilot data, a sample of approximately  $n=730$  would be sufficient to detect statistically significant differences.

#### 6.2.5.2 Actual samples: provider and respondent recruitment

Samples for surveys US1 and Ger were provided by the company *Usamp*. The sample for survey US2 was provided by *Knowledge Networks*, which maintains the US’ only nationally representative online access panel. Surveys US1 and Ger were fielded between 25 September and 5 October 2011, and US 2 between 22 September and 3 October 2011. In the case of US2, 52% of those who received an invitation to participate completed the survey (understood as clicking through all pages). Of those who started the survey, 91% per completed it. In the case of US1, of those who were invited, 9% completed it. Of those who started the survey, 96.9% completed it, with 1.5% declining to take part after reading the welcome page, and 1.6% dropping out at later stages. In the German sample, 5 % of all invitees completed, 2.7% declined to participate on reading the welcome page, and 1% dropped out later. Further details about the companies’ approaches to recruitment, use of incentives and comments on participation- and completion rates is provided at Appendix 6C.

#### 6.2.5.3 Actual samples: key demographics

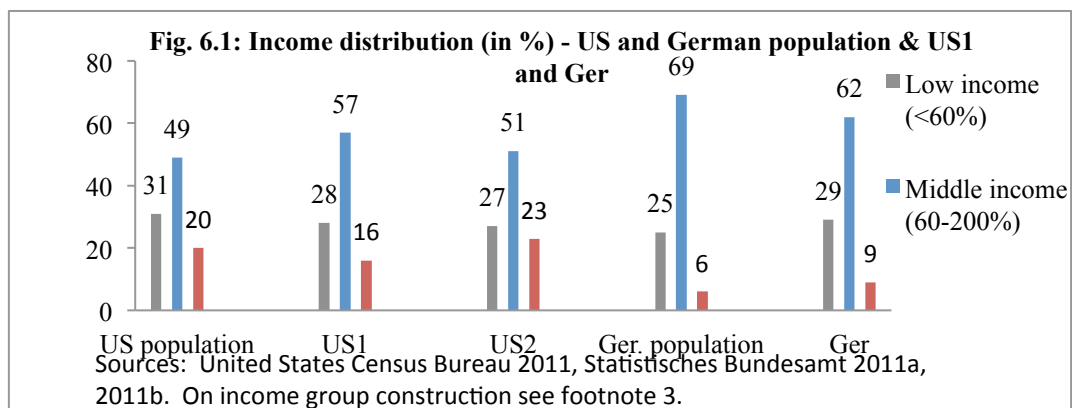
Data for sample US2 were provided by *Knowledge Networks* with weights based on a proprietary algorithm. To ensure representativeness of findings, all data reported here, demographic as well as evaluative, have been generated using these weights. US 2 comprised a total of 1,407 completed surveys. It was not possible to limit the age range for this survey before fielding it, and respondents below the age of 20 and above 65 were removed from the received dataset, leaving 1,115 respondents. US1 and Ger turned out to match census data rather closely with regard to the provided income quota sample definitions. However, due to an oversight by *Usamp*, US1 initially had a proportion of males to females of around 40 to 60. Since gender was likely to play a role for several of the questions, in particular relating to controlling

---

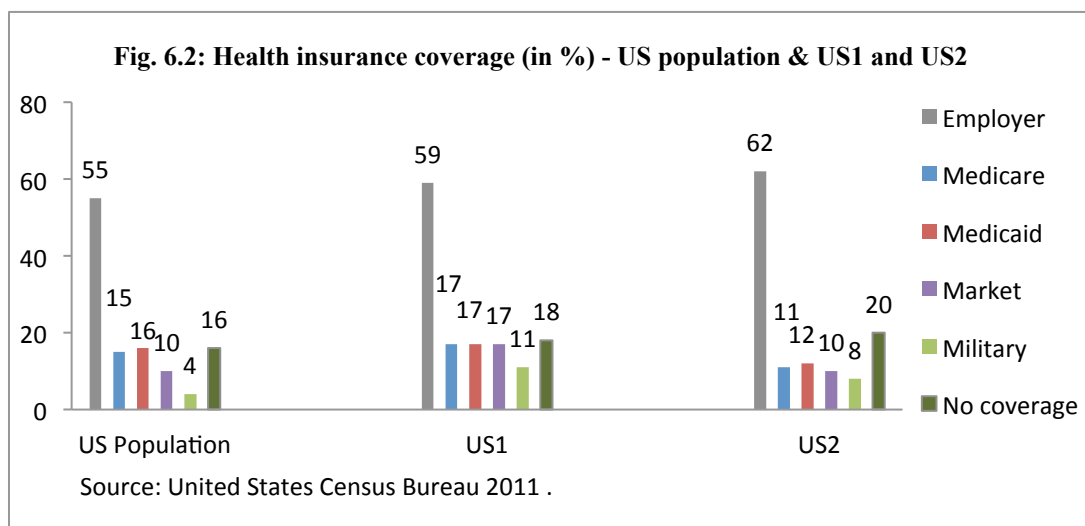
65,000€). Sources: (Statistisches Bundesamt 2011ca: 17,22; United States Census Bureau 2011: 6; Bundesministerium für Arbeit und Soziales 2008: 32). The subdivision within the three general income brackets were constructed using annual income statistics (Statistisches Bundesamt 2011ab: 8).

body weight, the sample was rebalanced by randomly removing 119 females and replacing them with males in equivalent income groups. As will be described below, the final sample US1 approximated the proportions of other key independent variables as available via census and other information in a satisfactory manner. Due to a similar oversight, sample Ger included considerably more young people. Since re-sampling was not possible at the stage at which this became apparent, post-stratification weighing was applied by decreasing the weight of those in the youngest age group (20–35,  $n=562$ ) by a factor of .58 (leading to  $n=305$ ), and increasing that of the oldest group (50–65,  $n=118$ ) by 2.87 (leading to  $n=339$ ), resulting in an overall age structure that was near identical to census data, and also more similar to sample US1. Appendix 6E provides a detailed breakdown of the effect of applying the weights on all other independent variables, and in comparison to census and other reference data to estimate population-level values. In the following, all data for Ger are provided using the age-weights. The final number of respondents in both US1 and Ger was  $n=1,000$ .

Due to the age frame of 20–65 years, the samples were relatively young, US2:  $M=42$   $SD=13$ ; US1:  $M=40$ ,  $SD=13$ ; Ger:  $M=42$ ,  $SD=12$ . Gender was practically equal (ratio female/male in US2: 51/49, US1: 50/50, Ger: 48/52). Figure 6.1 shows the income distribution for all three samples in the constructed groups alongside national data. Educational attainment and political orientation provide a broadly similar close match and are stated at Appendix 6E. As noted, educational data was not obtained in the survey itself, but provided by *Usamp*. Due to the different educational systems, the US categories are not straightforward to apply to the German situation, and there remain some questions about how the company ‘translates’ educational categories. Overall, the data do not seem robust enough to be considered here (see, however Appendix 6B listing the statistical significance of all variables that were tested, and including values for educational attainment in all samples).

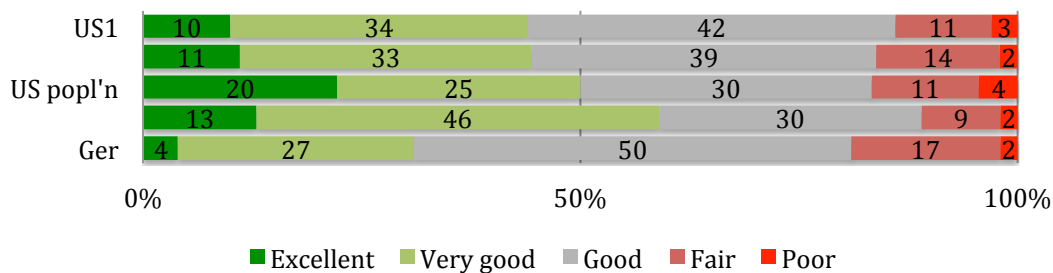


Health insurance status in all samples was very similar to census data. The majority of respondents in the US samples were covered through employer-based insurance and 20% in US2 and 18% in US without any coverage (nationally: 16%), see Figure 6.2. In the German sample, 83% were insured through statutory sickness funds, 15% privately, and 3% through other arrangements, again closely aligned to population level data.



In the US samples, poorer people were significantly less likely to have health insurance with 50% of the low-income group uninsured in US1, as opposed to 45% of the middle income group and 5% of the high-income group ( $\chi^2(2, n=966)=73.84, p<.0001$ , the respective proportions in US 2 were very similar: 56%, 38%, 6%,  $\chi^2(2, n=1,111)=132.51, p<.0001$ ). Self-assessed health status was almost normally distributed in all three samples, with an under-representation of respondents with excellent health, and an over-representation of those with good health in US 1 and US2. The German sample had a similar over-representation of good health, and an under-representation of those with very good or excellent health, see Figure 6.3.

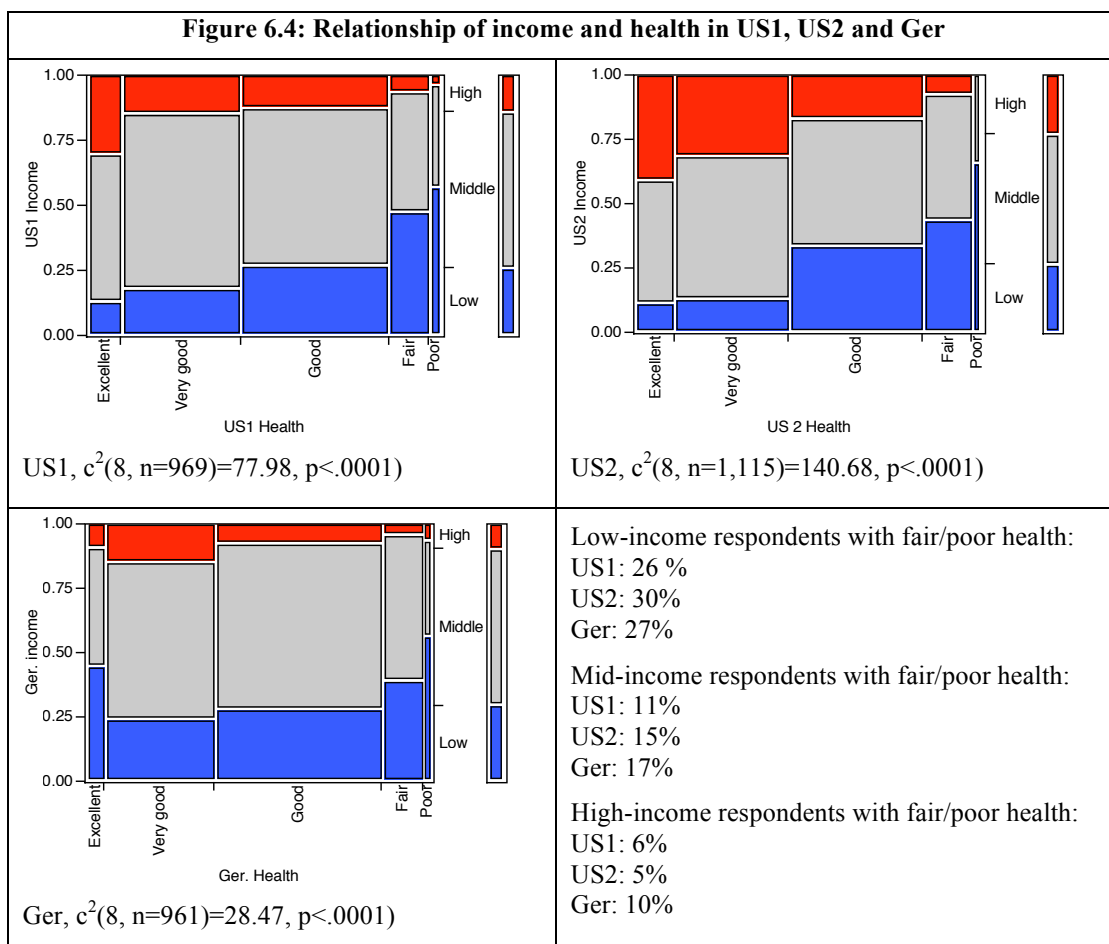
**Figure 6.3: Self-assessed health status (in %): US and German population & US1, US2, Ger**



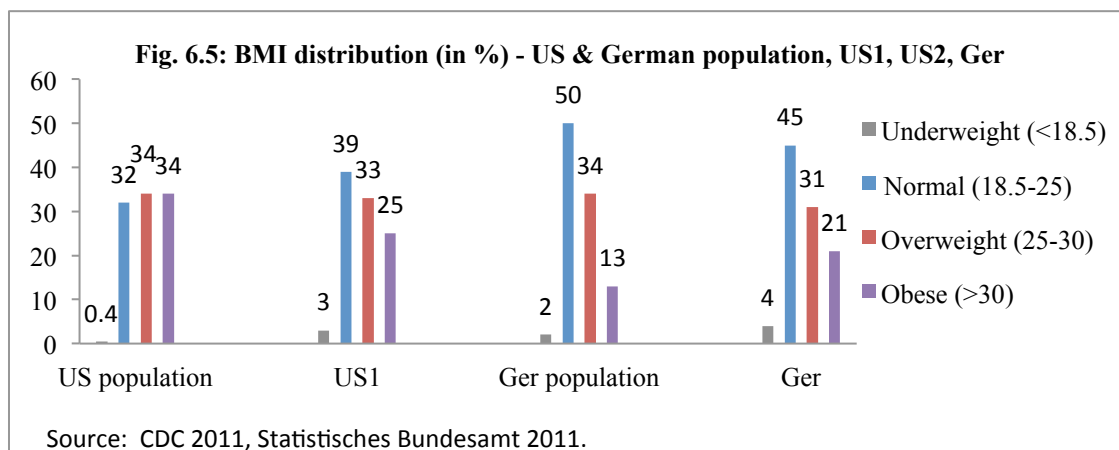
Source: CDC 2011, Statistisches Bundesamt 2011

In all three samples there was a clear and statistically significant relationship between income and health, with the better off generally enjoying better health. Figure 6.4 shows a very similar pattern in samples US1 and US2. Differences were somewhat less pronounced in the German sample, where, moreover, 43% of low-income respondents report excellent health (as opposed to 10% in the high-income group).

**Figure 6.4: Relationship of income and health in US1, US2 and Ger**



For samples US1 and Ger, Body Mass Indices were computed based on the self-reported height and weight. While the US sample includes around 8% fewer obese people and 2% fewer overweight people when compared to the population, it has 7% more people with normal weight. The German sample has 7% more obese people, 4% fewer overweight people, and 5% fewer people of normal weight, see Figure 6.5.



Overall, then, age and gender distributions were near identical to population level data. Income-wise the US samples had a slight over-representation in the mid-income group, while the German sample had more in the low-income group. The American samples had marginally higher levels of respondents without health insurance. In both samples respondents with excellent health were underrepresented, and the proportion of those with good health was higher, with the German sample having an overrepresentation of those with fair health. Poorer people were generally less healthy and in the US less likely to have insurance.

### 6.3 Findings

This section summarises responses regarding past, current and future use of incentives; the extent of approval and disapproval in relation to different incentive policies; what respondents thought about the role of opportunity of choice and effort; what levels were deemed appropriate for financial incentives for weight-control; and what might be learned about the possible impact of incentives on the doctor-patient relationship.

With three surveys, a five point Likert scale, typically more than three variable levels, and a considerable number of significant differences between levels (see Appendix 6B for an overview), a challenge in reporting the data is to strike the right

balance between enabling a clear picture of what the data amount to, on the one hand, and, on the other, avoiding confusion, or losing clarity, by reporting too much data. As noted in Chapter 3, the framework set out there draws to a significant extent on Thomas Scanlon's concept of reasonable rejectability (Scanlon 1998), according to which norms that no-one among a group of people reasonably rejects have particular robustness and in fact special moral value (see Sections 3.4 and 3.5). While it is an open question whether opposing views gathered here reflect grounds for reasonably rejecting a policy, the focus is nonetheless mainly on reporting, in the aggregate, views that strongly disagree and disagree vs. views that *strongly* agree, agree, or are neutral (again, aggregated<sup>3</sup>). 'Agree' and 'strongly agree' are generally only highlighted separately if there is a marked difference between the neutral midpoint position and the two types of agreement and disagreement. Only statistically significant differences between levels in the independent variables are reported. This approach should enable an overview of the most salient variations that are directly relevant to the research questions.

After restating the survey question and overall distributions, the first paragraph of each section provides an overview summary and broad response to the questions under which the survey responses are grouped (highlighted in a shaded box). The subsequent sections then provide more detail regarding the distributions for each of the relevant variables. A more fine-grained picture is available by referring to the detailed statistical reports for each question (see Volume B). Since, somewhat unavoidably, the description of the salience of the independent variables makes for rather dry reading, it can be useful to read all summary boxes first, for a general overview. This cursory approach should also be sufficient to follow the discussion section, although it may be useful to refer back to specific subsections for more context.

### 6.3.1 *To what extent do respondents use incentive programs?*

Respondents to surveys US1 and Ger were asked to comment on their past, current and intended future use of *participation* and *attainment* incentive programmes

---

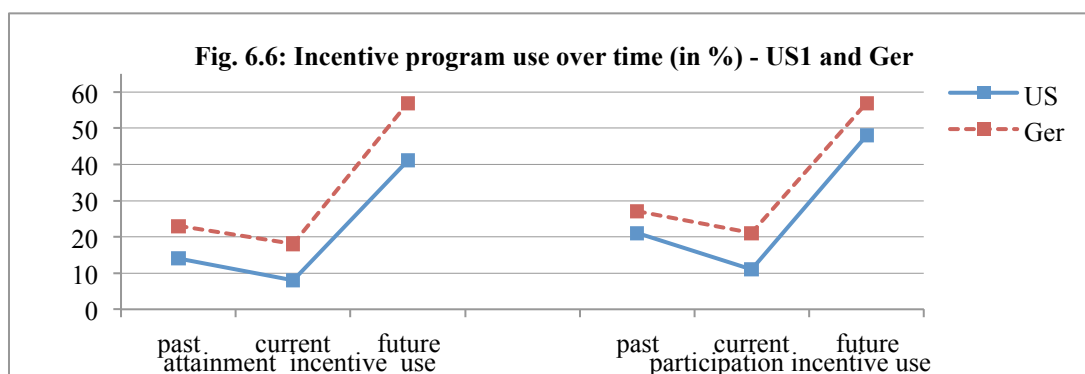
<sup>3</sup> There is some debate as to whether data from Likert scales should be reported in aggregate form, collapsing, for example, strongly agree and agree, or whether, instead, item by item should be reported, as the individual items do not necessarily constitute equal intervals. However, research also suggest that the assumption of equal intervals has some robustness, see (Dawes 2008).



(*participation* programmes, which are generally regarded as less demanding, merely require performing a behaviour that is deemed health conducive, while *attainment* incentives, require meeting specific targets, such as BMI or cholesterol values, which are not equally easy for all to achieve).

**Summary:** Around 10% in US1 and 20% in Ger currently use incentives. Across time, there were significant differences: past use was between 5–10%, and intended future use three to four times higher. Younger people, higher income, better health, and lower BMI groups generally express higher levels of interest, above 50%, although not all differences were significant. Levels among the poor and fair health, the middle and old age groups, as well as of the overweight and obese were moderate, with around a third of these groups showing interest in future use, and 40% of the low-income group showing this interest.

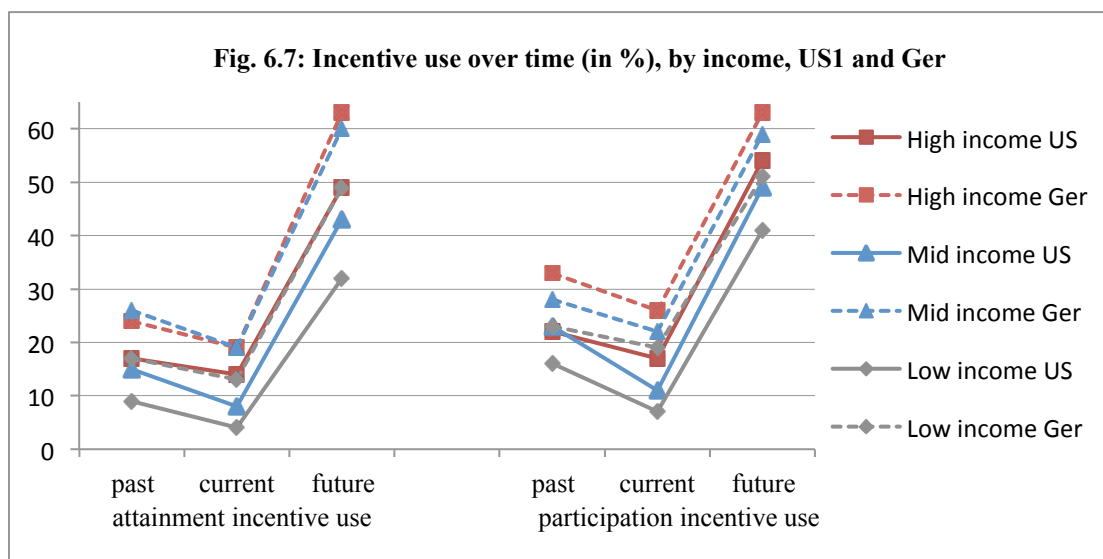
Figure 6.6 shows that overall, around 20 % of the German sample currently use incentives, with 10% in the American one. In both US1 and Ger, past use was 5–10% higher. In the German sample, around three times more people intended to use incentives in the future, and around four times more in the US sample. At all three time points, and for both types of incentives, levels in the German sample were approximately 10% higher than in the US sample. Levels for all three time points were marginally higher for participation than for attainment-incentives. (US1 n=980–988, Ger n=962–970).



**Age:** Younger respondents more frequently reported past and present use, but age was only significant for future attainment and participation -incentive use in US1, with 57% of the young, 39% of the middle age, and 35% of the old age group signalling intention to use attainments, and 53%, 46%, and 41% participation-incentives ([attainment]:  $c^2(2, n=980)=11.77, p=.0028$ ; [participation]:  $c^2(2,$

$n=986$ )=10.47,  $p=.0053$ ). Gender was only significant in the German sample and with regard to the use of attainment-incentives: 26% of males and 20% of females used them in the past,  $c^2(1, n=968)=6.06$ ,  $p=.0148$ , and 21% males as opposed to 14% females use them currently ( $c^2(1, n=969)=7.02$ ,  $p=.0081$ ).

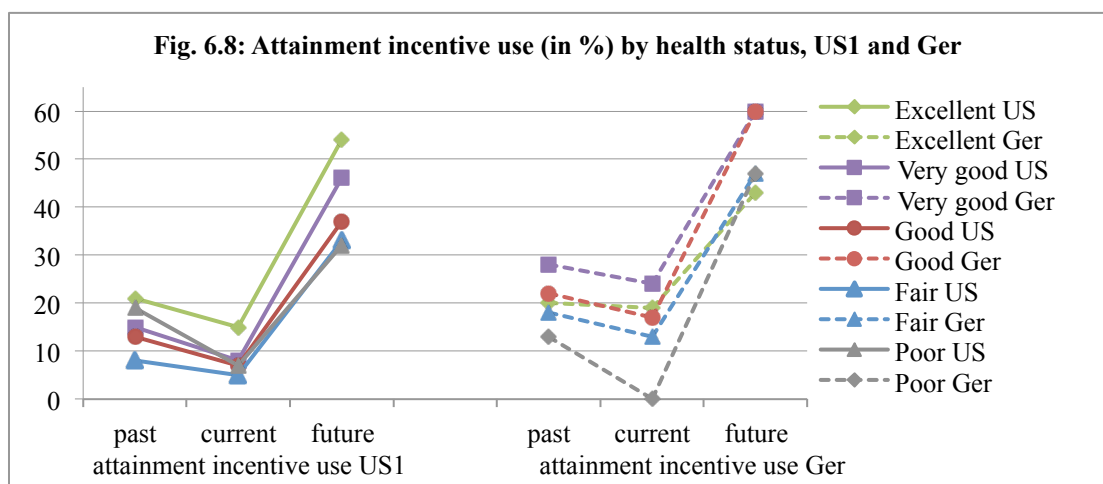
**Income:** Differences were only significant in US1, and here at all three time points for attainment-incentives, and regarding present and future participation - incentive use (Figure 6.7). In the past, 17% of the high-income group used attainment-incentives, 15% of the mid, and 9% of the low-income group. Currently, there are 14% high-income groups user, 8% mid, and 4% low-income group users, with future uptake declining from high to low-income at 49%, 43%, and 32% ([past]:  $c^2(2, n=988)=7.16$ ,  $p=.0279$ ; [present]:  $c^2(2, n=988)=12.29$ ,  $p=.0021$ ; [future]:  $c^2(2, n=980)=12.61$ ,  $p=.0018$ ). Regarding participation-incentives 17% of the high-income group use these currently, 11% of the mid-income group and 7% of the low-income group ( $c^2(2, n=987)=8.06$ ,  $p=.0177$ ). Intended future uptake across these groups declines at 54%, 49% and 41% ( $c^2(2, n=980)=7.08$ ,  $p=.0290$ ).



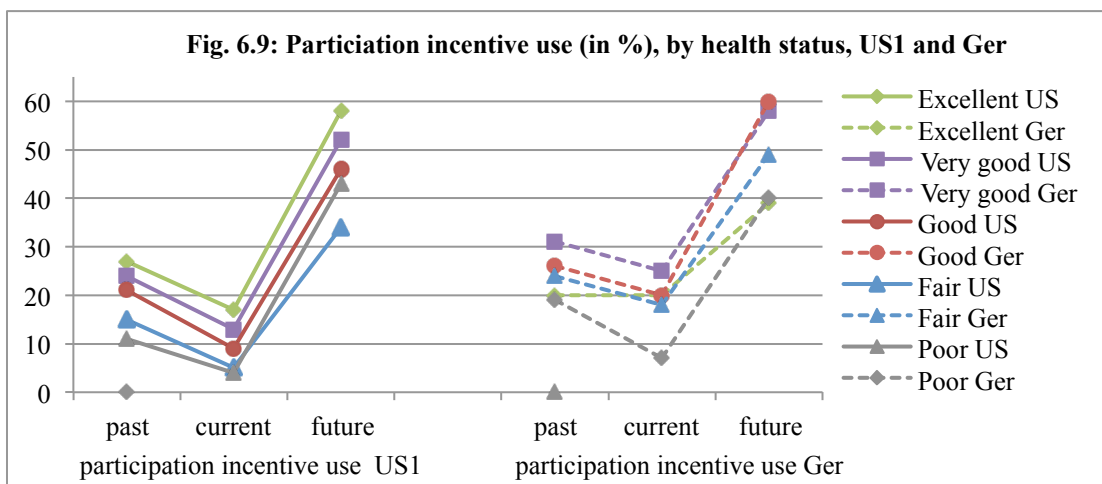
**Health:** Differences were significant in both US1 and Ger, for both types of incentives with regard to intended future use, and in addition for current use of participation-incentives by the American sample, and current use of attainment-incentive use by the German sample. Regarding future use, in US1, 54% of those with excellent health and 46% of those with very good health planned to use attainment-incentives, as opposed to 33% with fair, and 32% with poor health ( $c^2(4, n=958)=15.27$ ,  $p=.0042$ ). Regarding participation-incentives, 58% with excellent

health, and 52% with very good health planned to use them. The order at the bottom end was reversed, with 43% of the poor health group wanting to use them, and 34% with fair health ( $\chi^2(4, n=963)=14.95, p=.0048$ ). In the German sample, the excellent health group ranked lowest in both types of incentives. The highest intended use was here reported by the very good and good health groups, of whom 60% planned to use attainment-incentives. Poor and fair health were both at 47%, with excellent health at 43% ( $\chi^2(4, n=949)=14.27, p=.0065$ ). Of the very good and good health groups an almost equal proportion of 60% and 58% planned to use participation-incentives, with 40% of poor health, 49% of those reporting fair health, and 39% of those with excellent health ( $\chi^2(4, n=952)=14.26, p=.0065$ ).

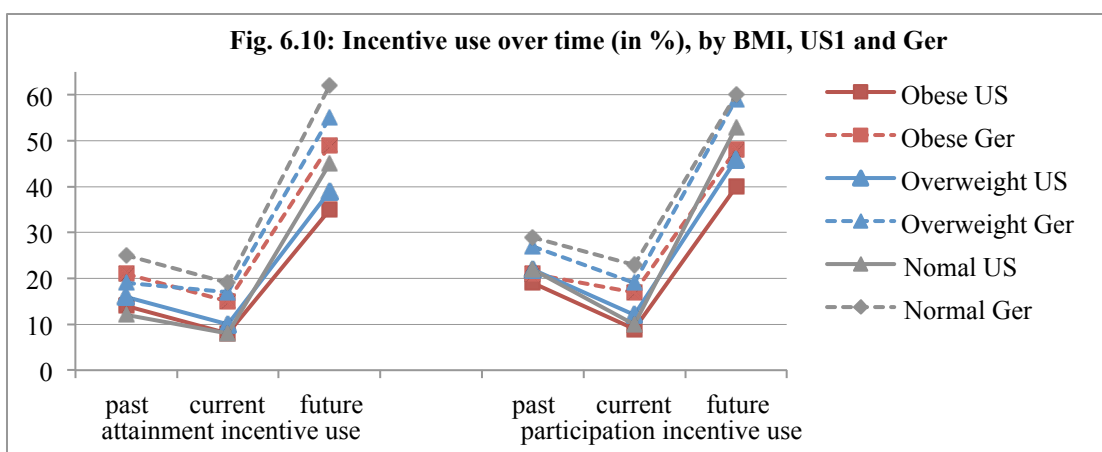
Health was also significant in Ger regarding current use of attainment-incentives: the group with very good health again represents the largest proportion of users (24%), but none of those with poor health use them. ( $\chi^2(4, n=954)=12.65, p=.0131$ ), see Figures 6.8 and 6.9.



In the American sample, health was also significant regarding current use of participation-incentives. Poor and fair health were at the bottom at 4 and 5%, with the excellent health group at a threefold higher level at the other end of the spectrum (17%), followed by those with very good health at 13%, and good at 9% ( $\chi^2(4, n=964)=11.97, p=.0183$ ).



With regard to body weight, BMI was significant for future attainment use in US1 and Ger, as well as for future participation -incentive use in the US sample, and current participation -incentive use in the German sample. Regarding future attainment use in US1, 45% of those with normal weight, 39% of overweight, and 35% of obese intended to use them ( $c^2(3, n=970)=12.94, p=.0048$ ), see Figure 6.10. In the German sample 62% of the normal weight respondents, 55% of the overweight, and 49% of the obese indicated that they would use this type of incentive in the future ( $c^2(3, n=950)=9.37, p=.0247$ ). Regarding future use of participation-incentives in the US, somewhat higher levels of 53% of those with normal weight, 46% of overweight, and 40% of obese respondents shared this objective ( $c^2(3, n=976)=13.00, p=.0046$ ). Figure 6.10 also shows that in the German sample an almost equal proportion of normal and overweight German respondents intended to use them (60% vs. 59%), with a non-significant 7% margin in the case of future attainment use. However, differences regarding current use of participation-incentives were significant, with 23% of the normal weight group using them, 19% of the overweight, and 17% of the obese ( $c^2(3, n=950)=10.16, p=.0172$ ).



Education was only significant in the German sample regarding present and future use of attainment-incentives; political orientation only regarding past attainment and future participation -incentive use in the German sample; and insurance status (in almost all cases) only in the US sample, see Appendix 6B.

### 6.3.2 *To what extent do respondents agree with rewards and penalties?*

Respondents provided approval ratings for five different types of incentive programmes, two relating to weight and three to colon cancer. This section reproduces the text for each scenario, provides a summary of distributions and describes, where relevant, significant differences that were found in analysing the role of independent variables.

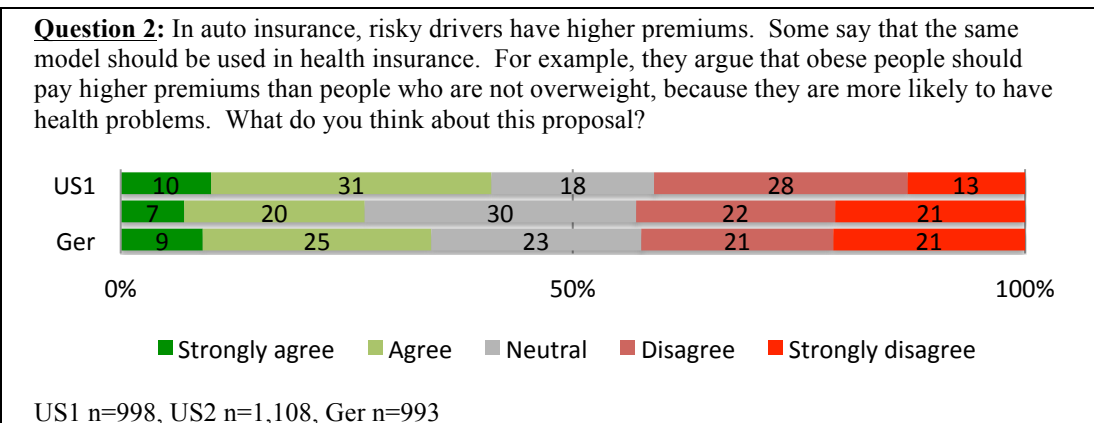
**Section summary:** Regarding the five incentive programmes<sup>4</sup> that involved both ‘carrots’ and ‘sticks’, the majority of respondents agreed with, or accepted them, even though the margin was small: no programme met with more than 42% disapproval. Between 20–35% in the three samples oppose an incentive programme in which those whose weight is normal over a year receive \$150, making it the scenario with the least disapproval. Between 30–40% in the German and US sample were opposed to the colon cancer prevention policy that requires attending counselling and accepting treatment, or else facing doubling of overall co-payments from \$500 to \$1,000, with around 5% higher disapproval in the target age population. Removing the compliance requirement, and changing the structure of the penalty did not significantly change these ratings, as almost identical assessments were found in all three samples regarding a penalty of \$250 for those who do not attend a counselling session, with higher disapproval only in the German sample (again around 5%). Very similar rates were also found regarding the policy that requires undergoing colon cancer screening at the appropriate age, or else facing a \$250 penalty (although in the US samples the target age subgroup had around 5% less opposition, while the German subgroup sample disagreed twice as much, and 5% more of the sample as a whole were opposed). Around 40% in all three samples

<sup>4</sup> 1. Should obese people have higher health insurance; 2. Should those whose weight is normal over a year receive \$150?; 3. Should the co-payment for colon cancer patients be increased from \$500 to \$1,000 if they have not attended counselling on screening at the relevant age and refuse treatment?; 4. Should those who refuse undergoing a colonoscopy face a \$250 surcharge?; 5. Should those who refuse counselling on colonoscopies face a \$250 surcharge?

disagreed with the car insurance analogy, according to which overweight people behave like risky drivers. Overall, by a margin of at least 8%, the majority of respondents therefore supported the policies, or were neutral towards them.

The perception of framing of incentives was tested in two scenarios, regarding the weight-control programme and the colon cancer policy requiring counselling and compliance. Despite considerable differences in demandingness, both programmes were viewed in roughly equal thirds as using either rewards, penalties, or both at the same time. Regarding the weight-control programme, in both samples, more of the better health groups generally took the reward view. In the American sample almost twice as many in the lowest group compared to the highest group viewed the programme as a penalty (36%). Those who had undergone a colonoscopy or other form of colon cancer screening were around ten per cent more likely to take the reward view.

### 6.3.2.1 The car insurance analogy (Question 2)



**Summary:** Overall, a majority of respondents in all three samples supported or accepted the policy. Opposition was around 40%, (41% in US1, 43% in US2, and 42% in Ger). Age was not significant in any sample; however, all other independent variables were. *Gender:* In all three samples women disagreed more than men. *Income:* Around half of the lower income groups were opposed, but only around a third of the highest income groups. *Political orientation:* More on the political right agreed with the policy than the left in the US, although in the German sample the Left and Green Party were closer to right wing parties than more liberal conservative positions. *Insurance status:* Almost twice as many insured through German sickness funds oppose the policy compared to those insured privately, and in the US around

5% more of uninsured disagreed than those with coverage. *Health*: While in all samples around a third of those with excellent health disagreed, the percentages in the poor health group ranged from 57–86%. *BMI*: More than half of the obese in US1 and Ger were opposed, and only a third of those with normal weight. However, between 14–41% of obese and overweight also agreed/strongly agreed.

**Gender:** In all samples, women disagreed more by a margin of around 10%:

- in US1, 45% of females strongly disagreed/disagreed, and 37% of males;
- in US2, 49% of females strongly disagreed/disagreed, and 38% of males;
- in Ger, 47% of females strongly disagreed/disagreed, and 37% of males;

(US1  $c^2(4, n=998)=17.83, p.0013$ ; US2  $c^2(4, n=1,108)=19.64, p.0006$ ; Ger  $c^2(4, n=993)=12.97, p.0114$ ).

**Income:** around half of the low-income group disagreed or strongly disagreed:

47% in US1, 56% in US2 and 50% in Ger. Only 11% agreed or strongly agreed in US1, in US2 it was 32%, and 29% in Ger. Of the high-income groups 31% disagreed or strongly disagreed in US1, 36% in US2, and 37% in Ger. Forty seven per cent agreed or strongly agreed in US1, 41% in US2 and 40% in Ger (US1  $c^2(8, n=998)=28.76, p.0003$ ; US2  $c^2(8, n=1,108)=71.50, p.<.0001$ ; Ger  $c^2(8, n=993)=37.71, p.<.0001$ ).

**Political orientation:** In US2, 38% of Republicans agreed/strongly agreed, but only 18% of Democrats, and 28% of Independents/No preference shared this view ( $c^2(8, n=1,108)=34.99, p.<.0001$ ). Republicans also agreed most in US1 (51%); democrats were at 39%, and Independents/no preference at 37%, however differences were not significant ( $p.0651$ ). In the German sample, 44% of those identifying as Right wing agreed/strongly agreed, 41% of those voting the Green party, 38% of the Left, with the remaining (Social Democrats, Liberal/Freedom Party, Christian/ Conservative and other/no preference) between 32 and 34%. Thirty three per cent of the Right and 31% of Christian/Conservatives disagreed or strongly disagreed, with all remaining between 41–43% (US2  $c^2(8, n=1,108)=34.99, p.<.0001$ ; Ger  $c^2(24, n=960)=68.47, p.<.0001$ ).

**Insurance status:** Forty-three per cent of insured in US1 agreed or strongly agreed, as opposed to 31% of the uninsured. In US2 the ratio was 29% vs. 16%. Disagreement was less strongly marked: 40% of insured in US1 vs. 45% of uninsured; with very similar figures in US2 (43% of insured in US2 vs. 48% of uninsured). In the German sample, 46% of those insured through statutory sickness funds disagreed/strongly disagreed, but only 25% of the privately insured. Of the latter 42% agreed/strongly agreed, and 33% of sickness fund insured do so (US1  $c^2(4, n=984)=13.93, p.0075$ ; US2  $c^2(4, n=1,105)=23.21, p.0001$ ; Ger  $c^2(8, n=965)=36.53, p.<.0001$ ).

**Health:** Differences were significant in all three samples and poorer health groups disagreed/strongly disagreed more: In US1 the ratios, from poor to excellent health, were (all in %): 57, 45, 34, 40% and in US2: 74, 61, 43, 37, 34, with respondents in Ger showing the highest opposition: 86, 61, 40, 34, 35. Around 50% of those with excellent and good health agreed/strongly agreed: 48/49 in US1, 43/36 in US2, and 52/42 in Ger. For those with poor or fair health the figures were 28/35 in US1, 19/11 in US2, and 9/25 in Ger. (US1  $c^2(16, n=967)=59.36, p.<.0001$ ; US2  $c^2(16, n=1,108)=120.18, p.<.0001$ ; Ger  $c^2(16, n=958)=75.76, p.<.0001$ ).

**BMI:** Data on body weight was only gathered in US1 and Ger. In US1 59% of obese disagreed/strongly disagreed, as do 40% of the overweight, 34% of the normal weight. More for completeness sake the category of underweight i.e. a BMI < 18.5 was also included and here 31% disagreed, although this group is generally less the focus of weight-control programmes. In the German sample, 68% of obese, 41% of overweight, 30% of normal weight and 44% of the underweight were opposed. However, 14% of obese in the German sample also agreed/strongly agreed, as did 35% of the overweight, and 44% of both the normal and underweight. In US 1 27% of obese, 41% of overweight, 49% of normal and 44% of underweight agreed/strongly agreed (US1  $c^2(12, n=980)=42.23, p.<.0001$ ; Ger  $c^2(12, n=958)=101.59, p.<.0001$ ).

### 6.3.2.2 Incentives for Weight Control (Question 3)

While the above policy proposal had an explicitly penalising character, a separate scenario relating to weight-control was framed in the opposite way.

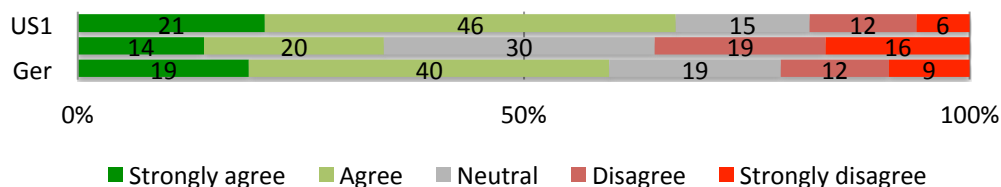


**Question 3:** Consider the following real policy. (Variations of this policy will also be used in the following questions. We will call it the “Incentives for Weight Control” policy)

A large employer wants to create incentives for people to be healthy

Do you think the health plan should use the policy?

- Employees are given the opportunity to be weighed twice a year, once every six months.
- If they are not overweight (in medical terms) on two consecutive weigh-ins, they get a reimbursement of \$150 at the end of the year.
- Participation is voluntary. Employees will not be weighed if they prefer not to be.



US1 n=996, US2 n=1,104, Ger n=992.

**Summary:** Overall, a clear majority supported or accepted the policy, although the proportion of those agreeing was less than half in US2 compared to the other samples. Only 18% in US1 oppose it, 35% in US2, and 21% in the German sample.

**Income:** In the US samples, lower income groups disagreed more than higher ones.

**Health:** In all samples poorer health groups agreed two to three times less than those with excellent health. **BMI:** Twice as many of the obese in both US 1 and Ger opposed the policy than those with normal weight. However, there was also a sizeable proportion of obese and overweight who agreed or strongly agreed (between 44–69%).

**Income:** While age and gender were not significant in any of the samples, income was significant in both US samples, with lower income groups showing less approval. In US2, 52% of the low-income groups disagreed/strongly disagreed, with 31% in the middle group, and 26% in the high-income group ( $\chi^2(8, n=996)=24.90, p<.0001$ ). In the same groups in US1 the figures were 22%, 15% and 16% ( $\chi^2(8, n=996)=24.90, p=.0016$ ).

**Health:** Health was significant in all samples, with the more unhealthy more likely to disagree/strongly disagree: in US2 58% of poor health and 53% of fair health were opposed, while only 27% of those with good health and 28% of those with excellent health took this view. In US1, the ratios were lower but followed a similar pattern: 25% of poor health disagreed/strongly disagreed, 23% with fair health, 15% of those with very good and 7% of those with excellent health. In the German sample, those with better health disagreed more than those with fair health, although the proportion of those opposed was highest among those with poor health

(poor: 51%, fair 23%, very good: 32, excellent 27%; US1:  $c^2(16, n=965)=52.34$ ,  $p<.0001$ ; US2:  $c^2(16, n=1,104)=69.73$ ,  $p<.0001$ . Ger:  $c^2(16, n=958)=38.00$ ,  $p=.0015$ ).

**BMI:** Weight was significant in both US1 and Ger, with higher weight groups being less supportive. In US1 23% of obese disagreed/strongly disagreed, as opposed to 16% of the overweight, 13% of those with normal weight, and 18% of underweight. In the German sample, 35% of obese were opposed, 21% of the overweight, 15% of people with normal weight, and 21% of the underweight. In both US1 and Ger the fraction of people holding the neutral position was between 15–20% in all groups, resulting in around 1/2 to 2/3 agreeing or strongly agreeing: 57% of obese in US1 (44% in Ger), 69% of overweight in US1 (59% in Ger) and 75% of normal weight in US1 (68% in Ger; US1:  $c^2(12, n=978)=31.28$ ,  $p=.0018$ ; Ger:  $c^2(12, n=958)=50.73$ ,  $p<.0001$ ).

### 6.3.2.3 Colon cancer prevention incentives: counselling & compliance (Question 10)

Three scenarios related to colon cancer prevention, with the first representing an abstracted version of actual German policy (see Section 2.3.1).

**Question 10:** Now some questions about another real example of an incentive policy. Please read the description carefully, as the policy is a little bit complicated.

This policy is used in a country where copays for things like prescriptions, hospital or doctor visits, are capped: employees with average income never pay more than \$1,000 annually in total. And because chronically ill face a higher burden of illness, they only need to pay half of that: a maximum of \$500.

To encourage better colon cancer care, a health plan wants employees to do two additional things, if they want to access the lower threshold of \$500.

1. Plan members who are older than 55 years need to attend a counseling session on the advantages and disadvantages of colon cancer screening (such as colonoscopies, or fecal occult blood test).

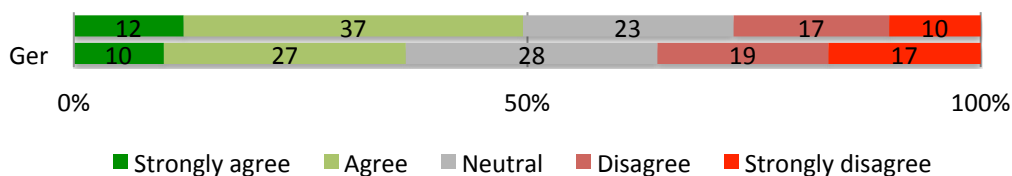
2. If they should later on require therapy for colon cancer, they should never refuse medically recommended treatment, for example, chemotherapy after an operation.

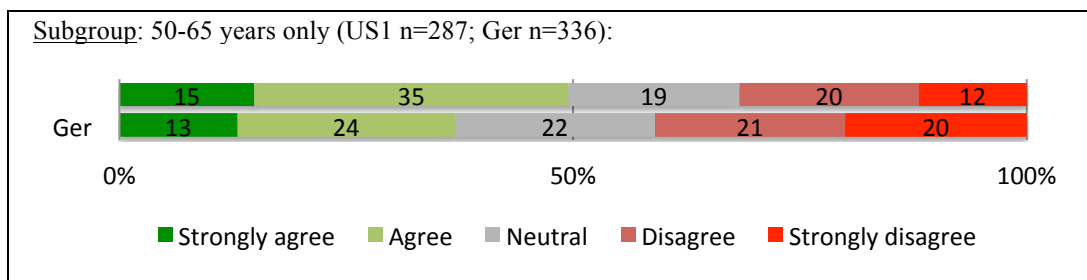
Employees who suffer from colon cancer and have not complied with the two conditions will therefore have a maximum of \$1,000 in copays. Employees who have complied with the two conditions have half of that (\$500).

The health plan says that the policy is justified because colon cancer patients who comply with the two conditions help control health care cost.

What do you think about this policy?

All age groups (US1 n=989, Ger n=983):





**Summary:** Overall, a clear majority accepted or supported the policy in both countries. Twenty-seven per cent were opposed in sample US1, and 32% in the target population subgroup of this sample. In the German sample, 36% of the complete sample were opposed, and again a slightly higher number in the target population subgroup (41%). *Age:* More than a third of the older age groups in both US1 and Ger were opposed, but this group also had the highest percentage of those agreeing strongly. *Gender:* Women were more likely to disagree. *Income:* The lowest income groups showed most disapproval in the American sample, although in Ger sample, and especially in the subgroup representing the target population, the middle income groups were least supportive (around one third in both cases, rising to 50% in the subgroup). *Insurance status:* Uninsured US respondents and those privately insured in Germany were less supportive than their counterparts. *Health:* Half or more in both US1 and Ger with poor health were opposed. Disapproval declined as health improved, although in Ger sample those with excellent health disapproved more than those with good or very good health. *Cancer screening:* In US1 more than a third of those who had colon screening other than colon cancer were opposed, but only around a quarter of those who had a colonoscopy or no screening at all. An inverse pattern was found in the German sample, where around a quarter of those who had colon cancer screening other than a colonoscopy were opposed, and more than two third of the remaining groups, with a more pronounced pattern in the target population subgroup.

**Age:** In both US1 and Ger there were higher levels of older respondents disagreeing/strongly disagreeing: across the three age groups 24% of in the youngest group were opposed, 28% in the middle group, and 32% in the oldest group. In the German sample, the levels were higher at 29, 35 and 41%. While the overall proportion of those agreeing or strongly agreeing was roughly constant across groups in both samples (around 50% in US1 and around 35% in Ger) the proportion of those

strongly agreeing mirrored those opposed, with the older groups showing higher levels of approval: in US1 15% in the oldest group agreed strongly, 13% of the middle and 10% of the youngest group; in the German sample 13% of the oldest agreed strongly, 11% in the middle group, and 7% in the youngest group (US1:  $\chi^2(8, n=989)=16.29, p=.0348$ ; Ger:  $\chi^2(8, n=983)=24.56, p.0018$ ).

**Gender:** Women disagreed more in both US1 and Ger and both subgroups, although gender was only significant in the complete German sample, where 40% of women as opposed to 31% of men disagreed strongly. Ger:  $\chi^2(4, n=983)=13.90, p.0076$ .

**Income:** Differences were significant in both complete samples and the German subgroup, although distributions in the US and German sample varied: in US1 poorer groups were more opposed, but in the German sample it was the middle income group: 30% of the low-income group in US1 were strongly opposed or opposed, 27% in the middle income group, and 22% in the high-income group. In Ger, 32% of the low-income group disagreed strongly, 38% in the middle group, and 25% of the high-income group. A more pronounced pattern was found in the target group in Ger: percentages of respondents opposed across the groups ranged from low to high-income at 28, 50, and 27% (US1:  $\chi^2(8, n=989)=22.60, p.0039$ ; Ger:  $\chi^2(8, n=983)=15.54, p.0495$ , Ger [50–65 years]:  $\chi^2(8, n=336)=21.28, p.0064$ ). Education and political views were significant in both German samples (see Appendix 6B).

**Insurance status:** Significant differences were found in both complete samples and the German subgroup. US1 showed only minor differences among those covered and those not covered: 26% vs. 27% for strong disagreement/disagreement, but 51% vs. 41% regarding agreement/strong agreement. In the full German sample, 36% of sickness fund insured disagreed/strongly disagreed, and 33% of privately insured (agreement/strong disagreement at 38% and 34%, respectively, US1:  $\chi^2(4, n=982)=13.19, p.0103$ ; Ger:  $\chi^2(8, n=967)=19.18, p.0139$ ).

**Health:** There were significant differences in both US1 and Ger, although Chi-Square square was suspect in both target-age-subgroups, as more than 20% of cells had an expected count of 5 or less, since the group with poor health was small in both samples (15 in US1 and 23 in Ger). In US 1, more of those with poor health

than those with better health disagreed or strongly disagreed (poor: 50%, fair: 32%, good: 33, very good and excellent both 25%). In the complete German sample the distribution was generally similar, although those with excellent health showed more disapproval (poor: 57%, fair: 43%, good: 35, very good: 29%, excellent: 42%; US1:  $c^2(18, n=965)=39.42, p.0009$ ; Ger:  $c^2(16, n=961)=53.33, p.<.0001$ ).

**Cancer:** Differences were significant in the complete German sample and the subgroup, however, Chi-Square square was suspect, as few respondents had a colon cancer diagnosis (and even smaller in the subsets): nine respondents in the complete samples of US1, and 12 in Ger had had a colon cancer diagnosis, while 31 in US1 had a diagnosis for a cancer other than colon cancer, and 69 in Ger, with the remainder never having had any cancer diagnosis (938 in US1 and 886 in Ger). A more detailed analysis is therefore not meaningful, but reports are in volume X.

**Cancer screening:** Having had colon cancer screening was significant in both complete samples and the German subgroup sample: in US1, 26% of those who had a colonoscopy disagreed strongly/disagreed, while 36% of those who had another form of colon cancer screening shared this view, and 27% of those who had no screening whatsoever, with a similar, although not statistically significant pattern in the subgroup of 50–65 year olds. In the complete German sample, an inverse pattern can be observed: 35% of those who had a colonoscopy were opposed, while 27% with other forms of colon cancer screening, and 37% of those with no screening disagreed/strongly disagreed. In the target population subgroup this pattern becomes more pronounced (40% vs. 27% and 48%; US1:  $c^2(8, n=979)=17.20, p.0281$ ; Ger:  $c^2(8, n=967)=17.98, p.0214$ ; Ger [50–65 years]:  $c^2(8, n=330)=16.45, p.0364$ ).

#### 6.3.2.4 Incentives for colon cancer prevention: counselling (Question 12)

Two separate versions explored whether attitudes differed if the compliance requirement was removed: one scenario required counselling only, the other went further and linked the incentive (same amount) to undergoing screening.

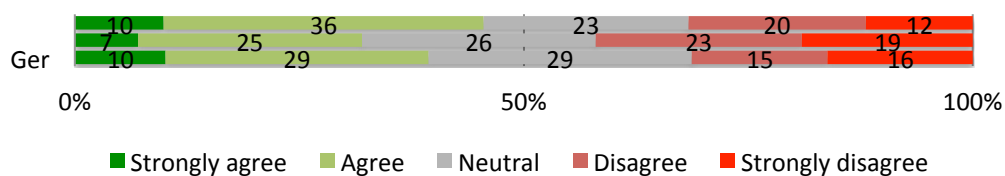
**Question 12:** A health plan determines that colon cancer screening for older employees would be good for their health and would save money. The plan wants to encourage employees who are 55 years old to be screened and tested.

The company proposes to send employees a letter with an invitation to attend a counseling session on the advantages and disadvantages of colon cancer screening (such as colonoscopies, or fecal occult blood tests). If employees do not attend a counseling session within a three year period, a one-time \$250 surcharge will be added to their insurance premiums.

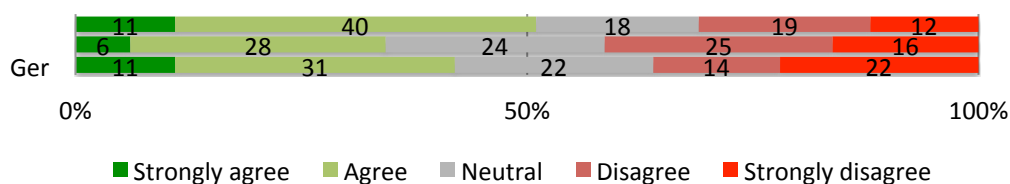
A spokesperson of the health plan says: “Colon cancer screening is good for the employee and

for the health plan, as we can cover other medical needs from savings that come from the early detection and treatment of cancer. Employees don't have to be screened if they prefer not to be. But both economic and medical considerations support the use of a one-time surcharge to encourage employees to attend an information session that will enable them to make informed decisions." What do you think about this policy?

All age groups (US1 n=989, US2 n=1,112 Ger n=976):



Subgroup: 50–65 years only (US1 n=287; US2= 392, Ger n=333):



**Summary:** Again, the majority either supported or accepted the policy. Opposition was very similar in the samples overall compared to the target population subgroups: 32% of both the complete sample and the subgroup in US1 were opposed; 42% of the complete sample in US2, and 41% in the subgroup; and 31% of the entire German sample, and 36% of this sample's subgroup. *Age:* In the American samples middle age groups disagreed the most, with half in US2 being opposed. In the German sample this pattern was not observed, and most in the oldest group were opposed, although only around a third. *Gender:* Differences were only observed in US2 and its target population subgroup, where women were around 10% more likely to be opposed. *Income:* Significant differences existed in the complete sample and subgroup of US2, where around a quarter of the low-income, but half of the high-income group were in favour, with opposition very similar across groups around 45%. By contrast, in the German target population subgroup, a third more of the higher income group were opposed compared to the lower groups, and they were also twice less likely to be in favour. In US2, almost half of those insured oppose the policy (in the complete sample and subgroup), and closer to a third of the uninsured. *Insurance status:* In the German sample those with sickness fund insurance were marginally more opposed than those with private insurance. *Health:* In the US samples healthier groups were generally more likely to welcome the policy, a trend not found in the German sample. Opposition across all health groups in the US

samples was around one third, and slightly higher in the German sample. *Cancer screening*: Colon cancer screening use was only significant in US1 where opposition across groups was around one third, but around half of those who had a colonoscopy or other colon cancer screen welcomed it. Approval among those who never had any form of screening was around ten per cent lower.

**Age:** In both US1 and US2 the middle age group disagreed the most, although differences were only significant in US2: 38% of the young, 50% of the middle, and 14% of the old-age group disagreed/strongly disagreed. In the German sample the elderly (36%) were more opposed (versus 29% of the young and middle-age group).

**Gender:** In all samples males were the most likely to agree strongly, but differences were only significant in the complete sample US2 as well as the subgroup of US2: 45% of females disagreed/strongly disagreed, as opposed to 35% of males. In the subgroup the figures were lower, but similar (disapproving females: 45%, males: 35%; US2:  $\chi^2(4, n=1,112)=14.10, p.0070$ ; US2 [50–65 years]:  $\chi^2(4, n=392)=12.35, p.0149$ )

**Income:** There were significant differences in US2 in the complete sample and the subgroup, as well as in the German subgroup. While disagreement was around 45% across all groups (low and middle: 42, high: 45), and around 10% higher than in US1 (low: 33, middle: 32, high 28), agreement/strong agreement increased with higher income in both American samples, although more in a more pronounced way in US2, where 25% of the low-income group, 33% of the middle and 51% of the high-income group were in favour. This pattern was not found in the complete German sample, and in the German subgroup higher income groups were more likely to disagree, and less likely to agree: 30% in the low-income group disagreed/strongly disagreed, and 38% of the middle, and 45% of the high-income group share this view. Forty-five per cent each of the low and middle income group agreed or strongly agreed, but only 18% of the high-income group do so. US2:  $\chi^2(8, n=1,112)=43.79, p.<.0001$ ; US2 [50–65 years]:  $\chi^2(8, n=392)=20.59, p.0038$ ; Ger [50–65 years]  $\chi^2(8, n=333)=21.41, p.0061$ )

**Insurance status:** Insurance was significant in Ger and its subgroup (suspect Chi-square, however), and in US2 and its subgroup. In US2, 45% disagreed/strongly

disagreed, (43% of the target population subgroup) and 37% of the uninsured (29% of the subgroup). Insured were also more likely to agree/strongly agree than the uninsured: 33% were in favour (35% in the subgroup) and 26% of the uninsured (29% of the subgroup). In US1 insured were equally more likely to agree, although the differences were not significant. In Ger, 33% of those with sickness fund coverage disagreed/strongly disagreed, and 28% of respondents with private coverage. 39% of both sickness fund users and those with private coverage agreed/strongly agreed (US2:  $\chi^2(4, n=1,109)=20.53, p.0004$ ; US2 [50–65 years]:  $\chi^2(4, n=392)=13.95, p.0074$ ; Ger:  $\chi^2(8, n=965)=29.92, p.0002$ ).

**Health:** Differences were significant in all three samples and US2's subgroup (although the subgroups' Chi-Squares were suspect). In US1 and US2 healthier groups were more likely to welcome the policy, with less differences regarding opposition. Disagreement/strong disagreement in US1 was around one third (poor–excellent, all in %: 32, 36, 35, 28, 36%). In US2 levels were higher, around 45% (poor–excellent, all in %: 44, 49, 40, 44, 47%). Ger centred around 35% (poor–excellent, all in %: 38, 38, 31, 29, 35). Groups in Ger showed minimal variation regarding agreement/strong agreement, where all groups ranged between 37–41%. In US1 however a clear gradient could be observed: 57% of excellent health approve, very good: 49%, good: 41%, and fair: 40%. But more of those with poor health agreed than the two prior groups: 46%. A similar pattern with those of poor health agreeing more was given in US2, where 31% agreed/strongly agreed, but only 21% of those with fair health, and 28% with good health. More of those with very good health (40%) than with excellent health (32%) shared this view (US1:  $\chi^2(16, n=965)=27.18, p.0395$ ; US2:  $\chi^2(16, n=1,112)=53.27, p.<.0001$ ; Ger:  $\chi^2(16, n=958)=28.40, p.0283$ )

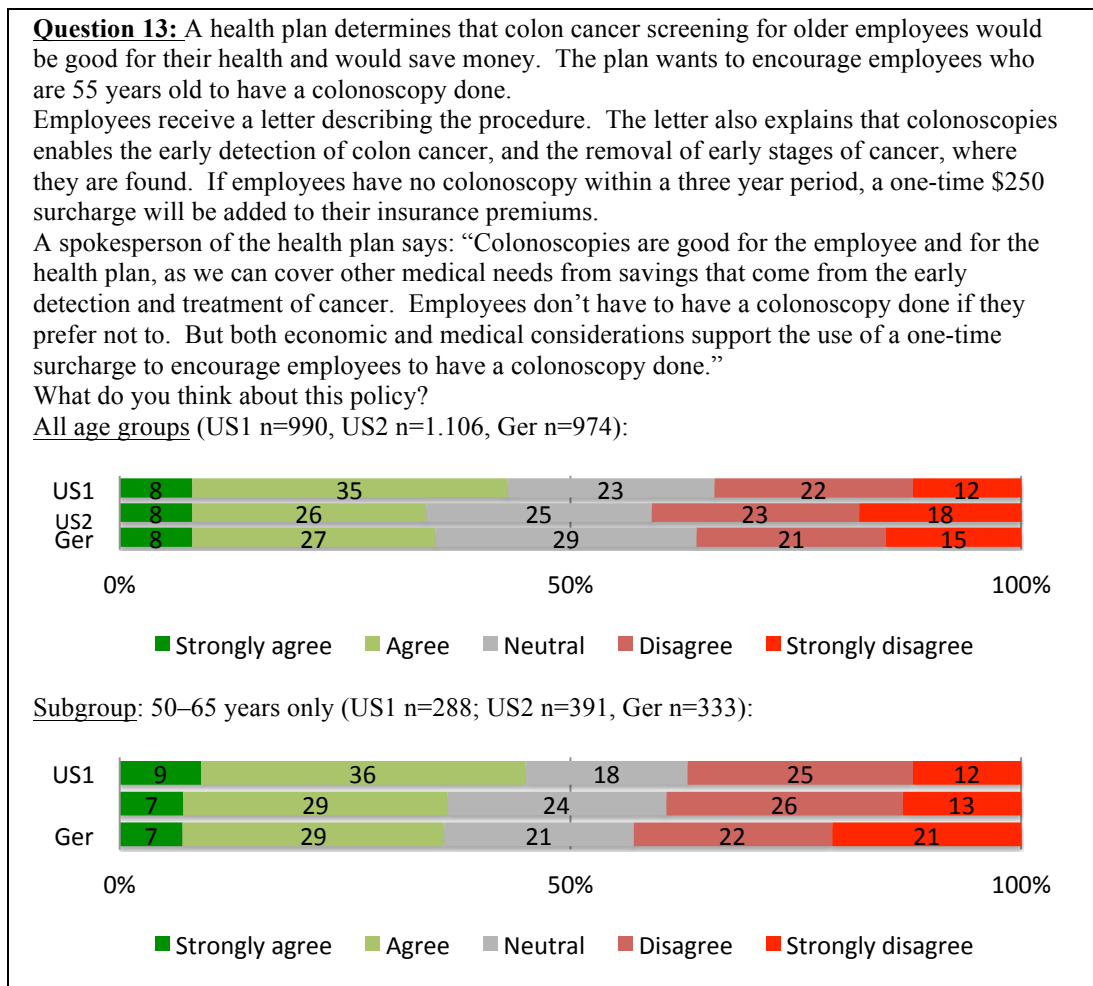
**Cancer screening:** Screening use was only significant in the complete sample US1 (but not in the target population subgroups). Disagreement/strong disagreement was around 30% for all groups (those who had had a colonoscopy: 29%, those who had colon cancer screens other than a colonoscopy: 33%, and those who had no screening: 32%). However, 55% of those who had had a colonoscopy agreed/strongly agreed, as did 51% of those who had screens other than a



colonoscopy, but some ten per cent less of those who had no screening at all (42%).  
 US1:  $\chi^2(8, n=979)=26.18, p.0010$ .

6.3.2.5 Incentives for colon cancer prevention: screening (Question 13)

Responses to a variation of the above scenario, requiring screening, rather than just attending counselling, were overall very similar, as the following analysis shows.



**Summary:** Again, in all samples a majority supported or accepted the policy. 34% of the entire sample US1 were opposed and 27% of the subgroup with the target population, age-wise. In US2, somewhat higher 41% were opposed, and 39% in the subgroup. In the German sample, 36% of all respondents were against the policy, and 43% in the subgroup. *Age:* Older respondents were generally more likely to support the policy than other age groups, but also had higher levels of those opposed. *Gender:* Women disagreed more and men were more in favour, but only US2 had a significant gender difference of around ten per cent. *Income:* In US1 and US2 higher income groups were more supportive, with roughly one third more endorsing the

policy in the highest as opposed to the lowest income group. Income was not significant in the full German sample, but in the subgroup an inverse relationship could be observed, with higher income groups opposed in a similar ratio. *Health*: In all samples, those with better health tended to have higher levels of disagreement, but also of agreement, than those with poorer health. *Cancer screening*: Use of screening correlated with approval: half of those screened in the US welcome the policy, and around ten per cent less in the German samples do so. Still, around a third of those who had been screened were opposed in US1, Ger, and Ger's subgroup.

**Age**: In all three samples the oldest group was more likely to agree/strongly agree, but differences were only significant in US2 and Ger. In the latter, as in US1, the oldest group also disagreed most. In US2, 33% of the youngest group disagreed/strongly disagreed, 32% of the middle group, and 43% of the oldest group. In the German sample the middle age group disagreed the most at 45%, with the youngest at 40%, and the oldest group at 39%. Regarding approval, 36% of the oldest group in the German samples agreed/strongly agreed, 33% of the middle-aged group, and 31% of the youngest. In US2 the same progression, and very similar percentages could be observed, with 36% of the old group in favour, 35% of the middle group, and 32% of the youngest (US2:  $\chi^2(8, n=1,106)=19.52, p.0123$ ; Ger:  $\chi^2(8, n=974)=29.12, p.0003$ ).

**Gender**: While marginally higher levels of females were opposed in all three complete samples, and higher levels of males tended to be in favour, gender was only significant in the target population subgroup of US2. Here, 43% of females disagreed/strongly disagreed, as opposed to 35% of males. Forty-three per cent of males agreed/strongly agreed, and this was the case in 30% of females (US2:  $\chi^2(4, n=391)=15.38, p.0040$ ).

**Income**: Differences were significant in the complete samples US1 and US2, and in the subgroup of the German sample. In both US samples, agreement varied with income with higher groups more supportive: in US1 36% of the low-income group, 44% of the middle, and 51% of the high-income group agreed/strongly agreed; in US2 the levels were lower but similarly spaced: 26% in the low-income group, 35% in the middle group, and 39% in the high-income group. Opposition in

US2 was around 40% (low: 38%, middle: 42%, high: 44%), while in US1 the middle income group has most reservations at 37%, with the low-income group at 32% and the high-incomers at 26%. In the German subgroup sample, a similar pattern regarding disapproval could be observed, with 46% of the middle income group disapproving, 39% in the low and 36% in the high-income group. However, approval was the inverse to the US samples: 44% of the low-income group agreed/strongly agreed, 35% of the middle, and 18% of the high-income group agreed (US1:  $\chi^2(8, n=990)=24.40, p.0020$ ; US2:  $\chi^2(8, n=1,106)=35.58, p.<.0001$ ; Ger [50–65]:  $\chi^2(8, n=333)=25.57, p.0012$ ).

**Insurance status:** Insurance was only significant in the full sample of US1 and Ger's subgroup (Chi-Square square however suspect in Ger). In US1, disapproval was similar between insured (33% disagreed/strongly disagreed) and uninsured (36%). However, more than 10% more of the insured agreed/strongly agreed (45%) compared to the uninsured (36%; US1:  $\chi^2(4, n=983)=10.37, p.0346$ ).

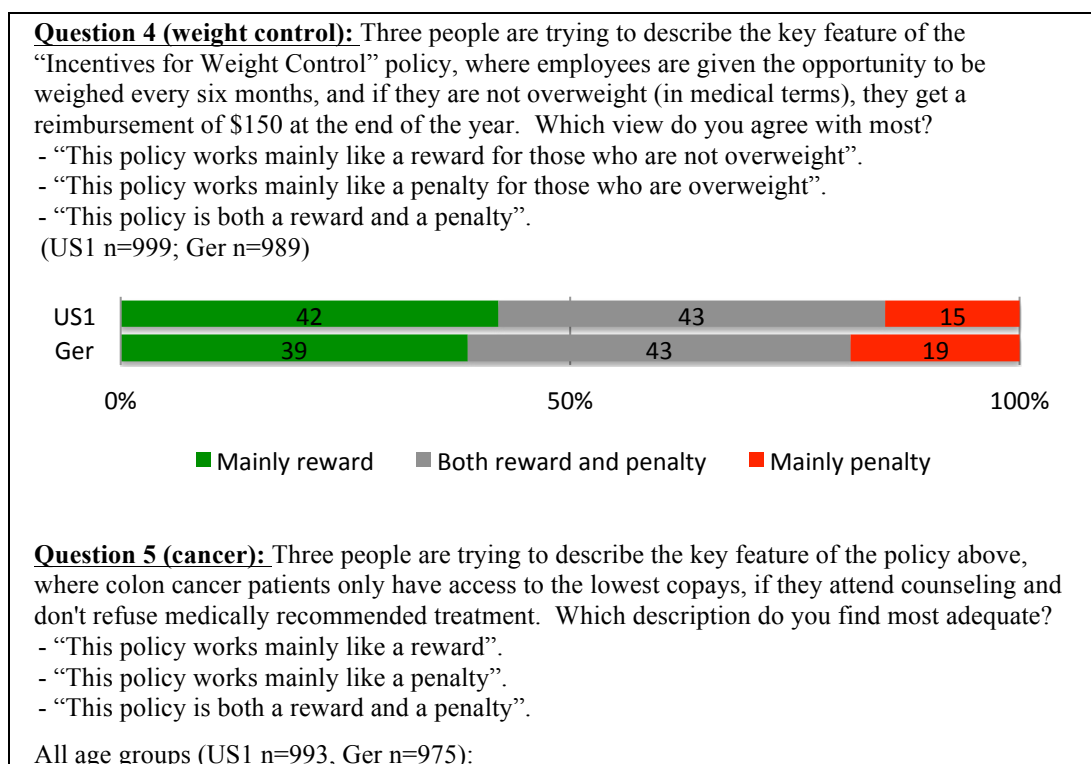
**Health:** There were significant differences in all three complete samples as well as in the subgroups, although Chi-Square square was suspect in all subgroups due to low numbers in some cells, even if overall distributions were very similar. In general, those with better health tended to have higher levels of disagreement, but also of agreement. Disagreement/strong disagreement in US1 progresses as follows: excellent health: 29%, very good, and good, both: 35%, fair: 38%, poor 43%. In the German sample 38% of those with excellent health were opposed, very good: 32%, good: 36%, fair: 42%, poor 43%. In US2, however, a more even pattern was given: excellent: 43%, very good: 38%, good: 43%, fair: 39%, poor 41%. Those with better health approved more in US1 and US2, where the progressions from excellent to poor health were as follows (all figures in %: US1: 55/45/40/37/32; US2: 40/40/28/25/39). The German sample was more even (all figures in %: 38/38/32/38/24). US1:  $\chi^2(16, n=966)=34.56, p.0046$ ; US2:  $\chi^2(16, n=1,106)=53.54, p.<.0001$ ; Ger:  $\chi^2(16, n=956)=40.30, p.0007$ ).

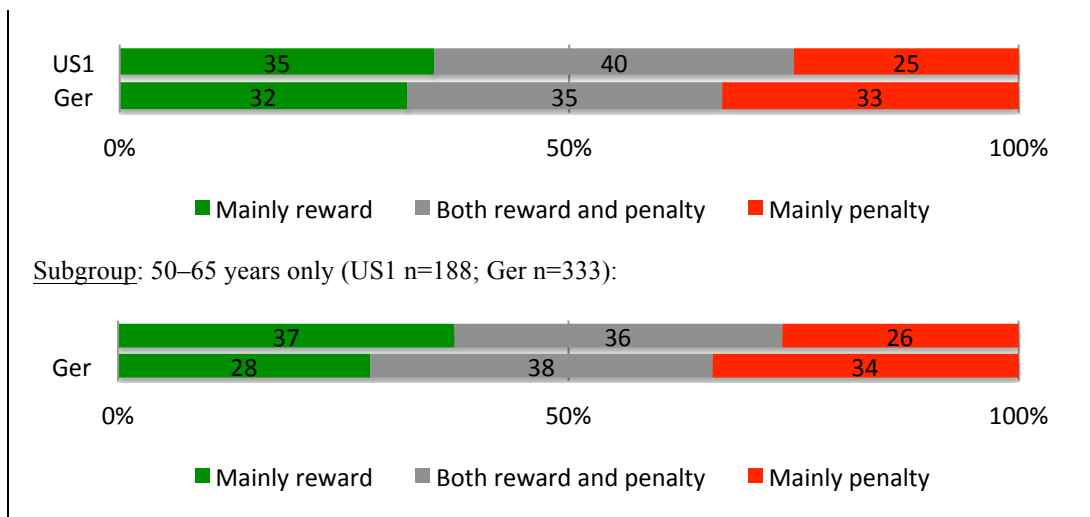
**Cancer screening:** While the analysis of cancer status (asked in US1 and Ger) was again not meaningful because of low numbers, screening use was significant in the full sample of US1 and Ger, as well as the German subgroup of the target

population. In US1, in all three groups (those who had had a colonoscopy; those who had colon cancer screens other than a colonoscopy, and those who had no screening at all) 34% disagreed/strongly disagreed. However, 50% of those who had had a colonoscopy agreed/strongly agreed, as did 47% of those who had colon cancer screens other than a colonoscopy and 40% of those who had no screening at all. In the German sample those who had used some form of cancer screening were less opposed: 31% of those who had had a colonoscopy; 35% with screenings other than colonoscopies, and 38% of those who never had any screening. In the target population subgroup, less of those who had a colonoscopy disagreed (27%), but more of those who had some other form of screening (43%) and more than half of those who had no screening (56%). Differences regarding approval were similarly marked between the full sample and the subgroup (subgroup figures are in square brackets): of those with colonoscopies, 41%[49%] agreed/strongly agreed, as did 36%[37%] with other forms of screening, and 32%[28%] of those with no screening (US1:  $c^2(8, n=980)=22.00, p.0049$ ; Ger:  $c^2(8, n=963)=15.56, p.0491$ ; Ger [50–65]:  $c^2(8, n=330)=36.89, p.<.0001$ ).

#### 6.3.2.6 Incentives and perceived framing: Carrots? Sticks? (Question 4)

Respondents were also asked to state to what extent they perceived the initial weight-control and colon cancer prevention policy as ‘carrots’, ‘sticks’ (or both).





**Summary:** Respondents' overall perspective on the two scenarios was strikingly similar in both US1 and Ger: neither was perceived by the majority as mainly a reward programme, with roughly a third viewing them as either as a penalty, or a reward, or as a penalty and reward at the same time. However, more of the healthier people tended to adopt the reward frame, with significant differences for both policies in the American sample, where two thirds of those who rated their health as excellent took this view for the obesity programme, but just a third of those with fair health, with a similar margin, though lower levels, regarding the cancer scenario.

**Gender:** While age was not significant in any sample, gender was significant in the complete German sample, where more males than females viewed both programmes as rewards: 43% males vs. 34% females perceived the weight-control programme in this way, with lower levels, but the same ratio for the cancer policy, 37% vs. 26% (Ger [obesity]:  $c^2(2, n=989)=8.29, p.0158$ ). Ger [cancer]:  $c^2(2, n=975)=12.23, p.0022$ ).

**Health:** Differences were significant in sample US1, and for both scenarios. Regarding the weight-control policy, 66% of those with excellent health perceived the programme as reward, 44% of those with very good health, 39% with good, 33% with fair, but also 43% of those with poor health (almost as many as those with very good health). The penalty view was most common in those with fair health (29%), good and poor health (both 14%), very good (11%), and excellent (8%; US1:  $c^2(8, n=968)=46.27, p.<.0001$ ). The German sample had a similar, but far more attenuated, (and non-significant) slope. Regarding the cancer policy, the penalty

view correlated with poorer health in US1, with almost twice as many in the lowest group compared to the highest group taking this view: poor health: 36%, fair: 33%, good: 26%, very good and excellent: both 20%. Forty-seven per cent of those with excellent health view it as a reward, but only 18% of those with poor health (US1:  $c^2(8, n=968)=18.75, p.0161$ ).

**BMI:** Body weight differences were significant in the German sample regarding the weight-control policy: 27% of the obese saw the programme as a reward, with the normal and overweight at 42%. Thirty per cent of the obese saw it as a penalty, with those of normal weight at 17%, and the overweight at 14% (Ger:  $c^2(6, n=955)=25.50, p.0002$ ).

**Cancer screening:** In the American sample those who had had colon cancer seeing were more likely to view the policy as a reward (42% of those who had a colonoscopy, 47% of those who had a colon cancer screen other than a colonoscopy, as opposed to 35% of those who had no such exam (the penalty view was taken by between 23–25% in all three samples; US1:  $c^2(4, n=982)=11.40, p.0224$ ).

**Other variables:** Income was not relevant in either Ger or US1. Education and political orientation for both policies were only significant in the German sample, and insurance status only for the weight-control policy.

### *6.3.3 To what extent do respondents value opportunity of choice, motivation and effort?*

Three sets of questions explored different aspects of the role of personal behaviour in relation to other factors that can determine health. The introductory question asked in general terms about the relative importance of different factors. A later question asked whether respondents found it fair or unfair if it was the case that lower income groups might be less likely to benefit from the aforementioned weight-control programme in which participants receive \$150 if their weight is normal. Five different variations of this policy then explored attitudes towards people who differ in their opportunities for behaviour change, preferences and success in reaching the target.

**Overall section summary:** Respondents in samples US1 and Ger demonstrated considerable optimism with regard to the role of personal behaviour,

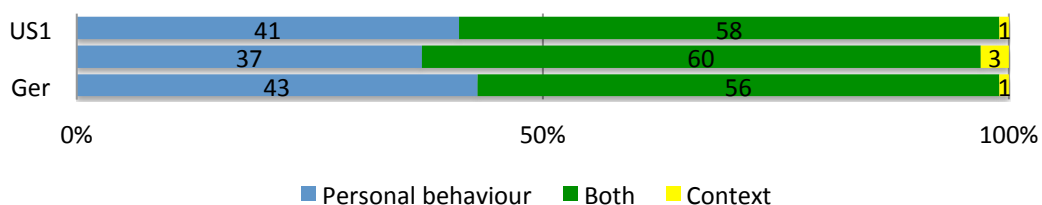
which between 37–42% judge to be the most important factor, with higher levels in higher income, and better health groups. Only 1–3% assigned this role to other factors exclusively, and the remainder choose the both/and option. Around 40% in both US1 and Ger found it unfair if lower income groups benefitted less from the weigh control incentive, which increased to one half in the low-income groups, but fell to one fifth in the high-income groups. Poorer health and higher BMI groups also disagreed more, with around one half of obese respondents disagreeing in both US1 and Ger and a third of the normal weight group. Regarding the five scenarios the overall ranking was consistent across both samples. Respondents found the most fair that people who had made genuine efforts to lose weight received the incentive—only around 10% had objections. The next most accepted scenario was that those who already show healthy behaviour should receive incentives, with similarly low levels of objections. Third came the vignette in which those who are overweight because they prefer food over exercise lose out on the incentive. Fourth, that people who made genuine efforts at weight loss, but were unsuccessful should not receive an incentive, although more than one fifth disagrees. The lowest approval was given to the scenario in which people received incentives even if their motivation ran counter to the programmes’ goals, with around 25% opposed.

### 6.3.3.1 Factors that influence health (Question 1)

**Question 1:** How healthy we are can depend on many things, including access to health care, the quality of our air and water, and our genes. Income and education can matter for good health as well. And personal behavior plays a role, including what (and how much) we eat and drink, and whether we exercise and get enough sleep.

When it comes to our health, how important is personal behavior in relation to other factors?

- How we behave is the most important influence on our health.
- Personal behavior is important, but so are other things: it is impossible to say which is most important.
- Compared to other factors influencing health, personal behavior is less significant.



(US1: 998, US2: 1,103, Ger: 993)

**Summary:** In all three samples, a majority was of the view that it is impossible to decide whether personal behaviour or other factors were more important. However,

between 37–43% were also of the view that personal behaviour is the most important factor. *Gender*: Less women tended to view personal behaviour as the most important factor in US1 and Ger, with a 10% difference in the latter sample. *Income*: In both US samples, higher income groups viewed personal behaviour as more important, a difference not observed in the German sample. *Health*: More than half of those with excellent health opted for personal behaviour as the most important factor, and only around half as many in the poor health group.

**Gender**: Age was not significant in any sample. Gender was only significant in the German sample, where 62% of women, but only 50% of males thought that personal behaviour was as important as other factors. Women were also more likely to decide on this option in US1, although the difference was not significant; in US2 the split across gender was exactly equal. Ger:  $c^2(2, n=993)=17.12, p.0002$ .

**Income**: In Ger marginally more middle income respondents viewed personal behaviour and other factors as equally important. In US1 and US 2 higher income groups valued personal behaviour more. Differences were only significant in US2: 32% in the low-income group viewed personal behaviour as most important, as opposed to 35% in the middle and 46% in the high-income group. The low-income group also had most respondents suggesting that other factors were more important than personal behaviour: 5% as opposed to 3% in both of the other groups. US2:  $c^2(4, n=1.103)=18.73, p.0009$ .

**Health**: Significant differences were found in all samples,<sup>5</sup> with more than half of those reporting excellent health selecting personal behaviour as the most important factor determining health. Percentages for personal health were as follows:

- US1: excellent: 54%, very good: 48%, good: 34%, fair: 40%, poor: 36%
- US2: excellent: 60%, very good: 40%, good: 31%, fair: 28%, poor: 33%
- Ger: excellent: 53%, very good: 50%, good: 39%, fair: 42%, poor: 22%

US1:  $c^2(8, n=976)=21.09, p.0069$ ; US2:  $c^2(8, n=1,103)=41.13, p.<.0001$ ; Ger:  $c^2(8, n=956)=17.93, p.0217$ ).

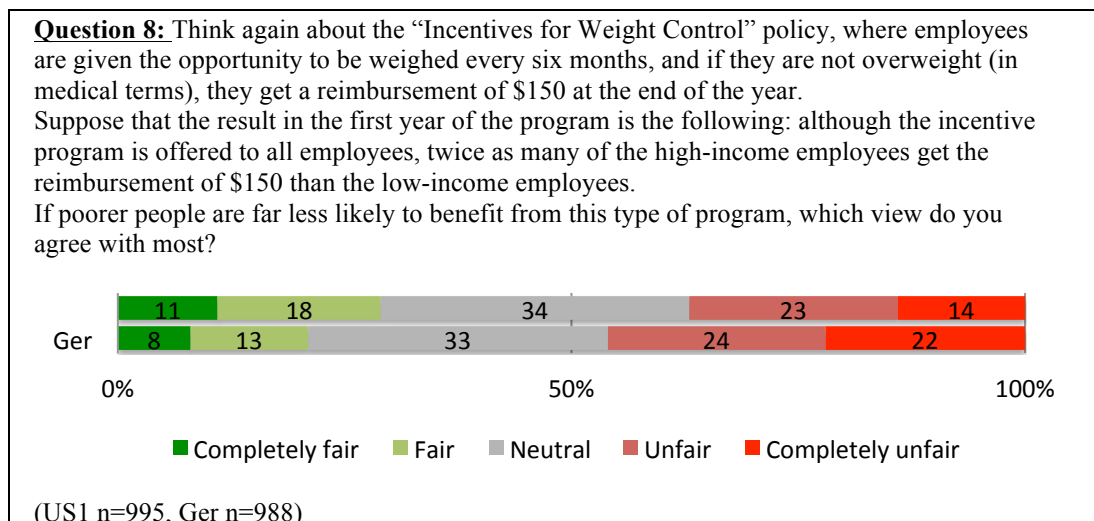
---

<sup>5</sup> Chi-square was suspect for US 1 and Ger; however this was chiefly because of low cell counts in the "Context" column across the five health groups: in US1, frequencies of respondents choosing this option from excellent to poor health were: 1/2/5/1/0, and in Ger 1/3/3/2/0. Numbers for the remaining two options were solid, however, see Volume B.



**Other variables:** BMI was only significant in US1, but only because 64% of the underweight group opted for personal behaviour: all other groups showed only marginal differences (US1:  $c^2(6, n=980)=12.65, p.0488$ ). Education was significant in US2, and political orientation in the German sample (although Chi-Square was suspect). Insurance status was not significant in any sample.

### 6.3.3.2 Incentives for weight-control and the social context (Question 8)



**Summary:** The majority of respondents agreed with the policy or accepted it. However, in the German sample almost half disagreed (46%) with 37% in US1. *Age:* Older respondents in both samples disapprove the most, and more women than men tended to find the policy unfair. *Income:* In both samples half of the low-income groups had fairness concerns, but only around a fifth of the high-income groups. *Political orientation:* Left leaning political position generally had most reservations, and those identifying with market-friendly or conservative positions found it mostly fair. *Health and BMI:* Fairness concerns also correlated with health and Body Mass Index in both US1 and Ger, with poorer health groups around twice as likely to find the policy unfair, compared to those with excellent health. Around a third of those with normal weight disagreed in both US1 and Ger but around half of the obese.

**Age:** Age was significant in both samples with older groups showing most disapproval, in the case of the German sample also most approval. In US1 35% of the young and middle age group found it unfair/completely unfair, and 42% of the oldest respondents took this view. In the German sample 44% were opposed in the

youngest group, 41% in the middle group and more than half, 53% in the oldest group. Around 29% of all groups in US1 found the policy fair/completely fair. In the German sample, 20% of the young and middle age group agreed in this way, and 23% of the oldest. US1:  $\chi^2(8, n=995)=29.56, p.0003$ ; Ger:  $\chi^2(8, n=988)=35.31, p<.0001$ ).

**Gender:** More women found the policy unfair/completely unfair in both US1 and Ger, but only in the German sample was this difference significant, where 51% of women took this view, but only 41% of men. Ger:  $\chi^2(8, n=988)=10.21, p.0370$ ).

**Income:** Differences were significant in both US1 and Ger: lower income groups disagreed the most, and agreed least. In US1 50% in the lowest group found the result unfair/completely unfair, but only 34% in the middle, and 26% in the high-income group. In the German sample 54% of the low, 44% of the middle, and 22% of the high-income group shared this view. Conversely, 29% in Ger found it fair/completely fair, as did 22% of the middle and 16% of the low-income group. In US1, 49% of the high-income group were in favour, 31% of the middle, and 19% in the low-income group. US1:  $\chi^2(8, n=995)=44.57, p.<.0001$ ; Ger:  $\chi^2(8, n=988)=18.06, p.0208$ ).

**Political orientation:** Education was significant in both groups, as was political orientation; in US1 43% of Democrats found the policy unfair/completely unfair, as opposed to 29% of Republicans. Of the latter, 37% opted for fair/completely fair, but only 27% of Democrats. In the German sample the three highest levels of objections were found in the Left Party (61%), Right Wing Parties (55%) and the Green Party (50%), with the remainder between 32% (Conservative/Christian Party) and 48% (Social Democrats). Liberal Democrats had the highest percentage of those who found it fair/completely fair (48%) followed by the Conservative/Christian Party (29%) and the Green Party (22%), with the remainder between 14% (Left Party) and 21% (Social Democrats; US1:  $\chi^2(8, n=968)=23.75, p.0025$ ; Ger:  $\chi^2(24, n=962)=50.80, p.0011$ ).

**Health:** Differences in US1 and Ger were significant and the less healthy tended to find the policy more unfair. This trend was particularly clear in Ger. The proportions of respondents finding the policy unfair/completely unfair were as follows:

- US1: excellent: 27%, very good: 30%, good: 39%, fair: 51%, poor: 65%
- Ger: excellent: 39%, very good: 43%, good: 44%, fair: 67%, poor: 80%

The percentages of respondents finding the policy fair/completely fair were inversed in the American sample, but only marginally different in the German sample:

- US1: excellent: 42%, very good: 36%, good: 25%, fair: 21%, poor: 11%
- Ger: excellent: 22%, very good: 23%, good: 20%, fair: 19%, poor: 13%

US1:  $\chi^2(16, n=968)=60.07, p.<.0001$ ; Ger:  $\chi^2(16, n=960)=77.40, p.<.0001$ .

**BMI:** Weight was also significant in both samples, with heavier respondents more likely to find the policy unfair. 32% of normal weight respondents in US1 disagreed; 37% of the overweight and 48% of the obese. In Ger rates were higher across the spectrum: 39% of the normal weight disagreed, 44% of the overweight and 63% of the obese. In the German sample 9% of the obese found the policy fair or completely fair, and 24% of the overweight and those of normal weight. In US1 19% of the obese share this view, 31% of the overweight and 32% of the normal weight respondents. US1:  $\chi^2(12, n=981)=39.77, p.<.0001$ ; Ger:  $\chi^2(12, n=959)=65.98, p.<.0001$ .

### 6.3.3.3 Incentives for weight-control: who benefits? (Question 9)

**Questions 9a–e:** *[note: in the survey, the question sequence was randomized. Responses are here organized in descending order of overall approval: the first scenario was found to be the most fair, and the last one the least fair, aggregating the ratings “completely fair” and “fair”]*

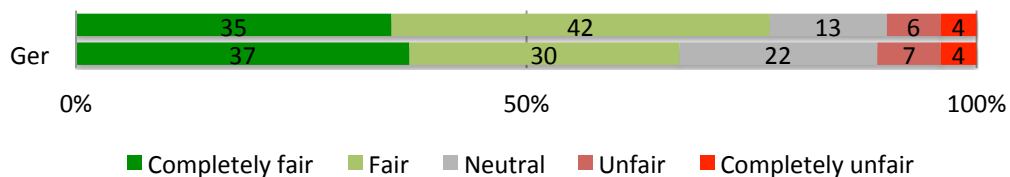
Again, think about the “Incentives for Weight Control” policy where employees are given the opportunity to be weighed every six months, and if they are not overweight (in medical terms), they get a reimbursement of \$150 at the end of the year.

Below, five different people describe what they do to be healthy:

- All are covered by a plan that offers such a program.
- All have average incomes and are around 45 years old.
- Some are overweight, some are not.

As a result, some get the incentive amount, and some don't. People may disagree about whether this is fair or not. Consider the people's situations. How fair or unfair, would you say, are the outcomes for each of them?

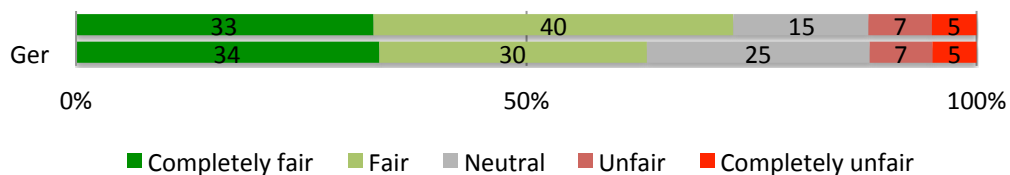
**9a:** “I like to eat, but dislike exercise, and have been overweight for a long time. However, I wanted to be slimmer. This program was a good opportunity to go on a diet and exercise twice a week. I am no longer overweight, and have been successful in keeping my current weight.”  
The statement is accurate. In line with the policy, this person gets \$150. Is this fair?



(US1 n=913, Ger n=981)

**9b:** “I am health-conscious by habit, always eat healthily, exercise regularly, and have not been overweight for years.”

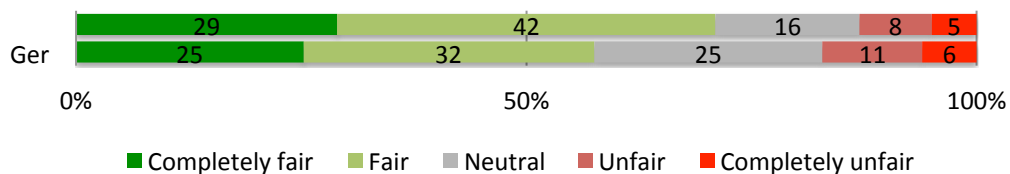
The statement is accurate. In line with the policy, this person gets \$150. Is this fair?



(US1 n=904, Ger n=980)

**9c:** “I like to eat, but dislike exercise, and have been overweight for a long time. But food is really important to me. I like living this way, and honestly don't mind the extra pounds.”

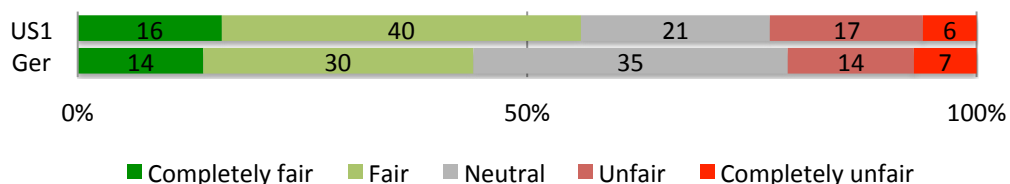
The statement is accurate. In line with the policy, this person does not get \$150. Is this fair?



(US1 n=910, Ger n=972)

**9d:** “I like to eat, but dislike exercise, and have been overweight for a long time. However, I wanted to be slimmer. This program was a good opportunity to go on a diet and exercise twice a week. I tried several diets, but didn't manage to stick to one. Some weeks there was time for exercise, but other weeks there just wasn't. I really tried and lost some weight, but am still overweight.”

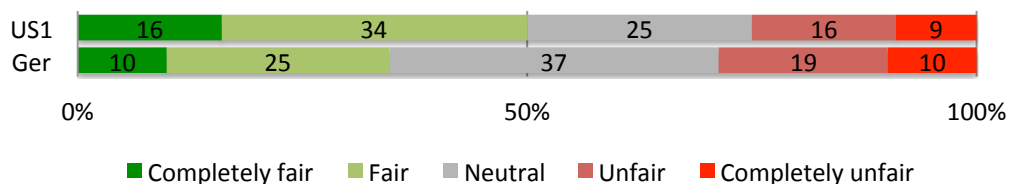
The statement is accurate. In line with the policy, this person does not get \$150. Is this fair?



(US1 n=910, Ger n=976)

**9e:** “I don't think about my health, eat mostly unhealthily, never exercise, but have not been overweight for years.”

The statement is accurate. In line with the policy, this person gets \$150. Is this fair?



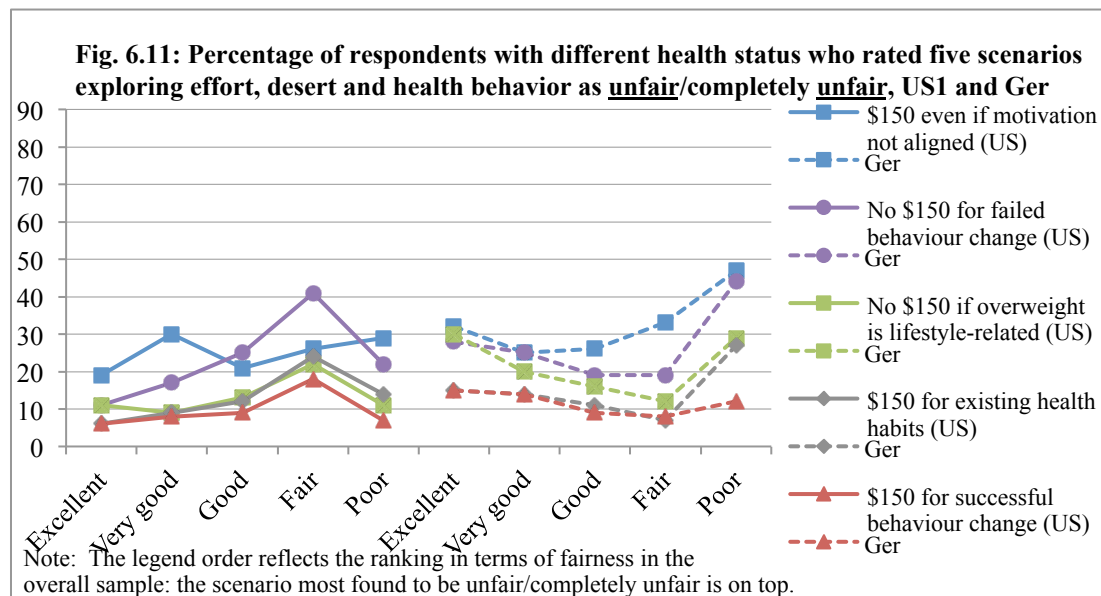
(US1 n=911, Ger n=976)

**Summary:** The five questions sought to explore attitudes towards appropriateness of incentive rewards for people who differed in their efforts at achieving the targets, as well as their underlying motivations. Overall, there were only minor differences between US1 and Ger regarding opposition. However, across all scenarios, the American sample had a higher percentage of those who found the scenarios fair/completely fair. Across all cases, and in both US1 and Ger, the scenarios that had the highest levels of approval also had the lowest levels of disapproval.

- The scenario that met with most approval related to the prototype of the responder group and concerns a person who successfully changes their behaviour, in line with the policy's goal: 77% in US1 and 67% in Ger found it fair/completely fair. Opposition was at 10% in US1 and 11% in Ger.
- The runner-up was the scenario where a person receives an incentive not for behaviour change, but for continuing to live healthily, with their motivation falling in line with the policy's goal. 77% in US1 found this fair/completely fair and 64% in the German sample; with 12% opposed in both samples.
- The third most accepted scenario was that a person who is overweight as a result of valuing food more than exercise should not receive the incentive amount: 71% in US1 agreed with this outcome, and 57% in Ger, with 13% opposed in the American and 17% in the German sample.
- Fewer people in both samples found it fair that someone who unsuccessfully attempted to lose weight, should *not* get the incentive: just over half in US1 found this fair/completely fair (56%) and just under half in the German sample (44%). Twenty-three per cent in US1 found this unfair/completely unfair, and 21% in Ger.
- The lowest approval and highest disapproval related to receiving an incentive without behaviour change, and motivation running counter to the policy's goal: 50% in US1 find this fair/completely fair but markedly less in the German sample (33%). Opposition was around a third in both samples (29% in Ger and 25 in US1).

Regarding independent variables, the focus of this section is on health status and body weight, as these are most central to the question of control that is being addressed here, and, moreover were significant in both US1 and Ger in all scenarios. An overview of other statistically significant variables can be found at Appendix 6B.

**Health:** Differences were less pronounced regarding perceived unfairness (Figure 6.11) than fairness (Figure 6.12). Both US1 and Ger generally followed the

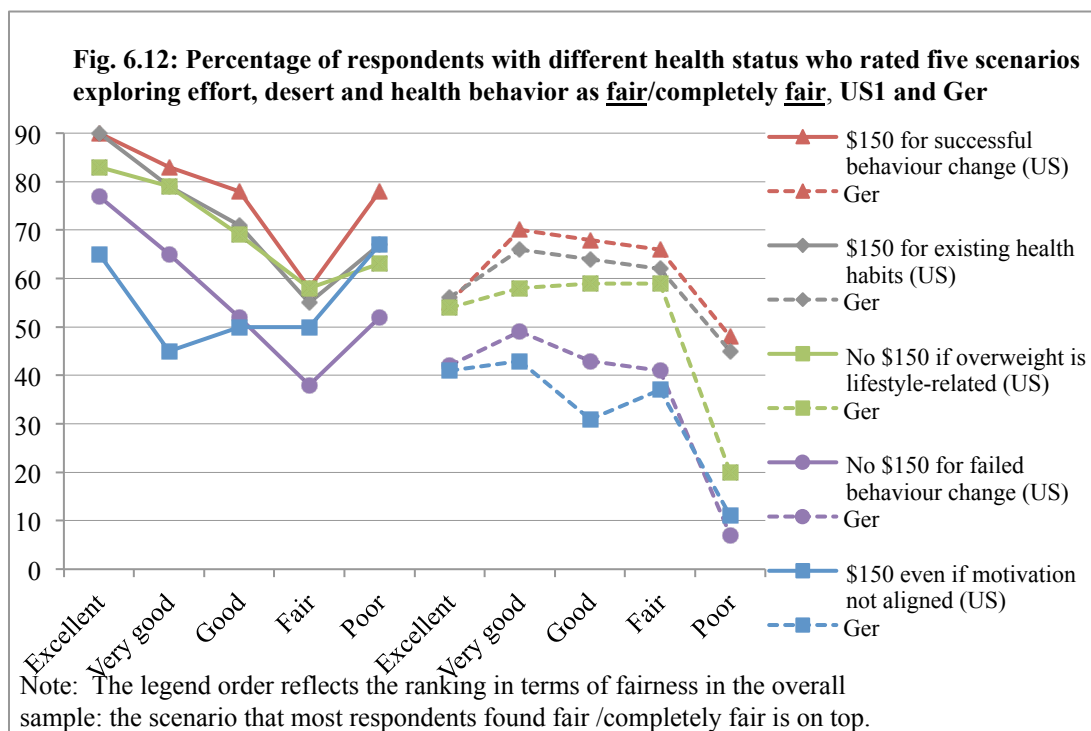


samples' overall fairness rankings, but an exception was that both the poor health groups approved least of the scenario in which someone fails to receive an incentive despite having attempted behaviour change: between 3–10% found this less fair than the case where someone gets an incentive although motivation runs counter the programme's goals. A consistent inversion in the order, mirroring the just mentioned flip, was also found regarding the rankings for disapproval: unlike the sample overall, in US1 the group with poor health found the most unfair the scenario in which those who attempt weight loss but are unsuccessful lose out on the incentive.

While most health groups fall in line with the overall ranking of unfairness, there were some further notable exceptions. In US1 the groups with good and fair health rank find most unfair that those who attempted behaviour change but failed lose out on the incentive: 25% of those with good health, 41% of those with fair health as opposed to 11% of those with excellent, and 17% of those with very good health. This pattern was not found in the German sample. In the American sample around twice as many of those with fair health compared to those with excellent or very good health disapprove of the scenario that those who change their behaviour successfully, as well as those who already have healthy habits receive \$150; and that those who like eating and dislike exercising lose out. Of those with fair and poor health, slightly higher levels of respondents find if unfair/completely unfair that

those who already have healthy habits receive the incentive, rather than those who prefer food over exercise (in the ‘fair’ health group 24 vs. 22%, and in the ‘poor’ health group 14% vs. 11% (US1 [successful behaviour change]:  $c^2(16, n=891)=53.64, [p.<.0001]$ ; US1 [failed behaviour change]:  $c^2(16, n=888)=57.93, p.<.0001$ ; US1 [healthy by habit]:  $c^2(16, n=882)=58.35, p.<.0001$ ; US1 [like eating]:  $c^2(16, n=889)=46.05, p.<.0001$ ).

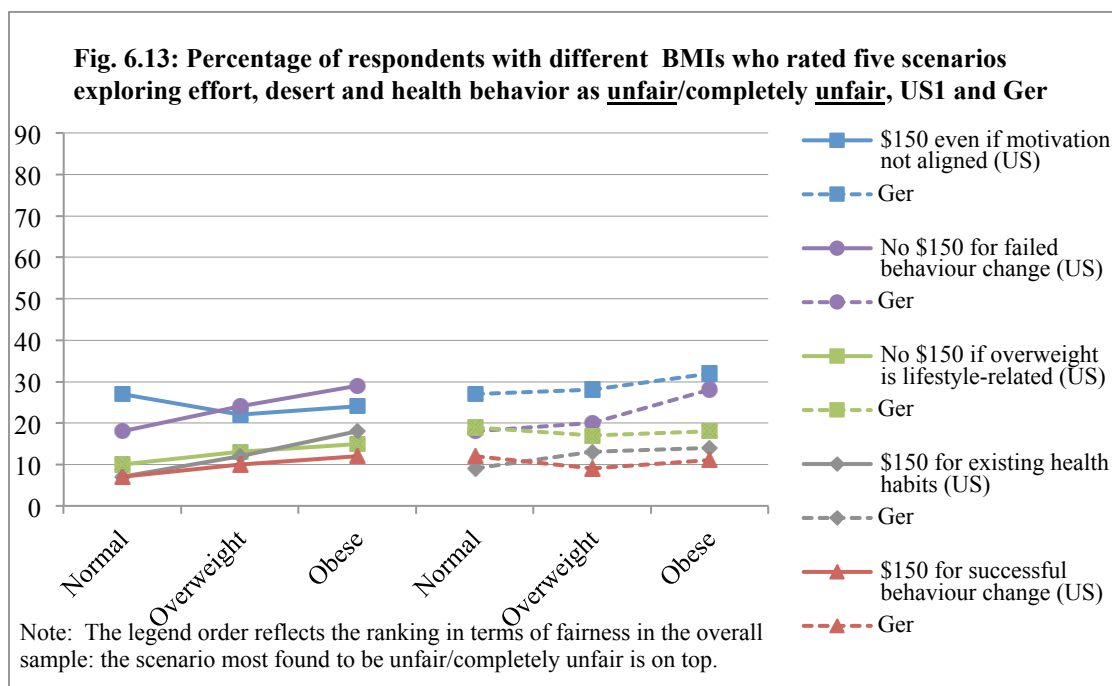
In the German sample, more respondents with excellent health appear to find unfair/completely unfair that those who prefer food over exercise lost out on the incentive compared to those who tried, but did not succeed in losing weight (30% vs. 28%. However, the margin was small: the overall sample averages were 17% and 21%). Two to three times more of those with poor health find it unfair/completely unfair that healthy people receive the incentive (27% as opposed to 7% of those with fair, 11% with good, 14% with very good, and, 15% of those with excellent health; Ger [healthy by habit]:  $c^2(16, n=954)=39.84, p.<.0008$ ; Ger [like eating]:  $c^2(16, n=946)=44.54, p.<.0002$ ; Ger [failed behaviour change]:  $c^2(16, n=888)=51.27, p.<.0001$ ). Differences between groups were most marked in the German sample with regard to perceived fairness of the scenarios, as shown in Figure 6.12.



The patterns for approval largely mirror those of disapproval, with the above observed dip in the fair health group in US1, and lowest levels in the poor health

group in Ger. Approval in most health groups also followed the pattern of the samples as a whole, with the exception that in Ger sample the scenario in which someone fails to receive an incentive despite having attempted behaviour change ranks lower than the one where a person gets the incentive although her motivation runs counter the programme's goals (7% vs. 11%), even if the margin was small. Those with excellent health also see the two cases as very similar, but 41% find the outcome fair. In the American sample, the inversion also occurred in the fair, and poor health group, although the margins were larger than 10%. In the poor health group, receiving the incentive despite misaligned motivation ranked second highest (67%), together with the scenario in which people receive the incentive for already existing healthy habits. (US1 [failed behaviour change]:  $c^2(16, n=888)=57.93$ ,  $p.<.0001$ ; US1 [misaligned motivation]:  $c^2(16, n=884)=35.63$ ,  $p.<.0001$ ; US1 [healthy by habit]:  $c^2(16, n=882)=58.35$ ,  $p.<.0001$ ; Ger [failed behaviour change]:  $c^2(16, n=888)=51.27$ ,  $p.<.0001$ ; Ger [misaligned motivation]:  $c^2(16, n=950)=60.80$ ,  $p.<.0001$ ).

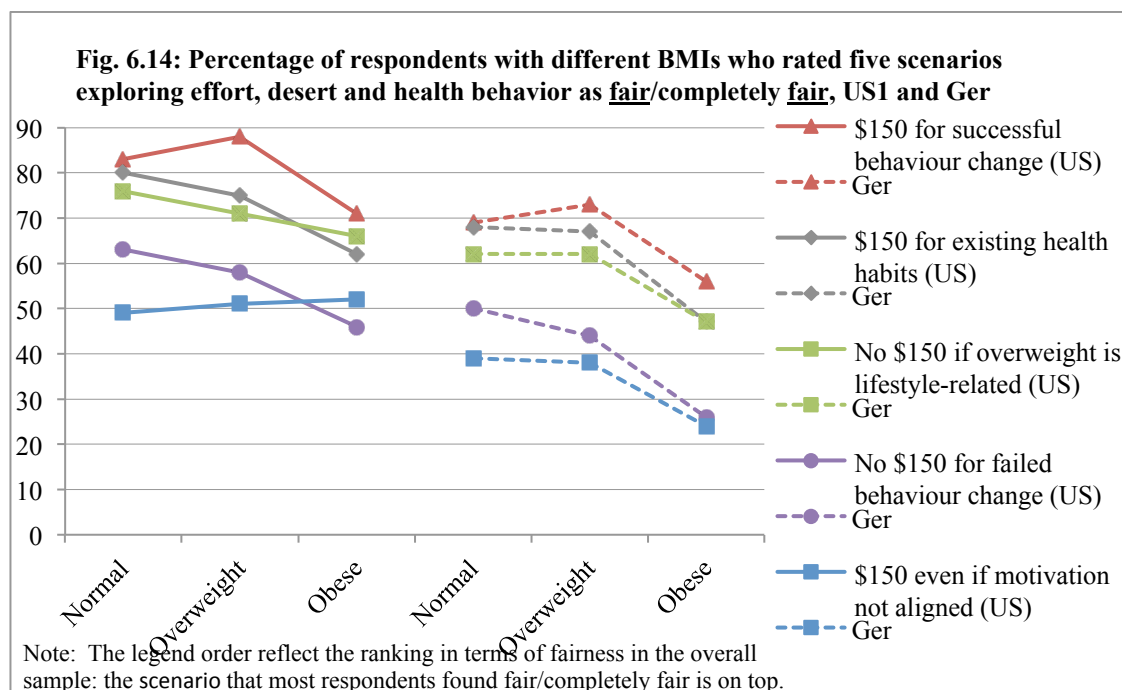
**BMI:** With regard to BMI, disapproval generally increased in both groups with weight. However, there were also some noteworthy inversions in the ranking order. Figure 6.13 shows the percentages of respondents who found scenarios unfair/completely unfair.





In both US1 and Ger, disapproval generally increased with body weight. An exception in US1 was the scenario of reward without behaviour change, whilst living unhealthily: normal weight respondents disagreed the most and 27% find this unfair/completely unfair (as opposed to 22% of the overweight, and 24% of the obese). Unlike the sample as a whole, both the overweight and the obese found most unfair that those attempting, but failing to achieve weight loss lose out (24% of overweight, 29% of obese, and 18% of those with normal weight). In Ger this policy was not felt to be the most unfair, but the same trend between groups can be observed: 28% of obese, 20% of overweight, and 18% of normal weight respondents find unfair/completely unfair that attempts are not rewarded. In US1, more of the obese find it unfair that those healthy by habit receive the incentive than the other weight groups, 18%, vs. 12% and 7% for the normal weight respondents. In the German sample, more respondents with normal weight appear to disapprove of incentives for those who succeed in weight loss than those who are healthy already, although the margin was small (12% vs. 9%; US1 [healthy by habit]:  $c^2(16, n=882)=58.35, p.<.0001$ ; US1 [failed behaviour change]:  $c^2(16, n=888)=57.93, p.<.0001$ ; Ger [failed behaviour change]:  $c^2(16, n=888)=51.27, p.<.0001$ ; Ger [successful behaviour change]:  $c^2(16, n=959)=40.97, p.<.0006$ ).

Regarding approval, the samples differed in that, overall, levels were 10–20% higher in US1, as Figure 6.14 shows.



Approval in the obese group in US1 differs compared to the other group and the sample overall in that the scenario that was the second highest rated was the one where those who prefer food over exercise receive no incentive. By a (small) margin of around 5%, those with normal weight and the overweight assigned this spot to those who already live healthily, while the obese inverse this with a similar differential. More of the obese than the other groups and the sample overall also find it fair/completely fair that people receive an incentive even if their motivation is not aligned. And, again unlike the other groups or the complete sample, the obese find this more acceptable than failing to provide an incentive for those who attempted, but did not succeed in behaviour change. In both the American and the German sample the overweight approved more than the two other groups of providing an incentive for those successful at behaviour change: in US1 88% of the overweight, 82% of the normal weight and 71% of the obese found this fair/completely fair, as did 73% of the overweight in the German sample, 69% of the normal weight, and 56% of obese respondents. US1 [like eating]:  $\chi^2(16, n=889)=46.05, p.<.0001$ ; US1 [healthy by habit]:  $\chi^2(16, n=882)=58.35, p.<.0001$ ; US1 [misaligned motivation]:  $\chi^2(16, n=884)=35.63, p.<.0001$ ; US1 [failed behaviour change]:  $\chi^2(16, n=888)=57.93, p.<.0001$ ; US1 [successful behaviour change]:  $\chi^2(16, n=891)=53.64, [p.<.0001]$ ; Ger [successful behaviour change]:  $\chi^2(16, n=959)=40.97, p.<.0006$ .

#### 6.3.4 *What are acceptable levels of incentives?*

The question of how high incentives should be is relevant both from an effectiveness and fairness point of view (Chapters 2–4). The surveys included three questions—presented in randomised order—in which respondents were asked to determine appropriate levels for different types of weight control incentive programmes, ranging from \$0–2,500\$. The upper level equals around 50% of the average cost of coverage in the US, constituting the maximum legally permissible amount.

**“Carrot”-weight incentives: how much, if insurance costs may be reduced? (Question 5a)**

A health plan wants to introduce an incentive programme that is similar to the “Incentives for Weight Control” policy. The policy would work like this:

- Employees are given the opportunity to be weighed every six months.
- If they are not overweight (in medical terms) on two consecutive weigh-ins, they get a partial rebate of their insurance premiums at the end of the year.
- If they are overweight (in medical terms) on two consecutive weigh-ins, or prefer not to be weighed, they have the same insurance premiums as before.

How much money should the plan give back to people who are not overweight, to create weight-control incentives? Please enter an amount between \$0 and \$2,500.

**“False carrot”-weight incentives: how much, if insurance may be increased? (Question 5b)**

A health plan wants to introduce an incentive programme that is similar to the “Incentives for Weight Control” policy. The policy would work like this:

- Insurance premiums for all employees are increased at the beginning of the year by a certain amount.
- Employees are then given the opportunity to be weighed every six months.
- If they are not overweight (in medical terms) on two consecutive weigh-ins, they receive a rebate equal to the increase at the end of the year.
- If they are overweight (in medical terms) on two consecutive weigh-ins, or prefer not to be weighed, they get no rebate.

By how much should the plan be allowed to increase insurance premiums for all at the beginning of the year, to create weight-control incentives? Please enter an amount between \$0 and \$2,500.

**“Stick”-weight incentives: how much, if insurance costs may be increased? (Question 5c)**

A health plan wants to introduce an incentive programme that is similar to the “Incentives for Weight Control” policy. The policy would work like this:

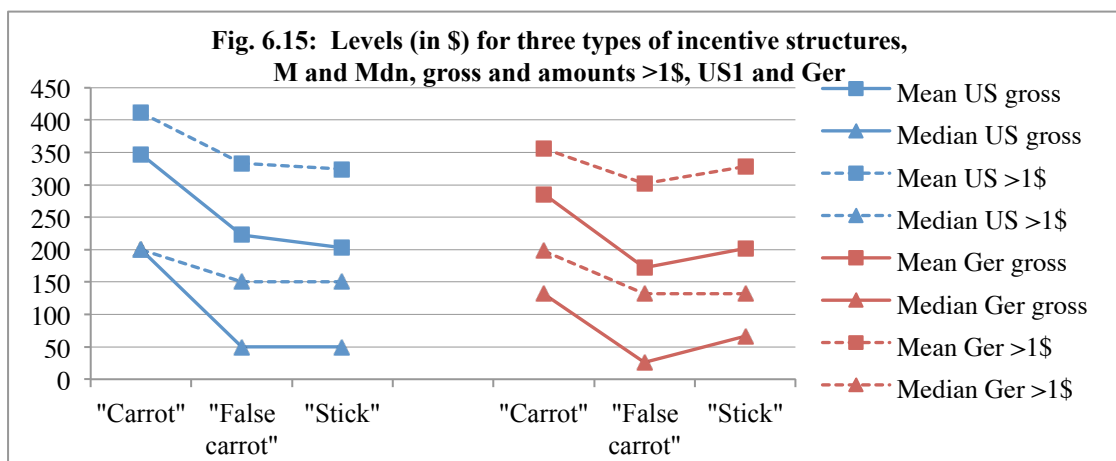
- Employees are given the opportunity to be weighed every six months.
- If they are not overweight (in medical terms) on two consecutive weigh-ins, their premiums stay the same.
- If they are overweight (in medical terms) on two consecutive weigh-ins, or prefer not to be weighed, their premiums are increased by a certain amount.

By how much should the plan be allowed to increase insurance premiums for overweight people, to create weight-control incentives? Please enter an amount between \$0 and \$2,500.

**Summary:** The gross median amount (inter-quartile range) for the ‘carrot’ scenario in US1 was \$200 (51–457) decreasing to \$50 (0–200) for ‘false carrot’ and ‘stick’. In Ger the ‘carrot’ gross median was lower: \$132 (66–330). ‘False carrot’ levels were s half of that of US1, at \$26 (0–159), but far higher for stick: \$66 (0–264).

*Age:* Younger people set higher levels. *Gender:* Not significant in US1 but in all scenarios in Ger males set higher levels. *Income:* Significant in all cases in US1: higher income generally correlated with higher incentive levels. *Health and BMI:* Significant in both US1 and Ger in all scenarios: better off health groups set higher levels, although the poor health group in US1 entered higher amounts than better off groups for ‘carrot’ and ‘stick’. Unlike US1, all German health groups except those with poor and fair health entered higher amounts for the ‘stick’, as opposed to the ‘false carrot’ option. For BMI, this pattern was observed in those with normal weight and the overweight, although not the obese. At the same time, the obese have set the highest amounts in all three scenarios. This was not the case in US1, where heavier respondents opted for lower amounts throughout, and those with normal weight were significantly above the other groups.

Data on overall levels can be presented and analysed in different ways, with different implications: (1) one could either analyse all sums entered (including \$0), or only those above a certain threshold, that can be regarded as genuine incentive amounts. (2) For both frames different approaches to determine central tendency can be used. Figure 6.15 shows for US1 and Ger, across the three scenarios, medians and means for the whole sample, in relation to all amounts entered, and regarding a subset of the sample in which respondents entered amounts larger than \$1.

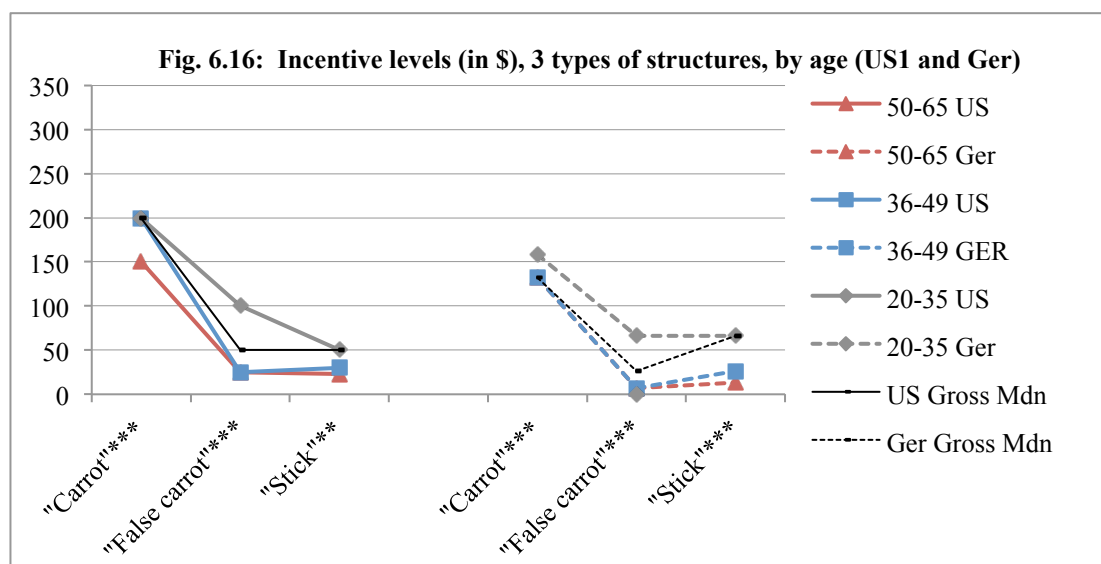


While amounts across the three scenarios were somewhat lower in Ger, they were overall in a similar range. In all depictions of the US1 amounts, 'false carrot' and 'stick' amounts were equal to, or lower than the 'carrot' amount. By contrast, three of the depictions in Ger show a V-shape, with higher amounts for 'stick' than for the 'false carrot' scenario. Table 6.1 shows the numbers of respondents who entered amounts overall, as well as those who entered amounts higher than \$1, which differed across scenarios and samples.

	US1			Ger		
	'carrot'	'false carrot'	'stick'	'carrot'	'false carrot'	'stick'
All amounts	984	983	983	990	993	992
Amounts >\$1	821	585	628	837	663	621

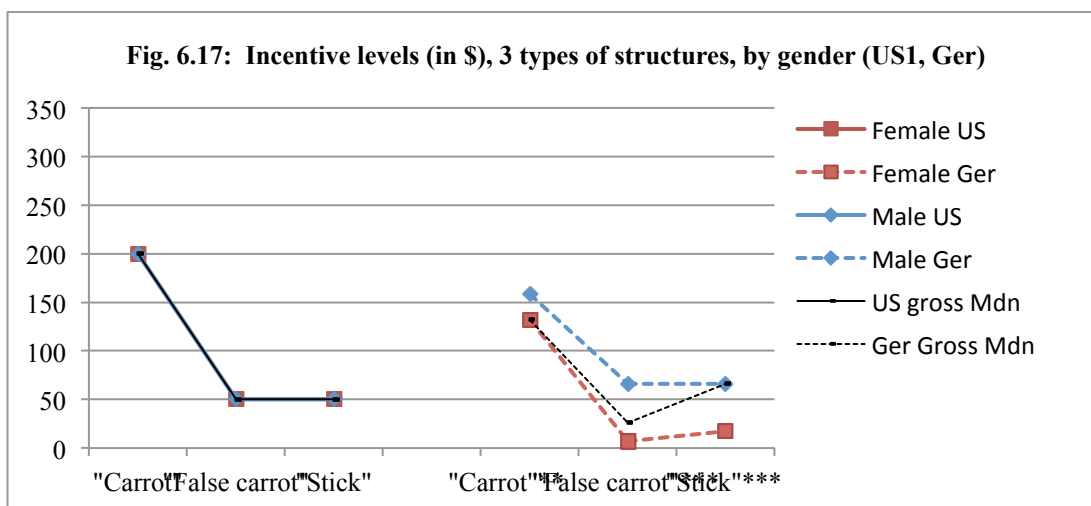
For the analysis of the independent variables age, gender, income, health and BMI, in the following the focus is on all amounts entered (including \$0), which is a better suited measure of overall approval (while a focus on amounts higher than \$1 can be of interest to understand those who approve of the respective approach). As Figure 6.15 suggests, the distribution in all three scenarios was non-normal with the vast majority at the lower end of the spectrum. It was therefore more appropriate to report and analyse medians.

**Age:** In US1 the median (inter-quartile range) for the oldest age group for ‘carrot’ was \$150 (20–269), with the two younger groups at \$200 (youngest group: 100–500, middle group: 50–300), see Figure 6.16. The youngest group placed the highest amounts in the two other scenarios, with \$100 (\$0–\$288) for the ‘false carrot’, and \$50 (0–250) for the ‘stick’. For these two scenarios the middle age group was at \$25 (0–200) and \$30 (0–200), and the oldest group at \$25 (0–200) and

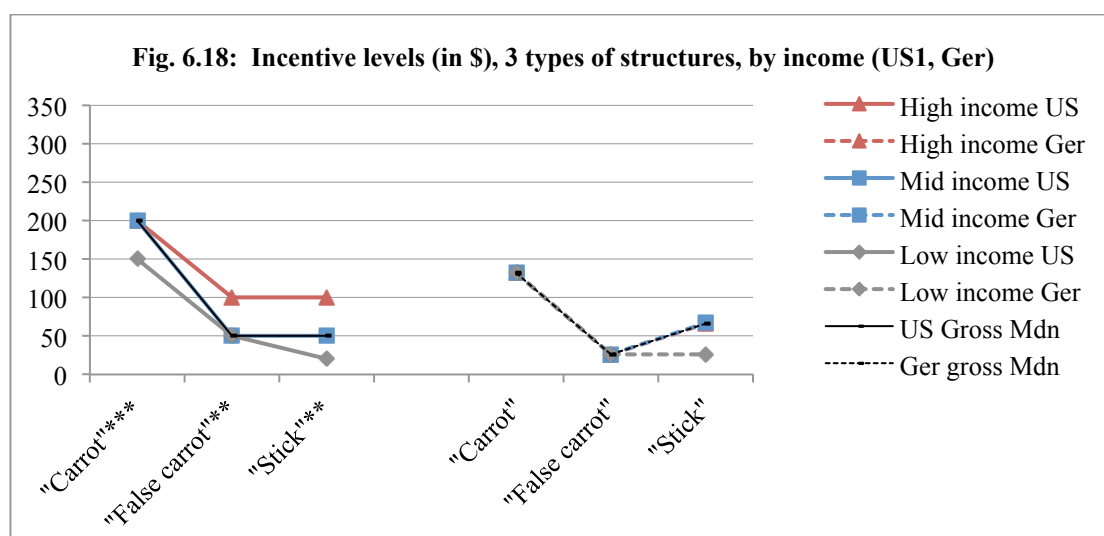


\$23 (0–150). Across all scenarios, the youngest group also entered the highest levels in the German sample: \$158 (66–290) for ‘carrot’ vs. \$132 in both other groups (middle age group: 66–290, oldest group: 0–198), \$66 (0–264) for ‘false carrot’ (vs. seven each in the other groups; middle age group: 0–132, oldest group: 0–106) and \$66 (0–264) for ‘stick’, with the middle age group at \$26 (0–158), and the old group at \$13 (0–139; US1 [carrot]  $H(2)=16.85$ ,  $p.0002$ ; US1 [false carrot]  $H(2)=20.61$ ,  $p.<.0001$ ; US1 [stick] ( $H(2)=11.67$ ,  $p.0029$ ; Ger [carrot]  $H(2)=19.56$ ,  $p.<.0001$ ; Ger [false carrot]  $H(2)=21.62$ ,  $p.<.0001$ ; Ger [stick]  $H(2)=17.72$ ,  $p.<.0001$ ).

**Gender:** Significant differences were only found in all German scenarios: \$158 (66–396) vs. \$132 (66–264) for ‘carrot’, \$66 (0–264) vs. \$10 (0–132) for ‘false carrot’ and \$66 (0–264) vs. \$16 (0–158) for ‘stick’ (Ger [carrot]  $H(1)=7.56$ ,  $p.0059$ ; Ger [false carrot]  $H(1)=23.18$ ,  $p.<.0001$ ; Ger [stick]  $H(1)=20.58$ ,  $p.<.0001$ ; Figure 6.17)

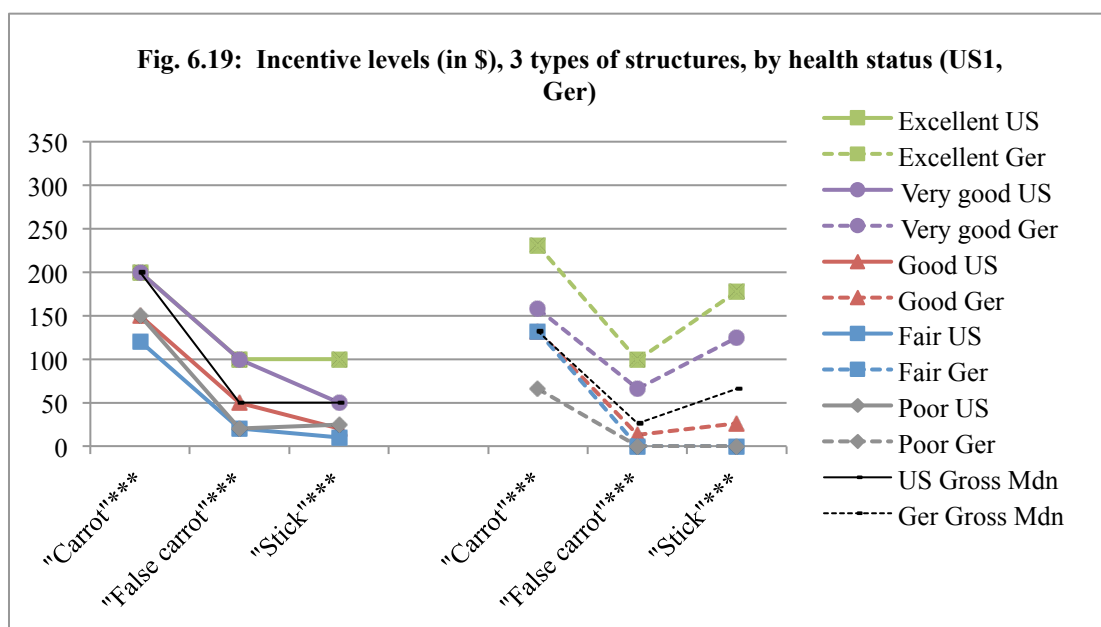


**Income:** Differences were significant in the American sample, where the low-income group provided the lowest amounts in two scenarios: while the best off and the middle income group entered \$200 for ‘carrot’ (inter-quartile range for both 100–500), the low-income group was at \$150. Low and middle income both opted for \$50 for the ‘false carrot’ (0–200 in both cases), with the high-income group doubling this at \$100 (0–500). The high-income group converged around \$100 (0–400) for ‘stick’, while the middle income was at \$50 (0–200), and the low-income group at \$20 (0–150). US1 [carrot]  $H(2)=20.66$ ,  $p.<.0001$ ; US1 [false carrot]  $H(2)=10.80$ ,  $p.0045$ ; US1 [stick]  $H(2)=12.72$ ,  $p.0022$ ). In the German sample the high-income group entered \$66 (0–363) for the ‘stick’ scenario, the middle group \$66 (0–264), and the low group \$26 (0–198), but this distribution was not significant. Unlike the American sample, the German distribution also showed the V pattern (with the high and middle income groups entering higher amounts for ‘carrot’ and ‘stick’ as opposed to the ‘false carrot’ scenario, see Figure 6.18).



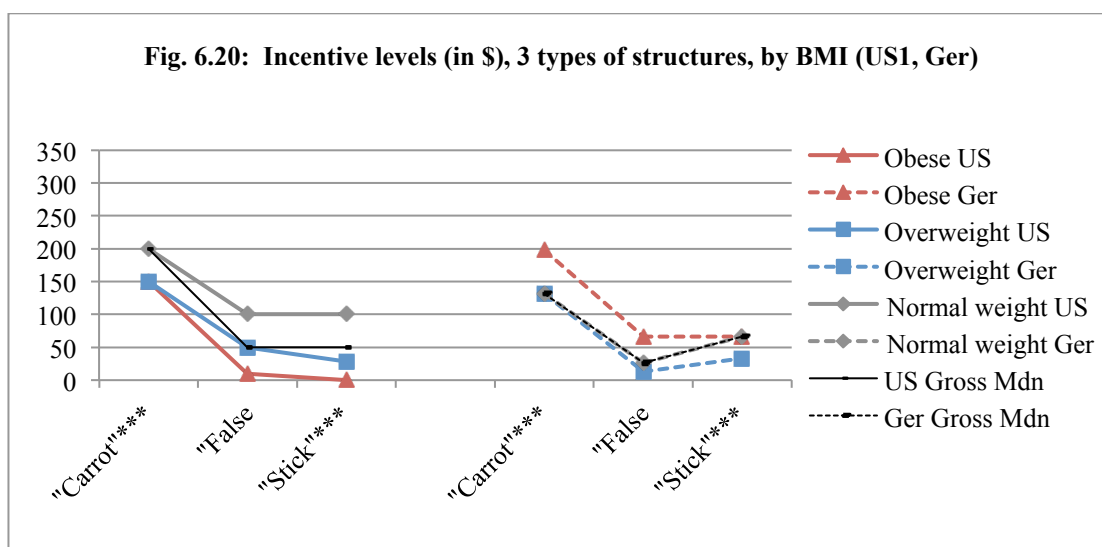
**Health:** There were significant differences in both US1 and Ger (Figure 6.19).

Across scenarios, Ger again showed the V-shaped pattern in all groups except for the poor health group. Overall, better health groups set higher levels in each of the scenarios, although this was not universally the case. In the ‘carrot’ scenario in US1, the poor health group, as those with good health, opted for \$150 (inter-quartile range for poor: 62–500, good: 29–300), and the fair health group for \$120 (0–250), with the remaining ones at \$200 (very good: 100–500, excellent: 100–500). Poor health groups also had higher amounts than the good and fair health group for the ‘stick’, although the differences were small: \$10 (0–150, fair health), \$20 (0–150, good) and \$26 (0–237, poor health), with those with excellent health at four times higher levels, at \$100 (1–242). US1 [carrot]  $H(4)=33.55$ ,  $p.<.0001$ ; US1 [false carrot]  $H(4)=27.73$ ,  $p.<.0001$ ; US1 [stick]  $H(4)=29.50$ ,  $p.<.0001$ .



In the German sample the “excellent” and “very good” health groups stand out as entering the highest levels in all three scenarios, and having the steepest V-shape. All groups except those with poor and fair health show this pattern, the latter two averaging \$0 in both ‘false carrot’ (poor: 0–0, fair: 0–84), and ‘stick’ (poor: 0–0, fair: 0–132). The excellent health group opted for \$178 (1–462) for ‘stick’, and those with very good health for \$125 (0–396). For the ‘false carrot’ scenario the excellent converged at \$99 (0–297) and the very good at \$66 (0–320). Ger [carrot]  $H(4)=22.43$ ,  $p.0002$ ; Ger [false carrot]  $H(4)=46.21$ ,  $p.<.0001$ ; Ger [stick]  $H(4)=55$ ,  $p.<.0001$ .

**BMI:** Weight was significant in both samples and all scenarios (Figure 6.20). In US1 those with higher weight generally set lower levels in all scenarios (except for the ‘carrot’, where both the overweight and the obese opted for \$150 (inter-quartile range for overweight: 50–300, obese: 8–272) and those with normal weight for \$200 (100–500). Obese and overweight in US 1 also set lower levels for the ‘stick’ as opposed to the ‘false carrot’: the obese’ levels for stick were \$0 (0–110) and the overweights’ \$27 (0–175), with the normal weight group three times higher at \$100 (0–250, US1 [carrot] H(3)=21.87, p.<.0001; US1 [false carrot] H(3)=30.72; p.<.0001; US1 [stick] H(3)=44.30, p.<.0001). By contrast, in Ger the obese opted for the highest amounts in the ‘carrot’ scenario: \$198 (66–330), overweight and normal weight: \$130 (overweight: 66–330, normal: 40–330). The obese also had the highest ‘false carrot’ levels: \$66 (0–264), normal weight: \$26 (0–159), overweight \$13 (0–132). In the ‘stick’ scenario the overweight were at \$33 (0–214), and the obese and normal weight jointly at \$66 (obese: 0–264, normal weight: 0–237). Ger [carrot] H(3)=34.63, p.<.0001; Ger [false carrot] H(3)=35.97, p.<.0001; Ger [stick] H(3)=59.17, p.<.0001.



### 6.3.5 Do incentive programmes interfere with the doctor patient relationship?

The impact of incentives on the doctor-patient relationship was explored with regard to the weigh control policy and the first version of the colon cancer prevention policy.

**Section summary:** For both scenarios, and in both US1 and Ger, respondents thought that around every second patient would ask their physician to lie in their favour, in order to obtain the incentive. More than half in both US1 and Ger



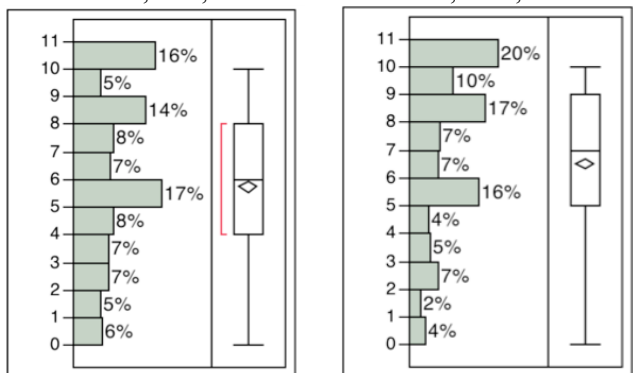
disapproved of a complying physician, rising to two thirds in the target population subgroup of the colon cancer policy. Since a sizeable proportion of patients is expected to request a lie it is plausible to assume that incentive programmes along the lines sketched in the scenarios do have a negative impact on the doctor-patient relationship.

6.3.5.1 Incentives for weight-control: The role of physicians (Question 6)

**Questions 6a and b:** Think again of the “Incentives for Weight Control” policy, where employees are given the opportunity to be weighed every six months, and if they are not overweight (in medical terms), they get a reimbursement of \$150 at the end of the year. To get the reimbursement, a doctor needs to confirm employees' weight to the health plan. Imagine that there is an employee who took the policy as an incentive to lose weight. By the time of the first 6 month weigh-in, he lost 10 pounds and met the target. But at the second six-month weigh-in he is 2 pounds too heavy.

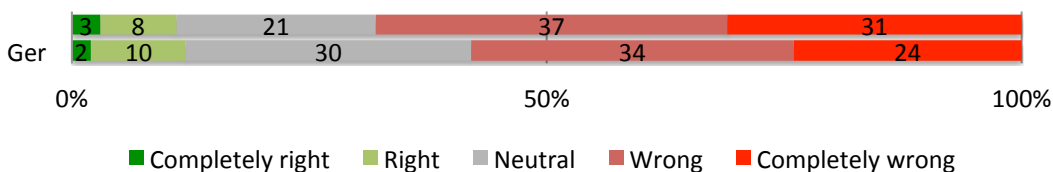
This is the last day of the qualifying period, and he cannot come back another day. Only the doctor and the employee are in the room, when it occurs to the patient, that he could ask the doctor to state that he has in fact met the target. Out of 10 patients, how many, would you think, would ask their doctor in such a situation to write down 2 pounds less?

US1: Mdn=6, M=6, SD=3.03 Ger: Mdn=7, M=7, SD=2.92



(n=978)  
7 lying doc is....

Suppose there is a doctor who agrees to note a lower weight. How would you rate the doctor's



action? The doctor acted...  
(US1: n=978, Ger: n=941)

**Summary:** Median and mean numbers for estimates of how many patients would request their physician to lie were 6 in US1 and 7 in Ger. Only age (US1), gender (Ger) and political orientation (Ger) were significant of the independent variables.

**Age:** Differences were significant in US1, where the mean for the youngest age group was 6.07 (SD=2.76), 5.76 for the middle group (SD=3.13), and 5.23 (SD=3.23) for the oldest group. (US1:  $F(2)=6.58$ ,  $p.0014$ ).

**Gender:** In Ger, the mean for women was 5.73 (SD=3.10), and for men 5.18 (SD=2.86; Ger:  $F(2)=7.96$ ,  $p.0049$ ). Political orientation is reported in Volume B.

Regarding assessments of the physicians' action, age was significant in both US1 and Ger, gender again in Ger, politics in the American sample, and health status in both. **Age:** Older age groups disapproved more: in sample Ger, 78% of the oldest, 67% in the middle group and 62% in the youngest group were opposed, in US1 lower levels but the same slope could be observed (66, 57 and 49%; US1:  $c^2(8, n=990)=29.51$ ,  $p.0003$ ; Ger:  $c^2(8, n=996)=31.96$ ,  $p.<.0001$ ).

**Gender:** In Ger 29% of males but only 20% of females found the action completely wrong: 60% of males overall opposed (versus 55%; Ger:  $c^2(4, n=990)=13.95$ ,  $p.0074$ ).

**Health:** Differences were significant, but no clear trend emerges in either sample, and the main difference in the American sample was that those with excellent health had the lowest level of opposition (but still 64%, all others between 67–71%), and the highest level of acceptance (19%, with all others between 8–13%). In the German sample four per cent of the poor health group finds the doctors action acceptable, 10% of the fair health, and all others at around 13%. Those with very good health disagreed the most, at 64%, the remaining at 65%, except for those with poor health, of whom 45% disapprove (US1:  $c^2(16, n=968)=35.95$ ,  $p.0029$ ; Ger:  $c^2(16, n=961)=30.66$ ,  $p.0149$ ).

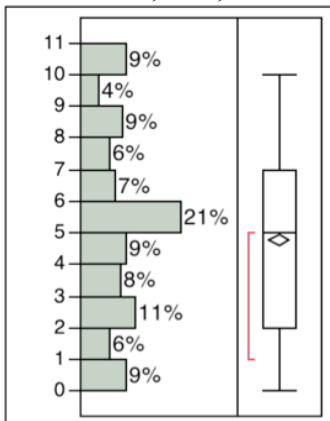
#### 6.2.5.2 Incentives for colon cancer prevention: the role of physicians (Question 14)

**Questions 14a and b:** Suppose someone has colon cancer. After an operation her physician discusses further treatment options, and recommends a therapy that is accepted as the most effective one by the professional bodies. However, the patient hears from a friend who underwent this treatment that it was very aggressive and helped little. The patient therefore plans to refuse the treatment that her physician recommends. The patient's health plan operates a policy like the one mentioned earlier, where patients who refuse treatment have overall copays of \$1,000 instead of \$500. She also knows that her doctor is required to inform her health plans about her decision.

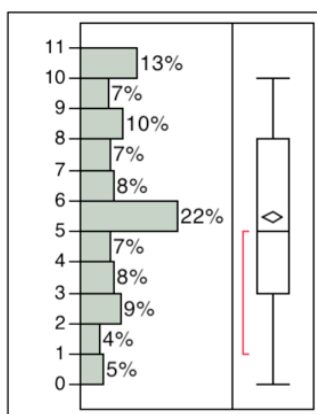
Only the patient and her doctor are in the room, when it occurs to her that she could ask the doctor not to inform the health plan. Out of 10 patients, how many, would you think, would ask their doctor in such a situation not to tell the health plan about her refusal?

US1: Mdn=5, M=5, SD=2.95

Ger: Mdn=5, M=5, SD=2.99



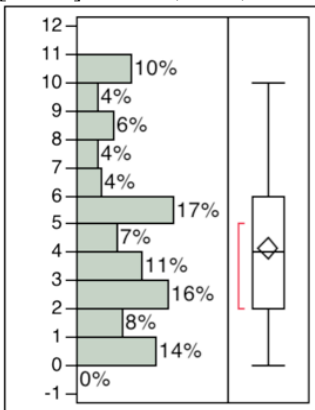
(n=978)



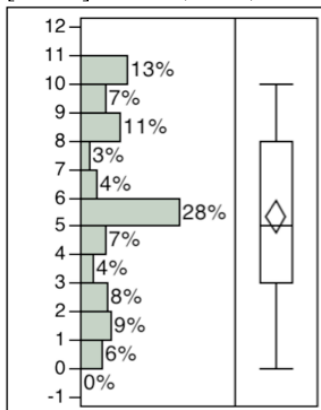
(n=941)

[50-65]: Mdn=4, M=4, SD=3.11

[50-65]: Mdn=5, M=5, SD=5.19



(n=282)

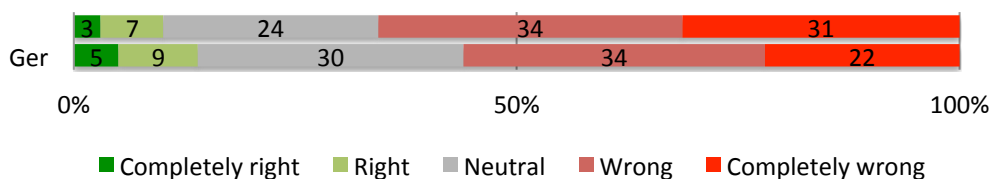


(n=114)

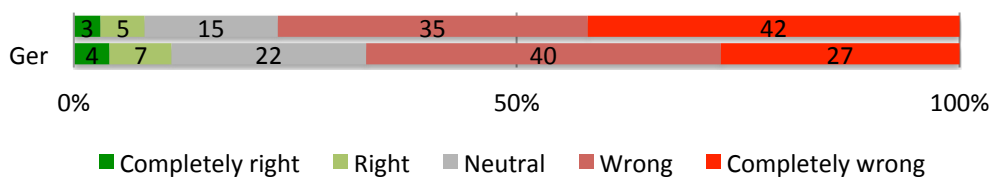
**15 lying doc is...**

Suppose there is a doctor who agrees not to let the health plan know. How would you rate the doctor's action?

All age groups (US1 n=975, Ger n=941):



Subgroup: 50-65 years only (US1 n=282; Ger n=114):



**Summary:** Median and mean numbers for estimates of how many patients would request their physician to lie were 5 in the complete samples of US1 and Ger; 5 in the target population subgroup of the German sample, and 4 in that of US1. As in the case of the obesity policy, age (US1), gender (Ger) and political orientation (Ger) were significant of the independent variables, in addition to the use of cancer screening (in US1).

**Age:** In US1, the mean for the youngest group was 5.08 (SD=2.70), 4.93 for the middle group (SD=3.05), and 4.14 (SD=3.11) for the oldest group. (US1:  $F(2)=9.24$ ,  $p.0001$ ).

**Gender:** The mean for women was again higher, at 5.73 (SD=3.10), that that for men: 5.18 (SD=2.86; Ger:  $F(2)=7.96$ ,  $p.0049$ ).

**Cancer screening:** In US1, the mean for those without any form of screening was 4.97 (SD=2.89), lower for those with colonoscopies (4.27, SD=3.04) and lowest for those with other forms of colon cancer screening (4.19, SD=3.14,  $F(2)=5.88$ ,  $p.0029$ ; political orientation is again not reported here, but see Volume B).

Regarding assessments of the physicians' action, age was again significant in both US1 and Ger, gender again in the German sample. Political orientation was significant in the German sample only, and use of cancer screening only in the American sample.

**Age:** Again, older age groups disapproved more, and in both samples the rates were within a margin of 3% compared to the obesity scenario. In sample Ger, 77% of the oldest, 67% in the middle group and 55% in the youngest group were opposed, in US1 lower levels, but the same slope could be observed (67, 53 and 47%; US1:  $c^2(8, n=991)=49.83$ ,  $p.<.0001$ ; Ger:  $c^2(8, n=970)=31.98$ ,  $p.<.0001$ ).

**Gender:** In the German sample 25% of males and 18% of females found the action completely wrong, with 55% of males overall opposed, and 51% of women, in all cases around 5% below the obesity scenario (Ger:  $c^2(4, n=970)=20.16$ ,  $p.0005$ ).

**Cancer screening:** In US1, 61% of those without any form of screening disapproved, and therefore less than those with other forms of colon cancer screening (71%) and those with colonoscopies (73%;  $\chi^2(8, n=982)=29.80, p.0002$ ).

## 6.4 Discussion

Rather than discuss the findings of each cluster of questions (or of each question) separately, the following section is structured around three main headings which concern (1) similarities and differences between the American and German samples; (2) similarities and differences between the views of survey respondents and physicians (as outlined in the previous chapter); and (3) a closer look at the winners and losers of incentive programmes.

### *6.4.1 Key similarities and differences between the American and German samples*

There were noteworthy similarities and differences regarding the use of incentives over time, the overall assessment of incentive programmes presented in the surveys, and the levels of financial incentives that were regarded as appropriate for weight-control incentives.

#### 6.4.1.1 Use over time, overall and by type of incentive

Strikingly, in both US1 and Ger, and for both attainment and participation-incentives, there was the same pattern of use over time, even if levels were consistently lower in the US (Figure 6.6). Current use in Ger was around 20%, and 10% in US1. Past use was around 5–10% lower in both, but intention for future use was three times higher than current use in Ger, and around four times higher in US1, with younger people, higher income, better health, and lower BMI groups expressing higher levels of interest, generally above 50%.

It should be noted that all data were, of course, self-reported and cannot be taken as true records of actual use. In addition, figures for present and past use are more dependable than those regarding intention for future use, which may, or may not, occur. There is more certainty, then, that past use was lower than current use, and less certainty that future use will be higher, even if the margin differs by a factor of three to four. In view of this situation it can still be plausible to interpret the markedly higher levels for future use as an expression of genuine interest in incentive programmes. At the same time, the somewhat harder evidence regarding current and

past use suggests that one should not be overoptimistic regarding actual future uptake and a certain mark-down would be appropriate, although it is not clear what formula would determine adjusted levels.

A related question is what prompted the higher levels of interest in future use. As fielded,<sup>6</sup> the examples used were framed as rewards and penalties, for obesity and colon cancer, with levels between \$150 and \$500 (unless higher amounts were entered by respondents). Throughout, the notions that incentives can be motivated by the desire to improve health, as well as to save cost to health care payers, and possibly also individuals, were emphasised. The nature of the data does not allow a deeper exploration of respondents' reasons for being interested in using incentives, however I return below to the questions of winners and losers more generally, which helps make some steps towards hypotheses of what might have been major drivers behind the increased levels of interest in future use.

While the overall trend of lower past use in comparison to present use was clear in both samples and both types of incentives, there were also two noteworthy variations regarding the extent of differences. As Table 6.2 shows, the largest difference among lower past use was in sample US1 for participation-incentives (+10%) with around half that (+6%) for attainment-incentives (in Ger the difference was +5% for attainment and +6% for participation-incentives). Given that participation-incentives are generally held to be less coercive one might have expected higher attrition for attainment-incentives, which, however, was not the case. Still, the margin of 4% was relatively small, and moreover it is also possible that respondents had a wider range of participation than attainment-incentives to choose from: the data for past use alone therefore provide no further leads as to possible reasons. The future-use data for US1, by contrast, could corroborate the view that respondents favour participation over attainment-incentives, as 7% more preferred them. However, in the German sample the inverse was the case, and by a similar margin (6%) respondents seemed to prefer the more robust attainment-incentives, despite a 5% difference in the opposite direction regarding current and past use.

---

<sup>6</sup> It should also be noted that the questions about use over time were asked in the last section of the survey, after participants had already responded to the weight-control and cancer incentive scenarios. Different answers regarding future use might have been given if another set of scenarios would have been used, or if, for example, the questions had been asked at the very beginning, in which case respondents would only have thought about programmes that they were already familiar with.

While, therefore, overall no clear picture emerged regarding a possible preference of respondents for participation over attainment-incentives, the absence of such a trend is also noteworthy, as it complicates assumptions that attainment-incentives are less welcomed.

	Current use	Current vs. past	Current vs. future
US1 attainment use	8	+ 6	+ 39
US1 participation use	11	+ 10	+ 46
Ger attainment use	18	+ 5	+ 39
Ger participation use	21	+ 6	+ 33

#### 6.4.1.2 Overall substantive agreement and disagreement regarding incentive programmes

The prominence of the value of solidarity in the German health care system, and the much more individual-centred focus of the US system could lead one to expect clear differences with regard to the overall assessment of the incentive policies used in the survey. However, this was not the case. In all three samples a majority either agreed strongly, agreed, or was neutral, with opposition ranging from 20–35% in the most accepted policy (\$150 for normal weight) and around 40% in the remaining three cancer policies as well as the car insurance analogy, which attracted the highest disagreement (between 41–42%). The differences in disagreement between US1 and Ger were between 1–3% (questions 2, 3, 12, 13) and 9% (the scenario modelled after the German cancer prevention policy, requiring attendance of counselling and compliance). Differences between US2 and Ger—with US2 showing higher levels of disagreement—were in a similar low range for two questions (question 2: 1%, question 13: 5%) and somewhat larger for question 12 (cancer/counselling only: 11%) and question 3 (\$150 for employees whose weight is normal over a year: 17%, still resulting in 42% agreeing with the policy, or accepting it).

The response in sample Ger can suggest a view of solidarity that is more focused on fair reciprocity than unconditional assistance in case of need: note that the support for the potentially penalising car insurance analogy differs by only 2% across the three samples. In view of the basic principles underlying social health insurance, it would not have been unreasonable to expect the German respondents to be more

opposed than their American counterparts. Equally it was noteworthy that in US1 and US2 the explicit reference to cost<sup>7</sup> as a reason for imposing a penalty did not seem to trigger concerns about rationing and insufficient respect for individual choice and autonomy, positioning the American respondents closer to their European counterparts than one might have expected.

Respondents in US1 and Ger also agreed by a margin of 8% on the ambiguity of the status of the two scenarios that explored framing. Forty-three per cent in both samples viewed as both ‘carrot’ and ‘stick’ the weight-control incentive where those who have normal weight over a year receive \$150. Forty per cent of respondents in US1, and 35% in Ger, thought the same about the scenario based on the German colon cancer policy (requiring attending counselling and complying in treatment or facing a co-payment threshold of \$1,000 instead of \$500). This suggests that the above-cited support for both policies is not based on a superficial impression according to which programmes were viewed as somehow ‘innocent’ offers for help, but respondents were clear about the dual nature of the conditionalities at work. The finding would also suggest that attempts by policy makers or employers to advertise as purely benevolent and voluntary initiatives along the lines of the weight-control programme are likely to meet with scepticism, which should be considered in the design of, and communication about, programmes. At the same time, it is also noteworthy that respondents differentiated only moderately between the framing of the weight-control and the colon cancer policy. The latter, moreover, had a far more overt ‘stick’ character, and was still viewed by around a third of respondents as “mainly reward”, highlighting, again the considerable complexities that underlie the perception and design of incentives.

#### 6.4.1.3 Appropriate levels for weight-control incentives

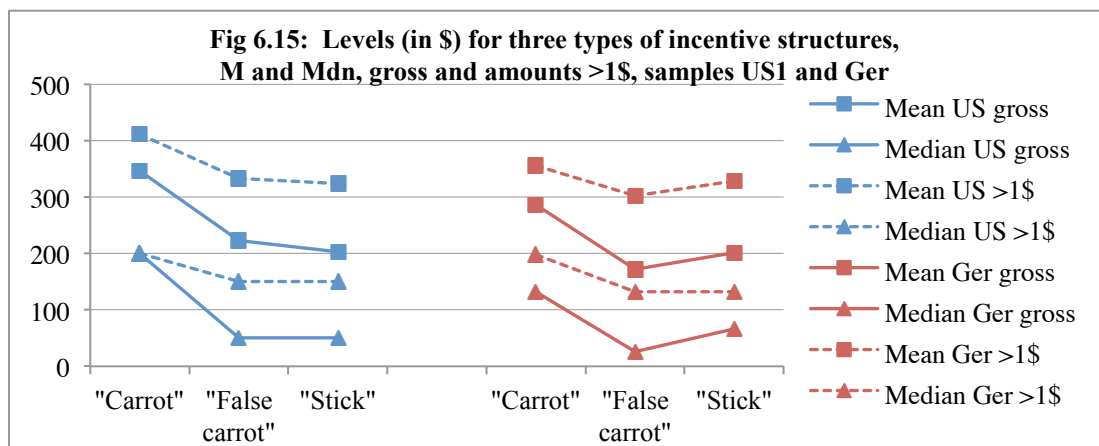
Chapter 4 noted that a common finding in surveys on personal responsibility for health is that there is some sympathy for penalising those who could have avoided poor health. Chapter 3 noted that the CEO of supermarket chain Safeway, who was

---

<sup>7</sup> “A spokesperson of the health plan says: ‘Colonoscopies are [colon cancer screening is] good for the employee and for the health plan, as we can cover other medical needs from savings that come from the early detection and treatment of cancer. Employees don’t have to have a colonoscopy [attend an information session] if they prefer not to. But both economic and medical considerations support the use of a one-time surcharge to encourage employees to have a colonoscopy done [to attend an information session].’”



influential in increasing incentive amounts under the 2010 health reform, viewed a surcharge of several thousand dollars as appropriate for smokers and obese.<sup>8</sup> The support for the car insurance analogy in US1 and Ger showed that respondents shared the basic assumption behind this push. Proposals for such policies therefore fall on fertile ground in principle. But the reproduction of Figure 6.15 also shows that considerably lower levels of disadvantage are viewed as appropriate. The mean of all those suggesting amounts of more than \$1, and therefore supporting the policy, was \$324 (SD=475) in US1 and \$328 in Ger (SD=477). The more adequate median in US1 was \$150 (inter-quartile range: 50–400) and \$132 in Ger (66–396). The median of all amounts entered—i.e. including \$0, reflecting overall approval—was \$50 (0–200) in US1 and \$66 (0–264) in Ger—around a 30<sup>th</sup> of the amount that proponents think should be passed on to obese employees.<sup>9</sup> Employers intending to implement this type of policy should therefore be aware that acceptance or tolerance of the principle says little about the accepted or tolerated amount. Moreover they should note that far higher amounts were entered for ‘carrot’ which may provide scope for health promotion programmes in which employee and employer interests are better aligned.



What is also striking about Figure 6.15 is that, again, despite the different health care systems and cultural backgrounds, amounts for the three scenarios ‘carrot’, ‘false carrot’, and ‘stick’ were in similar ranges in US1 and Ger. However, a clear difference was found in the V-shaped pattern in three of the German modes of

<sup>8</sup> In a presentation to policy makers given on Capitol Hill Steve Burd included a calculation which suggested that a family with two obese smokers generated incremental health care cost of \$4,480 (\$2,810 attributable to tobacco, and \$1,670 to obesity). Although HIPAA regulations enabled the company Safeway to shift \$1,560, there remained a “responsibility gap” of \$2,920. (Slide “HIPAA Incentive Allowance is Insufficient—Safeway example – Family with 2 obese smokers” from presentation: “Healthcare Solutions That Work”, June 11, 2009), slides available from HS.

analysis: while all patterns in US1 set the ‘stick’ amounts at identical, or marginally lower levels as the ‘false carrot’, three of the German analyses show higher levels for ‘stick’ than for ‘false carrot’. For employees with normal BMI, the outcome of ‘false carrot’ and ‘stick’ would be the same, as they have the same insurance cost as before, while others experience an increase. Insofar as one accepts the approach, one would think that it would be irrelevant which mode of implementation was chosen (as perhaps implied in the amounts entered in US1). However, the findings from the German sample suggest that this is not so. It is not easy to explain this finding. One form of reasoning might be that the prospect of a higher penalty is viewed as more effective in motivating behaviour change.<sup>9</sup> Another form might be fairness-based, and rest on a conception of solidarity according to which sickness fund enrollees should limit the burden they impose on the communally funded risk pool, which might be signalled more clearly through the ‘stick’ model. This would explain why the highest ratios between ‘false carrot’ and ‘stick’ are seen in those with the highest insurance contributions: the middle age group with a factor of 3.7, and, at around 2.5, females, the high and middle income group, and those with low BMI, see Table 6.3. A factor of two was seen in the old age group, and those with excellent, very good and good health. No change was found in the young age group, males, the low-income group, those with fair and poor health, and the obese. Somewhat jarring with this interpretation is, however, that the overweight also showed a factor of 2.5, although motivational factors may be at work here. In any case the high prevalence of the V-pattern across demographic groups can be surprising for those who would have assumed that the concept of solidarity underpinning German sickness funds would make German respondents more reluctant to accept overt penalties than their American counterparts.

---

<sup>9</sup> Respondents were not asked what amount would be most fair, or effective, but the question stem was deliberately open to obtain a measure of overall appropriateness (e.g.: “How much money should the plan give back to people who are not overweight, to create weight-control incentives?”)

<b>Table 6.3: Differences between demographic groups (in \$ and as factors) between “false carrot” (FC) and “stick” (S), US1 and Ger</b>									
		US1				Ger			
		FC	S	\$ +/-	Factor	FC	S	\$ +/-	Factor
Age	Young	100	50	-50	<b>0.5</b>	66	66	0	<b>1</b>
	Middle	25	30	+5	<b>1.2</b>	7	26	+19	<b>3.7</b>
	Old	25	23	-2	<b>0.9</b>	7	13	+6	<b>1.8</b>
Gender	Female	50	50	0	<b>1</b>	7	17	+10	<b>2.4</b>
	Male	50	50	0	<b>1</b>	66	66	0	<b>1</b>
Income	High	100	100	0	<b>1</b>	26	66	+40	<b>2.5</b>
	Middle	50	50	0	<b>1</b>	26	67	+41	<b>2.6</b>
	low	50	20	-30	<b>0.4</b>	26	26	0	<b>1</b>
health	Excellent	100	100	0	<b>1</b>	99	178	+79	<b>1.8</b>
	V good	100	50	-50	<b>0.5</b>	66	125	+59	<b>1.9</b>
	Good	50	20	-30	<b>0.4</b>	13	26	+13	<b>2</b>
	Fair	20	10	-10	<b>0.5</b>	0	0	0	<b>1</b>
	poor	20	25	+5	<b>1.25</b>	0	0	0	<b>1</b>
BMI	normal	100	100	0	<b>1</b>	26	66	+40	<b>2.5</b>
	overw.	50	28	-22	<b>0.6</b>	13	33	+20	<b>2.5</b>
	obese	10	0	-10	<b>0</b>	66	66	0	<b>1</b>

#### *6.4.2 Similarities and differences between views of survey respondents and physicians*

There were several similarities, but also striking differences, between the interviews with primary care physicians and oncologists in Berlin and Philadelphia, and the population level surveys US1 and Ger. While it is not particularly meaningful to make direct comparisons between the 20 responses of physicians and the more than 3,000 survey respondents, a contrasting analysis can nonetheless be useful to identify relevant hypotheses for follow-up research, and the discussion below should be read in this light.

##### 6.4.2.1 Appropriate levels for weight-control incentives

Physician’s views on using incentives in the ‘carrot’ format were, broadly, in line with those of respondents in samples US1 and Ger. By contrast, physicians were far less welcoming of the surcharge model. Only two American oncologists were unreservedly in favour of the ‘false carrot’ model<sup>10</sup> for weight-control incentives (suggesting around \$250), and one US primary care physician (proposing \$1,000), albeit with some qualifications. Those opposed viewed the premium increase as punitive or as sending the wrong signal, assuming also that people responded better to positive reinforcement. There were also concerns that higher premiums could result in barriers to care and that social cohesion would be strained, with the healthy

<sup>10</sup> As noted, the physician guide did not include the ‘stick’ scenario (Section 6.2.2 and Appendix 6A).

viewing unhealthy health plan members as directly responsible for the insurance premium increase. Such reasoning seemed to be less prominent in the survey samples US1 and Ger, as between 58–66% of respondents (see Table 6.1) entered amounts larger than \$1 for the ‘false carrot’ and ‘stick’ scenario and hence agreed with it, with levels as described above.

#### 6.4.2.2 Personal behaviour and other factors influencing health

Perhaps the starkest contrast between the views of physicians and the population level surveys was found in relation to the first question, which asked about the most important factor determining health. While 37% in US2, 41% in US1 and 43% in Ger settled for the first option, which suggested that personal behaviour played this role, none of the physicians shared this view. Instead, all but five chose the second option, which posited that it was impossible to identify a single most important factor. In the survey samples between 56–60% of respondents agreed with this view. One of the American oncologists who selected the third option—on which factors other than personal behaviour were generally the most important determinants of health—suggested that while the other views had some plausibility, ultimately the situation was similar to

the illusion of free will, that we think we have free will in a lot of aspects of our lives ... but there’s an enormous tilt on the table ... and background, and the environment [are] huge determinants” of good and poor health (US-4-ONC-M).

Several American physicians also highlighted unequal access to health care. However, only between 1–3% in the samples US1, US2, and Ger would be likely to agree with such assessments. The optimism regarding the potential of personal behaviour can be relevant in several ways. More than a third of respondents would be likely to think that incentive programmes that are directed at personal behaviour can work in principle, in addition to a sizeable proportion of those choosing the second option. The finding could also explain some of the endorsement of potentially penalising policies, which might be lower, if respondents were more aware of the social determinants of health: While 59% in US1, 57% in US2 and 58% in Ger supported the car insurance analogy, only four of the physicians were willing to agree (all added qualifications, and only one was a primary care physician).

Physicians also differed from the population level surveys with regard to the ranking of five scenarios that explored the role of effort, desert and health behaviour in relation to the weight-control. The overall rankings are reproduced in Table 6.4, showing two inversions in the physicians' assessments. The scenario in which someone tries to change their behaviour, but fails in achieving the target had the lowest approval by both US and German physicians. By contrast, in US1 and Ger this spot was given to the scenario in which people receive \$150 for meeting the target, despite their motivations running counter the programmes' goals. A less clear inversion was given regarding the second and third ranks. More survey respondents found the scenario fair in which those who already have healthy habits receive incentives, than the one in which those who were overweight as part of their voluntary lifestyle choices do not receive them. The US physicians shared this ranking, but the German ones did not, with the order being reversed (although by the smallest of margins, and in a sample that was overall very small). As noted, quantitative comparisons between the survey population and the physician interviews are far from straightforward. Still, the fact that physicians ranked as the least fair that someone does not receive \$150 despite attempted behaviour change, could be worthy of further exploration, not least because it might mean that they are more likely to bend the rules when it comes to ascertaining whether a patient should receive an incentive amount (see section 6.4.2.3).

Scenario/ranking	Survey samples			Physicians		
	US1*	Ger		Total†	US	Ger
1 \$150 for successful behaviour change	77	67	\$150 for successful behaviour change	19	10	9
2 \$150 for existing health habits	77	64	No \$150 if overweight is lifestyle related	17(16)	9	8
3 No \$150 if overweight is lifestyle related	71	57	\$150 for existing health habits	17(11)	10	7
4 No \$150 for failed behaviour change	56	44	\$150 as long as target is met, even if motivation is not aligned	14	8	6
5 \$150 as long as target is met, even if motivation is not aligned	50	33	No \$150 for failed behaviour change	11	6	5

\* Per cent of respondents opting for “completely fair” and “fair”.

† Frequencies out of a total of n=20, aggregating “fair” and “fair with qualifications” (in the case of the draw, the numbers in () state the subset of “fair” only).

Regarding the question of whether it was reason for concern if half as many in the low-income group benefited from incentive programmes in comparison with the

high-income group, a marginally higher proportion of physicians was uneasy than survey respondents: In sample Ger, 46% found this completely unfair or unfair, and 37% in US1. Twelve of the twenty physicians shared this view (two of these with qualifications), and questioned whether programmes that benefitted people who were already privileged were useful. Others however cautioned that ‘the perfect should not be the enemy of the good’ and that policies would always be somewhat coarse instruments.

Overall, the comparison between physicians and survey populations can suggest that the former are far more aware of the social determinants of health and opposed to overt penalties, more concerned about differential benefit, and also more concerned if those who make an effort at behaviour change lose out. In the survey populations, it is noteworthy that by a 10% margin, more of the German respondents have fairness concerns if poorer people benefit less from weight-control incentives.

#### 6.4.2.3 Colon cancer prevention

A clear contrast was found between physicians’ assessment of the three colon cancer prevention policies that featured in modified form in the surveys. Physicians were largely opposed to the current German policy requiring attending counselling and not refusing treatment. In particular, the compliance requirement met with opposition, and the fact that patients would face a financial burden at a time when they also have to cope with cancer treatment. Only three out of 20 physicians accepted the rule, and only with qualifications. Equally, there was opposition to the penalty for failing to undergo a colonoscopy at the relevant age, with concerns centring around the lacking evidence base and coerciveness. However, a far more complex picture emerged regarding the proposal to require attendance of a counselling session: nine physicians had similar concerns, but eleven accepted it (four with qualifications) believing that it could help improve health literacy and outcomes.

By contrast, respondents in sample Ger appeared to treat the three policies as largely equivalent, with 36% opposed to the first two scenarios, and 31% disagreeing with the counselling-only policy. In US1, 27% were opposed to the current German policy requiring counselling and compliance, 7% more disapprove of the screening policy, and around the same (5%) the counselling-only approach. It can be surprising that respondents in US1 and Ger make no major distinction between the

compliance requirement and the policies that do not entail this element, given the salience of this point in the physician interviews. In particular, it is striking that in the German sample 65% of respondents agreed with or accepted the policy, and would hence not be likely to see the need for reform that was implied by the physicians, and was also documented in position statements by professional groups, as noted in Chapter 2. One possible explanation is that the situation of a patient wishing to refuse treatment is more tangible to physicians than to the respondents, who might have focused more on the suggested benefits (health improvement, cost savings) than the burden to the patients.

#### 6.4.2.3 Incentives and the doctor-patient relationship

There were, however, also several areas where physicians and the survey respondents agreed. First, regarding the question of how many out of ten physicians would comply with a patient's request to record a lower weight so that an incentive can be secured, there was remarkable agreement regarding the median and mean in US1 (six) and Ger (five). Matching this situation, 15 of the 20 doctors thought that 50% of colleagues would comply with a patients' request. Regarding the parallel cancer scenario, the median and mean number in US1 and Ger was near identical to respondents' estimate for the obesity scenario (five in both samples). However, 17 physicians estimated a considerably lower number, between one and two. Key reasons for physicians in the obesity case included that not complying with the patient's request could undermine the effort they made so far; that it showed a lack of empathy; that weight thresholds were arbitrary to some extent; and that the professional consequences were minor at best. In the cancer scenario, several physicians equally observed that they felt sympathetic with a patient rejecting aggressive chemotherapy, but here, complying with the request would entail falsifying treatment records with potentially severe legal consequences. In both cases, physicians reported that being requested to lie made them feel uncomfortable. Physicians also reported that requests for lies came up in other contexts, and it in this sense the case of incentives is not unique. However, there is no need to add stress to the doctor-patient relationship if this can be avoided, and it would therefore be useful to consider options in which incentive can be implemented in ways that do not implicate physicians in this way, insofar as possible. The survey responses suggest that doing so would not merely represent a response to a theoretical possibility.

Leaving aside the negative impact on the doctor-patient relationship, a further important reason for finding alternative approaches was given by one of the physicians, as compliance with patients' requests risks interfering with the generation of robust evidence about the effectiveness of incentive programmes.

#### 6.4.2.4 Agreement in other areas

For the remaining scenarios there was broad agreement between physicians and survey respondents: hence, there was a mixed picture in both groups regarding the acceptability of the weigh control incentive of \$150, and equally both groups agreed that the two scenarios testing framing (in the case of obesity and colon cancer) were somewhat ambiguous, oscillating between 'carrot', 'stick' and "both carrot and stick".

#### *6.4.3 Winners and losers of incentive programmes*

Several survey questions aimed to elicit notions of fairness and deservingness, and there were noteworthy findings regarding the assessments of those more likely to lose out. The survey also offered insights into winners and losers of incentive programmes over time.

##### 6.4.3.1 Winners and losers: perceptions regarding the role of effort and motivation

In terms of perception, around two thirds of respondents in both US1 and Ger agreed that people who had either successfully changed their behaviour and lost weight, or were healthy by habit, were the fully deserving of incentives, and in this sense, legitimate winners. However, only half in US1, and a third in the German sample were willing to see as winners people who merely happen to meet the weight target, but in fact lived unhealthily. With regard to losers, just over half in US1 (57%) and just under half in Ger (44%) agreed that those who attempt behaviour change but do not achieve targets should not receive an incentive. Overall, the survey responses can therefore suggest that while results matter, intentions matter also. This can explain why the scenario of successful behaviour change and being healthy by habit were seen as equally fair in both samples, since both groups have the intention of achieving good health, even if only one of them changed their behaviour. Respondents do not seem to find that achieving behaviour change, rather than maintaining a healthy status quo, should receive a mark up in deservingness, nor that those who did not require the prompt of the incentive programme in the first place



were more deserving. At the same time, making an effort alone is not viewed as sufficient, as is clear from the views regarding those who attempted, but failed in achieving, behaviour change. This finding is of some relevance for concerns around cream-skimming, as it seems that respondents are not concerned if people who are healthier by habit—which is generally more common among the better off, see Figure 6.4—receive incentives even though they clearly do not need them for behaviour change.

#### 6.4.3.2 Winners and losers: perceptions regarding the role of income

However, cream skimming concerns appear to become more prominent when the weight-control scenario was geared towards financial status and suggested that twice as many in the high-income group benefited from incentives compared to the low-income group—which, as it turned out, is in fact the case with regard to current use of both participation and attainment-incentives in US1 (but not Ger).<sup>11</sup> Thirty-seven per cent in US1 and 46% in Ger found this completely unfair or unfair, however, there were marked differences across income groups: in both US1 and Ger half of the low-income groups had fairness concerns, but only around a fifth of the high-income groups. This situation suggests that half of the low-income group (and more than a third of all respondents) find the current incentive situation in the US unfair. Taking together the responses to this scenario and to the five questions regarding effort, motivation and opportunity of choice, such reasons would be especially relevant if it should be the case that a large number of people receive incentives who are high-income earners but behave in ways that are not aligned with the programmes goals.

#### 6.4.3.1 Winners and losers: victim blaming and self-victim blaming?

A striking finding regarding the car insurance analogy, which is widely viewed critically because of its potentially penalising characteristics, was that not all those likely to lose out under the policy were opposed. While in both samples around 60% of the obese and 40% of the overweight disagreed/disagreed strongly, in Ger 14% of obese agreed/strongly agreed, as did 35% of the overweight. In US1, 27% of the obese, and 41% of the overweight shared this view.

---

<sup>11</sup> In US1 differences between income groups were significant regarding current attainment and participation -incentive use. Participation -incentives: high-income=17% vs. low-income=7%. Attainment-incentives: high-income 14%, vs. low-income=4%. (In Ger the difference was not significant. Participation -incentives: high-income=26%, low-income=19%. Attainment-incentives: high-income=19%, low-income=13%, see Figure 6.7.

The text of the question was unambiguous in suggesting that the justification for the policy is to penalise voluntary risk taking. The fact that more than a third of the overweight appear to feel that this image is applicable to their situation can be surprising. As noted in chapter 2, a major concern in the ethics literature is victim blaming, in which case people are held responsible for factors that are, in fact, beyond their control. The survey findings seem to suggest that a sizeable proportion of the overweight either rejects such concerns as overly paternalistic, or engages in some form of self-victim blaming, seeing themselves as speeding drivers, but not realising—to stay in the metaphor—that their car is driving with considerable tail wind (or a mechanical defect).

In this context, it is noteworthy that responses to the question about what the most important factor determining health did not show a strong correlation with BMI in either sample, corroborating the notion that higher body weight is viewed as largely behaviour-dependent (health status, by contrast, did correlate with significant differences, with more than 50% of the best health group selecting personal behaviour as the paramount factor, and between 20–30% in the worst health group). While around a third of the overweight therefore appears to agree that they deserve a penalty, it also needs to be borne in mind that 10% more disagreed, and moreover, as noted, that the levels of penalties overweight respondents set for the ‘stick’ scenario in both panels were low (median [inter-quartile range] \$27 [0–175] in US1 and \$33 [0–214] in Ger).

#### 6.4.3.2 Winners and losers by actual and intended use and demographics

In terms of actual and intended future use, overall, younger, higher income, better health and lower BMI groups tended to use both participation and attainment-incentives more than those of other groups. However, as Table 6.5 shows, differences were not always significant and were most frequent regarding future use of participation and attainment-incentives in sample US1. There were also some exceptions to the overall pattern.

Table 6.5: Significance of independent variables for past, current and future use of participation and attainment incentives in samples US1 and Ger												
	Sample US1						Sample Ger					
	Participation			Attainment			Participation			Attainment		
	P	C	F	P	C	F	P	C	F	P	C	F
Age	-	-	x	-	-	x	-	-	-	-	-	-
Income	-	x	x	x	x	x	-	-	-	-	-	-
Health	-	x	x	-	-	x	-	-	x	-	x	x
BMI	-	-	x	-	-	x	-	x	-	-	-	x

Unequal uptake can be a matter of concern especially where incentive amounts are substantial. In this regard the significant differences regarding income in US1 are relevant. As noted, for participation-incentives more than twice as many of the high-income group (17%) than the low-income group (7%) are current users. For attainment-incentives (where significant future increases in levels of incentives are expected) the ratio was more than three-fold, and only 4% of the low, but 14% of the high-income group are users. At the same time intention for future use was considerably increased among the low-income group: 41% plan on using participation-incentives, and 32% attainment-based ones (the most well off group was at around 50% for both types). While caution about the dependability of future uptake has already been urged above, the margin of around 30% between current and future use was somewhat higher than in the high-income group (around 25%) and can suggest that low-income groups do not feel excluded by default, and moreover, are interested in using incentives (whether this interest is driven by health promotion potential or the proportionally higher incentive amounts).

The same striking difference between current and future use could also be observed regarding health status (Figures 6.8 and 6.9), and was perhaps most dramatic regarding current attainment-incentive use in the German sample, where 25% of those with very good health mark the highest level of users, but none of those with poor health use them, although 47% intend to do so in the future.<sup>12</sup> With one exception, in both US1 and Ger the margins between current and intended future were larger in lower health groups than in higher ones (US1, from low to high, in %: 39, 38, 30, 28, 25; Ger: 47, 34, 43, 36, 24). The anomaly in the German sample was due to a further interesting finding in overall usage. Whereas in the American sample uptake at all three time points increased with health status, in the German sample those with excellent health ranked lower than those with good and very good

<sup>12</sup> However, the group was small: n=23 (out of a total of n=23 with poor health).

health in terms of past use; lower than those with very good health regarding current use, and were in fact the lowest user group when it came to intended future use (see Figure 6.9). While these findings are not straightforward to interpret, one explanation would be that the health promotion aspect of incentive programme is salient for respondents: accordingly those with excellent health see little need to use programmes, whereas lower health groups do not feel that incentives are out of their reach, and moreover, feel that programmes can help improve health. The fact that this pattern was not seen in the US sample may be to do with the greater awareness of cost-sharing, and far higher levels of incentives.

Consistent with the pattern for health, intended future use by BMI also showed far higher levels compared to current use in both US1 and Ger. For both types of incentives in US1, those with normal BMI express the highest interest, but the levels for the overweight and obese were not low (as one might perhaps expect in particular for the obese): for participation-incentives percentages were, from low to high weight: 53, 46, 40, and for attainment 45, 39, 35. Current use for the three groups was between 8–12%.

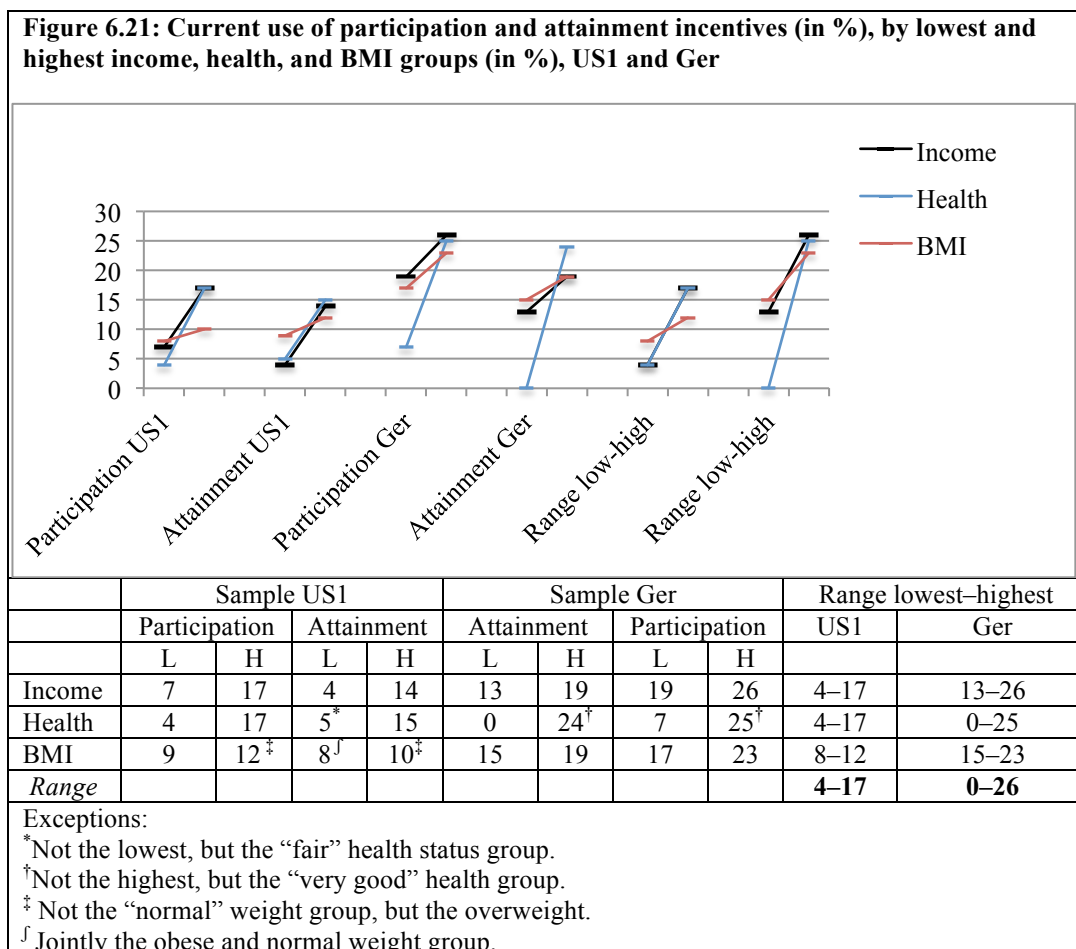
It is also interesting that the margins between intended use of participation and attainment-incentive become narrower as the weight of respondents increases: 8% more of those with normal weight prefer participation over attainment-incentives, with 7% more of the overweight, and 5% more of the obese sharing this preference. Given that participation-incentives are typically less demanding one might have expected an inverse progression, with more of those with normal weight willing to take more robust challenges posed by attainment-incentives, but the data suggest this is not so. A further interesting finding in the US sample is that for both past and current use of participation and attainment-incentives, weight was *not* significant. Moreover, unlike for future use, in which those with normal weight represented the highest users, the order was changed and the overweight, and not those with normal weight alone, were the highest user group.<sup>13</sup>

---

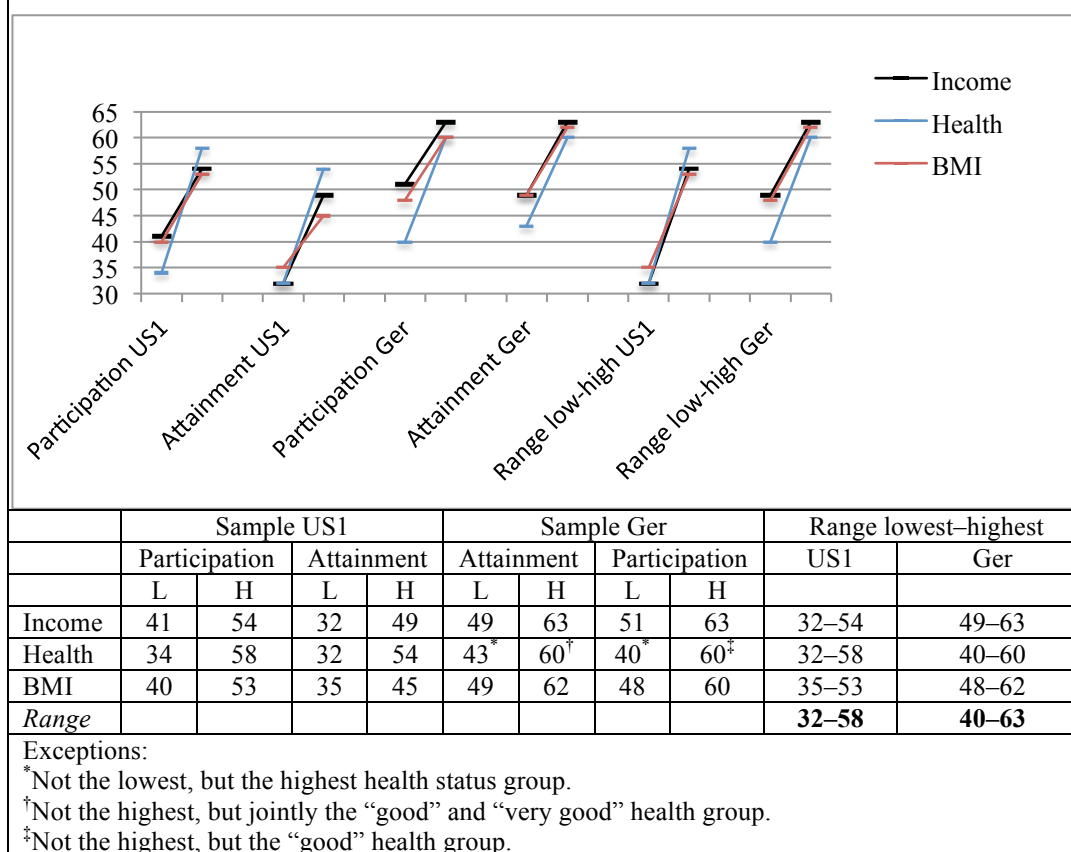
<sup>13</sup> Uptake was as follows (all figures in %, see Figure 6.10): past attainment-incentive use: overweight (16), obese (14), normal (12); current attainment use: overweight (10), obese and normal (8); past participation -incentive use: overweight and normal (22), obese (19); current participation -incentive use: overweight (12), normal (10), obese (9).

Overall then, while there were significant differences regarding age, income, health and BMI showing that the better off benefit more, these were not uniformly given. Levels among low-income, poor health and high body weight groups regarding future use of both participation and attainment-incentives were between 40–49% in Ger (best off: 60–63%); in US1 the range was 32–35% vs. 53–58% for the more privileged group (see Figures 6.21 and 6.22). While the levels of the worse off were therefore considerably lower, there is still more than a third who intend to use incentive programmes in US1, and more than 40% in Ger.

Again, intention for future uptake cannot be taken as prediction of what will happen, but the data do suggest that sizeable proportions of less privileged groups find incentives attractive on some grounds, possibly, because they are perceived as effective means of health improvement, and/or because the monetary value has traction.



**Figure 6.22: Intended future use of participation and attainment incentives (in %), by lowest and highest income, health, and BMI groups, US1 and Ger**



## 6.5 Limitations

Several limitations need to be noted. Leaving aside the advantages and disadvantages of online surveys as opposed to other survey forms, that have been discussed in section 6.2.1 and 6.2.5 it needs to be recalled that only US2 was a random sample, and although samples US1 and Ger were census-adjusted though proportional quota sampling by income, they cannot be taken to be representative of the population. While for questions 1 and 13 differences between US1 and US2 were within a reasonable margin of error due to incomplete sampling overall (below 5%), some of the variations in responses to questions 2, 3 and 12 differed by up to 26%, and there were a number of differences in the significance of independent variables (see Appendix 6B).<sup>14</sup> Some of these differences became less pronounced

<sup>14</sup> The discrepancies were as follows (for each question the answer options are listed, followed by the percentages for US1 and then US2): *Question 1*: factors determining health: personal behaviour (41/37); both personal and other factors (58/60); other factors (1/3); *Question 2*: car insurance: strongly agree (10/7); agree (31/20); neutral (18/30); disagree (28/22); strongly disagree (13/21); *Question 3*: \$150 for normal weight: strongly agree (21/14); agree (46/20); neutral (15/30); disagree (12/19); strongly disagree (6/16); *Question 12*: \$250 surcharge for not attending colon cancer screening: strongly agree (10/7); agree (36/25); neutral (23/26); disagree (20/23); strongly disagree

when Likert scale options were aggregated: variation between US1 and US2 regarding combined responses for “disagree” and “completely disagree” was 2% in question 2, 17% in question 3, 10% in question 12 and 7% in question 13. The only way of correcting these discrepancies would be to field the full surveys US1 and Ger to a random sample. Insofar as differences between US1 and US2 should be attributable to differences in the kind of people joining river-sampling-based online panels as opposed to accepting recruitment to a access panel as maintained by companies such as Knowledge Network, and to the extent that the same mechanisms are at work in the US and in Germany, the comparisons made here between US1 and Ger would be strengthened. These limitations are therefore relevant, in particular, for generalisations from the data. However, they are overall less relevant for the experiment regarding levels for the three types of incentives, and the ranking of the five different options exploring the extent to which effort and opportunity of choice can affect deservingness of \$150 for weight-control.

A separate limitation regarding these two sections of the survey is that although the respective questions were randomised, a stricter analysis would focus only on comparing the first options that respondents were presented with. This has not been done for two main reasons. First, due to the overall sample size Chi-Square was suspect in some of the contingency analyses, chiefly regarding health, because relatively few people reported either excellent or poor health. Working with a three or five fold smaller sample would have exacerbated this problem for other variables. Second, it was technically not straightforward to identify only the first responses, as the platform *SurveyMonkey*, that was used for the *Usamp* surveys, helpfully offers various options for question and page randomisation, but does not provide randomisation variables as part of the data files (instead, and very unhelpfully, it is necessary to download a report for each response individually which records the sequence as presented). The results as presented here therefore contain some noise and may vary in comparison with a purer analysis that focuses on first responses only. In addition, the levels provided for the three scenarios of ‘carrot’, ‘false carrot’ and ‘stick’ need to be viewed in light of the fact that the preceding question introduced as a real policy a weight-control incentive of \$150, which is likely to have

---

(12/19); *Question 13*: \$250 surcharge for not using colonoscopy: strongly agree (8/8); agree (35/26); neutral (23/25); disagree (22/23); strongly disagree (12/18); see also the variations in independent variables, Appendix 6B).

had an anchoring effect. However, the effect was uniformly applied to all respondents, and moreover there was significant variation around this amount, so it did not seem that the normativity of the factual constrained all respondents' subsequent selections for appropriate amounts. However, this question can only be answered by repeating the experiment without providing any prior examples of incentive levels.

A further limitation had to do with the currency conversion rate and cognitive anchoring around round numbers or sums. For example, several of the German figures for acceptable incentive amounts rather clearly reveal conversion factors from € s to \$. Thus, while we might say that the German sample opted for a higher level of \$132 for a particular scenario, whereas US respondents set, say, \$100, the German respondents did of course not actually think that \$132 would be the right sum, but thought that €100 would be adequate, i.e. \$132 at conversion rates at the time, etc. To some extent, these differences would therefore need to be factored in when making direct comparisons, perhaps along with differences in purchasing power parity; however, no such adjustments have been made here, as the current presentation seemed sufficient for the intended purposes.

Lastly, while the surveys asked about perceptions of framing of the weight-control and colon cancer prevention policy, different views might have been expressed had the scenarios been described differently. The same, of course, goes for the other questions and is a common place in survey design. However, it is of particular relevance to the topic here, as the acceptability of incentives can be very closely linked to the way incentives are framed.

## **6.6 Conclusion**

This chapter described the methodology and execution of three cross-sectional surveys with American and German respondents and discussed salient similarities and differences between respondents, as well as in relation to the more qualitative findings from the interviews with physicians in Philadelphia and Berlin. The chapter addressed sub-research question 3 and sought to shed more light on which aspects of cancer care and obesity policies that focus on individual behaviour change respondents found reasonable, and which ones they found objectionable.



Overall, differences between countries were far less pronounced than might have been expected, given the different health systems and implicit and explicit normative assumptions and cultural backgrounds underpinning these. Respondents in both samples used incentives more in the past than currently, but many more intend to use them in the future. Despite some variation, there was no clear indication that participation-incentives would be preferred over attainment-incentives, suggesting that respondents are not categorically opposed to harder targets (not even lower income, poorer health<sup>15</sup> or higher BMI groups, even if levels were lower). Nor was the ‘carrot’ vs. ‘stick’ distinction alone sufficient to determine acceptability, as a majority (even if by a slim margin of 8%) in US1, US2 and Ger supported, or was neutral with regard to, all incentive scenarios that were presented, only one of which was framed as a ‘carrot’ with four in the ‘stick’ format. Penalties for the overweight had support across all samples, including by a third of the overweight. However, high levels of penalties for overweight, as advocated by some, were not supported. The importance of effort and motivation was salient in both samples, and the weight-control scenario revealed high levels of acceptance for people who achieved weight loss or control, but low levels for those who just happened to meet targets. Regarding colon cancer prevention, all policies were framed to seek to promote health and control cost, and the majority accepted these policies, even if they entailed not refusing treatment, which led to considerable opposition in the physician interviews: so neither the extent of coercion nor the explicit rationale of cost saving was sufficient for a majority of respondents to reject the programmes.

The next chapter will draw together the findings from surveys US1, US2 and Ger and the physician interviews regarding colon cancer prevention, and apply the evaluative framework set out in Chapter 3 to this specific context, to illustrate how the approach advocated here, and the data generated, can help in the formulation of well-justified policies aimed at promoting personal responsibility in value pluralist societies.

---

<sup>15</sup> In Ger, none of those with poor health currently use attainment-incentives, however 47% intend to do so in the future: while the finding regarding current use may be cause to doubt the statement being made here, such doubt would be stronger if future use was equally low, or at least not as high as was the case.

## Chapter 7<sup>1</sup>

### Policy implications: the case of colon cancer policy

#### 7.1 Introduction

While general conclusions arising from the findings of the previous chapters will be presented in Chapter 8, this chapter seeks to draw conclusions for the specific area of colon cancer. The aim is to demonstrate the utility of the conceptual framework presented in Chapter 3 and the relevance of the findings from interviews with physicians (Chapter 5), and from surveys with the public (Chapter 6).

The general approach will be one of triangulation. In geometry, triangulation designates a process whereby the position of an object is identified by measuring at what angles it lies, with regard to known positions. In a metaphorical adaptation for the social sciences, triangulation means using a variety of methods to approach a particular phenomenon, recognising that a single method will often fail to capture important dimensions. Following Norman Denzin and Michael Patton, one can distinguish between four types of methods: using different data collection methods to ascertain consistency of findings (such as qualitative and quantitative data); examining different sources within the same method (for example, surveying similarly situated individuals with different views); using different analysts or observers to gather or interpret data; and different theoretical or normative frameworks to analyse findings (Patton 1999; Denzin 1978).

Chapter 3 already addressed the question of how competing normative frameworks with regard to health responsibility should be dealt with, and proposed a framework in which explicit justification can be provided in seven principal domains. This chapter will therefore build on that framework, rather than discuss findings against a range of other theoretical frameworks. While data gathered and

---

<sup>1</sup> Text in this chapter draws on, and in part reproduces verbatim, material published in Schmidt 2012. An earlier version of this chapter was also presented at the workshop “Public Health Ethics–Scientific methods, foundational concepts, and case analyses” Hanover Medical School, February 2011, and I am grateful to the participants for constructive feedback. Thanks are also due to Stephen John, for critical comment and suggestions for improvement, and to Dirk Horenkamp-Sonntag and Jürgen Wasem for advice on specific policy issues, as noted later in this chapter.

analysed in Chapters 5 and 6 did not benefit from additional interviewers, coders or interpreters, the two perspectives of physicians and the public still allow for important triangulation. For physician interviews enable a deeper insight into what particular policy alternatives mean for their roles and their understanding of the doctor-patient relationship. The quantitative data from the surveys enables a better understanding of what policy options mean for people of different social or health groups. As the discussion will show, there are further important perspectives that would need to be considered for an exhaustive analysis, chiefly, epidemiological and economic angles. The analysis here will therefore not be definitive—however, the perspectives and data brought together should hopefully help to sharpen significantly the focus of discussion, and separate less, from more relevant reasons.

## **7.2 Colon cancer prevention: denying the right not to think, but preserving the right to refuse treatment**

In the following I use the example of the current German colon cancer policy introduced in Chapter 2, Section 2.3.1, and taken up again in the instruments used for the physician interviews and population-level surveys. To summarise briefly: the policy aimed to promote public health, personal responsibility, and health system efficiency by, effectively, requiring people to think about their health and complying with treatment recommendations, or else face a potential financial disadvantage. The mechanism drew on existing co-payment arrangements. While the total burden is capped for all insurees at 2% of their gross annual income, for the chronically ill this level is reduced to 1%. Specific provisions in Article 62 SGB V and subsequent binding guidance by the *Gemeinsame Bundesausschuss* (G-BA, Federal Joint Commission) state that cancer patients will be eligible for the 1% threshold only, if they have attended counselling sessions on the advantages and disadvantages of the respective screens at the relevant age, and have not refused treatment should they require it (either cancer treatment or adjuvant care). If patients do not comply with one or both of these conditions, their co-payments cap is 2%, that is, the same as for all non-chronically ill.

While the actual German policy is applicable for three types of cancer (cervical, breast, and colon cancer), the focus is here on colon cancer only, as the two other areas raise separate and complex issues. In terms of the conditionalities, the

policy has two central components. One concerns compliance, and the intention here is to show that the compliance-requirement is highly questionable and should be abolished. The second can be put as: should there be a right not to think about one's health, when it comes to colon cancer? The answer to this question will be a cautious and qualified 'no'.

As noted in Chapter 3, in pluralist societies, consensus on the right normative theory of personal responsibility must remain elusive (Section 3.4–3.5). Moreover—and just as important—even if a society were to agree on a single normative theory, such a framework is typically at such a level of abstraction that it is not suited to provide concrete guidance to policy makers with regard to designing, implementing or evaluating a particular intervention. The framework I proposed here therefore established a number of 'tests' that concern the impact of a policy on key normative and structural values and components that are integral to practically all health care systems. These tests (see Box 3.2) mark the areas in which justification ought to be provided, and concern seven principal domains that will be used as a matrix to structure the discussion that follows.

### *7.2.1 Evidence, rationale, and feasibility*

This Section concerns a discussion of rationales related to health care and ethical and political aspects.

#### 7.2.1.1 Health care-related rationales

In both Germany and the US, cancer is the second most common cause of death. Colon cancer is the second most frequent metastatic malign tumour in both men and women in Germany, and the third most common one in the US. Incidence increases with age, and for Germany 73,000 new cases were projected for 2011 alone, with an estimated 26,000 deaths (Robert Koch-Institut 2010: 36). In the US, 141,000 new cases were predicted for the same timeframe, and 49,400 deaths (American Cancer Society 2011: 1). Mortality progresses relatively slowly: survival rates for persons with colorectal cancer are 83% after one year after diagnosis, 65% after five years, declining to 58% at 10 years. If (localised) colon cancer is detected and treated at an early stage, five-year survival is 90%, although due to relatively low screening

uptake rates, comparatively few cancers are detected at this stage (39% in the US, American Cancer Society 2011: 7).

Colonoscopies involve the insertion of a hollow, lighted, flexible tube through the rectum into the colon. The device enables a visual examination for malign stages of cancer and the removal of precancerous tissue (adenoma) in one procedure. A recent strategic policy document by an expert group of the German National Cancer Plan (*Nationaler Krebsplan*), notes that although preventative colonoscopies have been offered since 2002, uptake is low, with 15% of men and 17% of women (in total roughly 500.000–600.000 of 24 million eligible individuals) undergoing the procedure, which is covered by all sickness funds. To improve uptake, the *Nationale Krebsplan* emphasises that a concerted effort of all stakeholders is required to ensure that eligible patients make “an informed decision for or against participating in screening” (Bundesministerium für Gesundheit 2011: 8). In the US uptake is significantly higher, and there is an inverse usage with regard to gender, with 52% of men and 48% of women having had a colonoscopy within the past 10 years, or a sigmoidoscopy<sup>2</sup> within the past five years (American Cancer Society 2011: 14).

Recent evaluations from Germany estimated that for an eight year period between 2002 and 2010, colonoscopies prevented around 99,000 cases of colon cancer and in around 47,000 cases led to the detection of early stage cancer, amenable to treatment (Brenner, Altenhofen, and Hoffmeister 2010: 756). The *Nationale Krebsplan* cites research according to which the vast majority of deaths from colon cancer could be averted, if target populations would take up screening (Winawer et al. 1993; Sieg and Brenner 2007). However, sceptics point out that there is none of the data on the benefits of colonoscopies is from randomised controlled trials, and assert that, even on the assumption of high uptake rates, only three to four deaths could be averted per 1,000 patients undergoing screening (Mühlhauser 2007: 1806, see also; Mühlhauser 2011). A similar spectrum regarding documented benefits and remaining open questions was reflected in the comments oncologists made in the interviews, and G-2-ONC-L and G-3-ONC-M pointed to difficulties in terms of the evidence base (Sections 5.3.9–5.3.12). I am not qualified to settle the

---

<sup>2</sup> A sigmoidoscopy is similar to a colonoscopy in principle, but only examines around 1/3 of the colon, and cannot remove large polyps. A colonoscopy is required if abnormalities are found.

question of the magnitude of the strength of the empirical evidence, but proceed here on the assumption that routine colonoscopies are a procedure that can significantly reduce the risk of death from colon cancer for at least a substantial proportion of the population.

#### 7.2.1.2 The ethical and political case

It is equally not straightforward to assess the economic case, which depends to a significant extent on questions around how many people need to be treated to avert death, or prolong life. The *Nationale Krebsplan* emphasises with regard to treatment that cancer therapy is resource intensive and costly: six month expenditure for pharmaceutical therapy are around \$50.000, and it is expected that the availability of new drugs will lead to a significant further increase (Bundesministerium für Gesundheit 2011: 1). The extent to which these expenditures can be avoided through screening clearly depends on the balance between the cost of drugs and cost of implementation of screening. It is noteworthy that neither the commentary to the 2007 law that introduced the new cancer policy, nor the *Nationale Krebsplan* set out explicit cost estimates, but simply appear to assume that the economic case is given. The idea that there may be savings is not seem implausible, and insofar as they manifest themselves, they can, in principle, be cited in support of a version of the personal responsibility argument that expects the insured population to consider the implications of their actions on the resources that are available to meet everyone's needs. As noted in Chapter 2, this argument played a key role in the Government's overall justification—one that did not lack clarity in this regard—and also featured centrally in a statement by the then Minister of Health, Ulla Schmidt, who, as we saw earlier, welcomed the new regulations, by asserting that:

Early diagnosis enables better treatment. This is why the statutory sickness funds offer a range of screening and disease-prevention measures. I would like more people to make use of them, and the most recent reforms have created new incentives. It is important that all take more responsibility: for their own health, and also towards the community of people who jointly fund health care and ultimately pay for the treatment of chronic diseases.

For now it can be noted that a crucial element with regard to the justification—viz.: robust evidence on the economic potential—has not been provided by the proponents of the German policy, but could reasonably be expected. If it should turn out that colonoscopies do not, in fact, result in net savings, this argument falters, and would

collapse into a more one-dimensional paternalistic one, according to which insurees should behave accordingly because doing so is regarded as good for their individual health. Such a situation would have implications, in particular, for the extent to which measures might be justified that go beyond a voluntary offer of the procedure. In the absence of evidence from either policy documents or the academic literature on the magnitude of cost implications associated with a broader uptake of colonoscopies, in the following, I assume that the implicit economic case is given, but clearly, if that were not the case, another assessment would need to be made.

It is furthermore noteworthy that no explicit justification was provided for the compliance requirement, the benefits of which, it seems, were simply regarded as self-evident. Overall, as noted in the discussion in Section 2.3.3, both the debate in Parliament and the subsequent work of the G-BA was rushed, and did not involve a consideration of the full extent of relevant evidence, or involvement of all relevant stakeholders, to ensure that all relevant evidence and reasons could be duly considered. This situation undermines significantly the legitimacy of the initial decision, as well as of the G-BA's amendment, which had to operate within the constraints of the law that had already been passed.

In terms of practical feasibility, implementing the counselling requirement is indeed possible, as long as systematic documentation of who attended counselling and who did not can be provided. By contrast, feasibility with regard to the compliance requirement faces significant obstacles, which will be taken up below, and again in the discussion of the policy's impact on the doctor-patient relationship.

### *7.2.2 Intrusiveness and coerciveness*

It is noteworthy that not a single physician who participated in the qualitative interviews endorsed the German policy in its entirety, whereas 37% of the German survey sample agreed with it, and 49% of sample US1 (with 36% disagreeing in the German case and 27% in the US. Proportion in the target age group were higher both with regard to support and opposition, see Section 6.3.2.3). Physicians objected, in particular, to the compliance requirement. Their objection did not focus on a violation of patient autonomy, as, of course, the policy does not force patients to accept treatment. Instead, physicians felt very strongly that the higher cost

constituted an unhelpful and unacceptable ‘double whammy’ for people who are in a highly distressing situation to start with. As US-1-ONC-M said, in view of the situation of often aggressive adjuvant therapy and ambitious physicians who want to pursue every treatment option available: “You should be able to say ‘I’ve had enough’, without any strings attached”. Equally G-4-ONC-M noted that she found it “really extreme” to connect counselling and compliance (Section 5.3.9).

These views have much plausibility. Decision making in these situations is challenging enough, and burdening insurees additionally by having to factor in the financial impact of their decisions can reasonably be viewed as lacking respect for their situation. Moreover, even from a purely economic perspective it is far from clear on what grounds patients should be persuaded to undergo therapy if they judge it not worth the pain, and, as a consequence, are likely to die earlier than (presumably) would otherwise have been the case. On both cost and *a priori* ethical grounds, then, with regard to the situation that is given for patients as a result of the policy, there is little that speaks in favour of connecting the lower co-payment threshold with the condition of not refusing treatment. The finding that more than a third of the German respondents, and almost half of the US respondents nonetheless appear sympathetic to the policy is of concern and deserving of further exploration.<sup>3</sup>

With regard to the informational element, one might argue that, since participation is voluntary, one would be misguided to perceive any coercion at play here. People are relatively free to take up the incentive: the co-payment is doubled, but still comparatively modest. If it had been set at, say, 30% of gross annual income, then clearly most people could not afford the risk of such a burden, and it would become quasi-mandatory. But there is no canonical understanding of the concept of coercion (Ashcroft 2011: 194) and extent of the financial burden is only one dimension in which a policy can be more or less intrusive or coercive. Two

---

<sup>3</sup> In particular, it is noteworthy that the US sample was more supportive than the German one. A reading of the solidarity argument as offered in the above-cited statement by the former Minister of Health might suggest that Germans would be more inclined to agree with it on efficiency grounds, and it would not be unusual to expect that Americans would be more opposed than their German counterparts to the explicit reference to cost in the scenario. But this was not the case, and the data alone provide no explanation why this is so. One possible explanation might be that respondents lack familiarity with the side-effects of interventions such as chemo-therapy, and, unlike the physicians, therefore did not visualise the situation of a patient wishing to refuse treatment. But even if this was the case, it is not clear what underlies the difference between the American and the German sample.



further ones relate to how those affected will judge the mechanism by which the incentive has been created (a penalty is imposed if people do not comply with the condition), and the nature of the activity that is being incentivised (people are required to think about their health, and the relevance of cancer screens).

With regard to the structure used to create the incentive, it is noteworthy that, in principle, the same effect could have been achieved had the co-payment threshold been reduced for those complying with the conditions, rather than increased for those not complying. Thus, the proposal could have been made that those who attend a counselling session would have a 0.5% co-payment, rather than 1% applicable to all other chronically ill (and 2% for the non-chronically ill). Arguably, choosing between this and the currently implemented option is not just a matter of presentation, as the potential financial burden differs, and the framing will be experienced differently. While, from a behavioural economics point of view there can be good reasons to use ‘sticks’, rather than ‘carrots’, as, for example the literature shows that loss aversion can often be a more motivating factor than offering gains of some kind (Kahneman and Tversky 1979, see also Section 3.2) it is noteworthy that no such justification was provided in the documentation in which the bill was introduced (Bundesregierung 2006) nor in subsequent official documentation preceding the passing of the law. Leaving aside the general question of lacking evidence on the policy’s economic impact, it is also to be noted that no argument was presented that would have demonstrated that a gain-sharing approach is unfeasible, or that would have rationalised any possible higher cost of those not undergoing screening to the higher co-payment threshold. Instead, the reference point was simply the co-payment threshold for all non-chronically ill insures, ignoring that other chronically ill patients were unconditionally eligible for the 1% threshold. In part, it appears that political issues would also stand in the way of lowering the threshold for 0.5%. Currently around 40% of insurees are exempt from further co-payments in any year, as they reach the 2% (or 1%) threshold, and there are influential health policy commentators who argue that this threshold, already, has detrimental impact on the distribution of cost, and, if anything should be increased.<sup>4</sup> Proposals for further reductions would hence be unlikely to be welcomed.

---

<sup>4</sup> I am grateful to Professor Jürgen Wasem for elucidating this point (personal communication, 22 July 2011).

While five of the interviewed physicians agreed with the above-cited health minister and nonetheless saw the policy as a ‘carrot’, twelve viewed it as a ‘stick’, and two as both. In the population-level surveys, German respondents were almost equally split between these options, with 33% sharing the ‘carrot’ view, 34% the ‘stick’, and 34% the option of ‘both’. In line with the stronger overall endorsement of the policy among US respondents, the penalising aspect was less prominent here: 35% took the ‘carrot’ view, 25% the ‘stick’ and 40% opted for ‘both’. Physicians were also asked to prioritise three options in which either the current incentive structure would be used (with a 1% threshold available if insured comply with both conditions), or the scenario in which compliance results in a 0.5% threshold, or an alternative, in which no difference would be made between colon cancer patients and all other chronically ill. Only one physician chose the current option, and here with qualifications; four chose the 0.5% scenario, and seemed to feel that the incentive structure required tweaking. Fifteen opted to make no difference at all between different types of chronically ill: not surprising, given the general objection to the policy, based on, chiefly, the compliance requirement.

However, a very different situation arose when it came to the information element only. Seven physicians firmly opposed the implied right not to think, and seven approved of it, with six in an undecided in-between position. US-4-ONC-M, negating the right not to think noted:

all they’re doing, all they’re being forced to do, is ponder their mortality, which is not so awfully coercive... I just don’t think it meets that standard of interference with people’s lives—I think it’s ok.

G-5-PC-H also suggested that “this is at such a low level... it’s not expecting too much. I think you exaggerate liberty if you say you can’t [attend counseling].” All proponents agreed that the insured should consider the pros and cons of screening options, but also thought that “you have to stop somewhere”, proposing appealing to aspirations instead (G-2-PC-L). US-4-PC-H saw counselling sessions as similarly patronising as making people who drive a vehicle under the influence of alcohol attend special classes, and suggested that, instead, physicians should simply talk to their patients about screening as part of their routine interactions.

The survey of the German sample showed that 39% agreed with the policy, whereas 31 disagreed. In the US sample 46% agreed and 32% disagreed, showing, again, higher, and somewhat surprising levels of endorsement.

It is true that to a significant extent the policy is novel and unique. There are currently no other areas in which the insured population is required to document that they have considered the pros and cons of taking up a particular health related initiative, with the exception of counselling rules that are in place for abortions. However, here, the rationale is that the action in question is not merely self-regarding (as both going ahead with a pregnancy or terminating it will have implications for the woman concerned), but can also affect the interests of an embryo or foetus. In the protracted debates that led to the current abortion policy it was deemed an acceptable compromise to request that women demonstrate that they have considered all relevant arguments, even if the ultimate decision about an abortion should be left up to them (Funke 1996: 36). The motivation behind promoting reflection about colon cancer screening equally mixes potential self- and other regarding motivations, as outlined regarding the possible health- and cost-containment benefits. However there is no other party in a *direct* dependency relationship that would be harmed as a result of someone not undergoing cancer screening (even though refusal to do so, and subsequent later detection of cancer may result in avoidable cost that curtail available resources and therefore indirectly affect other people, and might be seen as other-regarding in this respect).

The rather exceptional nature of the approach also becomes clear if we transpose the policy into a different context. For example, taking out a private pension plan in addition to state pension arrangements is prudent behaviour in practically all countries. However, not everyone does so, even though there are low-risk and affordable products on the market. A proposal to require citizens to attend a counselling session on the pros and cons of private pensions, in order to access certain levels of state pension might be justifiable, but would no doubt meet with considerable opposition.

At this stage it needs to be noted that the revisions to Article 62 SGB V were, of course, not driven by an eagerness to deny the right not to think about one's health

as an insure of the statutory sickness fund. Rather, they represented the fallout of the amendments' genesis as it passed through parliament and into the hands of the G-BA, that had to come up with arrangements that satisfied the letter of the law as well as a the views of a wide range of different stakeholders.

Still, it seems defensible to suggest here that insofar as the evidence regarding colonoscopies stands up to scrutiny in terms of health and cost-containment benefits, the arrangement represents not merely a tolerable political compromise, but can, in principle, be an acceptable—albeit overtly paternalistic—way of promoting health literacy, personal responsibility and public health. It could also be argued that the right not to think about one's health is not as exceptional as it might seem, a view based on the following reasons. Mandatory schooling as a basis of citizenship implements the concept at a broader and perhaps the most fundamental level. Again, the rationales are twofold and concern both self and other-regarding interests, as it is assumed that basic education is a good that forms part of a flourishing life of people as individuals, and furthermore, that it is of instrumental value to furthering the prosperity of a society as a whole. While mandatory schooling has no perfectionist tendencies, and people only need to attend a certain number of years, the idea is that to function as a citizen, certain baseline knowledge and skills should be provided to all.

One problem with this analogy is of course that the eligibility age for colon cancer screening is well past that of those in secondary education. But in view of the promising health promotion and cost-saving potential; data from health literacy research pointing to relatively low awareness of the health benefits (Wuppermann, Wuppermann, and Riemann 2009); the fact that the features of colonoscopies are not such that they will ever become something people enjoy undergoing; and the resultant aforementioned low uptake rates of colon cancer screening, the principle remains the same: viz. that denying a right not to think about one's health can make sense if the evidence in prudential and economic terms is robust; the incentive structure is appropriate (I will return to this point below); the ultimate decision on whether or not to follow through with screening remains with the individual concerned; and all people have reasonable access to counselling sessions, to which I turn next.

For now the interim conclusions regarding coerciveness and intrusiveness of the policy are that it can plausibly be described as being both coercive and intrusive, on the grounds that it interferes with a sphere of decision making that many would regard as falling in the personal domain exclusively, and because the incentive is implemented in a way that it concerns evading a penalty, rather than securing an economic benefit, compared to the *ex ante* situation.<sup>5</sup>

### 7.2.3 *Attributability/opportunity of choice*

According to the dictum ‘ought implies can’ it is of little use, if not cynical, to establish responsibilities if people cannot reasonably comply with them. This is relevant in several senses to the policy considered here. First, it would be Kafkaesque to implement the penalty retrospectively for, say, a 70 year old patient requiring urgent colon cancer surgery, but who failed to attend any form of counselling 20 years ago. For this reason the inclusion criteria for the policy are such that it only applies to people who had the opportunity to chose to comply with it (or not) at the age at which screening is recommended (roughly, people born after the early 1960s). As obvious as such a requirement might seem, it can nonetheless not be taken for granted, as is clear from Article 52 SGB V, that was also revised under the same reforms. As noted in Section 2.3.1, this provision enables sickness funds to request full or partial shouldering of cost by insurees where they request treatment for complications arising from cosmetic surgery, tattoos, or piercings, or another “non-medically indicated” measure. The tension that is at work here is that it can be unfair to hold people accountable for actions that may have occurred a considerable time in the past.

---

<sup>5</sup> Of course, a further question is at what stage, and in comparison to what alternatives, such a policy might be implemented (and evaluated). It is again striking that the German government, in introducing the policy, did not make the case that all other alternative methods had been attempted first, but plateaued, or failed to achieve the goal, altogether. For example, it would be possible to evaluate the effect of information and awareness campaigns, or invitation letters from health ministries or sickness funds. The *Nationale Krebsplan* cites findings from pilot research in the Saarland, which found that an invitation letter with a list of qualified physicians led to a doubling in colonoscopies, and a similar scheme by the sickness fund AOK in Hamburg and the Rhineland led to an annual increase of colonoscopies of 75% (in neither case were absolute frequencies given, Bundesministerium für Gesundheit (BMG) 2011: 10). It is also striking that the average uptake in the US is almost 40% higher than in Germany – although all screening is entirely voluntary. However, even levels of above 50% of the target population leave room for improvement, and the suggestion above, that attaching financial benefits to attending a counselling session can be reasonable must be seen in this context, i.e., it should be pursued once other options are no longer suitable to increase or sustain uptake levels. At the same time, I would suggest that these measures, too, can be conceived of as negating the right not think about ones health, and the principal remaining difference is then the incentive mechanism that is attached.

Therefore, while the age cut off is important, the somewhat parallel case of Article 52 SGB V also points to another potential problem with Article 62 SGB V. Consider a 55 year old male who is aware of the policy, but feels little inclination to attend counselling, perhaps because he feels generally healthy, is not aware of any cases of cancer in the family, and is also sure that he could cope with the higher financial burden, should the unlikely case arise that he should require cancer surgery. Now, fast-forward 20 years and assume that with age the same man has become far more health conscious, that his economic situation has changed dramatically for the worse, and that he requires colon cancer surgery. Since he knowingly opted not to attend the counselling session he faces the 2% instead of the 1% co-payment threshold. Insisting that he pay up would certainly be within the law. But no doubt many people would at the same time feel that there is something wrong about imposing the penalty in this case. Indeed this is obvious from the quotes from physicians, who perceived a ‘double whammy’, and perhaps also from the survey findings which showed that in both samples around one third disagreed with the policy (even if it is not possible to distil from the findings which aspect they disagreed with). In general, this situation can raise the question of how stringent the requirement of attributability ought to be implemented, and whether there might be better ways of achieving the intended goal.

While there is no public documentation as to the policy’s underlying reasoning, on a charitable interpretation the rationale of requiring only those patients who actually suffer from cancer to face higher costs can be based on certain fairness considerations. For, had they acted otherwise earlier, they might have been able to avoid the cost. But it would also be feasible to focus less on the actual consequences of choices, and more on the risks implied by certain choices (such as not attending counselling), as, for example Alexander Cappelen and Ole Norheim emphasised recently (Cappelen and Norheim 2005: 478), and as Julian Le Grand noted earlier regarding tobacco policy (Le Grand 1991: 121). For there are always likely to be complications around determining the exact degree to which certain health outcomes can be said to be under the control of individuals, and it could be fairer to penalise risk taking *per se*, based on the expected cost associated with the behaviour in question, rather than discrete choices that led to undesirable outcomes.

Moreover, not delaying the psychological and financial effects associated with the incentive (be they genuine rewards or penalties), can also make sense with regard to what is known as present bias or hyperbolic discounting (Loewenstein, Brennan, and Volpp 2007; Frederick 2002), as noted in Section 2.3. For the context here, it is plausible to assume that most people would take an immediate financial disadvantage more seriously than one that may only materialise ten, twenty, or more years later. Hence, while the rationale of linking the incentive strictly to whether or not someone in fact suffers cancer has some plausibility, it can conflict with the results of behavioural economics research and can lead to penalising people in a situation of heightened vulnerability.

A further and related point has to do with the assumption, made so far, that all insurees are in fact aware of the policy, and of the availability of counselling services. While the rules are set out clearly in the law, pilot projects aside, current data protection legislation means that sickness funds are not allowed to write directly to eligible insurees to invite them to attend counselling (and remind them of the consequences of failing to do so). In particular they cannot target particular groups, for example those, who failed to use preventive services over the past two years.<sup>6</sup> Instead, they can only provide information passively and in general terms, for example, through news items in their (typically monthly) member newsletters.

As noted above, structured invitation approaches, as currently explored in pilots, can make sense with a view to promoting awareness of Art 65 SGB V. More importantly, it is reasonable to argue that they are a necessary condition for the overall acceptability of the policy. For the relevant target population should be fully informed about the rules, and be reminded about approaching deadlines in ways that make it feasible for them to attend screening in the envisaged period of two years after becoming eligible. Equally, it should go without saying that arrangements need to be in place for patients to be seen by their physicians for the counselling session within the prescribed timeframe. Furthermore, counselling needs to be conducted in

---

<sup>6</sup> As per article 284 (3) SGB V, sickness funds may only use data on health care usage for purposes of ascertaining correct accounting, and may not use them for other potentially interesting purposes. I am grateful to Dirk Horenkamp-Sonntag for this clarification.

non-directive ways, aiming to assist people in making informed decisions,<sup>7</sup> enabling the insurees to make up their own mind, rather than being steered one way or the other (as overly directive counselling can, of course, also raise coercion and intrusion issues).

The discussion of the issues raised by considerations around attributability and opportunity of choice helps identify several key features that programmes should meet, viz., that appropriate means are chosen and implemented that enable insurees to know about the programme and avail themselves of the opportunities that it can offer, in principle. Provided this is the case, it may then be permissible to follow through with the sanctions, although the discussion has also shown that, on fairness grounds, it can be difficult to hold people responsible for failing to attend counselling decades after the fact. Evidence from behavioural economics would suggest that present preference bias would favour more immediate feedback, even if this would change the nature of the incentive from imposing part of the cost of disease one actually has, to being held responsible for accepting a certain risk of undetected colon cancer.

#### *7.2.4 Equity*

Incidence and prevalence of diseases often varies with factors such as region, gender, age, household income or, perhaps, educational attainment. For example, in the US, colon cancer incidence and mortality rates are about 35% to 40% higher in men than in women, most likely due to gender-related differences in hormone and risk factor exposure. With regard to race, incidence and mortality rates are highest in African American men and women, with incidence 20% higher than in Whites, and mortality around 45% higher (American Cancer Society 2011: 3).

Uptake of preventative measures equally varies, and very few, if any, are taken up equally across different groups of people. In a recent national survey (n=582) in a random sample of African American, White, and Hispanic adults, aged 50–75 without cancer, Chanita Halbert and colleagues found that 59% of all respondents reported having had a colonoscopy. Respondents with health insurance were

---

<sup>7</sup> Effect of evidence based risk information on “informed choice” in colorectal cancer screening: randomised controlled trial



significantly more likely to report colonoscopies, as were older respondents (Halbert et al. 2011: 1315). However, no differences were found with regard to gender, educational attainment or employment, nor regarding race, in contrast to earlier work, that suggested significant differences, drawing on considerably larger samples in the area of  $n=30,000$  (Crawford et al. found that 45% of Whites reported endoscopy, 41% of Blacks and 32% of Hispanics; and focusing on Medicare recipients, White et al. found that Blacks were less likely to receive colonoscopies than Whites with a reported odds-ratio of 0.8 [0.68–0.95, CI 95%], and Hispanics with an odds-ratio of 0.73 [0.54–0.99] (Crawford, Jones, and Richardson 2010: 362; White et al. 2011: 811-812). Halbert et al. hypothesise that this may be because these earlier studies also included sigmoidoscopy and other types of colon cancer screening. In a small study of  $n=104$ , Alex Mathews and colleagues found that 77% of insured as opposed to 33% of uninsured eligible patients had completed any form of colon cancer screening (Mathews, Anderson, and Nattinger 2005: 735).

For Germany, Olaf von dem Knesebeck and Andreas Mielck carried out a survey of the population of 50 years and older ( $n=1,921$ ), and analysed differences of prevention uptake across three educational groups, as well as three income and three financial assets groups (including, for example, investments and savings). For colon cancer they found that whereas odds for low educational attainment groups were 1.0, they were 1.20 (0.89–1.62, 95% CI) for middle income groups, and 1.48 (1.06–2.07) for those in the high attainment group. Regarding equivalised net income, while the low-income group had odds of 1.0, they were 1.02 (0.80–1.30) for the middle group and 0.94 (0.74–1.20) for the high-income group. Including assets, odds were 1.0 for the low asset group, 1.32 (1.04–1.69) for the middle, and 1.36 (1.07–1.74) for the high asset group (von dem Knesebeck and Mielck 2009: 45-48).

It would generally be illusory to expect that strict equality in uptake and/or outcomes would be possible. Still, it is of crucial importance to assess whether programmes differ in terms of the extent to which different groups of people use them: while some inequality may be acceptable and a result of people's voluntary choices, others may be based on structural biases and be unfair. Incentive programmes can raise special issues in this respect. With regard to the programme considered here, von dem Knesebeck and Mielck's research can suggest that those

with above average health literacy are more likely to be aware of the programme, and many in this group would probably have informed themselves along the lines envisaged in the counselling sessions anyway—for them, strictly speaking, the programme provides no incentive to behave differently.

Others, however, might not have done so, and the programmes' existence nudges them into action, and provides occasion to attend counselling (even if this availability and the financial incentive may not be the sole reason for their action). This would be the core target group of programmes, although it cannot be assumed that all people who are offered them will fall into this category. For some people may face obstacles that make it more difficult for them than for others to use the programme. Perhaps they have a harder time understanding its relevance, or objective. Or they may have work schedules that make attending counselling sessions more difficult. Whether or not these and other possible obstacles are in fact significant depends in many ways on the precise implementation, and we have seen that some aspects, such as the rather crucial invitation mechanism, still remain to be developed further to ensure adequacy. A last group of people may object in principle to the goals and may refuse to take part for this reason.

A conceptual outline of some of the differences between these five groups was provided in Section 3.5.3, and while due to sampling constraints the survey data analysed here is insufficient to gauge reliably patterns of uptake across different groups, I note merely, that it would be overly simplistic to assume that because the programme is largely voluntary, no significant equity issues are raised, as people either decide to use it, or not. In Germany, Article 20 SGB V explicitly requires sickness funds to contribute to reducing socially determined inequalities, and it would certainly not be ideal if the programme exacerbated, rather than reduced health inequalities. Of course, even if only the (financially) better off benefitted, a net increase in health improvement (and possibly a reduction in cost) would be achieved. However, the aforementioned 'double whammy' for those who do not attend counselling, but end up suffering cancer, and face higher co-payments would be disproportionately harder on those with less disposable income. Therefore, improving uptake across social groups would clearly be desirable, in both Germany and the US (where, as noted, significant differences in incidence and mortality also

exist among Whites and Blacks), and progress towards this end should be considered in monitoring the programmes uptake and effectiveness, or modified versions thereof.

#### 7.2.5 *Solidarity/risk-pooling*

As noted, in Chapter 2, the value of solidarity, tightly intertwined with that of personal responsibility, has overarching function for the provision of health care in Germany, meaning both that individuals are entitled to claims against the community for health care related support, and *vice versa*, that the community has claims against individual insurees, as these should not make unreasonable demands on the system and contribute to its efficient operation. The concept plays no similar explicit role in the US. However, it was striking that in the interviews with physicians it was also invoked by American doctors, in particular in relation to the different cancer scenarios. As US-1-ONC-M noted regarding the policy option of imposing a penalty for not attending counselling on the advantages and disadvantages of screenings:

We're talking about health plan, we're talking about a finite number of dollars, and trying to save money and trying to keep a population healthier...you know, I don't disagree with the statement.

US-5-ONC-M went further and endorsed a financial penalty for those who refuse a colonoscopy, reasoning that this was acceptable, as the message that was conveyed to patients was merely that “if you want health insurance, you have to actually do things that we think will keep you healthy”. US-2-ONC-M equally had sympathy for the approach “if it's going to have impact on the costs of running a health plan”. US-4-ONC-M and US-5-ONC-M were the only American physicians who supported the entire current German policy, albeit with some reservations, and focussed centrally on cost in doing so. G-1-PC-H, shared their view, arguing that “those who aren't behaving in a solidaristic way have to pay more. That's quite clear.” However, G-3-ONC-M illustrated that the solidarity can cut both ways. Commenting on the three different options for different levels of incentives, the argument was made that all insurees should have the same contributions, even if this meant that healthy insurees had higher contributions: “You know, *that's* the solidarity principle, but not to penalize chronically ill.” The population-level surveys did not explore the concept of solidarity in more detail, although, as noted in Section 7.2.2 and footnote 4, the relatively high levels of support of the three cancer policy scenarios may be indicative of a view according to which solidarity can justify penalties.

Insofar as cost-savings can in fact be realised through the policy, it may enable more efficient use of resources and in this sense support solidarity. At the same time, and as a direct result of the nature of the conceptual interlocking of personal responsibility and solidarity in the law, there are also some aspects in which solidarity can be undermined. First, by not treating all groups of chronically ill equally, and increasing the co-payment for colon cancer patients who do not attend counselling, the web of solidarity has become weaker for this patient group. Second, insofar as it would be the case that those not attending counselling would include disproportionately higher numbers of the least well off, it would have become weaker in particular for those generally most in need of support.

#### *7.2.6 Affected third parties*

The German and US surveys included two scenarios in which respondents were asked what percentage of patients would ask their physician to lie, if doing so conferred a financial advantage to them. In the case of the first scenario, a \$150 bonus for weight-control was assumed, and in the second case, securing a co-payment rate that is capped at \$500 instead of 1,000. For the weight-control scenario, German respondents estimated that 50% of patients would make such a request, with US respondents centred around 60% (median). For the cancer scenario both samples were at 50%: across scenarios and countries, there were hence only minor differences. Respondents were also asked to rate how they viewed a physician who complied with a patient's request. For the obesity case 54% of German survey respondents found such behaviour wrong, whereas 12% found it right. Among US respondents, the proportions were 68% (wrong) versus 10% (right). A similar picture is given for the cancer case, where 58% of the German sample viewed a complying physician as acting wrong, and 12% as right, with the US sample showing higher rate of disapproval, with 68% opting for 'wrong', and only 11% for 'right' (Section 6.3.5.1 and 6.3.5.1). Therefore, around half of the German sample disapproved of cooperating physicians in both scenarios, and approximately 15% more in the US sample did so. The picture also suggests that, from a patient perspective, asking a physician to lie is broadly equally likely in the obesity and the cancer case, with every other patient in these situations estimated to make such a request. In two important ways these data contrast with the estimates of physicians of how many of their colleagues would, in fact, comply, with patients' requests.

First, for the obesity case, fifteen of the twenty doctors thought that more than 50% of their colleagues would comply—suggesting an almost perfect match with the survey data (Section 5.3.6). However, in the cancer scenario, only one physician opted for this ratio, with 17 estimating between 10 and 20%, leaving a far higher number of patients to themselves (Section 5.3.13).

Second, for both scenarios, several physicians pointed out that being put in the situation of deciding whether to lie or not made them highly uncomfortable, in particular in the cancer case, where many understood the patient’s interest in declining therapy, as G-1-PC-H: “I wouldn’t feel like chemotherapy—and we have to admit, many chemotherapies don’t lead to a cure but are really only done palliatively”. Still, practically all pointed to the serious legal consequences they would face if they falsified treatment documents. Conflict is therefore likely in such scenarios, as is a negative impact on the doctor-patient relationship. This situation, together with the noted element of intrusion for patients therefore presents further reason to reject the compliance element of the policy in its current form.

#### 7.2.7 Coherence

While, cumulatively, therefore, only one part of the German policy can withstand scrutiny, a further question remains regarding the informational element. Above, it was noted that there appeared to be no direct parallel of a policy that would require one to think about one’s health. In this sense the policy certainly breaks new ground. However, it can still be seen as flowing from the general provisions regarding the relationship of personal responsibility and solidarity as set out in SGB V, or, more implicitly, given in risk-sharing arrangements that also underlie most US health plans.

### 7.3 Conclusion

Regarding rationales and evidence there is little doubt that colonoscopies can prevent premature mortality, while not being the kind of intervention people look forward to undergoing. This situation, coupled with cost estimates that suggest that higher uptake could lead to more efficient resources can certainly point in the direction of *prima facie* obligations of insurees to consider whether they should

undergo a colonoscopy or other form of screening. The obligation would be the stronger, the harder the evidence is that the potential for health promotion and cost-saving are real.

Whether or not such obligations should be seen as justifiable in view of the implementation mechanism chosen in the current policy is an entirely different question, however. The practice of generating the incentive by lifting the baseline co-payment of some chronically to that of all non-chronically ill is problematic with regard to both the lack of supporting evidence from behavioural economics, and also in view of solidarity considerations that should ensure that all chronically ill are afforded a status that differs from that of the non-chronically ill. More fundamentally, it can also be questioned whether the mechanism of experiencing the financial consequences of one's actions several years, if not decades, after one was supposed to undergo screening is adequate and goal-oriented. It would therefore seem that the policy's goal could be achieved more fairly and effectively with shorter-term, genuine incentives (Section 2.2).

The discussion has hopefully also helped to illustrate what work the proposed approach for evaluating health responsibility policies more generally can do, and what not. Irrespective of the fact that the discussion here must remain incomplete because of some missing medical and economic evidence, even if we had comprehensive data on cost-saving potential, uptake of different socio economic groups, impact on the doctor patient relationship and so on, some weighing would still be required, and a solution cannot be computed in some mechanical way. However, it is hoped that the framework has thrown into sharper relief the issues that are at stake, which can help identify more clearly areas of consensus and disagreement. While it is unlikely that one area will trump all others, changes in assessments in one domain can still affect what we should think of other domains. For example, the discussion around the cost-saving rationale and the health promotion rationale has shown that proponents defended some of the coercive aspects by reference to the cost implications: but if these are not given, their case needs to be reconsidered. Furthermore, while it may be that other ethical issues are raised in further domains, the ones set out here concern the most central ones and, hopefully, enable a focused, yet nuanced discussion of the issues at stake.

## Chapter 8<sup>1</sup>

### Conclusion

#### 8.1 Introduction

It is uncontroversial that smoking has severe health risks and a clear social gradient, with more of the less privileged being smokers than of the better off. At least in the medium term smoking is associated with higher health care cost. Many people want to quit smoking, but few interventions are effective. A randomised controlled trial that was awarded the 2009 *British Medical Journal's* Research Into Practice Award as is highlighted in a recent Cochrane review as the only incentive study showing longer term successes in achieving smoking cessation (Cahill and Perera 2011) showed that a combination of incentives amounting to a \$750 reduction in health care expenditure for users led to three times higher cessation rates than alternative approaches among employees of a large American company (Volpp, Troxel et al. 2009). Persuaded that the expenditure for incentives represented good value for money, the company, General Electrics, decided to provide the programme to all its 152,000 employees. However, the proposal attracted a backlash from non-smoking employees, whose concern was that they, after all, had behaved responsibly all along, but were losing out on the cheaper cost of health insurance that was effectively offered to their smoking colleagues. The company neither stood up to the non-smokers, nor did it fold the incentive programme: instead, it turned the 'carrot' into a 'stick' by replacing the \$750 reduction with a \$625 surcharge for smokers.

This brief vignette illustrates succinctly that it is one thing to explore the effectiveness of incentives in the most rigorous scientific manner, and quite another to put policies into practice, as they can be caught in powerful moral cross-winds, buffeted by business interests and different views of fairness and the ethics of personal responsibility. Many of these aspects have been explored in the previous

---

<sup>1</sup> Text in this chapter draws on, and in part reproduces verbatim, material of two forthcoming manuscripts: Schmidt, Halpern, Asch 2012 and Schmidt, Madsion, Halpern, Volpp 2012. All sections used here were initially contributed by me. The chapter also includes material presented at the 6<sup>th</sup> Annual International Bioethics Conference of the Harvard University Program of Ethics in Health April 28–28 2011, and I am grateful to my coauthors and participants of the meeting for discussion and important pointers.

chapters. This closing section provides a brief general summary of the thesis' findings in relation to the initial research questions, and comments on overarching policy implications, including those emerging from the above scenario, as well as on limitations and next steps. Overall, the focus is less on restating findings and more on reflections arising from them.

## **8.2 Summary of findings**

The thesis' main research question had three constituent sub-questions which will be reviewed before returning to the general scope of inquiry circumscribed in the main question.

### ***8.2.1 Sub research question one: Policy context***

Chapter 2 sought to address the first sub research question:

In what way do principal US and German health care policies take into consideration, in cost-of-care decisions, the behaviour of individuals regarding health maintenance and contribution to efficient services, and what are the underlying implicit and explicit values?

As many other post-industrial countries, the US and Germany are experiencing rising trends of chronic diseases and are struggling to maximise efficiency in the provision of health care. Both countries increasingly adopt policies that focus on the individual level in seeking to align the behaviour of people covered by health insurance with the interests of those providing and paying for health care. The rising interest in the US that was noted in Chapter 2 is particularly striking: during the time that this thesis has been written, the use of wellness incentives has more than doubled. In 2009, 36% of large employers (with at least 1,000 employees) used incentives; in 2011 54% used them and 80% plan to do so in 2012. The use of penalties over these time points has increased almost fivefold: it more than doubled from 8% in 2009 to 19% in 2011, and, within only one year, doubled again, to 38% in 2012 (National Business Group on Health and Towers Watson 2011).

Chapter 2 showed that in Germany, the Social Security Code sets out detailed provisions on a number of policies aiming to promote personal responsibility that are directly binding for the sickness funds. These include incentives for regular use of dental care; no-claims discounts for minimised health care use; doubled co-payments for insurers who suffer from colon cervical, or breast-cancer and who have not used



the respective prevention programmes; and so-called bonus programmes for the use of a wide range of other preventive and wellness programmes, where sickness funds have much freedom in designing initiatives. Regarding bonus programmes, sickness funds need to report on cost-savings at least every three years, as incentives may only be financed if savings result from programme participation: the model of financing is therefore based on gain-sharing, not cost shifting. The reporting requirement does not extend to health improvements, which do not need to be documented. In the US, incentives have been used since 2006, and 2010 health reform legislation, fundamentally driven by the intention to provide affordable coverage to all citizens, has substantially increased the permissible levels of health incentives, with the counteracting potential for medical underwriting by the back door. US law and policy moreover explicitly permit cost shifting and, unlike the German situation, do not require systematic reporting on the economic consequences (nor on health improvement).

Four main rationales with different underlying values for using individual-level incentives were identified: health promotion, cost-saving, enhancing competition, and improving workplace productivity. All played a role in both countries, but some had particular relevance in only one of them, due to the specific policy context.

Incentives are attractive to both countries because of their potential for health promotion, and the implicit foundational value here is that being healthy is a good thing, whether intrinsically, or because of its instrumental worth (for example, being able to pursue one's life plans, or being a productive member of society). If other rationales are not given, incentive programmes would face traditional health-paternalism challenges, and these challenges become stronger the weaker the justification of other rationales is.

A further prominent rationale was the shared belief by policy makers in both countries that better health will curb, if not save, health care expenditure, although it was also shown that the evidence for this assumption is not clear when one takes a full lifespan view. Fairness considerations underpinned the cost argument in both countries: in Germany the basis was centrally that resources were necessarily limited, and that solidarity meant insureds are obliged to limit demands on the system

(although, and because, they are also entitled to assistance from the community). In the US, by contrast, the tension was framed differently and surfaced more as a conflict of interests between employers as the payers on the one side, and the collective of health care users on the other.

In the US context, the rationale of improving productivity is also a far more central one, as unhealthy employees will generally be perceived by employers as costing more and contributing less, adding a rather different facet to the fairness argument, shaped far more by business interests. Such interests played at best a peripheral role in the German context, mainly because employers—with the exception of very large employers who operate their own sickness funds—are more removed from the administration and implementation of health care than their American counterparts.

Still, German incentive programmes do have a function that relates to market dynamics, as they can be used to help attract and retain ‘good risk’ insurees, i.e. people with high-income and low morbidity. Although the same mechanism has recently been documented in the context of Medicare (Cooper and Trivedi 2012) this function is generally less relevant in the US, where employees are usually restricted to a small number of health plans that their employer offers. By contrast, in the German context, where, in principle, people can choose from more than 150 sickness funds, and where the Government is persuaded that fostering competition will lead to better health care, selection effects of programmes can play a larger role. A major difference between the incentive programmes used in both countries were the levels of monetary value: while the German programmes rarely exceeded £120 (€150), US levels could amount to up to 50% of the cost of coverage, and even levels of 20% amount to four figure sums, based on average income.

### ***8.2.2 Sub research question two: Ethical framework***

Chapter 3 sought to address the second sub research question:

To what extent can the philosophical literature on personal responsibility and the justification of norms in pluralist societies help clarify the normative principles underlying the policies in the US and Germany, as identified in the analysis following Sub RQ 1?

It would be a category mistake to expect health policy to conform to coherent philosophical frameworks. Still, it can be useful to explore on what basis policy in the planning, implementation or evaluation phase might be justified, and by reference to what theory or framework one might discuss, and ideally resolve, possible tensions. Chapter three argued that established approaches are not overly well suited for either purpose. First, while a number of normative theories exist that address distributive justice problems, and might set out general arguments that defend solidarity with the imprudent (as perhaps some versions of communitarianism), or penalising risk takers (as some versions of luck-egalitarianism), none of these are fine grained enough to address the specific issues arising in health policy in a workable manner. Second, because normative positions typically make assumptions about substantive foundational values, in pluralist societies agreement on these assumptions remains elusive. Hence, even if more nuanced variations of particular normative theories were to be developed, they would fail to command universal respect. As a response to both problems I set out a procedural justice based account that developed Norman Daniels' and James Sabin's Accountability for Reasonableness approach. Drawing also on Thomas Scanlon's contractualism, I suggested that much progress can be made by analysing which parties might reasonably reject policies, and revising, if necessary, programmes so that either the policies, or else the reasons for not revising policies are accepted.

I also sought to show that the work on the concept of personal responsibility in the literature was disjointed, and that a wide variety of different meanings were found. I aimed to synthesise these into a plausible concept of *health responsibility as co-responsibility*, which concerned one's reasonable prospective and retrospective obligations as a healthy person, patient or convalescent, to lead a healthy life, to respect the health of others, and to contribute to an efficient health care system, insofar as available choices and external factors permit this (Section 3.4).

I sought to show that overly zealous proponents of personal responsibility failed to take seriously the evidence from the social determinants of literature. This body of work casts considerable doubt on all frameworks that focus in a narrow way on individual choice, by demonstrating that not all people start from a level playing field. Equally, I tried to show that those overly opposed to the concept of personal

responsibility ignored that there was practically value-free, functional version of the responsibility argument that focused on causal responsibility, which was of central importance in basic health promotion programmes and interactions with primary care and other physicians. Failing to take this notion seriously risks conveying a sense of fatalism and resignation.

By contrast, the concept set out here maximises the benefits associated with the concept of personal responsibility and minimises the harms. Moreover, it is capable of eliciting the salient features of established laws and policies relating to personal responsibility. Together with seven tests concerning key normative and structural values and components that are integral to practically all health care systems, it can help clarify the normative principles underlying the policies in Germany, the US, and elsewhere, and provide a unique framework for their justification. The tests require an analysis of the acceptability of reasons provided by those proposing or conducting programmes, with acceptability being determined by the degree to which reasonable rejection of programmes by those affected by a policy is minimised. The seven areas related to: (1) Providing the evidence, rationale(s), and feasibility of policies seeking to promote personal responsibility; (2) explaining the extent to which programmes present the least intrusive option; (3) commenting on the cost and benefits in terms of promoting equity; and (4) solidarity/risk-pooling; (5) setting out in what way it is required for behaviours or targets to be attributable to the actions of those offered incentives, and how possible constraints arising from differences regarding opportunity of choice are taken into account in specifying target behaviours or biometric targets; (6) explaining in what way possible impacts on affected third parties have been minimised; and (7) contextualising the policy in the wider framework of health and social policy by showing that it stands in a coherent relationship.

### ***8.2.3 Sub research question three: Views of physicians and the public***

Chapters 4 and 5 sought to address the third sub research question:

In the views of members of the public and physicians in the US and Germany: which aspects of cancer care and obesity policies focussing on individual behaviour and contribution to efficient services are reasonable, and which ones are not?

Physicians represent an important party in the implementation of responsibility policies and their views mattered on several grounds: they are uniquely able to comment on the health improvement potential of policies; their reasons for accepting or rejecting policies are particularly important for the legitimacy of programmes; and, on a more practical note, lacking cooperation can mean that policies exist simply on paper and will not be implemented. One of the seven proposed tests therefore also entails exploring the impact of incentive programmes on the doctor patient relationship, as it would be short-sighted to view such programmes merely as transactions between payers and users of health services.

Although the sample was purposive, the analysis of the interviews with 20 primary care physicians and oncologists in Berlin and Philadelphia is clearly not sufficient to claim any kind of representativeness of findings, nor are quantitative explorations useful. These limitations need to be borne in mind in reviewing the chapters' findings as well as this summary. Regarding the role of personal behaviour, none of the physicians viewed personal behaviour as the most important determinant of health, with the majority finding it impossible to decide on a single paramount factor, although five (American) physicians settled for the primacy of a range of other factors (environmental conditions, income, education, access to the health care system or inherited genetic factors). This finding, and the corresponding non-universal belief in the effectiveness of individual action—and hence also individual-level incentives—was echoed in further assessments of the reasonableness of incentive programmes. Concerns about their impact on affordability and access to care were particularly salient among American physicians, and a major reason for rejecting incentive programmes that had this effect. Many expressed frustration with non-compliant patients and acknowledged that personal behaviour did have a role to play: in this sense they welcomed programmes insofar as they were able to improve patient behaviour. But there was also a common theme that inflexible one-size-fits-all incentive programmes were not ideal, and would be likely to result in undesirable inequalities in effectiveness and fairness. Some therefore suggested modifications, but most appeared to accept such outcomes as an inevitable consequence of policies that are never likely to be tailored exactly to individual needs and circumstances.

A similar common theme among physicians was the belief that ‘carrot’ incentives would be more effective than ‘stick’ versions, although not a single physician pointed to evidence that would have supported this view. Moreover, penalties were not universally rejected: the case of the modified cancer policy in which a surcharge would be imposed if people failed to attend a counselling session on the advantages and disadvantages of colon cancer screenings was accepted by just over half of the interviewees, with underpinning reasons focussing on the potential to improve health literacy, health outcomes and cost reduction. By contrast, the current German cancer policy met with almost no approval. Objections centred around the lacking evidence base, the disproportionate coerciveness, and the fact that patients were penalised financially at a point when they had to cope with a cancer diagnosis. Physicians in both countries also found that the policy put them in a very uncomfortable position in cases where patients requested them to cover up non-compliance in treatment, the latter point also surfacing in relation to the obesity policy.

The ‘carrot’ version of the obesity policy was broadly welcomed as a possible tool in helping patients control their weight, even if there was some scepticism that it would be effective for all people wishing to lose weight. There were also questions about how rigidly penalties for failure to meet targets should be enforced, as this could undermine potentially fragile motivational structures. The collective estimate of interviewees was that half of the physicians who would be asked to confirm that the target weight had been reached (even if this had not been the case) would comply with the request—matching precisely the estimates by survey respondents, who estimated that every other patient would make such a request. In contrast, while the surveys suggested that just as many patients would request a cover-up in the parallel cancer scenario, here physicians estimated that only around one or two colleagues would comply, because of the more severe health and legal consequences.

The car insurance analogy and ‘false carrot’ incentives were not supported unconditionally by any physician, and reasons for rejecting these approaches centred around the potential of a ‘double whammy’ for disadvantaged populations; the unfairness of singling out one risk factor, but ignoring others (such as high risk sports); the potential that such measures may lead to barriers to care; and the risk that

social tension might result from non-overweight people blaming their insurance increases on the overweight.

It was noteworthy that the country's different policy contexts did not translate into marked differences in attitudes and reasons of the two groups of physicians, nor were there very clear differences between the professional specialisms, with the exception noted above of more emphasis on the impact of incentives on affordability of care among US physicians. An unexpected and striking similarity between the American and the Germans sample was the shared reference to solidarity. In the German case the actual term was used, while in the American context paraphrases were found, to the effect that one person's actions mattered in the sense that they could curtail (or enable) use of medical resources for others. All physicians were clear though that failing to comply with such obligations ought never to mean that people should face substantial difficulties in accessing care.

Population level views were elicited to explore how lay people related to key issues arising from the policies and the physicians' assessments, and in particular, to ascertain whether particular vulnerable groups might have reasons to reject specific aspects of policies. Overall, differences between the samples in both countries were also less marked than one might have expected. In both the US and German samples, respondents had used incentives more frequently in the past than currently, but interest in future use was three to fourfold higher. Margins between current and intended future use were larger in lower health groups than in higher ones, and in US1 the low-income group had a slightly higher margin than the high-income group, with 30% signalling interest in future use. Between 35–46% of the overweight and obese in both samples shared this view, which can suggest that less privileged groups do not feel excluded by default, and moreover, are interested in using incentives. However it is not clear whether such interest is driven by health promotion potential (as the levels of low health and high BMI groups would suggest), or the proportionally higher incentive amounts that might have more traction with lower income groups.

Respondents were not categorically opposed to harder attainment-incentives, compared to participation-incentives. The framing of 'carrot' versus 'stick' alone

was also not decisive in terms of acceptability, as in all samples no programme met with more than 42% of disapproval (i.e.: “disagree strongly” and “disagree”). Respondents were broadly undecided about whether the weight-control incentive and the German colon cancer prevention policy should be viewed as ‘carrot’, ‘stick’ or “both carrot and stick” with roughly equal thirds opting for each option, suggesting considerable awareness of the ambiguity of the ‘carrot-stick’ distinction.

Unlike the physicians, a majority of respondents in samples US1 and Germany approved of the car insurance analogy and penalties for the overweight, including a sizeable proportion of the overweight; however, there was no support for high levels of penalties that are advocated by some. Compared to the US sample, respondents in the German sample set higher levels for the ‘stick’ scenario than the ‘false carrot’ one, perhaps because direct penalties are seen as more effective, or because this approach is seen as communicating more clearly that being overweight imposes a burden on the solidaristic community: the highest ratios between the two scenarios were seen in those groups with higher insurance contributions and better health (however, the overweight also had a high ratio, possibly because a higher penalty was seen as more motivating). The support of the car insurance analogy by the overweight (in US1 41% agreed or strongly agreed, and 35% in Ger) was somewhat puzzling and suggests either that victim blaming concerns found in the literature are without a basis for this group, or that some form of mistaken victim self-blaming occurs here (even if, as noted, the levels of penalty were comparatively low, overall and especially among the overweight).

Effort and motivation mattered for respondents’ assessment of the fairness of incentives. Levels of approval were lowest in both samples where people just happened to meet targets, and their motivation and actions ran counter the programmes’ goals. By contrast, levels were highest in the case where people achieved weight loss or control, strikingly with almost identical levels in cases where people had not changed their behaviour, but met targets because of a healthy lifestyle. While respondents therefore did not seem to mind if the healthier—and typically financially better off (see Fig 6.4)—benefited more, the analysis of differential use by income groups did reveal some concerns about cream skimming.



Thirty-seven per cent in US1 and 46% in Ger found completely unfair or unfair a scenario in which twice as many of the high-income employees of a company receive a weight-control incentive of \$150 for meeting targets. In both samples, levels were higher in the lower income groups, of whom half had fairness concerns, but only a fifth of the high-income group shared this view. As the data on current usage showed, the scenario used in the survey was not merely hypothetical as it in fact reflected current US incentive usage (see Figure 6.7). Therefore, half of the low-income group, and more than a third of all respondents, are likely to find the current incentive situation in the US unfair.

Regarding colon cancer prevention, all policies were framed such that health promotion and control cost were central goals, and all had an overt ‘stick’ format. The majority of respondents in both samples did not oppose any of the three policies. Unlike the physicians, they made no difference between the compliance requirement and policies without this provision. Neither the extent of coercion, nor the explicit goal of cost containment was sufficient to prompt a majority to reject the policies.

#### **8.2.4 Main research question**

The discussion and research findings in Chapters 2–7 collectively sought to answer the main research question:

In promoting health and striving towards efficient delivery of health care, to what extent should German and American health care systems adopt policies that incentivise personal responsibility and behaviour change at the level of the individual?

At this point, the answer to the main research questions may appear obvious, but it is nonetheless non-trivial: German, American and other health care systems should adopt policies that incentivise personal responsibility and behaviour change at the level of the individual to the extent that no one—in particular vulnerable subpopulations—reasonably rejects:

- (1) the evidence that is cited suggesting that such policies are likely to work; the rationales that support them; and the feasibility of the means chosen;
- (2) the extent of intrusion;
- (3) the likely or actual propensity that policies lead to health or income inequities;
- (4) possible effects on the principles of solidarity or risk-pooling;

- (5) that penalties or rewards are attributable to the behaviour of people with due consideration of their opportunities of choice;
- (6) possible impacts on third parties; and
- (7) the coherence with the wider policy context.

If no one rejects adequately comprehensive justifications in these seven areas, policies can be seen as genuinely accepted. Clearly, such uniformity will hardly ever be achieved, and the point of the approach presented here is clearly not to offer a single person the opportunity to veto a policy that is accepted by all other affected people. Rather, the purpose of the phrase that policies are accepted to the extent that no one reasonably reject them is to draw attention to the importance of addressing all relevant reasons that people present—irrespective of how many people second the view. The fewer people present reasonable arguments to reject a policy, the more robust and accepted it will be, but that does not mean that only those policies are permissible to which—literally—no one objects.

Health is a result of both individual behaviour and a range of other factors. It is hence both legitimate, and in fact required, that health systems explore the extent to which people can be enabled to change behaviour when they experience poor health, through cost-effective individual-level incentives. This is not to ignore the evidence from the social determinants of health literature, which strongly suggests that the health care system narrowly construed has limited potential in this regard. However, the emphasis of the social determinants must not lead to dogmatism and require that population level health improvement can be achieved exclusively through traditional public health measures that typically do not involve individual level incentives, but focus on improving the enabling conditions and opportunities of choice.

The plea is, then, to acknowledge that in principle, incentives for health can be complementary to traditional public health measures (Schmidt 2009b), and that only a review of the research carried out to date, and openness with regard to the potential of future research, can settle the question of what the most effective and accepted approach for particular areas and conditions is. While the work presented here has identified a number of reasonable grounds to object to specific forms of incentive

policies, and while there are hence important risks that need to be considered, there are also risks associated with the option of not using incentives.

The evidence from the surveys conducted here suggested significant interest in using incentives in both sample US1 and Ger, with poorer health and higher BMI groups in the US and German sample, as well as lower income groups in the US being more interested in using incentives than better off groups, despite lower current use. The interviewed physicians accepted the principle behind, if rarely the form of, the incentive policies that formed part of the instrument. To the extent that uptake of effective and cost-effective policies can be achieved through programmes that meet with low levels of reasonable opposition, they should therefore be welcomed as a further tool available to improve health of people individually and collectively.

The focus in the above paragraphs was on health promotion only. However, as is clear from Chapter 2, rewarding maintenance of good health, or behaviour change, is only one aspect of incentive programmes. The concept of health responsibility as co-responsibility includes further aspects, chiefly to do with contributing to efficient use of resources. Just as health promotion is a proper and legitimate goal of health care systems, so is efficiency with regard to maximising the utility of available resources. Hence, again, it is *prima facie* reasonable for policy makers to adopt programmes that incentivise personal responsibility at the individual level to the extent that no one—in particular vulnerable subpopulations—reasonably rejects them, as outlined above, and the framework presented provides a structure for providing explicit justification in the relevant domains.

As acknowledged explicitly in Chapter 2, none of the seven tests establishes a “master” test, nor are they supposed to be seen as standing in a fixed hierarchical order. Rather, the overall justification will be holistic, by responding to the issues raised by policies in each of the domains. Much faith then, is placed in the claim that mapping clearly the ground in which justification is required; providing justifications in these areas in ways that are appropriate for the understanding of those affected by them; and responding to reasonable objections of those affected will advance policy and practice. To those sceptical of this claim I suggest that the highly problematic German colon cancer policy would have been unlikely to have come about in the

current form, had such an approach been followed. Equally, the American push towards higher penalties for the overweight has admittedly fallen on fertile ground, but the data gathered here would suggest that once it becomes more widely implemented, it is unlikely to meet with much approval, and at this stage it is entirely unclear what dynamics will ensue from this significant change in policy. Likewise, the well intentioned smoking cessation study by General Electric might have had a different fate had the company engaged employees earlier in the process and explained the rationales, including that the cost of the ‘carrot’ incentives would be offset by savings in health care expenditure, and would therefore either not affect, or lower, individual contributions of non-smokers.

In all of these cases, a fuller consideration of the relevant sets of reasons would have helped better align interests. Without a question, it would not, by necessity, have guaranteed the satisfaction of all stakeholders, but it would have helped structure both reasoning and justification in an enormously complex area. I comment more on how this might be done in practice in the section on policy implications and next steps below.

Before that I also wish to emphasise that the approach set out here is not limiting in terms of providing more comprehensive of substantive justification, and should not be seen as an attempt to present the last word on the matter. I asserted above that none of the established normative frameworks present a nuanced enough approach to frame the issues presented by health incentives in a workable manner, and while I maintain that this is an accurate assessment, it should also be noted that the seven tests set out here can be combined with substantively rich(er) normative theories, that go beyond Scanlon’s contractualism on which I drew here.

That is, egalitarians may use the full set of tests and, when it comes to assessing equity, “plug in” the normative defences established by distinct schools of thought. So, for example, on a strict equality view, situations in which some people are worse off than others are bad in themselves, which would be likely to prohibit all ‘stick’ incentives, and possibly also a wide range of ‘carrot’ incentive programmes. On a priority view, the moral value of improving the situation of a person is the greater the lower that person’s level of welfare is, which, by contrast, could help

justify a range of incentives, even if some inequalities exist. Or one could take a sufficiency view and focus on absolute levels rather than relative distribution: in this case all is well (with regard to the equity test) as long as all people are enabled to enjoy certain threshold levels. This might entail, for example, that there is no problem at all with the current German situation in which some sickness fund members lose out of incentives that have the effect of leading to a net reduction in the cost of health care, because overall insurance contribution levels might be viewed as reasonable (For the principal distinction between these three views see: Wilson 2009). Thus, it would appear that the approach presented here has the advantage that it is compatible with all substantively rich theoretical accounts, in offering a “docking station” for the application to the context of assessing the adequacy of health responsibility policies. At the same time, it does not require them in making progress in policy discourse, that may often be less sophisticated, but driven by the urgent need for decisions. I contend that the process of decision making under suboptimal circumstances is improved more by urging discourse parties to provide each other with justifications in the seven areas set out here, rather than deciding first, among other things, which normative theory truly captures what fairness amounts to; then converting those not yet believing to this view; and then applying the victorious theory to the context of health responsibilities.<sup>2</sup>

### **8.3 Policy implications**

The research presented here has numerous different implications for policy and practice at different levels of importance. Below the focus is on three general policy implications that are applicable to both countries, two with special relevance for the US only, and three for the German situation.

#### *8.3.1 Policy implications for the US and German context*

The findings here imply that it is particularly important to identify ways in which findings from behavioural economics can be brought to bear onto actual policy; consumers can be actively involved in the design of programmes and robust and suitably comprehensive evaluations of existing programmes are carried out.

---

<sup>2</sup> See also footnote 4 in Chapter 3, responding to Richard Ashcroft’s polemic against procedural justice accounts, as set out in Ashcroft 2008.

### 8.3.1.1 Behavioural economics and the Realpolitik of health policy

A key challenge in securing the health promotion potential, that, at least overtly, is at the centre of most incentive programmes, is to design programmes that effectively exploit the basic principles of behavioural economics and, at the same time, are compatible with the practical and administrative constraints of health insurance design. As noted in the previous chapter and in Chapter 2, evidence from the literature suggests that principles such as present preference bias and loss aversion can be used effectively to promote behaviour change. But as the discussion of the colon cancer policy in the previous Chapter showed, due consideration of this evidence is not universal, and I suggested ways in which the current situation could be improved considerably with both enhanced effectiveness and fairness. A different, but related problem is illustrated by the smoking cessation programme that was introduced at the beginning of this chapter. The study used incentives at three stages, as participants in the incentive arm received \$100 for completion of a smoking cessation programme, \$250 for smoking cessation after six months, and \$400 for not smoking for an additional six months, confirmed by a biochemical test. Administration of these staged payments which had much salience for participants, was straightforward when the ‘carrot’ format was chosen, but not feasible when the switch to the ‘stick’ option was made, as the penalty was rolled into payroll deductions, which was more readily feasible. Administrative requirements can therefore interfere with the effectiveness of programmes and need to be considered carefully in the design of studies, as well as in assessments of actual programmes, which may perform sub-optimally because of such constraints.

Moreover, a theme that came up in the interviews with physicians and that was implicit in the conceptual analysis of the five groups presented in Chapter 3 is that incentives for behaviour change typically need to be highly person specific. A single format may work for some subsection of a population, but not for all in the target group. Again, it may be that one simply needs to accept this fact as an inevitable consequence of policy—as some of the physicians reluctantly did—but it may also be that several different programmes could be offered that respond the specific needs of the target population. Appendix 8A sets out a randomised controlled trial for a study in which in which the personalisation of incentive programmes could be

explored, but I turn next to the more general need to better understand consumer preferences, which forms a necessary part of this proposal.

#### 8.3.1.2 Engaging consumers

Incentive programmes are typically developed by large employers, for-profit benefit-design consultants, government officials or academic researchers, with or without involvement of health plans. As noted in the summary of consumer engagement in other areas of health policy in Chapter 4, views of consumers, especially those of poorer health and wealth, have been largely absent from the process. This is also the case in an on-going US study on the impact of workplace health and wellness programmes for which the Department of Health and Human Services recently gave notice of a proposal. This research intends to gather the views of 3,000 employers, but only 48 employees would be involved in a small number of focus groups (Department of Health and Human Services 2011).

The plea for consumer engagement should not be mistaken for a crude form of direct democracy, in which enrollees of a health plan would simply vote on acceptable and unacceptable plans. One of the main problems with such an approach would be that most plausible forms of engagement, such as consultations, surveys, focus groups or citizens juries, only provide access to a limited subset of the affected population that can typically not claim to have a mandate of the wider population from they have been recruited (or self-selected), and participants can therefore neither claim genuine representativeness, nor the ability to provide proxy consent (Daniels 2009a). Rather, the plea stands in direct continuation of Scanlon's contractualism and Daniels' and Sabin's Accountability for reasonableness approach. The principal goal of engagement as envisaged here would be to elicit the broadest set of relevant reasons that matter for the design of effective and fair programmes.

In many ways, fairness and effectiveness considerations will be raised simultaneously by programme design features. For example, the tenant of behavioural economics that losses loom larger than gains could favour 'sticks' over 'carrots' when it comes to effectiveness, although we have seen that this approach can result in unfairness. Similarly, several studies regarding substance abuse suggest that attainment-incentives are significantly more effective than participation-

incentives (Higgins et al. 2010; Lussier et al. 2006; Petry et al. 2006), even though the potential of such policies becoming unduly penalising due to different socio-economic background conditions has also been highlighted. The effectiveness of actual programmes can only be assessed empirically, but analysing the acceptability of different formats through public engagement before, during and after implementation, can help tailor programmes to the needs of particular populations and limit unfairness, where those planning programmes should be unaware of significant obstacles that user groups identify in the process.

Moreover, as Richard Ashcroft concluded in a recent conceptual contribution, the extent to which programme are in fact coercive or infringing unduly on the autonomy of individuals cannot be decided from the armchair, and “the initial moral reactions widely reported in the literature and the mass media turn out on closer inspection to be neither so compelling nor so straightforward as they at first appear” (Ashcroft 2011: 199). The survey findings presented here also indicate that ‘sticks’ and the more demanding attainment-incentives are not universally viewed as negative by respondents—including groups traditionally viewed as likely to lose out on incentives—as one might expect. A clear and important policy implication is, then, that consumer preferences need to inform more thoroughly the design of incentive programmes, both to improve acceptability and effectiveness, and to better to understand the determinants of acceptability and effectiveness.

#### 8.3.1.3 The need for robust evaluations

A closely related, but also to some extent independent consequence for policy is to emphasise the need for robust and suitably comprehensive evaluations of programmes that have been put into practice. The need for evaluations follows naturally for those agreeing that consumer involvement has considerable benefits, as such assessments are required to determine whether parameters identified through engagement have indeed been salient, or whether assessments have changed for the better or worse.<sup>3</sup> However, evaluations can clearly also be carried out without prior public engagement, and illuminate several central issues, in particular in relation to

---

<sup>3</sup> Moreover, the effectiveness of engagement itself, which, it needs to be emphasised, cannot be taken for granted, also needs to be evaluated, not least because different forms may be effective to different extents (Kreis and Schmidt 2013).



the actual effects in terms of central rationales such as health promotion or cost saving.

The dramatic increase in the use of incentives, especially in the US, provides a unique set of natural experiments from which important lessons could be learned. However, there is currently no reporting requirement, and in the German situation only cost-savings need to be demonstrated. Moreover, sickness funds may choose among the programmes they offer, whichever ones seem appropriate, and do not need to provide exhaustive evidence for all programmes they offer. While it is easy to make the call for more comprehensive data, generating and analysing them faces several non-trivial methodological, administrative and political challenges.

Regarding the format of incentives, there is a very wide range of factors that can determine the design of a programme. First, the type of behaviour can differ and may relate to relatively simple one-off activities, such as receiving a flu vaccine, or completing a health risk assessment, to more complex chronic disease related behaviours, such as smoking or exercise. Second, incentives can be aimed at individuals, or groups of individuals (for example a family or work team). Third, incentives can be provided as ‘carrots’, ‘false carrots’ or ‘sticks’ for participation and attainment programmes. Fourth, incentives can be provided in cash (as fix sums, or perhaps lottery wins), in kind (such as sports or concert vouchers), or in non-materialistic ways (such as achievement or honour awards). Fifth, rewards and penalties can be provided with immediate impact and salience (such as the daily option of lottery wins for people in weight loss programmes who are on track regarding their projected progression), or over a medium term time span (such as the annual smoker surcharge), or an even longer term (such as the current German cancer policy that may imposes a penalty a decade or so after the activity that was incentivised). Sixth, in terms of levels, incentives (or the value of in-kind goods) can vary from a marginal amount to, in the US, the equivalent of half of the cost of health care coverage. Seventh, in terms of determining health improvement, programme design and data analysis can follow a range of different formats and, for example, follow gold standards of research, including randomization, control groups, and comprehensive baseline and exit analysis of key health related parameters. Alternatively, more pragmatic approaches could be adopted, such as matched pairs

analysis; one could use existing datasets such as claims-data as proxies for health-status, or simply analyse how many people accessed a programme, and how many did not, assuming that the health related activity resulted in some benefit, without exploring in more detail the precise extent to which health of those participating has in fact improved.

It is clear that these options that all determine the effectiveness of programmes provide an extraordinarily complex set of combinations that can lead to different programme designs and different options for analysis, posing challenges, especially, for comparisons between interventions and systematic or meta-reviews. Even small variations, for example in relation to the incentive amount, and whether a ‘carrot’ or ‘stick’ approach is adopted, can make comparisons between two (or more) programmes challenging.

Regarding affordability and fairness, it would be important to know whether programmes that use high levels of incentives that are implemented in the ‘stick’-mode lead to affordability problems for employees. However, it is not clear how employers would gather data on the impact of a programme on affordability. Conducting a survey among the workforce would be possible in principle, but this can probably not be expected to be chosen as a routine measure.

A number of administrative, economic and political issues can also interfere with the willingness or ability of employers to gather and share information. Information about wellness programmes by necessity involves personal health information. Sharing such data is subject to comprehensive data protection rules, and employers may be concerned about the compliance burden. While some employers are open about an aggressive approach in terms of cost-shifting, others may prefer to attract less attention, and may be reluctant to share data. Rigorous design and evaluation measures typically add cost to a programme, and are likely to have an effect on short-term return on investment. This can lead employers or health plans to be reluctant to explore different designs, as long as there is a sufficient likelihood that some positive effect on health care cost can be achieved. They may also favour cost-shifting approaches, that are more likely to guarantee them short term benefits, over gain-sharing ones, where benefits may only accrue more longer-

term. While a number of templates for wellness programmes are available at no cost, others have been developed by benefit design consultants, and proprietary issues may stand in the way of sharing information on specific design features and outcomes. As in the case of known publication bias, employers may be willing to share results from programmes that have proven to be successful, but less so, in the case where they had no, or certain negative effects, for fear of reputational damage—even though avoiding the duplication of unsuccessful incentive programmes should be in everyone's interest. Employers and, to a lesser extent, sickness funds, may also perceive a detailed reporting requirement as another instance of 'big government' interfering in health care delivery, and experience reporting requirements as stifling their innovative efforts.

This is not the place to discuss in detail ways in which these concerns can be addressed. However, it is clear that addressed they must be, and that agreeing which basic data on implemented programmes can be generated and shared in what ways, with involvement of all relevant stakeholders, is timely and urgently needed. Poorly designed programmes that focus on cost-shifting only may taint the approach as a whole, undermine its health promotion potential, and, understandably, alienate employees. Equally, half-hearted or patchy evaluations may satisfy employers or sickness funds who merely seek short term gain, and may be inevitable if, in view of the complexity of the task, government agencies gather data in an insufficiently structured way. But they do little to further important insights into the assumed and actual health promotion and behaviour change potential of programmes. By contrast, learning from best practice, and avoiding unsuccessful strategies, can help harness the potential of this approach and help maximise in a sustainable way the benefits from wellness incentives, which is particularly relevant with regard to the recent significant increase in permissible levels of incentives in the US.

### *8.3.2 Policy implications for the US context*

A central element with relevance to the fairness of US wellness incentives is the legal provision regarding the so-called alternative standard. As noted, in the case of attainment-incentives this requires employers to offer an alternative way of securing the incentive value for those employees for whom it is unreasonably difficult due to a medical condition to meet a target, or medically inadvisable. Information about what

alternative standards were provided, and what number of people used them could, among other things, help illuminate which programmes are challenging for which people, and could inform the design of tailored incentive programmes for these groups of people.

However, on the grounds cited above, employers may be reluctant to share such information, and there is no requirement to provide such data, nor to be proactive in providing alternative standards. Employers have complied with the law as long as they advertise the availability of alternative standards. This situation is likely to be insufficient to understand whether eligible people use it, or whether they are content with the arrangements (and do not, for example, find such petitioning humiliating). There is also anecdotal evidence that employees who provided a certification from their physician testifying that that they qualified for an alternative standard were made to see a company employed physician who overruled such attestations,<sup>4</sup> and the adequacy of the provision in promoting fairness can only be assessed if data on these and further aspects of implementing the alternative standard provision are generated and analysed by parties that do not have any conflict of interest.

Even more difficulties can arise regarding the assessment of the impact of incentive programmes on the structure of a company's risk pool. For it is not inconceivable that actively advertised wellness incentive programmes can affect the decisions of persons looking for employment, with healthier people more likely to chose companies that offer them cheaper insurance, and unhealthy people more likely to shun those employers who use penalties, in the case where a choice between similar jobs at two different employers exists. Such evaluations would need to compare whether the health status of employees has changed over time, and whether changes can be attributed to knowledge about incentive programmes, which would clearly require cooperation of employers. It is also not clear what should be done if it should turn out that there are indirect selection effects, as one could not force employers not to offer them, if they complied with the applicable law and policy. However, it is also clear that having these data would be of relevance for a proper

---

<sup>4</sup> Personal communication, Sue Nelson, Vice President, Federal Advocacy, American Heart Association, 16 November 2011.

understanding of the effects of wellness incentive programmes in health and social policy more broadly.

### *8.3.1 Policy implications for the German context*

Further to the negative response from professional organisations regarding the current German colon cancer policy that was noted in Chapter 2, the interviews with physicians carried out here corroborated that there are good reasons to reject, in particular, the compliance requirement of this policy in its current form. The survey findings suggested that tensions in the doctor-patient relationship were very likely, while also indicating that a majority of the public, unlike the physicians, was not opposed to the current policy. The latter finding can explain why there was no broader public backlash when the policy was introduced, and also illustrates again the purpose and function of the framework advocated here, which may well be used to argue against a majority view, and is merely intended to provide a structure for areas in which justification is owed. In this case, the structure strongly suggests that there are very good reasons to reject the policy in its current shape, and that reforms are needed.

The focus of the discussion here was on obesity and colon cancer, as these were particularly good candidates for a comparative approach, and a range of other policies were addressed at best implicitly. Of note in this regard is the no-claims bonus programme that was introduced at the same time as the cancer prevention policy, and puts at stake considerably higher levels of incentives for those who do not see a physician for a prescription and do not require hospitalisation over a year. A subsequent law that entered into force in January 2011 (the GKV-Finanzierungsgesetz (GKV-FinG)) stipulated that programmes of this kind needed to demonstrate *ex ante* evidence that the programme could be implemented in a cost-saving way (as for these programmes too, financing through insurance increases or cost-shifting to less healthy enrolees is not permissible). GKV-FinG also reduced the minimum enrolment time from three to one year (Meißner 2011). These provisions go some way towards addressing possible equity concerns, but still fall short of investigating other salient aspects, chiefly, whether such programmes lead some enrolees to take “health gambles” by delaying needed care, with potential welfare and economic costs.

A further policy that was not discussed after its introduction in Chapter 2 related to the cost-sharing policy that required insurees to refund the sickness funds all or part of the cost of treatment or sick pay in the case of services that were provided for complications arising from non-medically indicated interventions, such as tattoos piercings or non-medically indicated plastic surgery. While there are good reasons to assume that the policy will simply not be implemented, a better and cleaner way forward would be to repeal the provisions. This is not the place to provide a comprehensive justification, but salient aspects would relate to the attributability test, on which it is reasonable to reject the rule because people falling under it did not have the opportunity to act otherwise. For, unlike in the case of the cancer policy, where the explicit age cut-off means that it only applies to people who are eligible for cancer check-ups after the rule had been introduced, the policy regarding non-medically indicated interventions applies retrospectively to all insurees. An examination of the rationale would also throw up questions about the actual cost-impact for the conditions that are listed in the law, as well as about the general reasons for the inclusion of these conditions (and not, for example, alcoholism or overweight). The third party impact test would suggest, as for the cancer policy, that the rule puts unnecessary stress on the doctor patient relationship. It would therefore be better to abandon the rule, and, if policy makers still insist that it has some justification, engage in a broader debate in society that can help elicit relevant reasons that are required in a proper justification of a policy along these lines.

#### **8.4 Limitations and next steps**

The research conducted has several important limitations. Further to those already acknowledged in previous chapters, this section also suggests how constraints in relation to key limitations of the overall framework, the interviews and the surveys can be addressed in future research.

##### *8.4.1 Limitations in relation to the overall framework*

The framework provided here provided no off the shelf answers as to which policies that incentivise personal responsibility and behaviour change at the level of the individual can be justified. The illustration of its application to the case of colon

cancer prevention policy in the previous chapter also showed that the approximation to the method of triangulation was able to broaden the set of relevant reasons that need to be considered in comprehensive justifications, but it did not pin point to a single process or set of values held by either physicians or the public that would serve as ultimate standards for arbitration. A key challenge is, then, to demonstrate the utility of the overall framework.

One way of doing this is by exposing it to the criticism of the academic community at conferences and applying it in appropriate publications to specific cases, as I have attempted in relation to incentive policies in general (Schmidt 2008a), surcharges for obesity (Schmidt 2009c) and colon cancer (Schmidt [2012]) and several publications setting out different aspects of the conceptual framework (Schmidt 2009a, 2011a, 2010). Feedback in debate, from reviewers and editors can help ascertain whether important questions remain uncovered by the approach, and to what extent it helps in providing a framework for coherentist justifications. A further step would be reconstruct or test the framework in empirical research with the public: would other tests be added? Would the proposed ones be regarded as appropriate?

A different type of limitation can result from the form of the framework, and it might be advantageous to develop a more policy and practitioner friendly version. This could, for example, take the form of a checklist, following the approach popularised by Atul Gawande, which led to significant improvements in surgical outcomes through the more systematic use of brief manuals to review the preparations for key stages (Gawande 2009). The seven tests could be presented in a more succinct version in the form of brief questions that policy makers could consider in contemplating new or existing programmes. Such a list may also be of use to journalists, whose work can be influential, but, understandably, is not always reflective of the range of different issues raised by incentive programmes. A further version could also be developed for health plan enrollees, to sharpen awareness of the central features of programmes that are offered to them, and to ask the right kind of

questions of employers regarding, for example, the provision of alternative standards.<sup>5</sup>

#### *8.4.2 Limitations in relation to the interviews with physicians*

The interviews with physicians were coded and analysed by me only, and repeating the process by a separate researcher would no doubt enhance the robustness, even though the questions were rather specific and answers largely unambiguous. A more significant limitation resulted from the constraints arising from the small sample of 20 physicians in terms of quantitative analyses. For example, while it was striking that not a single physician thought that personal behaviour was the most important determinant of health, it can, of course, not be assumed that a survey with a large enough number to allow for a representative sample would reveal the same scepticism (and the same goes, clearly, for all other questions). The only way to approach this question is, then, to carry out quantitative research on physician views too.

Such work could also allow one to explore in more detail themes that emerged from the interviews and that are of particular relevance to the overall justification of programmes. One such theme was the thought that there are no prêt-à-porter incentives, largely because motivation and the scope and limitations of behaviour change are highly person-specific. Different scenarios could explore the extent to which modifications to standard programmes are viewed as likely to be effective and acceptable. A separate strand could examine how widespread, in fact, the preference for ‘carrots’ over ‘sticks’ is, and, related, explore physicians’ views regarding the acceptability and likely effectiveness of participation as opposed to attainment-incentives (the latter was not explored to the same extent in the physician interviews and the surveys). Especially regarding the question of participation and attainment-incentives, physicians may be privy to important insights as to which approach is likely to be more effective for particular conditions or types of patients, and current knowledge is patchy at best.

---

<sup>5</sup> I initially presented this idea in a presentation at the sixth annual international bioethics conference of the Harvard University Program in ethics and health, *New Strategies for Health promotion: steering clear of ethical pitfalls*, 28–28 April 2011, Boston and am currently following up this idea.



#### *8.4.3 Limitations in relation to the population-level surveys*

As noted, only one of the surveys carried out here (US2) could claim to be representative of the population, and the instrument used in surveys US1 and Ger would need to be fielded to a probability-based sample in order to allow for similar extrapolations. In addition, as noted above in Section 8.3.1.3, because the design of incentive programmes can depend on a great multitude of factors, the findings from US2 are strictly speaking also only valid for exactly those scenarios that were presented there.

Moreover, several of the scenarios had more than one driving criterion that respondents were likely to pick up on: all three colon cancer policies emphasised that the purpose was both health promotion and cost saving, and the data alone do not allow one to distinguish whether one of these rationales was more salient than the other in prompting approval or disapproval (in addition, the first cancer scenario included the compliance requirement, that could plausibly be objected to). In a similar vein, the three scenarios that asked respondents to identify how high levels for ‘carrot’, ‘false carrot’ and ‘stick’ incentives should be, did not provide respondents with a rationale such as “in order to be effective” or “in order to be fair”, but simply asked: “By how much should the plan be allowed to increase/reduce insurance premiums ... to create weight-control incentives?”, without further qualifications. This naturalistic framing was deliberate and adequate for the purposes pursued here, but a further interesting strand of research for follow up surveys would therefore concern a wider set of scenarios that included other incentive formats, and also varied identical scenarios, better to understand the salience of different aspects that are likely to play a role in assessments.

In particular, it would be interesting to explore whether programmes that followed more closely principles of behavioural economics are viewed more or less sympathetically than those that are more modelled on standard policy proposals (such as the ones that were used in the survey). Such research would not be trivial, however. The example of the smoking cessation policy introduced in this chapter already illustrated that it requires a considerable amount of stage-setting and explanation which is not necessarily ideally suited for the survey format. An alternative approach might therefore be to break down the constitutive elements of

each incentive approach and explore, for example, whether respondents would find an annual \$150 incentive for two passed weigh-ins more effective, or a twelfth of the sum provided monthly for those who are of normal weight (or achieve a certain reduction in weight).

While such and more sophisticated work can provide useful insights into the design of programmes, it also needs to be acknowledged that, of course, data on intentions and attitudes are less reliable than actual use, as already observed on the occasion of the discrepancy between past, present and future use of incentive programmes. An ideal format would therefore be to field an attitudinal survey first, to develop, on the basis of this, incentive programmes that are aligned as closely as is appropriate with the views that have been expressed, and then to repeat surveys at later stages, combining data on actual use of programmes with attitudinal research. This would, first, get around the problem of analysing mere intentions, and second, could also help understand whether attitudes change over time, as already noted above.<sup>6</sup>

A better understanding of the determinants of successful and unsuccessful use of incentive programmes might itself be sufficient to tailor programmes to the needs of people wishing to change behaviour. It could also lead to the identification of some general principles that underpin the effectiveness which may not always be apparent to the users themselves. One of the basic assumptions behind behavioural economics is that people are typically far less rational than they think, and not always their own best judges. Possibly, this also applies when it comes to deciding between different incentive structures. Gaining a better understanding of the determinants of the effectiveness of incentives for particular types of people could lead to the development of a screening tool, perhaps in the form of a short questionnaire, that can help identify which incentive structure is most likely to work for a given individual. Appendix 8A outlines a four-armed randomised controlled trial that could then help shed light on the differential effect of such an intervention.

---

<sup>6</sup> Together with colleagues I have recently submitted a grant application for funding to the US' Patient Centered Outcomes Research Institute for a two year study that in which such a project would be carried out in collaboration with three large employers, and also secured support for this work from the National Business Group on Health, and the American Heart Association.

## 8.5 Closing

The need to respond to a significant rise in chronic diseases and to improve the efficiency of health care is not only found in the US and Germany, on which this study concentrated, but in a considerable number of other OECD countries (OECD 2011) and increasingly, the emerging economies of Brazil, Russia, India, and China, in which rising income levels among large parts of the population and changes in cultural practice are associated with an increase in “life-style” associated conditions (Miller 2011).

In many of these societies the question of whether higher levels of morbidity are to blame primarily on the individual, or the context in which individuals live is likely to be discussed in as polarised a fashion as is often the case in the US and Germany. A central point of this thesis was to argue that such a tug-of-war is fruitless, and that it is more productive to acknowledge the complex interplay of personal behaviour and environmental factors when it comes to both the co-production (Forde and Raine 2008) of, and normative co-responsibility for, health.

The framework set out here is intended to assist policy makers, practitioners, ordinary citizens as well as expert philosophers in scrutinising and evaluating policies that seek to promote personal responsibility for health. The empirical work has shown that several central ethical questions cannot be resolved from the armchair, but require quantitative and qualitative research. While considerable work lies ahead, together, these different elements have hopefully begun to bring more clarity to central aspects of the debate around personal responsibility.

## References

- Abelson, J., M. Giacomini, P. Lehoux, and F. P. Gauvin. 2007. Bringing 'the public' into health technology assessment and coverage policy decisions: From principles to practice. *Health Policy* 82 (1):37-50.
- American Heart Association 2009. *Close the Loophole to Medical Underwriting in the Senate Health Care Reform Bill (Letter to Members of Congress), 21 Dec.* Vol. December 21. Washington: American Heart Association.
- Allensbach (Institut für Demoskopie). 2011. Die Sonntagsfrage. 27 Oktober 2011, available at: <http://www.wahlrecht.de/umfragen/allensbach.htm> accessed 27 October 2011.
- American Cancer Society. 2011. *Colorectal Cancer Facts and Figures, 2011-2013.* Atlanta, Georgia: American Cancer Society.
- Anderson, E. 1999. What is the Point of Equality? *Ethics* 109:287-337.
- Arneson, R. J. 1997. Equality and Equal Opportunity for Welfare. In *Equality: Selected Readings*. Now York, Oxford: Oxford University Press.
- . 1989. Equality and equal opportunity for welfare. *Philosophical Studies* 56:159–94.
- . 2000. Luck-egalitarianism and Prioritarianism. *Ethics* 110:339-49.
- Ashcroft, R. 2006. Individual freedom versus collective responsibility: an ethicist's perspective. *Emerging Themes in Epidemiology* 3 (1):11.
- . 2008. Fair Process and the Redundancy of Bioethics: A Polemic. *Public Health Ethics* 1 (1):3-9.
- . 2011a. Personal financial incentives in health promotion: where do they fit in an ethic of autonomy? *Health Expectations* 14:191-200.
- Ashcroft, R., A. Campbell, and S. Jones. 2000. Solidarity, Society and the Welfare State in the United Kingdom. *Health Care Analysis* 8 (4):377-394.
- Ashford, E., and T. Mulgan. 2009. Contractualism. In *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, edited by E. N. Zalta.
- Atreja, A., N. Bellam, and S. R. Levy. 2005. Strategies to Enhance Patient Adherence: Making it Simple. *MedGenMed* 7 (1):4-8.
- Baicker, K., D. Cutler, and Z. Song. 2010. Workplace wellness programs can generate savings. *Health Aff (Millwood)* 29 (2):304-11.
- Bauer A. 2005. Die Medizin ist eine soziale Wissenschaft - Rudolf Virchow (1821-1902) als Pathologe, Politiker, Publizist. *GMS Medizin Bibliothek Information* 5:16–20.
- Bishop, G., Brodkey, A. C. 2006. Personal responsibility and physician responsibility--West Virginia's Medicaid plan. *N Engl J Med* 355 (8):756-8.
- Black, D. 1980. *Inequalities in Health*. London: Department of Health and Social Security.
- Bödeker, W., H. Friedel, and M. Friedrichs. 2008. Ökonomischer Nutzen der BKK Bonusprogramme. *Die BKK* 4:214-218.

- Boivin, A., K. Currie, B. Fervers, J. Gracia, M. James, C. Marshall, C. Sakala, S. Sanger, J. Strid, V. Thomas, T. van der Weijden, R. Grol, and J. Burgers. 2010. Patient and public involvement in clinical guidelines: international experiences and future perspectives. *Qual Saf Health Care*.
- Bowling, A. 1996. Health care rationing: the public's debate. *BMJ* 312 (7032):670-4.
- Brace, I. 2004. *Questionnaire Design: How to Plan, Structure and Write Survey Material for Effective Market Research*, Market Research in Practice Series. London: Kogan.
- Branigan, T. 2010. Beijing workers shape up for return of compulsory exercises. *The Guardian* Tuesday 10 August.
- Braun, B., H. Reiners, M. Rosenwirth, and S. Schlette. 2006. *Anreize zur Verhaltenssteuerung im Gesundheitswesen*: Bertelsmann Stiftung.
- Braye, S., and M. Preston-Shoot. 2005. Emerging from out of the shadows? Service user and carer involvement in systematic reviews. *Evidence & Policy* 1 (2):173-193.
- Brenner, H., L. Altenhofen, and M. Hoffmeister. 2010. Zwischenbilanz der Früherkennungskoloskopie nach acht Jahren – Eine Hochrechnung und erste Ergebnisse. *Dtsch Arztebl Int* 43 (43):753-9.
- Bridgelal R., M., P. R. Grocott, and H. C. Weir. 2008. Issues and challenges of involving users in medical device development. *Health Expect* 11 (1):63-71.
- Bringedal, Berit, and Eli Feiring. 2011. On the relevance of personal responsibility in priority setting: a cross-sectional survey among Norwegian medical doctors. *Journal of Medical Ethics* 37 (6):357-361.
- British Medical Association. 2008. No question of not treating the elderly, says BMA [press release, Monday 28 Jan 2008].
- Brown, Alexander. 2005. If We Value Individual Responsibility, Which Policies Should We Favour? *Journal of Applied Philosophy* 22 (1):23-44.
- Brown, K., J. Dyas, P. Chahal, Y. Khalil, P. Riaz, and J. Cummings-Jones. 2006. Discovering the research priorities of people with diabetes in a multicultural community: a focus group study. *British Journal of General Practice* 56 (524):206-213.
- Bundesärztekammer. 2007. Beschlussprotokoll des 110. Dt. Ärztetages vom 15.–18. Mai 2007 in Munster. Berlin: Bundesärztekammer.
- Bundesministerium für Arbeit und Soziales. 2008. *Armutbericht - Dritter Armuts- und Reichtumsbericht der Bundesregierung*. Berlin: Bundesanzeiger Verl.-Ges.
- Bundesministerium für Gesundheit. 2007. *BMG-Pressemitteilung Nr. 113 vom 17.12.2007 - Medizinische Vorsorge wird gestärkt*.
- . 2011. *Nationaler Krebsplan, Handlungsfeld 1: Ziele-Papier 2 b Weiterentwicklung der Darmkrebsfrüherkennung*. Bonn.
- Bundesregierung. 2006. *Entwurf eines Gesetzes zur Stärkung des Wettbewerbs in der Gesetzlichen Krankenversicherung (GKV-Wettbewerbsstärkungsgesetz – GKV-WSG)*, 25. Oktober 2006.

- Burd, S. 2009. How Safeway Is Cutting Health-Care Costs - Market-based solutions can reduce the national health-care bill by 40%. *Wall Street Journal* June 12.
- Busse, R., C. Krauth, H. P. Wagner, F. W. Schwartz, C. Claes, and J-M von der Schullenburg. 1999. Lebensalter, Sterben und Kosten - Was wissen wir über den Zusammenhang? . In *Public-Health-Forschung in Deutschland*. Bern: Huber, S. 393-396, edited by Dt. Gesell. f. Publ. Health. Bern: Huber.
- Busse, R. 1999. Priority-setting and rationing in German health care. *Health Policy* 50 (1-2):71-90.
- . 2002. *Health care systems: Germany and Britain compared*. London: Anglo-German Foundation for the Study of Industrial Society.
- Busse, R., Riesberg, A. 2005. *Gesundheitssysteme im Wandel: Deutschland*. Kopenhagen: WHO Regionalbüro für Europa im Auftrag des Europäischen Observatoriums für Gesundheitssysteme und Gesundheitspolitik.
- Buyx, A. M. 2008. Personal responsibility for health as a rationing criterion: why we don't like it and why maybe we should. *J Med Ethics* 34 (12):871-4.
- Cahill, K., and R. Perera. 2011. Competitions and incentives for smoking cessation. *Cochrane Database Syst Rev* (4):CD004307.
- Callahan, D. 1981. Minimalist Ethics. *The Hastings Center Report* 11 (5):19-25
- . 1998. *False hopes: Overcoming the obstacles to a sustainable, affordable medicine*. Piscataway, NJ: Rutgers University Press.
- . 2008. Europe and the United States: Contrast and Convergence. *Journal of Medicine and Philosophy* 33 (3):280-293.
- Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin. 2003. Regulation for Conservatives: Behavioral Economics and the Case for 'Assymetric Paternalism. *University of Pennsylvania Law Review* 1 (151):1211-1254.
- Cappelen, A. W., and O. F. Norheim. 2005. Responsibility in health care: a liberal egalitarian approach. *J Med Ethics* 31 (8):476-480.
- Caron-Flinterman, J. F., J. E. Broerse, J. Teerling, and J. F. Bunders. 2005. Patients' priorities concerning health research: the case of asthma and COPD research in the Netherlands. *Health Expect* 8 (3):253-63.
- Centers for Disease Control and Prevention. 2011. Behavioral Risk Factor Surveillance System, available at: <http://www.cdc.gov/brfss>.
- Chamaz, K. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Cohen, J., T., Neumann, P., J., Weinstein, M., C. . 2008. Does preventive care save money? Health economics and the presidential candidates. *N Engl J Med* 358:661-3.
- Commission on Social Determinants of Health. 2008. *Closing the Gap in a Generation: Health Equity through Action on the Social Determinants of Health*. Geneva: World Health Organization.

- Cooper, Alicia L., and Amal N. Trivedi. 2012. Fitness Memberships and Favorable Selection in Medicare Advantage Plans. *New England Journal of Medicine* 366 (2):150-157.
- Crawford, N. D., C. P. Jones, and L. C. Richardson. 2010. Understanding racial and ethnic disparities in colorectal cancer screening: behavioral risk factor surveillance system, 2002 and 2004. *Ethnicity and Disease* 20:359-65.
- Dalal, M., R. Skeete, H. L. Yeo, G. I. Lucas, and M. S. Rosenthal. 2009. A Physician Team's Experiences in Community-Based Participatory Research: Insights into Effective Group Collaborations. *American Journal of Preventive Medicine* 37 (6, Supplement 1):S288-S291.
- Daniels, N., Sabin, J. 1999. Decisions about access to health care and accountability for reasonableness. *Journal of Urban Health* 76 (2):176–191.
- Daniels, N. 2000. Accountability for reasonableness. *BMJ* 321 (7272):1300-1301.
- . 2009a. Accountability for Reasonableness and the Citizens Council. In *Patients, the Public and Priorities in Healthcare*, edited by P. Littlejohns and M. Rawlins. Oxford: Radcliffe Publishing Ltd.
- . 2009b. Accountability for reasonableness and the Citizens Council. In *Patients, the public and priorities in healthcare*, edited by P. Littlejohns and M. Rawlins. Abingdon: Radcliffe Publishing.
- . 2007. *Just Health: Meeting Health Needs Fairly*. Cambridge: CUP.
- . 2011. Reflective Equilibrium. In *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), edited by E. N. Zalta.
- Dawes, J. 2008. Do Data Characteristics Change According to the Number of Scale Points Used - an Experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research* 50 (1):61-77.
- DeFulio, A. 2012. The use of incentives to reinforce medication adherence. *Preventive Medicine* [epub ahead of print].
- Denzin, N. K. 1978. *Sociological Methods*. New York: McGraw-Hill.
- Department of Health and Human Services. 2011. Agency Information Collection Request. 60-Day Public Comment Request. *Federal Register* 76 (59):17130.
- Departments of Health and Human Services, Department of Labor, Department of the Treasury. 2001. Notice of Proposed Rulemaking for Bona Fide Wellness Programs, Part II, 66 FR 1421.
- Departments of Labor, Treasury, and Health and Human Services 2006. Nondiscrimination and wellness programs in health coverage in the group market; final rules, *Federal Register* 2006;71:75014-55, December 13.
- Deutscher Bundestag 2007. Bundestagsplenarprotokoll, Stenografischer Bericht, 80. Sitzung, 16/196. Berlin, Freitag, den 2. Februar 2007.
- Devers, K. J., and R. M. Frankel. 2000. Study design in qualitative research--2: Sampling and data collection strategies. *Educ Health (Abingdon)* 13 (2):263-71.

- Diederich, A., and M. Schreier. 2010. Einstellungen zu Priorisierungen in der medizinischen Versorgung: Ergebnisse einer repräsentativen Bevölkerungsbefragung. Bremen: Jacobs University.
- Dixon, A., Pfaff, M., and J. Hermesse. 2004. Solidarity and competition in social health insurance countries. In *Social health insurance systems in western Europe*, edited by R. Saltman, Busse, Reinhard, Figueras, Josep. Milton Keynes: Open University Press.
- Dolan, P., R. Cookson, and B. Ferguson. 1999. Effect of discussion and deliberation on the public's views of priority setting in health care: focus group study. *BMJ* 318 (7188):916-9.
- Doyle, Y. G., A. Furey, and J. Flowers. 2006. Sick individuals and sick populations: 20 years later. *J Epidemiol Community Health* 60 (5):396-8.
- Dreier, M., and H. Linger. 2006. Bonus für gesundheitsbewusstes Verhalten—erste Ergebnisse einer Gesetzlichen Krankenkasse aus 2004. *Gesundheitswesen* 7:68.
- Dworkin, G. 1981. Voluntary health risks and public policy. *Hastings Cent Rep* 11 (5):26-31.
- . 2010. Paternalism. In *Stanford Encyclopedia of Philosophy*, , edited by E. N. Zalta.
- Dworkin, R. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: Harvard University Press.
- Emanuel, E. 2012. What We Give Up for Health Care. *New York Times*, 21 January.
- Engelhardt, H. T., Jr. 1981. Human Well-Being and Medicine: Some Basic Value Judgements in the Biomedical Sciences. In *Science, Ethics and Medicine*, edited by J. Z. TA Mappes. New York:: McGraw-Hill Book Company.
- English, D. C. 2005. Moral Obligations of Patients: A Clinical View. *Journal of Medicine and Philosophy* 30 (2):139 - 152.
- Ensign, J., and T. Carper. 2008. Ensign-Carper Amendment #C8 to Americas Healthy Future Act of 2009, available at: <http://finance.senate.gov/sitepages/leg/LEG%202009/102009%20America%207s%20Healthy%20Future%20Act%20Modified%20Amendments.pdf> , accessed on 15 Dec 2009.
- Facey, K., A. Boivin, J. Gracia, H. P. Hansen, A. Lo Scalzo, J. Mossman, and A. Single. 2010. Patients' perspectives in health technology assessment: a route to robust evidence and fair deliberation. *Int J Technol Assess Health Care* 26 (3):334-40.
- Feinberg, Joel. 1986. *The Moral Limits of the Criminal Law. Vol. 3, Harm to Self.* . New York: Oxford University Press.
- Feiring, E. 2008. Lifestyle, responsibility and justice. *J Med Ethics* 34 (1):33-36.
- Forde, I., and R. Raine. 2008. Placing the individual within a social determinants approach to health inequity. *Lancet* 372 (9650):1694-6.



- Frederick, S., G. Loewenstein, and T. O'Donoghue. 2002. Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* 40 (2):351–401.
- Friedrichs, M., H. Friedel, and W. Bödeker. 2009. Teilnehmerstruktur und ökonomischer Nutzen präventiver Bonusprogramme in der betrieblichen Krankenversicherung. *Gesundheitswesen* 71 (10):623-627.
- Funke, N. 1996. Abortion Counselling and the 1995 German Abortion Law. *Connecticut Journal of International Law* 12 (1):33-65.
- Gauthier, C. C. 2005. The Virtue of Moral Responsibility and the Obligations of Patients. *Journal of Medicine and Philosophy* 30 (2):153 - 166.
- Gauvin F. P., J. Abelson, M. Giacomini, J. Eyles and J. N. Lavis. 2010. "It all depends": conceptualizing public involvement in the context of health technology assessment agencies. *Soc Sci Med* 70 (10):1518-26.
- Gawande, A. 2009. *The Checklist Manifesto: How to Get Things Right*. New York: Metropolitan Books, Holt and Company.
- Gericke, C., M. Wismar, and R. Busse. 2004. Cost-sharing in the German health care system. *Diskussionspapier 2004/4*, Technische Universität Berlin, Fakultät Wirtschaft und Management, Berlin.
- Gemeinsamer Bundesausschuss. 2007. *Beschluss des Gemeinsamen Bundesausschuss über eine Änderung des Chroniker-Richtlinie: Feststellung therapiegerechten Verhaltens*. Siegburg: G-BA.
- Gillon, R. 1985. Autonomy and the principle of respect for autonomy. *British Medical Journal (Clinical research ed.)* 290 (6484):1806-1808.
- Giuffrida, A., and D. J. Torgerson. 1997. Should we pay the patient? Review of financial incentives to enhance patient compliance. *BMJ* 315 (7110):703-7.
- Gollust, S., and J. Lynch. 2011. Who Deserves Health Care? The Effects of Causal Attributions and Group Cues on Public Attitudes about Responsibility for Health Care Costs. *J Health Polit Policy Law*.
- Gooberman-Hill, R., J. Horwood, and M. Calnan. 2008. Citizens' juries in planning research priorities: process, engagement and outcome. *Health Expect* 11 (3):272-81.
- Gray, R. E., M. Fitch, C. Davis, and C. Phillips. 2000. Challenges of participatory research: reflections on a study with breast cancer self-help groups. *Health Expect* 3 (4):243-252.
- Groves, R. M. 2006. Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly* 70 (5):646-675.
- Halbert, C. H., F. K. Barg, C. E. Guerra, J. A. Shea, K. Armstrong, M. Ferguson, B. Weathers, J. Coyne, and A. B. Troxel. 2011. Cultural, economic, and psychological predictors of colonoscopy in a national sample. *J Gen Intern Med*. 2011 Nov;26(11):1311-6. Epub 2011 Jul 6.
- Halpern, D., C. Bates, and G. Beales. 2003. *Personal Responsibility and Behaviour Change*. Prime Minister's Strategy London: Prime Minister's Strategy Unit, Cabinet Office.

- Handelsblatt. 2010. CDU nimmt bei Suche nach Geldquelle Dicke ins Visier. Handelsblatt Online, 22 July.
- Heath, I. 2008. The emperor's new constitution. *British Medical Journal* 337:787.
- Helfand, M. 2010. AHRQ Series Editorial: Public involvement improves methods development in comparative effectiveness reviews. *Journal of Clinical Epidemiology* 63 (5):471-473.
- Higgins, S. T., I. M. Bernstein, Y. Washio, S. H. Heil, G. J. Badger, J. M. Skelly, T. M. Higgins, and L. J. Solomon. 2010. Effects of smoking cessation with voucher-based contingency management on birth outcomes. *Addiction* 105 (11):2023-30.
- Higgins, S. T., Y. Washio, S. H. Heil, L. J. Solomon, D. E. Gaalema, T. M. Higgins, and I. M. Bernstein. 2012. Financial Incentives for Smoking Cessation Among Pregnant and Newly Postpartum Women. *Preventive Medicine* [epub ahead of print].
- Hilzenrath, D. 2010. Misleading claims about Safeway wellness incentives shape health-care bill, Sunday. *Washington Post*, January 17.
- Houtepen, R., and R. ter Meulen. 2000. New Types of Solidarity in the European Welfare State. *Health Care Analysis* 8:329-340.
- Iarossi, G. 2006. *The Power of Survey Design: A User's Guide for Managing Surveys, Interpreting Results, and Influencing Respondents*. Washington, DC: The World Bank.
- Institute of Medicine. 2003. *The Future of the Public's Health in the 21st century*. Washington, D.C.: National Academies Press.
- . 2012. *Essential Health Benefits: Balancing Coverage and Cost*. Washington, DC: The National Academies Press.
- Jallinoja, P., P. Absetz, R. Kuronen, A. Nissinen, M. Talja, A. Uutela, and K. Patja. 2007. The dilemma of patient responsibility for lifestyle change: perceptions among primary care physicians and nurses. *Scand J Prim Health Care* 25 (4):244-9.
- Jarrett L. and the Patient Involvement Unit. 2004. *Patient Involvement Unit: A report on a study to evaluate patient/carer membership of the first NICE Guideline Development Groups*.
- Jeffery, R. 2012. Financial Incentives and Weight Control. *Preventive Medicine* [epub ahead of print].
- Jetten, J., and T. Postmes. 2006. "I did it my way" Collective Expressions of Individualism. In *Individuality and the Group*, edited by J. Jetten, Postmes, Tom., London: Sage.
- Jochelson, K. 2007. *Paying the patient: improving health using financial incentives*. London: King's Fund.
- Kahneman, D., and A. Tversky. 1979. Prospect theory: an analysis of decision making under risk. *Econometrica* 47:263-91.
- Kaiser Family Foundation, Health Research & Educational Trust 2010. *Employer Health Benefits 2010 Annual Survey*. Oakland: Kaiser Family Foundation.

- Keeter, S., C. Miller, A. Kohut, R. Groves, and S. Presser. 2000. Consequences of Reducing Nonresponse in a National Telephone Survey. *Public Opinion Quarterly* 64 (2):125-148.
- Kelley, M. 2005. Limits on Patient Responsibility. *Journal of Medicine and Philosophy* 30 (2):189 - 206.
- King, D., and A. Maynard. 1999. Public opinion and rationing in the United Kingdom. *Health Policy* 50 (1-2):39-53.
- Kitto, S. C., J. Chesters, and C. Grbich. 2008. Quality in qualitative research. *Med J Aust* 188 (4):243-6.
- Knaack, N. 2008. Chancen und Grenzen der Bonifizierung von Gesundheitsverhalten in der Gesetzlichen Krankenversicherung, Fakultät Organisationspsychologie, University of Dortmund, Dortmund.
- Knowledge Networks. 2011. Response to ESOMAR 26 Questions to Help Research Buyers of Online Samples. San Francisco: Knowledge Networks. <http://www.knowledgenetworks.com/insights/docs/Knowledge-Networks-Response-to-ESOMAR-26-Questions.pdf>.
- Koops, L., and R. I. Lindley. 2002. Thrombolysis for acute ischaemic stroke: consumer involvement in design of new randomised controlled trial. *BMJ* 325 (7361):415.
- Kreis, J., and H. Schmidt. [2013]. Public Engagement in Health Technology Assessment and Coverage Decision: A Study of Experiences in France, Germany, and the United Kingdom. *Journal of Health Politics, Policy and Law* 38:1 [pp tbc].
- Kühn, H. 2005. Demographischer Wandel und GKV. *Die Krankenversicherung* 6 (7):178–223.
- Laibson, D. 1997. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics* 112 (2):443–477.
- Langston, A. L., M. McCallum, M. K. Campbell, C. Robertson, and S. H. Ralston. 2005. An integrated approach to consumer representation and involvement in a multicentre randomized controlled trial. *Clin Trials* 2 (1):80-7.
- Le Grand, J. 1984. Equity as an Economic Objective. *Journal of Applied Philosophy* 1 (1):39-51.
- . 1991. *Equity and choice. An essay in economics and applied philosophy.* London: Harper Collins Academic.
- . 1999. Competition, cooperation, or control? Tales from the British National Health Service. *Health Affairs* 18 (3):27-39.
- . 2007. *The Other Invisible Hand: Delivering Public Services through Choice and Competition.* Princeton: Princeton University Press.
- Le Grand, J., and D. Srivastava. 2009. *Health England Report No 3: Incentives for Prevention.* London: Health England.
- Lehoux, P., G. Daudelin, O. Demers-Payette, and A. Boivin. 2009. Fostering deliberations about health innovation: What do we want to know from publics? *Social Science and Medicine* 68 (11):2002-2009.

- Levin, D. 1987. *Pathologies of the Modern Self*. New York: New York University Press.
- Liedtke, A. 2009. *Priorisierung im Gesundheitswesen. Eine Umfrage der Allianz Deutschland AG.* . München: Allianz Private Krankenversicherungs-AG.
- Loewenstein, G, T Brennan, and K Volpp. 2007. Asymmetric paternalism to improve health behaviors. *JAMA* 298:2415-2417.
- Long, Judith, Marie Helweg-Larsen, and Kevin Volpp. 2008. Patient Opinions Regarding 'Pay for Performance for Patients'. *Journal of General Internal Medicine* 23 (10):1647-1652.
- Lussier, J. P., S. H. Heil, J. A. Mongeon, G. J. Badger, and S. T. Higgins. 2006. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction* 101 (2):192-203.
- Madden, G. J., N. M. Petry, G. J. Badger, and W. K. Bickel. 1997. Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology* 5 (3):256-262.
- Marckmann, G. 2005. *Eigenverantwortung als Rechtfertigungsgrund für ungleiche Leistungsansprüche in der Gesundheitsversorgung? In Gleichheit und Gerechtigkeit in der modernen Medizin*, edited by O. Rauprich, G. Marckmann, J. Vollmann. Paderborn: Mentis.
- Marckmann, G., and B. Gallwitz. 2007. *Gesundheitliche Eigenverantwortung beim Typ-2-Diabetes*. *Zeitschrift für medizinische Ethik* 53:103-116.
- Marmot, M. G., M. J. Shipley, and G. Rose. 1984. Inequalities in death--specific explanations of a general pattern? *Lancet* 1 (8384):1003-6.
- Marmot, M. G., G. D. Smith, S. Stansfeld, C. Patel, F. North, J. Head, I. White, E. Brunner, and A. Feeney. 1991. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 337 (8754):1387-93.
- Marsden, J., and J. Bradburn. 2004. Patient and clinician collaboration in the design of a national randomized breast cancer trial. *Health Expect* 7 (1):6-17.
- Marteau, T. M., R. E. Ashcroft, and A. Oliver. 2009. Using financial incentives to achieve healthy behaviour. *BMJ* 338.
- Marteau, T. M., A. Oliver, and R. E. Ashcroft. 2008. Changing behaviour through state intervention. *BMJ* 337.
- Mason, Jennifer. 1996. *Qualitative Researching*. London: Sage.
- Matschinger, H., and M. C. Angermeyer. 2004. The public's preferences concerning the allocation of financial resources to health care: results from a representative population survey in Germany. *Eur Psychiatry* 19 (8):478-82.
- Matthews, B.A., R. C. Anderson, and A. B. Nattinger. 2005. Colorectal cancer screening behavior and health insurance status. *Cancer Causes Control* 16 (6):735-42.
- McCullough, Laurence B., John H. Coverdale, and Frank A. Chervenak. 2007. *Constructing a Systematic Review for Argument-Based Clinical Ethics*

- Literature: The Example of Concealed Medications. *Journal of Medicine and Philosophy* 32 (1):65-76.
- McNamee, S. J., and R. K. Miller. 2009. *The meritocracy myth*. Plymouth: Rowman & Littlefield.
- Medical Hospital Doctor Magazine. 2008. NHS Rationing Survey. Surrey.
- Meißner, Marc. 2011. GKV-Finanzierungsgesetz - Gegenwind für Wahltarife. *Deutsches Ärzteblatt* 108 (31-32).
- Merzel, C., and J. D'Afflitti. 2003. Reconsidering community-based health promotion: promise, performance, and potential. *Am J Public Health* 93 (4):557-74.
- Metz, D. 2001. Changing health status—a demographic time bomb? [Paper presented at IPPR seminar]. London: IPPR.
- Miles, M. B., and M.A. Huberman. 1994. *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks, CA: Sage.
- Mill, J. S. 1989 [1859] *On Liberty*. Edited by S. Collini, *On Liberty and other essays* Cambridge: Cambridge University Press.
- Miller, T. 2011. World's Rising Powers Face Growing Economic Threat of Diabetes, Heart Disease <http://www.pbs.org/newshour/rundown/2011/09/brics-face-off-with-economic-threat-of-diabetes-heart-disease.html>.
- Minkler, M. 1986. "Generational equity" and the new victim blaming: an emerging public policy issue. *Int J Health Serv* 16 (4):539-51.
- . 1999. Personal responsibility for health? A review of the arguments and the evidence at century's end. *Health Educ Behav* 26 (1):121-40.
- . 2000. Personal Responsibility for health: context and controversies. In *Promoting Healthy Behaviour: How much freedom? Whose responsibility?*, edited by D. Callahan. Washington DC: Georgetown University Press.
- Mossialos, E., and D. King. 1999. Citizens and rationing: analysis of a European survey. *Health Policy* 49 (1-2):75-135.
- Mühlhauser, I. 2007. Früherkennung und Prävention: Ist Vorbeugen besser als Heilen? *Dtsch Arztebl* 104 (25):A-1804 / B-1589 / C-1529.
- . 2011. Vollständige und verständliche Daten sind nötig [Complete and Comprehensible Data Are Required]. *Dtsch Arztebl Int* 108 (10):171.
- Müller, S. 2003. *Bonusprogramme als Instrumente des Beziehungsmarketing – Eine Bestandsaufnahme bisheriger Studien*. Erlangen – Nürnberg: Universität; Betriebswirtschaftliches Institut, Lehrstuhl für Marketing.
- Muntz, D. 2011. *Population-Based Survey Experiments*. Princeton: Princeton University Press.
- Narveson, J. 1988. *The Libertarian Idea*. Philadelphia: Temple University Press.
- National Business Group on Health. 2011. *Survey: Large Employers' 2011 Health Plan Design Changes*. DC: National Business Group on Health.
- National Business Group on Health, and Towers Watson. 2011. *2011/2012 Staying@Work Report*. New York: Towers Watson.

- Nord, E. 2006. Severity of illness and priority setting: worrisome lack of discussion of surprising finding. *J Health Econ* 25 (1):170-2; discussion 173-4.
- O'Donnell, M., and V. Entwistle. 2004. Consumer involvement in decisions about what health-related research is funded. *Health Policy* 70 (3):281-90.
- OECD. 2011. *Health at a glance 2011: OECD Indicators*: OECD Publishing.
- Oliver, A, and L. Brown. 2011. A Consideration of User Financial Incentives to Address Health Inequalities. *Journal of Health Politics, Policy and Law* Advance Publication, published on December 5, 2011.
- Oliver, A, A. Healey, and J. Le Grand. 2002. Addressing health inequalities. *The Lancet* 360 (9332):565-567.
- Patton, M. Q. 1999. Enhancing the quality and credibility of qualitative analysis. *Health Services Research* 34 (5):1189-1208.
- Paul-Ebhohimhen, V., and A. Avenell. 2008. Systematic review of the use of financial incentives in treatments for obesity and overweight. *Obes Rev* 9 (4):355-67.
- Petry, N. M., S. M. Alessi, K. M. Carroll, T. Hanson, S. MacKinnon, B. Rounsaville, and S. Sierra. 2006. Contingency management treatments: Reinforcing abstinence versus adherence with goal-related activities. *J Consult Clin Psychol* 74 (3):592-601.
- American College of Physicians. 2010. *Ethical Considerations for the Use of Patient Incentives to Promote Personal Responsibility for Health: West Virginia Medicaid and Beyond*. Philadelphia: American College of Physicians.
- Prainsack, B., and A. Buyx. 2011. *Solidarity: Reflections on an emerging concept in bioethics*. London: Nuffield Council on Bioethics.
- Promberger, M., R. C. Brown, R. E. Ashcroft, and T. M. Marteau. 2011. Acceptability of financial incentives to improve health outcomes in UK and US samples. *J Med Ethics*.
- Raikka, J. 1996. The social concept of disease. *Theor Med* 17 (4):353-61.
- Resnik, D. B. 2005. The Patient's Duty to Adhere to Prescribed Treatment: An Ethical Analysis. *Journal of Medicine and Philosophy* 30 (2):167 - 188.
- . 2007. Responsibility for health: personal, social, and environmental. *J Med Ethics* 33 (8):444-445.
- Riesberg, A., and R. Busse. 2003. Cost-shifting (and modernization) in German health care. *EuroObserver* 5 (4):4-5.
- Robert Koch-Institut. 2010. *Krebs in Deutschland 2005/2006. Häufigkeiten und Trends*. 7. Ausgabe. Berlin.
- Roemer, J.E. 1993. A pragmatic theory of responsibility for the egalitarian planner *Philosophy and Public Affairs* 22:146–66.
- . 1994. *Egalitarian Perspectives*. Cambridge: Cambridge University Press.
- Roemer, J. 1995. Equality of opportunity. *Boston Review*, April-May.
- Rose, G. 1985. Sick individuals and sick populations. *Int J Epidemiol* 14 (1):32-8.

- Ross, Andrew. 2008. Safeway plan part of Senate health care debate. *SF Gate*, October 8.
- Russell, L. B. 2009. Preventing chronic disease: an important investment, but don't count on cost savings. *Health Aff (Millwood)* 28 (1):42-5.
- Saldana, J. 2011. *The Coding manual for qualitative researchers*. Thousand Oaks: Sage.
- Saltman, R. 2004. The historical and social base of social health insurance systems. In *Social health insurance systems in western Europe*, edited by R. Saltman, Busse, R., Figueras, J. Maidenhead: Open University Press.
- Savage, C. L., Y. Xu, R. Lee, B. L. Rose, M. Kappesser, and J. S. Anthony. 2006. A case study in the use of community-based participatory research in public health nursing. *Public Health Nurs* 23 (5):472-8.
- Scanlon, T. 1998. *What we owe to each other*. Cambridge, MA: Belknap Press.
- Scheller-Kreinsen, D., M. Blümel, and R. Busse. 2009. Chronic disease management in Europe. *EuroHealth* 15 (1):1-5.
- Schmidt, H. 2007. Personal responsibility for health: Developments under the German Healthcare Reform 2007, *European Journal of Health Law* 14 (2007) 241–250
- . 2007b. Patients' charters and health responsibilities, *BMJ* 2007;335(7631):1187–1189;
- . 2007c. [Public consultation and "reasonable override,"] Commentary on the role of public consultation in Furger's and Fukuyama's *Beyond Bioethics*, with reply from Furger and Fukuyama. *Hastings Center Report* 37 (6):6.
- . 2008a. Bonuses as Incentives and Rewards for Health Responsibility: A Good Thing? *J Med Philos* 33 (3):198-220.
- . 2008b. Germany Institutes 'Incentives' for Cancer Patients, *Hastings Center Bioethics Forum*, Jan 2008;
- . 2009a. Just health responsibility. *J Med Ethics* 35 (1):21-6.
- . 2009b. Personal responsibility in the NHS Constitution and the social determinants of health approach: competitive or complementary? *Health Econ Policy Law* 4 (Pt 2):129-38.
- . 2009c. Sickness funds, healthy people, obese people: are higher contributions for those who 'cannot control themselves' justified? In *Whose weight is it anyway: essays on ethics and eating*, edited by I. de Beaufort, S. Vandamme and S. van de Vathorst. *Gent Acco (Peer Reviewed Series)*
- . 2010. Personal Responsibility for health: a proposal for a nuanced approach. In *Improving health and health care: Who is responsible? Who is accountable?*, edited by B. Rosen, A. Israeli and S. Shortell. Jerusalem: The Israel National Institute for Health Policy.
- . 2011a. Wellness Incentives, Equity, and the 5 Groups Problem. *Am J Public Health*. 2012 Vol. 102, No. 1, pp. 49-54.
- . 2011b. Public Health Ethics. In: *Encyclopedia of Applied Ethics*, Ruth Chadwick (ed). Oxford: Elsevier

- . 2012. Public health ethics and the duty to think. In *Ethics in public health and health policy. Concepts, Methods and Case Studies*, edited by D. Strech, I. Hirschberg and G. Marckmann. Dordrecht: Springer [forthcoming].
- Schmidt, H., S. D. Halpern, and D. A. Ash. 2012. Fairness and wellness incentives: what is the relevance of the process-outcome distinction? *Preventive Medicine* [forthcoming].
- Schmidt, H., Madsion, K, Halpern, S.D., Volpp, K. [2012] Improving Wellness Incentive Programs Through Reporting and Evaluation [currently in preparation for submission]
- Schmidt, H., S. Stock, and T. Doran. 2012. Moving forward with wellness incentives under the affordable care act: some lessons from Germany Commonwealth Fund Issues Brief. New York: The Commonwealth Fund.
- Schmidt, H., S. Stock, and A. Gerber. 2009. What can we learn from German health incentive schemes?, *BMJ* 2009;339:725–728.
- Schmidt, H., K. Voigt, and D. Wikler. 2010. Carrots, sticks, and health care reform--problems with wellness incentives. *New England Journal of Medicine* 362 (2):e3;
- Schomerus, G., H. Matschinger, and M. C. Angermeyer. 2006. Alcoholism: Illness beliefs and resource allocation preferences of the public. *Drug Alcohol Depend* 82 (3):204-10.
- Schreier, M., and A. Diederich. 2008. Kriterien der Priorisierung: Praxis anderer Länder und erste Ergebnisse einer qualitativen Befragung in Deutschland. In *Priorisierung in der Medizin*, FOR 655 Nr. 14 / 2008. Bremen: Jacobs University.
- Schroeder, S. A. 2007. We Can Do Better — Improving the Health of the American People. *New England Journal of Medicine* 357 (12):1221-1228.
- Segall, S. 2007. In Solidarity with the Imprudent: A Defense of Luck-Egalitarianism. *Social Theory & Practice* 33 (2):177-198.
- Serrano-Aguilar, P., M. M. Trujillo-Martín, J. M. Ramos-Goñi, V. Mahtani-Chugani, L. Perestelo-Pérez, and M. Posada-de la Paz. 2009. Patient involvement in health research: A contribution to a systematic review on the effectiveness of treatments for degenerative ataxias. *Soc Sci Med* 69 (6):920-925.
- Sherwin, S. 1998. *The Politics of Women's Health: Exploring Agency and Autonomy*. Philadelphia: Temple University Press.
- Sieg, A., and H. Brenner. 2007. Cost-saving analysis of screening colonoscopy in Germany. *Z Gastroenterol* 45 (9):945-51.
- Statistisches Bundesamt. 2008. *Sozialleistungen - Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus) Fachserie 13 Reihe 1.1*. Wiesbaden: Statistisches Bundesamt.
- . 2009. *Bevölkerung und Erwerbstätigkeit – Bevölkerungsfortschreibung, Fachserie 1 Reihe 1*. Wiesbaden: Statistisches Bundesamt.



- . 2010. Bevölkerung 2009 nach Altersgruppen und Bildungsabschluss nach ISCED97 [spreadsheet provided on request]. Wiesbaden: Statistisches Bundesamt.
- . 2011a. Jährliche Einkommensteuerstatistik 2007, Fachserie 14, Reihe 7.1.1. Wiesbaden: Statistisches Bundesamt.
- . 2011b. Körpergröße, Körpergewicht und Body-Mass-Index 2009 nach Geschlecht und Altersgruppen [spreadsheet, provided on request]. Wiesbaden: Statistisches Bundesamt.
- . 2011c. Wirtschaftsrechnungen – Leben in Europa (EU-SILC), Einkommen und Lebensbedingungen der Europäischen Union, Fachserie 15, Reihe 3, DeStatis Fachserie. Wiesbaden: Statistisches Bundesamt.
- Steinbrook, R. 2006. Imposing personal responsibility for health. *N Engl J Med* 355 (8):753-6.
- Stemplowska, Z. 2008. Making Justice Sensitive to Responsibility. *Political Studies* 57 (2):237-259.
- Stock, S., H. Schmidt, G. Buscher, A. Gerber, A. Drabik, C. Graf, M. Lungen, and B. Stollenwerk. 2010. Financial incentives in the German Statutory Health Insurance: new findings, new questions. *Health Policy* 96 (1):51-6.
- Stratton-Lake, P. 2004. *On What We Owe to Each Other: Scanlon's Contractualism*. Oxford: Blackwell.
- Strech, D., M. Synofzik, and G. Marckmann. 2008. Systematic reviews of empirical bioethics. *Journal of Medical Ethics* 34 (6):472-477.
- Sunstein, C., and R. Thaler. 2003. Libertarian Paternalism is not an oxymoron. *The University of Chicago Law Review* 70 (4):1159-1202.
- Taylor, R., and A. Rieger. 1984. Rudolf Virchow on the typhus epidemic in Upper Silesia: an introduction and translation. *Sociology of Health & Illness* 6 (2):201-217.
- ten Have, H. 1994. The distinction between prospective and retrospective responsibility. *Health Care Analysis* 2 (2):119-123.
- Thaler, R., and C. Sunstein. 2003. Libertarian Paternalism. *The American Economic Review* 93 (2):175-179.
- . 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven: Yale University Press.
- Tong, A., K. Howard, S. Jan, A. Cass, J. Rose, S. Chadban, R. D. Allen, and J. C. Craig. 2010. Community preferences for the allocation of solid organs for transplantation: a systematic review. *Transplantation* 89 (7):796-805.
- Ubel, P. A., C. Jepson, J. Baron, T. Mohr, S. McMorro, and D. A. Asch. 2001. Allocation of transplantable organs: do people want to punish patients for causing their illness? *Liver Transpl* 7 (7):600-7.
- United States Census Bureau. 2011. *Income, Poverty, and Health Insurance Coverage in the United States: 2010*. Edited by C. DeNavas-Walt, B. Proctor and J. Smith, Current Population Reports. Washington DC: United States Census Bureau,.

- USamp. 2011. Reply to ESOMAR 26 Questions. Dallas, London, Los Angeles: Usamp. available at: [http://directory.esomar.org/moreinfo/2923/uSamp\\_DMS--ESOMAR-26.pdf](http://directory.esomar.org/moreinfo/2923/uSamp_DMS--ESOMAR-26.pdf)
- van Baal, P. H., J. J. Polder, G. A. de Wit, R. T. Hoogenveen, T. L. Feenstra, H. C. Boshuizen, P. M. Engelfriet, and W. B. Brouwer. 2008. Lifetime medical costs of obesity: prevention no cure for increasing health expenditure. *PLoS Med* 5 (2):e29.
- Venkatapuram, S., and M. Marmot. 2009. Epidemiology and social justice in light of social determinants of health research. *Bioethics* 23 (2):79-89.
- Volpp, K. G., A. Gurmankin Levy, D. A. Asch, J. A. Berlin, J. J. Murphy, A. Gomez, H. Sox, J. Zhu, and C. Lerman. 2006. A randomized controlled trial of financial incentives for smoking cessation. *Cancer Epidemiol Biomarkers Prev* 15 (1):12-8.
- Volpp, K. G., A. B. Troxel, M. V. Pauly, H. A. Glick, A. Puig, D. A. Asch, R. Galvin, J. Zhu, F. Wan, J. DeGuzman, E. Corbett, J. Weiner, and J. Audrain-McGovern. 2009. A randomized, controlled trial of financial incentives for smoking cessation. *New England Journal of Medicine* 360 (7):699-709.
- Volpp, K.G., L. John, A.B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein. 2008. Financial Incentive-based Approaches for Weight Loss: A Randomized Trial. *JAMA* 300:2631-7.
- Volpp, K.G., G. Loewenstein, A.B. Troxel, J. Doshi, M. Price, M. Laskin, and S.E. Kimmel. 2008. A test of financial incentives to improve warfarin adherence *BMC Health Services Research* Page 8:272.
- Volpp, KG, Troxel AB, Pauly MV, Glick HA, Puig A, Asch DA, Galvin R, Zhu J, Wan F, DeGuzman J, Corbett E, Weiner J, and Audrain-McGovern J. . 2009. A Randomized Controlled Trial of Financial Incentives for Smoking Cessation. *New England Journal of Medicine* 360:699-709.
- von dem Knesebeck, O., and A. Mielck. 2009. Soziale Ungleichheit und gesundheitliche Versorgung im höheren Lebensalter. *Z. Gerontol. Geriatr.* 42 (1):39-46.
- Wallen, J., and S. R. Williams. 1982. Employer-based health insurance. *J Health Polit Policy Law* 7 (2):366-79.
- Weide, U. 2005. Law and the German Universal Healthcare System: A Contemporary Overview. *GERMAN LAW JOURNAL* 6 (8):1143-1172.
- Welsh, E. 2002. *Dealing with Data: Using NVivo in the Qualitative Data Analysis Process*. Vol. 3, 2002.
- Wendt, C., H. Rothgang, and U.Helmert. 2005. The self-regulatory German health care system between growing competition and state hierarchy. *TranState Working Papers, University of Bremen, Collaborative Research Center 597: Transformations of the State* 32.
- Werner, M. H. 2002. Verantwortung. In *Handbuch Ethik*, edited by M. Düwell, Hübenenthal, Christoph, and M. H. Werner. Stuttgart: J. B. Metzler.
- White, A., S. W. Vernon, L. Franzini, and X. L. Du. 2011. Racial and Ethnic Disparities in Colorectal Cancer Screening Persisted Despite Expansion of

- Medicare's Screening Reimbursement. *Cancer Epidemiology, Biomarkers & Prevention* 20 (5):811-17.
- Wikler, D. 1978. Persuasion and coercion for health - ethical issues in government efforts to change lifestyles. *Millbank Memorial Fund Quarterly/Health and Society* Vol 56 (No.3):303-338.
- . 2004. Personal and Social Responsibility for Health. In *Public Health, Ethics and Equity*, edited by S. Anand, Peter, F., Sen, A. Oxford: OUP.
- Wilson, J. 2009. Towards a normative framework for public health ethics and policy *Public Health Ethics* 2 (2):184-194.
- Winawer, S. J., A. G. Zauber, M. N. Ho, M. J. O'Brien, L. S. Gottlieb, S. S. Sternberg, J. D. Wayne, M. Schapiro, J. H. Bond, J. F. Panish, and et al. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med* 329 (27):1977-81.
- Wittenberg, E., S. J. Goldie, B. Fischhoff, and J. D. Graham. 2003. Rationing decisions and individual responsibility for illness: are all lives equal? *Med Decis Making* 23 (3):194-211.
- Woolf, S., H. 2009. A closer look at the economic argument for disease prevention. *JAMA* 301:536-8.
- Wright, K. B. 2005. Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. *Journal of Computer-Mediated Communication* 10 (3):00-00.
- Wuppermann, D, I Wuppermann, and JF Riemann. 2009. Aktueller Wissensstand der Bevölkerung zur Darmkrebsvorsorge – eine Untersuchung der Stiftung LebensBlicke mit dem Institut für Demoskopie in Allensbach *Z Gastroenterol* 47:1132–6.
- Yeager, D. S., J. A. Krosnick, L-C Chang, H. S. Javitz, M. S. Levendusky, A. Simpser, and R. Wang. 2011. Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *Public Opinion Quarterly* 75 (4):709-747.

## Appendix 2A: Health-related activities incentivised for insurees of the Betriebskrankenkassen

<b>Core activities offered by all Betriebskrankenkassen (BKKs)</b>
Health check ups pursuant §§ 25 u. 26 SGB V: biennial screening for +35 year olds for cardiovascular conditions and diabetes; cancer screening for women +20, men +45; children health check-up
Quality assured services pursuant § 20 Abs. 1 SGB: primary prevention programmes, in particular those that help reduce socially determined health inequalities
Non-smoker bonus
Regular sport—active sports club membership
Regular sport—membership in quality assured/accredited gym
Regular sport—German Sports Certificate ( <i>Sportabzeichen</i> : age and gender-specific timed runs, sit ups, push ups, etc)
Annual dental check ups pursuant §§ 22 u. 55 SGB V
<b>Other programs (offered by at least one of the BKKs)</b>
<b>Workplace activities</b>
Participation in occupational health promotion schemes
Health check up by company-employed physician
<b>Sport &amp; wellness</b>
Sports certificates (other than <i>Sportabzeichen</i> ), e.g. swimming certificate
Participation in sports programmes in context of rehabilitation or disability
Wellness weekends
Health courses
Attending health-relevant lectures
Individual (health and) nutrition assessments
<b>Vaccinations</b>
Vaccinations at own cost
Travel-vaccinations at own cost
Attending counselling session on travel vaccinations
Flu vaccination
<b>Screening and check-ups</b>
Prenatal tests and checks
Measuring key health indicators e.g. BMI, blood-sugar, blood-pressure, cholesterol
Variation of above: bonus only if key indicators are in acceptable range (or: have improved) over a year
Melanoma, breast, colon, cervical cancer screening
PSA value assessment
Osteoporosis check up
Glaucoma check up
Eye-pressure measurement
Pre-pregnancy and prenatal
<b>Other activities</b>
Personal primary prevention initiatives
Professional cleaning of teeth
Participation in quality assured sexually transmitted disease counselling
Keeping of a personal health record
Participation in preventative outpatient services
Source: Evaluation matrix used by the <i>Bundesversicherungsamt</i> , the relevant authority to which BKKs need to report financial data in relation to implemented Bonus programmes pursuant SGB V Article 65a (data file made available on request, May 2008).

## Appendix 2B: Health care premiums of covered US workers by plan type, company size and employer/employee contribution

As Table 2B1 and Figure 2B1 show, in the US, more comprehensive or flexible coverage is typically more expensive than in Germany and employers contribute a larger share of healthcare cost. While the differences in cost of coverage between the different plan options is relatively moderate, the amount workers contribute directly varies more significantly.

	Annual single coverage	Annual family coverage
<b>HMO</b>		
All Small Firms (3–199 Workers)	\$5,133	\$13,285
All Large Firms (200 or More Workers)	\$5,169	\$14,492
<b>PPO</b>		
All Small Firms (3–199 Workers)	\$5,169	\$13,735
All Large Firms (200 or More Workers)	\$5,104	\$14,161
<b>POS</b>		
All Small Firms (3–199 Workers)	\$5,145	\$12,825
All Large Firms (200 or More Workers)	\$5,402	\$13,850
<b>HDHP/SO</b>		
All Small Firms (3–199 Workers)	\$4,454	\$12,022
All Large Firms (200 or More Workers)	\$4,482	\$12,640
<b>All plans</b>		
All Small Firms (3–199 Workers)	\$5,046	\$13,250*
All Large Firms (200 or More Workers)	\$5,050	\$14,038*

Source: Kaiser/Health Research & Educational Trust Survey of Employer-Sponsored Health Benefits, 2010: 15

**Fig 2B1: Average annual health care premiums (total, and by employer/employee contributions) by health plan type**



## Appendix 2C: Cost shifting through wellness incentives: BeniComp®

### Advantage

The text below is copied directly from the BeniComp website, though some sections have been omitted for reasons of space. Available at:

[http://www.benicompadvantage.com/index.php?option=com\\_content&view=article&id=8&Itemid=4](http://www.benicompadvantage.com/index.php?option=com_content&view=article&id=8&Itemid=4), accessed 18 February 2012.

[From page: "Detailed overview"]

#### What haven't you tried? – BeniComp® Advantage!

You've increased employee contributions, raised copays and coinsurance, required generic prescriptions, implemented wellness awareness campaigns, what next? We all have tried to shift the cost of health care with plan design changes and by having our employees share in the cost. It simply has not worked. The cost of health care continues to escalate at an alarming rate and eat at our profits and our ability to invest in our companies and employees.

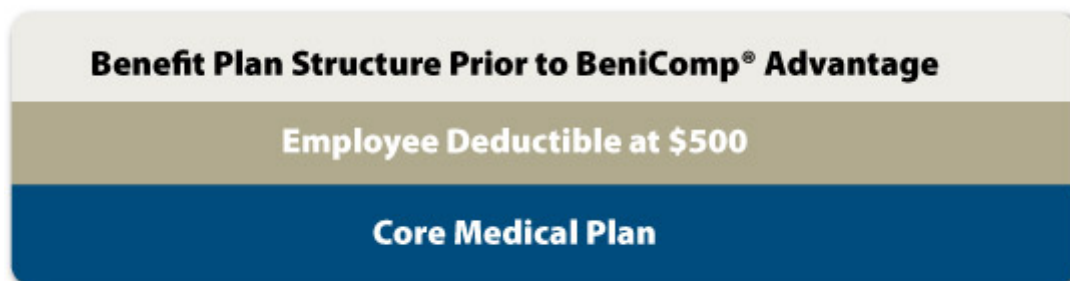
BeniComp® Advantage works in conjunction with a high deductible health plan and is the first plan that rewards employees for managing their own health. If an employee maintains or improves their health, BeniComp® Advantage offers a financial reward. For those employees who choose not to modify their lifestyle, they do not receive these rewards and will incur higher out of pocket cost.

Either way, costs are allocated more fairly between employees with high and low utilization patterns, a much better way to share health care costs. BCA is voluntary – employees who do not participate simply have a higher deductible.

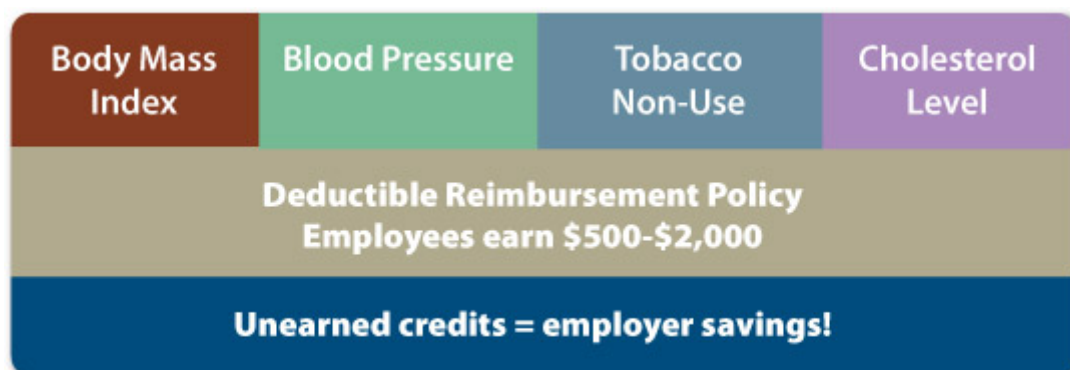
#### Here's How It Works

The BeniComp® Advantage plan does shift cost, but only to those who have lifestyles that threaten controlled premiums for the rest of the population.

Using BeniComp® Advantage, the base medical plan is the same for all employees. BeniComp® Advantage overlays the base medical plan with a policy that reimburses portions of an employees base plan deductible.



Separate Fully Insured BeniComp® Advantage Policy:



This example assumes that the base medical deductible is \$500/\$1,000 for the employee/family. When adding the Advantage product, the base medical plan deductible is increased by \$2,000 making the deductible \$2,500/\$5,000 per employee/family. Employees can qualify for credits worth \$500 each for the lifestyle choices shown. Employees, based on results from a medical screening, can qualify for any or all of the available credits by demonstrating appropriate body mass, blood pressure, and cholesterol levels, as well as non-tobacco use.

[From page: “Employers FAQ”]

**Is this allowed under HIPAA?**

BCA has been structured to comply with non-discrimination requirements for wellness plans. Refer to our HIPAA Position Statement. Does everyone have to participate?

No one is required to participate in this program. It is entirely voluntary and up to each individual. However, by not participating, individuals will be responsible for the entire deductible, no credits will be awarded.

**How much money can employers expect to save?**

Savings are found by:

- Reduction in health care claims due to healthier employees and the higher deductibles 12%-30%
- Lower aggregate attachment points and premium, many will experience 12%-18% net savings including the cost of the program
- Employees with significant risk factors identified during the screening will receive a “critical notification” and be advised to see their physician.
- Discounted specific premium rates, if the employer is on an ASO the savings is directly reflected in the employer’s bottom line
- Employees who choose other health care options

**How soon will I realize savings?**

Immediately. Since deductibles will be raised, the amount of money that the employer pays will decrease beginning day one. Once employees improve their lifestyles, you will see a significant savings in reduced claims. [...] On average plans will benefit from a 12-18% net savings.

**This seems aggressive. How do most employees react?**

If you have been encouraging consumerism in your health plan by raising deductibles, eliminating co-pays and increasing coinsurance, this policy is a natural progression. Employees understand that the company cannot continue to provide rich benefits with plan expenses outpacing profits 5:1. Some employers have suggested that like auto insurance, those employees who are doing what they can to lower their risk of serious illness receive the equivalent of a “safe driver discount”. Employees understand this concept and receive it very well.

[From page: “Employees FAQ”]

**How are employees rewarded for healthy lifestyles?**

Under this plan, a health screening will take place to assess basic health parameters. You will be evaluated on Blood Pressure, Cholesterol, Body Mass Index and Tobacco use. If your results fall within a specific range, you will qualify for wellness credits to be used to reimburse portions of your deductibles.

What are the ranges for each of the above lifestyle qualifiers?

The National Institute of Health publishes guidelines that BeniComp® Advantage uses as benchmarks. NIH guidelines are:

- Blood Pressure: <120/80
- Cholesterol LDL: <100
- Body Mass Index (BMI): <24.9

While these are established national guidelines, your employer may choose to modify these parameters.

**Why are these categories “singled out”? Shouldn’t we address accidents and serious diseases instead?**

The basic intent of an insurance plan is to provide benefits that promote employee’s self-management and preventative care and to help employees afford care if a catastrophic injury or illness occurs. Your base core medical plan still does this. Recent studies have shown that 70% of illness is lifestyle related and thus may be preventable. The Centers for Disease Control have estimated that one-third of all US deaths are related to diet, physical inactivity and smoking and that 30% of cancers in the US are related to obesity and diet.

While employees may not be able to control the fact that they have a disease, illness or injury they can control their own lifestyle choices. BeniComp® Advantage is a tool that rewards employees who choose to manage their own personal health and reduce their risks for future health related issues. This is similar to the concept of a “good driver discount” you enjoy with auto insurance.

**Appendix 3A: Implications of different policy responses to the five groups problem**

Policy options/ groups	'Lucky ones'	'Yes I can' group	'I'll do it tomorrow' group	'Unlucky ones'	'Leave me alone' group	Comments
Offer universally	Benefit	Benefit	Don't benefit	Don't benefit	Don't benefit	'Unlucky ones' lose out 'I'll do it tomorrow' group, treated identical to 'leave me alone' group Some 'lucky ones' reap benefits even if they do not change behaviour, or comply with spirit of policy
Offer universally, modified	Benefit	Benefit	May benefit	May benefit	Don't benefit	Create alternative standards for 'unlucky ones' & 'I'll do it tomorrow;' can improve fairness, but faces practical and arbitrariness challenges Shift from offering alternative standards in response-mode to proactive-mode, to reduce negative aspects of petitioning Shift focus from attainment to participation-incentives
Targeted, not universal	Don't benefit	Benefit	Benefit	May benefit	Don't benefit	No incentives for 'lucky ones' as they do not require further encouragement Potential for curbing cost, if focus is on improving health status of worst off Minimises potential for exacerbating existing inequalities in terms of health and wealth Reverses financing of benefits where benefits result from cost-shifting: instead of the poorer and unhealthy financing the benefits of the better off and more healthy, here, controversially, the



						opposite is the case
Abandon	Don't benefit	Don't benefit	Don't benefit	Don't benefit	Don't benefit	No unfairness from different use of incentives, but also no potential to use incentives complementary to action at the level of the social determinants of health Strongest case if it can be shown that there are other measures to improve population health that are equally or more effective

## Appendix 4A: Initial draft instrument for survey and semi-structured interviews<sup>1</sup>

	<b>Survey questions</b> (population level)	<b>Data/type of question</b>	<b>Interview questions</b> (physicians, prim. Care / cancer)	<b>Purpose of question</b>
	<b>General responsibility attitudes</b>			
1	<p>It can be argued that whether or not people are healthy depends on many different things. Environmental factors such as clean air and water matter, as do genetics which may lead to robust health or inherited diseases. Also important are access to health care, and social circumstances such as where people live and how much money they have. Personal choices are another factor, as health is influenced by what we eat, or drink, whether we exercise or smoke or not, and how we look after ourselves.</p> <p>Broadly speaking, when it comes to our own health, how important are personal choices in relation to these other factors? Choose one of the following four options:</p> <ol style="list-style-type: none"> <li>1. Personal choices are the most important factor that determines health: we are largely responsible for how healthy we are</li> <li>2. The role of personal choices is an important factor, and we always bear some responsibility for how healthy we are. But sometimes other factors can be more important.</li> <li>3. The importance of personal choices is overstated: many of the factors above influence quite strongly what kind of choice we can make. Therefore it does not make much sense to say that people are responsible for their health.</li> <li>4. Don't know/none of the above.</li> </ol>	Categorical/ Closed	Same, if (4) is chosen, ask for what other option they think of	<p>'Calibrate' respondent/see where they are 'coming from'.</p> <ul style="list-style-type: none"> <li>- correlate assessment made here with views about auto-insurance analogy and penalizing those who make 'bad' choices (e.g.: do all who are for high penalties also believe that personal choice has much power?)</li> <li>- test notion of voluntariness, and health as 'co-production'</li> <li>- correlate assessment here with demographics, esp. income and health status</li> </ul>
2	<p>In car insurance, risky drivers have higher premiums. Some say, the same model should be used in health insurance, so that, for example, obese people pay higher premiums than healthy people.</p> <p>On a scale of 1 to 5, how strongly do you agree with this proposal?</p> <p>(5 means that you fully agree, 1 means that you disagree completely)</p>	Ordinal /Likert-scale	Same, ask also for reasons	<p>'Calibrate' respondent/see where they are 'coming from'.</p> <p>Test:</p> <ul style="list-style-type: none"> <li>- extent of sympathy with views a la Burd</li> <li>- implicitly: solidarity</li> <li>- voluntariness (as in question above)</li> </ul>

<sup>1</sup> Note that the instrument has changed considerably after pre-testing (as was to be expected). It might therefore be more instructive to review the instruments as used for the physician surveys (Appendix 5D and E) and population level surveys (as reproduced in Chapter 6).

				- compare general attitude expressed here with responses to more specific questions below - correlate assessment here with demographics, esp. income and health status
3	If you disagree with the car-insurance analogy: what are your reasons?	Open	Same	As above
4	How would you rate the following proposal: Unhealthy enrollees of a health plan don't pay more, but those who show that they are healthy get \$500 a year? Please rate on a scale of 1 to 5 (5 means that you fully agree, 1 means that you disagree completely)  [adjust level for Germany; EUR100]	Ordinal /Likert-scale	Same, ask also for reasons	'Calibrate' respondent/see where coming from  Relate general assessment of attitudes towards 'carrots' and 'sticks' to responses to specific policies below
	<b>Questions about obesity incentives</b>			
	I want to ask your views about a real policy. A large employer offers his workers a high deductible health plan. The normal deductible is \$500. To reward healthy behavior, the employer increases the deductible for all employees to \$1000. However if employees show that their Body Mass Index (BMI) is in the normal range over a year, they can reduce the deductible to the previous level of \$500.			
5	What do you like about this policy, what do you not like?	Open	same	Direct question on RQ3: which "aspects of policies are reasonable, which not?"
6	Would you describe this approach as a 'carrot' (because it rewards healthy behavior) or as a 'stick' (because it penalizes unhealthy behavior)?	Categorical/ Closed	Same, ask also for reasons	Attitudes on framing and usefulness of 'carrot'/'stick' classification  Coerciveness
7	Suppose you are in a plan like the above. So you can get \$500 if you have a normal BMI. You have lost 10 pounds, but on the day you are weighed to find out if you qualify, you are 1 pound over the specified weight target. Only you and a doctor is in the room when the measurement is recorded.	Ordinal /Likert-scale	Suppose a patient is enrolled in a plan like the above. You are involved in determining whether a patient has met the	Doctor patient relationship (from both sides)  Compare assessments here with those

	<p>(a) On a scale of 1-5, how likely would you ask the doctor to write down 1 pound less? (5 means certainly would, 1 certainly would not)</p> <p>(b) Suppose you do ask, and the doctor agrees to note one pound less. On a scale of 1-5, would you say this makes her/him a good doctor? (5= excellent doctor, 5 = not a good doctor at all)</p> <p>[adjust premium and scenario for German case to EUR100]</p>		<p>weight target. The patient has lost 10 pounds, but on the day he/she is 1 pound over the specified target. Only you and the patient are in the room when the measurement is recorded.</p> <p>(a) Would you find it understandable if the patient would ask you to record a normal weight?</p> <p>(b) Suppose a doctor accepts the patient's requests to record normal weight. On a scale of 1-5, would you say this makes her/him a good doctor? (5= excellent doctor, 5 = not a good doctor at all)</p> <p>[adjust premium and scenario for German case to EUR100]</p>	<p>below on doc role in refusing treatment</p>
8	<p>Think again about the plan above (you can reduce your deductible from \$1000 to \$500 if you have a normal BMI). Now think of 3 people:</p> <p>Anna likes to eat a lot and dislikes exercising. She does not mind being overweight, but she will not get the \$500.</p> <p>Mary would like to lose weight, often tries, but always fails. She also does not get the \$500.</p> <p>Jane never exercises, or thinks about what she should eat. She meets the BMI target without any effort, and so gets \$500.</p> <p>On a scale from 1 to 5 (5 means that you fully agree, 1 means that you disagree completely) how much do you agree that:</p> <p>It is OK that Anna, who does not like exercise and likes eating, can't reduce her deductible to \$500</p>	Ordinal /Likert-scale	Same, ask also for reasons	Question about Equity/Fairness/ who benefits

	<p>It is OK that Mary, who tries to lose weight but is unsuccessful, can't reduce her deductible to \$500</p> <p>It is OK that Jane, who meets the target without effort can reduce her deductible to \$500</p> <p>[adjust premium and scenario for German case to EUR100]</p>			
9	<p>Poorer people are often less healthy and more overweight than those with more money. Think again about the policy above where you can reduce your deductible from \$1000 to \$500 if you have a normal BMI.</p> <p>Suppose that research shows that almost twice as many of the most well off benefit from the programme than the poorest. On a scale from 1 to 5 how concerned are you that the poor are less likely to benefit? (5 is very concerned and 1 is not concerned at all)</p>	Ordinal /Likert-scale	Same	Question about Equity/Fairness/ who benefits
<b>Questions about cancer incentives</b>				
	<p>Now some questions about another real example of an incentive policy.</p> <p>A health plan wants to improve colon cancer care. The plan offers an incentive scheme that requires the members to do two things. First, after the age of 40, members should attend counseling on the advantages and disadvantages of getting screened—it is up to them to decide whether they follow through with getting a screen. Second, if they should suffer from colon cancer, they should never refuse treatment.</p> <p>If members do both things, their annual copays are capped at 1%.</p> <p>If they comply only with one, or none, the co-payment is 2%.</p> <p>All other chronically ill have a 1% co-payment, and the plan says that colon cancer patients have special responsibility.</p>			
10	What do you like about this policy, what do you not like?	Open	Same	Direct question on RQ3: which “aspects of policies are reasonable, which not?”
11	Would you describe this approach as a ‘carrot’ (because it rewards good behavior) or as a ‘stick’ (because it penalizes bad behavior)?	Categorical/Closed	Same, ask also about reasons	Attitudes on framing and usefulness of ‘carrot’/‘stick’ classification  Coerciveness
12	Do you think it is better policy if those who come to counseling and never refuse treatment had a cap of 0.5% and all other chronically ill 1%?	Categorical/Closed	Same, ask also about reasons	Ditto

13	<p>Suppose a policy maker suggests to link the lower co-payment not just to attending counseling, but to undergoing the screen. So, in order to get the lower co-payment people need to be tested. On a scale of 1-5, how strongly do you agree with the following statement of a critic ( 5 means: fully agree, 1 means disagree completely.)</p> <p>“There should be a right not to know about whether one will suffer from diseases in the future, including cancer. This means that if someone does not want to take a test, there should be no financial disadvantage or penalty.”</p>	Ordinal /Likert-scale	Same, ask also about reasons	solidarity vs. responsibility autonomy, coercion
14	<p>A policy maker defends the current policy, which, as you know, does not require undergoing screening, but attending counseling on the advantages and disadvantages of screens, by saying the following:</p> <p>“We accept that there is a right not to know about whether one will suffer from cancer in the future. But we do not accept that there is a right not to have to think about whether one should be screened. This is why a financial disadvantage is justified for those who refuse to think about the advantages and disadvantages of colon cancer screens.”</p> <p>On a scale of 1-5, how strongly do you agree with the policy maker’s statement (5 means: fully agree, 1 means disagree completely)</p>	Ordinal /Likert-scale	Same, ask also about reasons	Ditto
15	<p>Suppose you had colon cancer but had reason not to take the prescribed treatment. You know that refusing treatment means that you have higher copays.</p> <p>(a) On a scale of 1-5, how likely would you ask your doctor not to tell your health plan (5 being the most likely)?</p> <p>(b) Suppose you do ask, and the doctor agrees not to tell the health plan. On a scale of 1-5, would you say this makes her/him a good doctor? (5= excellent doctor, 5 = not a good doctor at all)</p>	Ordinal /Likert-scale	<p>Suppose a patient has colon cancer but has reason not to take the prescribed treatment. He/she knows that refusing treatment means higher copays.</p> <p>(a) Would you find it understandable if the patient would ask you not to tell the health plan?</p> <p>(b) On a scale of 1-5, would you say this makes her/him a good doctor? (5= excellent doctor, 5 = not a good doctor at all)</p>	Doctor patient relationship (both ways) Compare if differences between this case and meeting obesity targets, above
16	<p>A question about how high incentives should be.</p> <p>Suppose the total average annual cost of health care for people is \$5000. A health plan wants to increase the employees’ contributions and offer them the incentive of reducing it to previous</p>	Categorical/Closed	Same	Key to acceptability of scope of legal provisions and framing

	<p>levels by taking part in programmes that require meeting targets, for example in relation to body weight.</p> <p>What should be the maximum by which the employer should be allowed to increase contributions? (Select only one of the following six options]:</p> <ul style="list-style-type: none"> <li>- I am against such programmes so no incentives should be given</li> <li>- Maximum 5% of the cost of insurance, so \$250</li> <li>- Maximum 10% of the cost of insurance, so \$500</li> <li>- Maximum 20% of the cost of insurance, so \$1000</li> <li>- Maximum 30% of the cost of insurance, so \$1500</li> <li>- Maximum 50% of the cost of insurance, so \$2500</li> <li>- There should be no limit: whatever health plans think is appropriate</li> </ul> <p>[adjust premium and scenario for German case to EUR100]</p>			
	<b>Demographics and non-evaluative questions</b>			
	6 standard ones: Gender, age, household income, health status, weight, educational attainment	Continuous/ ordinal	N/A	
17	<p>Your use of incentive programmes [multiple selections possible]:</p> <p>(1) I have participated in incentive programmes that require meeting a target (such as weight or cholesterol) in the past</p> <p>(2) I currently participate in incentive programmes that require meeting a target (such as weight or cholesterol)</p> <p>(3) I intend to participate in incentive programmes that require meeting a target (such as weight or cholesterol) in the future</p> <p>(4) I have participated in incentive programmes that merely require doing something (such as filling in a health assessment, or going to a health lecture) in the past</p> <p>(5) I currently participate in incentive programmes that merely require doing something (such as filling in a health assessment, or going to a health lecture)</p> <p>(6) I intend to participate in incentive programmes that merely require doing something (such as filling in a health assessment, or going to a health lecture)</p>	Categorical/ Closed	N/A	

## **Appendix 5A: Method of recruiting physicians via professional organisations, snowball sampling, and practice area profiling**

For the German sample, an initial long list of potential research participants was established by using the online database of the *Kassenärztliche Vereinigung Berlin*,<sup>2</sup> an association of all physicians working with the providers of statutory health insurance, whose website permits a search of all practicing physicians by specialism and borough. Boroughs reflecting low, medium and high socio-economic status were identified by means of the *Sozialstrukturatlas*.<sup>3</sup> Issued annually by the state of Berlin, the Atlas describes areas by relevant criteria, including household income, unemployment rates, and educational attainment, as well as an aggregate value of 'social class', which was used as a proxy for socio-economic status. The long list was established based on the abovementioned inclusion and exclusion criteria and comprised 53 physicians. Of these, a random selection was made, and in order to recruit five primary care physicians, an initial set of 18 emails was sent in mid-January 2011. The email described the purpose and envisaged mode and format of the research, required time of the physician, data management arrangements, and brief biographical information about myself. A reminder email was sent two weeks later, offering also payment of loss-of-earnings related expenses, if necessary, as the initial email did not include such provisions (see sample at Appendix 5D; Appendix 5E provides the equivalent information in English). No positive responses, and only one negative response were received.

Advice from experienced researchers suggested that further cold-emailing was likely to be of similar limited success, and that a personal introduction would be preferable, while not, by itself, compromising the research objectives. A different approach was therefore adopted in the next wave. Academic colleagues, family contacts, friends, and individuals who had assisted in pilot testing of the instrument were contacted to ask about physicians (both primary care and oncologists) who would fit the inclusion and exclusion criteria and would be willing to participate. A physician working in a medium-size hospital that employed no oncologists as permanent staff members, but had regular weekly meetings with external consultants to discuss cases, enabled me to present the research in outline to the oncologist present. As recruitment continued in parallel to interviews being carried out,

---

<sup>2</sup> Internetportal "Willkommen bei der Arztsuche der Kassenärztlichen Vereinigung Berlin" [Welcome to the online search of the society of sickness fund contracting physicians of Berlin], available at: <http://www.kvberlin.de/60arztsuche/index.html>, accessed 21 Februar 2012.

<sup>3</sup> Senatsverwaltung Gesundheit und Soziales, Berlin, Referat Gesundheitsberichterstattung, Epidemiologie, Gemeinsames Krebsregister, Sozialstatistisches Berichtswesen, Gesundheits- und Sozialinformationssysteme, Spezialberichte, [Special reports issues by the health department of the state government], available at: <http://www.berlin.de/sen/statistik/gessoz/gesundheit/spezial.html>, accessed 21 February 2012.



snowballing also identified further physicians who were recommended by those already interviewed. Contact was also made with the Director of the *Deutsche Krebsgesellschaft* (the German Cancer Society), the largest scientific society with a focus on oncology, and the Chairman of the *Berlin-based Verein der niedergelassenen Hämatologen*, a membership association of internal oncologists.

In total 22 physicians were approached in this way, directly or indirectly, and all were being supplied with the information as reproduced at Appendix 5D; of these, ten accepted. The selection criterion of gender was not adhered to stringently, as it was more important to have interviewees with patient populations that had the desired socio-economic profiles. However in the event, equal overall gender representation was achieved across both countries, with almost equal numbers in each of the countries, and professional specialisms. For primary care physicians in both countries it was possible to implement the above sketched distribution goal regarding low middle and high socio-economic status patient populations. However, the schema was adhered to less stringently in the case of oncologists, as most cancer care is provided in hospital settings with mixed patient populations.

For the Philadelphia counterparts, enquiries were made with colleagues conducting similar research on a regular basis. Regarding primary care physicians it was recommended to post a circular via an email mailing list of *Clinical Care Associates (CCA)*, the primary care network of the *University of Pennsylvania Health System*, consisting of community-based primary care physician practices with around 175 physicians at roughly 40 locations throughout Philadelphia, New Jersey, and Eastern Pennsylvania. The circular is at Appendix 5F. Four positive responses were received, of which one had to be disregarded, as the author had published on the topic in the peer reviewed literature, and it the goal was to interview normatively 'naïve' professionals. However, this person was able to recommend a suitable colleague; one remaining primary care physicians was identified through further snowballing; and the last via a departmental colleague. Oncologists were initially identified through a departmental colleague carrying out related research on physician's attitudes towards cancer screening. Thirteen contacts were identified, six were emailed, and four positive replies were received. Snowballing contacts led to three further invitations to potential interviewees, and one positive reply. Ultimately, no financial incentives were provided to any physicians in either Berlin or Philadelphia.

## **Appendix 5B: Notes on use of NVIVO as computer-assisted qualitative data analysis software**

NVIVO was chosen as it is one of the established software packages that I had used previously, and I had familiarised myself with it in a two-day training course as part of a research fellowship in the 2009–10. The training took place in November 2010 and was facilitated by *QRS International*, the company producing the programme.

Established qualitative researchers, rightly caution that the potential of CAQDAS must not be overestimated. Elaine Walsh summarised the concerns of some commentators by noting fears that “software may ‘guide’ researchers in a particular direction [and that it] could serve to distance the researcher from the data, encourage quantitative analysis of qualitative data, and create a homogeneity in methods across the social sciences” (Welsh 2002: 2). Even if functions such as auto-coding, in principle, enable automated analysis, clearly the device cannot and must not replace a deeper engagement with the material.

The principal advantages of NVIVO (and other similar programmes) for the type of research pursued here are that they can help develop, monitor and trim efficiently evolving code-lists, as hierarchy relationships can be made visible and rearranged easily. Equally, changes in codes are applied automatically to all coded sections, which can improve consistency. Comprehensiveness can be improved, as un-coded text can be made visible easily. Search functions can enable quicker *ad hoc* and in-depth interrogations than would otherwise be possible, and can, for example, help show differences in salient concepts between the German and American physicians (or between oncologist and primary care physicians) without conflating qualitative with quantitative research methods.

Electronic memos and annotations can help organise one’s thinking as the data are coded, and in cases where one should have lost oversight of the number of insights gained in the process, they can enable electronic searching and probably quicker retrieval of what are deemed important thoughts than a review of notes written in the margins of print outs, or on post-it notes. Electronic code lists also enable other researchers to review the assignment of individual codes the structure of the list as a whole. In short, NVIVO was used principally as a tool of organising the data, and has hopefully helped improve its rigor.<sup>4</sup>

---

<sup>4</sup> However, in the same way that *Microsoft Word* was used as a tool for organising this thesis, by themselves neither Word nor NVIVO (for better or worse) did the work of generating the substance of it.

### **Appendix 5C: Development of the code-list**

In the first (pre-coding cycle), handwritten notes from 14 interviews that had been completed at that point in time were reviewed (eight German and six American ones)<sup>5</sup>. During the interviews, I had marked points that were made particularly strongly, or were otherwise “summative, salient, essence-capturing, and/or evocative attribute[s]” (Saldana 2011: 3), alongside other notes. In reviewing these notes, comments that initially seemed of less relevance, but now represented repetitions, or part of patterns, also stood out more clearly.

A total of 71 codes were identified in this way, which included a few cases of closely overlapping codes. The pieces of paper were then clustered into groups of codes that shared some conceptual family-resemblance, with a view to ensuring that groups were exclusive, as far as possible, such that items within the groups would not be likely to feature in another group. The following 8 clusters emerged (numbers in brackets indicate the number of individual items, i.e. excluding semantic redundancy; for a full list, see Appendix 5I): Unintended and bad consequences (14); Alternatives and modifications (11); Exclamations (9); Drawing lines in policy (8); Why responsibility matters (4); Evaluation and evidence (3); Framing (2); Other (1).

In the second cycle, the 14 full transcripts were coded in NVIVO, beginning with two physicians in each category and country who provided particularly rich responses. Sections that raised issues that could be subsumed under the codes identified in the first cycle were linked accordingly, and new codes were added in the process. In the third cycle, the remaining 6 transcripts were coded in NVIVO. New codes that emerged were applied to the previously analysed transcripts, and the final code list is at Appendix 5J.

---

<sup>5</sup> Of these, there were five German primary care physicians, three German oncologists, three American primary care physicians and three American oncologists.

## **Appendix 5B: Physician recruitment: email to German primary care physicians (version for oncologist adapted accordingly)**

Sehr geehrter Herr/Frau Dr [X]

Vor einiger Zeit wandte ich mich mit der Anfrage unten an Sie und wuerde mich freuen von Ihnen zu hoeren, ob ein kurzes Gespraech eventuell moeglich sein sollte. Bitte lassen Sie mich auch wissen, wen hierfuer eine Aufwandsentschaedigung erforderlich sein sollte. Dies ist zwar eine rein-akademisch- wissenschaftliche Studie ohne Drittmittel, aber ich wuerde mich selbstverstaendlich an uebliche Braeuche halten.

Mir ist voellig klar, dass der betriebsame Arztalltag generell wenig Zeit fuer derartige Projekte laesst, aber ich hoffe dennoch, dass Sie die Studie vielleicht interessiert, oder dass ich Sie sonst zur Teilnahme gewinnen kann: Ihre Meinung zu diesem Thema waere fuer meine Arbeit ausserst relevant.

Mit freundlichen Gruessen,

Harald Schmidt

---

Sehr geehrter Herr/Frau Dr [X]

Ich schreibe um zu fragen, ob Sie eventuell bereit waeren, an einer ca 30 minuetigen Befragung im Rahmen eines Forschungsprojekts zur Eigenverantwortung teilzunehmen.

Um ein wenig Hintergrund zu meiner Person zu geben: Ich bin seit 2002 Assistant Director am Englischen Nuffield Council on Bioethics (dem Deutschen Ethikrat vergleichbar), und seit 2007 mit einer (Teilzeit)Dissertation innerhalb LSE Health an der London School of Economics and Political Science beschaeftigt. Bei dem Projekt geht es insgesamt um die Frage, wie man Eigenverantwortung fuer Gesundheit sinnvoll gestalten kann. Meine Studie vergleicht Ansaetze in den USA und Deutschland, und ich beschaeftigte mich mit dem Amerikanischen Teil im letzten Jahr, waehrend eines Fellowships an der Harvard School of Public Health. Etwas mehr zum Ansatz, ein kurzer Lebenslauf, und eine Liste relevanter Publikationen ist hier zu finden: <http://personal.lse.ac.uk/schmidth>

Ein Teil meines Dissertationsprojektes besteht in einer Befragung von Hausaerzten in Berlin, und ich fragte mich daher, ob ich Sie eventuell fuer die Studie gewinnen koennte?

Es geht um ca. 15 Fragen, von denen fuenf mit Einstellungen zu Eigenverantwortung allgemein zu tun haben, und die uebrigen mit Ansichten zu zwei speziellen (realen) Regelungen: (1) finanzielle Anreize fuer Gewichtskontrolle, wie zB Bonusprogramme, bei denen Versicherte Geldbetragee bekommen, wenn sie bestimmte BMI Ziele erreicht, und (2) die Chronikerrichtlinie, die hinsichtlich Darmkrebs vorsieht, dass man die niedrigste Schwelle an Zuzahlungen nur dann bekommt, wenn man zu einem Beratungsgespraech zu den Vor-und Nachteilen des screeing kommt, und, sollte man spaeter an Darmkrebs leiden, die Behandlung nicht ablehnt.

Insgesamt wird in keinem Fall Wissen (etwa ueber die Regelungen) abgefragt, sondern es geht lediglich um Meinungen und Einstellungen zu bestimmten praktischen und theoretischen Aspekten, die sich aus den Regelungen ergeben. Aus methodologischen Gruenden kann ich den Fragebogen leider nicht vor dem Interview zur Ansicht schicken, kann aber versichern, dass die Fragen keinerlei Vorbereitung erfordern.

Daten werden weder in der Dissertation, noch in moeglichen Publikationen den Befragten namentlich zugeschrieben, noch werden Namen ueberhaupt genannt, da alle Daten anonymisiert werden. Vom Zeitraum wuerde ich das Gespraech gerne Ende Januar oder Anfang Februar fuehren. Idealerweise im direkten persoenlichen Gespraech, aber wenn es einfacher sein sollte, telefonisch (oder ueber Skype) zu sprechen, so ginge dies selbstverstaendlich auch.

Ich wuerde mich freuen, von Ihnen zu hoeren, und stehe selbstverstaendlich fuer Rueckfragen jederzeit zur Verfuegung.

Mit freundlichen Gruessen,

Harald Schmidt

**Appendix 5E: Physician recruitment: email to US oncologists**

Dear Dr A

Your colleague AAA AAA a kindly provided me with your email address as she though that I could approach to you ask if you might be willing to participating in a brief one-on-one interview regarding your views on incentive programs to promote colorectal cancer screening.

For context, I have recently joined Penn's Center for Health Incentives and am currently wrapping up a PhD at the London School of Economics' Health Policy Dept on issues around personal responsibility for health. A brief outline and some relevant publications can be found here: [personal.lse.ac.uk/schmidt/](http://personal.lse.ac.uk/schmidt/) (this is slightly out of date, but should hopefully give an impression anyway).

One part of my thesis concerns semi-structured interviews with oncologists. Interviews last no more than 30 mins, should not require any kind of preparation, and concern general attitudes regarding the role of personal responsibility, and views on features of incentive programs that seek to promote health and uptake of colon cancer checkup.

The protocol has been determined to be exempt from IRB review and neither the thesis nor any publication relating to the interview data will attribute any quotes, or mention names of interviewees (or institutions).

If you might be willing to assist me I should note that timing is flexible, although I was hoping to schedule this in the next couple of weeks—the sooner the better. I would be happy to visit you at your practice, or to speak over the phone, whatever is preferred.

Thank you for considering my request. If you should have any questions please email, call me on the number below, or mobile: [xxx-xxx-xxxx].

Many thanks again and best wishes,

Harald Schmidt

**Appendix 5F: Physician recruitment: email to US primary care physicians, as edited and circulated by CCA staff (*Clinical Care Associates*, the primary care network of the *University of Pennsylvania Health System*)**

Dear CCA Physicians: We are writing to ask if you might be interested in participating in a one-on-one interview regarding your views on incentive programs related to weight loss and colorectal cancer screening. I am wrapping up a PhD at the London School of Economics' Health Policy Dept on issues around personal responsibility for health.

One part of the thesis concerns semi-structured interviews with primary care physicians. Interviews last around 30 mins, should not require any kind of preparation, and concern attitudes regarding the role of personal responsibility and views on incentive programs for colon cancer checkups. The protocol has been determined to be exempt from IRB review and neither the thesis nor any publication relating to the interview data will attribute quotes or mention names of interviewees (or institutions).

If you might be willing to assist me by participating in this important project, we are only asking for 30 minutes of your time and that would be most helpful. Timing is flexible but we were hoping to schedule this in the next couple of weeks—the sooner the better. I would be happy to visit you at your practice, or to speak over the phone, whatever is preferred.

Thank you for considering participation. If you should have any questions please email, call me on the number below, or mobile: [xxx-xxx-xxxx].

Many thanks again and best wishes,

Harald Schmidt

PS

Should you be interested, a bit more detail and a list of relevant publications can be found here: <http://personal.lse.ac.uk/schmidth>  
(last updated in late 2010, but should hopefully still give an idea).

### Appendix 5G: Physician interview guide (English version)

DOCs	[Intro: HS, Upenn, HSPH, NCOB. Ok to record? Anonymity. 3 part questions, general, obesity, colon cancer. 25 min. Key is: reasons.]
1	<p>Overall, when it comes to our own health, how important are personal choices in relation to other factors such as environmental conditions (clean air and water), income, education, access to the health care system or inherited genetic factors? Please choose one of the following options (and expand on why, if you wish):</p> <ol style="list-style-type: none"> <li>1. Personal choices are really the most important factor that determines health: we are largely responsible for how healthy we are.</li> <li>2. The role of personal choices is an important factor, and we always bear some responsibility for how healthy we are. But sometimes other factors can be more important: it's therefore not possible to say which factor is the most important one.</li> <li>3. The importance of personal choices is overstated: many of the factors above such as the environment, how much money we have, and what genes we inherited influence quite strongly what kind of choice we can make, and they are practically always far more important.</li> <li>4. Don't know/none of the above.</li> </ol>
2	<p>In auto insurance, risky drivers have higher premiums. Some say that the same model should be used in health insurance. They argue that, for example, obese people should pay higher premiums than people with normal weight. What do you think about this proposal?</p>
3	<p>Consider the following real policy. A large employer wants to create incentives for people to be healthy. Employees who show that over a year they are not overweight (in medical terms), would get a cash reward of \$150. Participation is voluntary, that is, no one is forced to be weighed. Do you think the health plan should use the policy?</p>
4	<p>Two people are trying to describe the key feature of the policy just described. Which description do you find most adequate?</p> <ol style="list-style-type: none"> <li>1. "This policy works mainly like a 'carrot': it rewards those with normal weight"</li> <li>2. "This policy works mainly like a 'stick': it penalizes those who are overweight"</li> </ol>
5a	<p>A question about how high incentives should be.</p> <p>An employer's health plan wants to introduce an incentive program like the one above, where employees who show that over a year they are not overweight (in medical terms), would get a cash reward of a certain amount.</p> <p>The plan has to decide what the maximum level of incentives should be, for employees who are on average national income. Please select of the following options the one that seems most fair to you.</p> <ul style="list-style-type: none"> <li>- I am against such programs so no incentives should be given</li> <li>- Maximum \$100</li> <li>- Maximum \$250</li> <li>- Maximum \$500</li> <li>- Maximum \$1000</li> <li>- Maximum \$1500</li> <li>- Maximum \$2500</li> <li>- There should be no limit: whatever the health plans think is appropriate</li> </ul>

5b	<p>Another health plan also wants to offer incentives to people whose weight is normal. However, this plan wants to use a different approach. The employers' idea is this: they increase insurance contributions for all plan members by a certain amount. This amount is then offered to employees as an incentive. That means, employees whose weight is normal get this amount reimbursed and avoid the higher premiums.</p> <p>So the question is, by how much should insurers be allowed to increase insurance contributions for employees who are on average income? Again, please select from the following options the one that seems most fair to you:</p> <ul style="list-style-type: none"> <li>- I am against such programs so no incentives should be given</li> <li>- Maximum \$100</li> <li>- Maximum \$250</li> <li>- Maximum \$500</li> <li>- Maximum \$1000</li> <li>- Maximum \$1500</li> <li>- Maximum \$2500</li> <li>- There should be no limit: whatever the health plans think is appropriate</li> </ul>
6a	<p>Assume you are seeing a patient who is enrolled in a plan as described above, where you can get \$150 if you are either not overweight, or succeed in not being overweight (in medical terms), over a period of one year.</p> <p>A doctor needs to confirm the patients weight to the health plan.</p> <p>The patient who has come to see you has lost 10 pounds over a year, but on the day he is weighed to record his weight for the health plan, he is 2 pounds over the specified weight target. This is the last day of the qualifying period, and the patient cannot come back another day. Only you and the patient are in the room, when the patient asks you to state that he has in fact achieved the target.</p> <p>Out of ten doctors, how many do you think would comply with the patients' request to write down 2 pounds less?</p>
6b	<p>Let's suppose there is a doctor who accepts the patient's requests to record that he has met the target.</p> <p>How would most of your peers judge the behavior of their colleague who puts the patient's interest above complying with the policy's rules?</p>
7	<p>Poorer people are often less healthy and more overweight than those with more money. Think again about the policy above where you can get \$150 if you are not overweight.</p> <p>Suppose that research shows the following: although the incentive program is offered to all employees of a company, it turns out that almost twice as many of the high-income earners benefit from the program than the low-income earners.</p> <p>Does it concern you in any way that the poor are less likely to benefit?</p>
8	<p>Now some questions about another real example of an incentive policy. My apologies, but the policy is a bit complicated. Please let me know if there is something that should not be clear.</p> <p>A health plan in another country wants to improve colon cancer care. The plan offers an incentive program that works as follows: Chronically ill on average income never have to pay more than \$550* per year in co-payments to all health care (co-payments are generally low in this country). However, for colon cancer patients on average income this threshold is increased to \$1100 per year, unless they comply with two conditions. If they comply, they too, have the lower co-payment-cap of \$550.</p> <p>The two conditions are as follows:</p> <ol style="list-style-type: none"> <li>1. Plan members who are older than 50 years need to attend a counseling session on the advantages and disadvantages of colon cancer screening.</li> <li>2. If they should require treatment for colon cancer, they should never refuse medically recommended interventions.</li> </ol> <p>So, the level of co-payments is doubled for colon cancer patients, unless they comply with these two conditions.</p> <p>The health plan says that the policy is justified because colon cancer patients have special personal responsibilities to contribute to cost control.</p> <p>Is there anything that you like or dislike about this policy?</p>



9	<p>Two people are trying to describe the key feature of the policy. Which description do you find most adequate?</p> <p>1. "This policy works mainly like a 'carrot', because it rewards responsible behavior"</p> <p>2. "This policy works mainly like a 'stick', because colon cancer patients who do not satisfy the conditions are faced with double the co-payment rate than other chronically ill"</p>
10	<p>Another health plan thinks about implementing the above policy, but has received some letters from colon cancer patient groups who find it unfair. The patient groups make two alternative proposals.</p> <p>The health plan does not want to be accused of being unfair. Executives therefore consider the policy alongside the two alternative proposals, that are sketched out below. Which option would you find most fair?</p> <p>Policy 1: Higher co-payment cap for colon cancer patients As above, all chronically ill have a maximum of \$550 in co-payments, except colon cancer patients who have higher cap of \$1100 Colon cancer patients can only reduce their co-payment cap to \$550 % if they comply with the two conditions (i.e., they come to counseling and don't refuse treatment)</p> <p>Policy 2: Lower co-payment cap for colon cancer patients On this policy, all chronically ill, including colon cancer patients, have a maximum of \$550 in co-payments. However, colon cancer patients have the option to pay less. If they comply with the two conditions (i.e., they come to counseling and don't refuse treatment) they have a cap of \$275.</p> <p>Policy 3: No difference between colon cancer patients and other chronically ill On this policy, all chronically ill, including colon cancer patients, have the same cap of \$550 in co-payments. But there is no difference between colon cancer patients and other chronically ill: all are treated the same.</p>
11	<p>Suppose a health plan wants plan members to do more than just attend a counseling session. Instead, the plan intends to offer the lower co-payment cap only for those who have undergone recommended colon cancer screening. So, in order to reduce their copays they need to come to the screening.</p> <p>However, the proposal has some opponents. What do you think about the following statement?</p> <p>"There should be a right not to know about whether one has a serious disease such as cancer, or will suffer from it in the future. This means that if someone does not want to take a test, there should be no financial disadvantage or penalty."</p>
12	<p>Another health plan defends the alternative policy introduced above as the first option, which, as you know, does not require undergoing screening, but attending counseling on the advantages and disadvantages of screenings in order to access the lower co-payment threshold.</p> <p>The plan therefore accepts that there is a right not to know about whether one will suffer from serious diseases such as cancer now or in the future. But a spokesperson also emphasizes the following point:</p> <p>"We do not accept that there is a right not to think about whether one should take part in a colon cancer screening. This is why a financial disadvantage is justified for those who refuse to attend a counseling session in which they can think about the advantages and disadvantages of colon cancer screens."</p> <p>What do you think about this statement?</p>

13a	<p>Suppose a person has colon cancer. The physician recommends post-operative treatment that is accepted as the most effective by the professional bodies. However, the patient thinks that a natural remedy would be better, that has been recommended by a friend who read a lot about alternative medicine. She therefore plans to refuse the treatment that the physician recommends.</p> <p>The patient knows that refusing treatment means that his overall co-payments for all health care cost are \$1100 instead of \$550. She also knows that doctors are required to inform health plans about patients who refuse treatment.</p> <p>Only you and the patient are in the room, when the patient asks you not to tell the health plan about her refusal.</p> <p>Out of ten doctors, how many do you think would comply with the patients' request not to notify the health plan?</p>
13b	<p>Let's suppose that there is a doctor who accepts the patient's requests not to tell the health plan about her refusal.</p> <p>How would most of your peers judge the behavior of their colleague who puts the patient's interest above complying with the policy's rules?</p>
14a	<p>Think again about the plan above (you get \$150 if you are not overweight). Below, five different people are described, who are covered by a plan that offers such a scheme. All of them have average incomes and are around 45 years old. Because of their different circumstances and preferences, they differ in how they are affected by the policy. Please give us your ratings of the fairness of this policy for each of the below individuals:</p> <p>Person A is not overweight. He/she is health-conscious by habit, always eats healthily, exercises regularly, has had a normal body weight for years.</p> <p>In line with the policy, Person A gets \$150. Is this fair?</p>
14b	<p>Person B is not overweight. He/she is not health-conscious by habit, eats mostly unhealthily, never exercises, but nonetheless has had normal body weight for years.</p> <p>In line with the policy, Person B gets \$150. Is this fair?</p>
14c	<p>Person C: likes to eat, dislikes exercise, and is overweight. However, person C would like to be slimmer and views the program as an opportunity to go on a diet and exercise twice a week. As a result, person C loses weight and is successful in maintaining normal body weight.</p> <p>In line with the policy, Person C gets \$150. Is this fair?</p>
14d	<p>Person D: likes to eat, dislikes exercise, and is overweight. However, person D would like to be slimmer and takes the program as an opportunity to go on a diet and exercise twice a week. Person D ends up trying several diets, but can't stick to one. Some weeks there is time for the exercise programs, but other times there isn't. As a result D loses some weight, but is not successful in losing weight and keeping it off and remains overweight.</p> <p>In line with the policy, Person D does not get \$150. Is this fair?</p>
14e	<p>Person E: likes to eat, dislikes exercise, and is overweight. Person E eats a lot, and hardly ever exercises, but does not mind the extra pounds.</p> <p>In line with the policy, Person E does not get \$150. Is this fair?</p>
15	<p>Is there anything else that you thought of, while we spoke, that I should be aware of?</p>

### Appendix 5F: Physician interview guide (German version)

DOCs	[Intro: HS, Upenn, HSPH, NCOB. Ok to record? Anonymity. 3 part questions, general, obesity, colon cancer. 25 min. Key is: reasons.]
1	<p>Insgesamt, wenn es um unsere Gesundheit geht, wie wichtig ist persönliches Verhalten im Verhältnis zu anderen Einflüssen wie zB Umweltfaktoren (saubere Luft und Wasser), Einkommen, Bildung, Zugang zum Gesundheitssystem, oder vererbten genetische Faktoren? Wählen Sie bitte eine der folgenden Optionen:</p> <p>1 Persönliches Verhalten ist der allerwichtigste Faktor, der unsere Gesundheit bestimmt: die Hauptverantwortung für Gesundheit liegt daher bei uns selber.</p> <p>2. Persönliches Verhalten spielt eine wichtige Rolle, und wir tragen immer ein gewisses Maß an Eigenverantwortung für unsere Gesundheit. Aber in manchen Fällen können andere Faktoren mehr Einfluss haben: man kann daher nicht sagen, welches der wichtigste Faktor ist.</p> <p>3. Die Rolle des persönlichen Verhaltens ist überbewertet: Viele der Faktoren die oben genannt wurden bestimmen überhaupt erst, welche Verhaltens-Möglichkeiten wir haben: Dinge wie die Umwelt, der finanzielle Status und die Gene sind eigentlich immer wichtiger.</p> <p>4. Weiss nicht/keine der Optionen scheint richtig.</p>
2	<p>In der Autoversicherung haben risikofreudigere Fahrer höhere Beiträge. Manche sagen, dass dieselbe Regelung im Gesundheitssystem angewendet werden sollte. Sie fordern zum Beispiel, dass stark Übergewichtige höhere Versicherungsbeiträge zahlen sollten als Versicherte mit normalem Gewicht. Was halten Sie von diesem Vorschlag?</p>
3	<p>International gibt es verschiedene Versuche, Gesundheitsverhalten von Versicherten zu fördern. Eine gesetzliche Krankenkasse in einem anderen Land möchte Versicherten Anreize bieten, gesund zu sein. Wenn Versicherte sich wiegen lassen und sie über ein Jahr hinweg medizinisch gesehen nicht übergewichtig sind, bekommen sie einen Bonus von €100. Die Teilnahme ist freiwillig, das heißt, niemand muss sich wiegen lassen. Denken Sie, das die Krankenkasse die Regelung umsetzen sollte?</p>
4	<p>Zwei Personen versuchen, das wesentliche Merkmal der eben beschriebenen Regelung zu beschreiben. Welcher Formulierung ist Ihrer Meinung nach die zutreffendste Beschreibung?</p> <p>1. „Diese Regelung ist eher ein Anreizprogramm: diejenigen deren Gewicht normal ist, werden belohnt“</p> <p>2. „Diese Regelung ist eher ein Bestrafungsprogramm: diejenigen, die übergewichtig sind, können es nicht benutzen und sind benachteiligt“</p>

5a	<p>Eine Frage dazu, wie hoch finanzielle Anreize sein sollten.</p> <p>Eine Krankenkasse möchte ein Anreizprogramm wie das oben anbieten, bei dem Versicherte sich wiegen lassen können und einen Bonus von €100 bekommen, wenn sie über ein Jahr hinweg medizinisch gesehen nicht übergewichtig sind.</p> <p>Die Krankenkasse muss entscheiden, wie hoch der Bonus für Versicherte mit Bundesweitem Durchschnittseinkommen maximal sein soll. Bitte wählen aus den folgenden Optionen die aus, die Ihnen am gerechtesten erscheint:</p> <ul style="list-style-type: none"> <li>- Ich bin gegen Bonus oder Anreizprogramme, und sie sollten daher abgeschafft werden</li> <li>- Bonusbeträge sollten nie höher als € 70 sein</li> <li>- Bonusbeträge sollten nie höher als € 175 sein</li> <li>- Bonusbeträge sollten nie höher als € 350 sein</li> <li>- Bonusbeträge sollten nie höher als € 700 sein</li> <li>- Bonusbeträge sollten nie höher als € 1750 sein</li> <li>- Es sollte kein Limit geben: was immer die Krankenkassen für richtig erachten</li> <li>- Weiß nicht</li> </ul>
5b	<p>Eine andere Krankenkasse will ebenso ein Bonusprogramm einführen, bei dem Mitglieder die normales Gewicht haben, einen Geldbetrag bekommen. Allerdings möchte sie einen anderen Ansatz wählen.</p> <p>Was die Krankenkasse vorhat, ist folgendes: für alle Versicherten wird der Beitrag um einen bestimmten Betrag angehoben. Dieser Betrag wird dann als Anreiz angeboten, das heißt, Versicherte, deren Gewicht normal ist, bekommen den Betrag, und müssen keinen höheren Versicherungsbeitrag bezahlen. (Siehe Fig 3, Plan B)</p> <p>Die Frage ist also, um wieviel die Krankenkasse die Beiträge für Versicherte mit Durchschnittseinkommen erhöhen dürfen sollte, um Anreize für normales Gewicht zu schaffen.</p> <p>Bitte wählen Sie aus den folgenden Optionen wieder die aus, die Ihnen am gerechtesten erscheint:</p> <ul style="list-style-type: none"> <li>- Ich bin gegen Bonus oder Anreizprogramme, und sie sollten daher abgeschafft werden</li> <li>- Beiträge sollten nie um mehr als € 70 angehoben werden</li> <li>- Beiträge sollten nie um mehr als € 175 angehoben werden</li> <li>- Beiträge sollten nie um mehr als € 350 angehoben werden</li> <li>- Beiträge sollten nie um mehr als € 700 angehoben werden</li> <li>- Beiträge sollten nie um mehr als € 1750 angehoben werden</li> <li>- Es sollte kein Limit geben: was immer die Krankenkassen für richtig erachten</li> <li>- Weiss nicht</li> </ul>

6a	<p>Denken Sie wieder an die oben vorgestellte Regelung, unter der man €100 bekommt, wenn man entweder nicht übergewichtig ist, oder es schafft innerhalb eines Jahrs Normalgewicht (im medizinischen Sinn) zu erreichen. Ein Arzt muss feststellen, ob Versicherte das Ziel erreicht haben, und der Krankenkasse eine Meldung ueber das gemessene Gewicht machen.</p> <p>In der Sprechstunde sehen Sie einen Patienten der ueber ein Jahr hinweg ungefaehr 5 Kilo abgenommen hat. Am Tag als er gewogen wird ist er jedoch ein Kilo zu schwer. Er kann nicht spaeter nocheinmal zurueckkommen, da dies bereits der letzte Tag des Berechnungszeitraums ist. Nur Sie und der Patient sind im Raum, als der Patient sie bittet ob sie nicht notieren koennten, dass er den Zielwert erreicht hat.</p> <p>In ihrer Einschaeztung, wieviele von 10 Aerzten wuerden der Bitte des Patienten nachgeben?</p>
6b	<p>Nehmen wir an, ein Arzt willigt ein, ein Kilo weniger zu notieren. Wie wuerden die meisten Ihrer Kollegen die Handlung des Arztes bewerten der den Interessen des Patienten mehr folgt, als der Absicht der Regelung?</p>
7	<p>Aermere Menschen sind haeufiger von Uebergewicht betroffen als Wohlhabendere. Denken Sie bitte wieder an die Regelung in der Versicherte € 100 bekommen, wenn sie nicht übergewichtig sind.</p> <p>Nehmen Sie nun an, die Forschung zeigt Folgendes: obwohl die Regelung allen Einkommensgruppen gleichermassen angeboten wird nutzen fast doppelt so viel der hohen Einkommensgruppen die Regelung und bekommen €100, als diejenigen, die in den nidrigsten Einkommensgruppen sind.</p> <p>Was halten Sie davon, dass die Aermsten am wenigsten von der Regelung profitieren?</p>
9	<p>Nun einige Fragen zu einer anderen Anreiz-Regelung.</p> <p>Die Chronikerrichtlinie hat neue Anreize für (u.a.) Darmkrebs Praevention geschaffen. Die Regelung funktioniert so, dass für chronisch Kranke mit Durchschnittseinkommen Zuzahlungen zu allen Gesundheitsleistungen auf einen Schwellenwert von € 400 begrenzt werden. Allerdings wird diese Schwelle für Darmkrebspatienten mit Durchschnittseinkommen allerdings auf € 800 angehoben, ausser sie erfuellen zwei bestimmte Bedingungen. Nur wenn sie diese Bedingungen erfuellen, fallen sie unter die Hoechstgrenze von € 400, die alle anderen chronisch Kranken auch haben.</p> <p>Die beiden Bedingungen sind die folgenden:</p> <ol style="list-style-type: none"> <li>1. Versicherte die aelter als 50 Jahre sind, muessen zu einem Beratungsgespraech kommen, in dem Sie ueber die Vor- und Nachteil des Darmkrebsscreening lernen.</li> <li>2. Sollten Versicherte eine Darmkrebs-Therapie benoetigen, duerfen Sie medizinisch angeratene Behandlung nie ablehnen.</li> </ol>

	<p>Die Zuzahlungsgrenze für Darmkrebspatienten ist also gegenüber anderen chronisch kranken verdoppelt, es sei denn, dass Sie diese beiden Bedingungen erfüllen.</p> <p>Die Krankenkasse sagt, dass die Regelung angemessen ist, da Darmkrebspatienten besondere Eigenverantwortung haben, zur Kostenkontrolle beizutragen.</p> <p>Gefällt Ihnen an dieser Regelung etwas, oder gibt es etwas, das Ihnen nicht gefällt?</p>
9	<p>Zwei Personen versuchen, das wesentliche Merkmal der Regelung zu beschreiben. Welche Formulierung ist Ihrer Meinung nach die zutreffendste Beschreibung?</p> <p>1. „Diese Regelung ist eher ein Anreizprogramm: diejenigen, die Eigenverantwortung zeigen, werden belohnt“</p> <p>2. „Diese Regelung ist eher ein Bestrafungsprogramm: diejenigen, die die Bedingungen nicht erfüllen haben die doppelte Zuzahlungsgrenze und sind benachteiligt“</p>
10	<p>Eine andere gesetzliche Krankenkasse erwägt, dieselbe Regelung anzuwenden. Sie hat jedoch Briefe von Interessengruppen von Darmkrebspatienten bekommen, die die Regelung ungerecht finden. Die Gruppen machen zwei Vorschläge für alternative Regelungen.</p> <p>Die Krankenkasse möchte nicht beschuldigt werden, ungerecht zu sein. Leitende Angestellte erwägen daher die Regelung der anderen Krankenkasse im Vergleich mit den beiden alternativen Vorschlägen, die ich gleich skizzieren werde. Bitte sagen sie mir, welche Option Sie am gerechtesten finden.</p> <p><b>Regelung 1: Höhere Grenze für Darmkrebspatienten</b>–Wie oben bereits dargestellt:</p> <ul style="list-style-type: none"> <li>• alle chronisch Kranken ausser Darmkrebspatienten haben eine Zuzahlungsgrenze von € 400*, ausser Darmkrebspatienten, die eine höhere Schwelle von € 800 haben</li> <li>• Darmkrebspatienten koennen die € 800 nur dann auf € 400 reduzieren, wenn sie die beiden Bedingungen erfüllen (Beratungsgespraech, keine Behandlungsablehnung).</li> </ul> <p><b>Regelung 2: Niedrigere Grenze für Darmkrebspatienten</b></p> <ul style="list-style-type: none"> <li>• alle chronisch Kranken, einschliesslich aller Darmkrebspatienten, haben eine Zuzahlungsgrenze von € 400.</li> <li>• Darmkebspatienten haben jedoch die Moeglichkeit weniger als andere chronische Kranke zu zahlen.</li> <li>• Wenn sie die beiden Bedingungen erfüllen (Beratungsgespraech, keine Behandlungsablehnung) werden ihre Zuzahlungen auf die Haelfte begrenzt (€ 200)</li> </ul> <p><b>Regelung 3. Keine Unterschiede zwischen Darmkrebspatienten und anderen chronisch kranken</b></p> <ul style="list-style-type: none"> <li>• Alle chronisch Kranken, einschliesslich aller Darmkrebspatienten,</li> </ul>

	haben dieselbe Zuzahlungsgrenze von € 400. Es gibt aber keine Sonderregelungen: alle zahlen das gleiche.
11	<p>Stellen Sie sich vor, dass eine gesetzliche Krankenkasse vorhat, mehr von Versicherten zu verlangen, als bloss zu einem Beratungsgespraech zu kommen. Die Kasse moechte, die niedrigere Zuzahlungsgrenze nur dann vergeben, wenn Versicherte auch das Darmkrebs scening mitmachen. Das heisst, um die Zuzahlungen zu reduzieren muessen Versicherte auf Darmkrebs getestet werden.</p> <p>Das Vorhaben hat jedoch Kritiker. Was halten sie von der folgenden Meinung:</p> <p>„Ich bin gegen diese Regelung. Es sollte ein Recht auf Nichtwissen geben in Bezug darauf, ob man schwere Krankheiten wie zum Beispiel Darmkrebs hat, oder in der Zukunft bekommen wird. Das heisst, wenn jemand ein screening oder einen Test ablehnt, dann sollte dies nicht bestraft werden, und es sollte daraus kein finanzieller Nachteil entstehen.“</p>
12	<p>Eine Vertreterin einer anderen gesetzliche Krankenkasse verteidigt die Regelung die oben zuert eingefuehrt wurde, derzufolge die niedrigere Zuzahlung nicht an eine Test oder Screening Teilnahme gebunden ist, sondern daran, dass man an einem Beratungsgespraech teilnimmt. Die Krankenkasse akzeptiert also, dass es ein Recht auf Nichtwissen gibt, in Bezug auf die Frage, ob man jetzt oder spaeter schwere Krankheiten wie Krebs hat. Die Sprecherin beton jedoch auch folgenden Punkt:</p> <p>„Wir akzeptieren nicht, dass es ein Recht auf nicht-Nachdenken gibt. Daher ist es unserer Auffassung nach gerechtfertigt, wenn diejenigen, die es ablehnen, zum Beratungsgespraech ueber die Vor-und Nachteile des Darmkrebs-Screenings zu kommen, einen finanzieller Nachteil haben.“</p> <p>Was halten Sie von dieser Aussage?</p>
13a	<p>Stellen Sie sich eine Patientin vor, die Darmkrebs hat. Ihr Arzt empfiehlt ihr die Behandlungsmethode die von Experten als die effektivste angesehen wird. Die Patientin hoert jedoch von einem Freund, der viel ueber Alternativmedizin gelesen hat, das eine natuerliche Behandlung besser als Schulmedizin sei. Die Patientin hat daher vor, die vom Arzt empfohlene Behandlung abzulehnen.</p> <p>Sie weiss, dass eine Ablehnung bedeutet, dass sie € 800 statt € 400 ihres Bruttojahreseinkommens an Zuzahlungen zu allen Gesundheitsleistungen hat. Und sie weiss auch, dass Aerzte verpflichtet sind, die Ablehnung von Behandlungen der gesetzlichen Krankenkasse zu melden.</p> <p>Nur die Patientin und Sie sind im Raum, als die Patientin Sie bittet, der Krankenkasse keine Mitteilung zu machen.</p> <p>In ihrer Einschaetzung, wieviele von 10 Aerzten wuerden der Bitte des Patienten nachgeben?</p>
13b	Nehmen wir an, ein Arzt hat Verstaendnis fuer die Situation des Patienten und willigt ein, der Krankenkasse keine Mitteilung ueber die Ablehnung der Behandlung zu machen.

	Wie wuerden die meisten Ihrer Kollegen die Handlung des Arztes bewerten (der sich zugunsten seines Patienten ueber die Regelung hinwegsetzt)?
14a	Denken Sie bitte noch einmal an die Regelung oben, bei der Versicherte sich wiegen lassen koennen und einen Bonus von €100 bekommen, wenn sie ueber ein Jahr hinweg medizinisch gesehen nicht uebergewichtig sind. Ich beschreibe gleich funf verschiedene Personen, die bei einer gesetzliche Krankenkasse versichert sind, welche eine solche Regelung anbietet. Alle Personen haben durchschnittliches Einkommen und sind ca 45 Jahre alt. Aufgrund Ihrer Umstaende und Praeferenzen hat die Regelung jedoch unterschiedlichen Stellenwert für sie. Bitte sagen Sie mir, wie gerecht sie die Regelung für jede dieser Personen finden.  Person A: ist nicht uebergewichtig; aus Gewohnheit Gesundheitsbewusst, ernaeht sich immer gesund, treibt regelmaessig Sport, und hat seit Jahren Normalgewicht. In Einklang mit der Regelung bekommt sie € 100. Finden Sie das gerecht?
14b	Person B: ist nicht uebergewichtig; kuemmert sich aus Gewohnheit nicht um die eigene Gesundheit, ernaeht sich meistens sehr ungesund, treibt nie Sport, hat aber dennoch seit Jahren Normalgewicht. In Einklang mit der Regelung bekommt sie € 100. Finden Sie das gerecht?
14c	Person C: isst sehr gerne, macht ungerne Sport, und ist uebergewichtig. Sie wuerde allerdings gerne schlanker sein. Sie nimmt das Progam als Anlass eine Diaet zu machen, und zweimal woeentlich zum Sport gehen. In der Folge nimmt sie ab und schafft es, ihr Normalgewicht auf Dauer zu halten. In Einklang mit der Regelung bekommt sie € 100. Finden Sie das gerecht?
14d	Person D: isst sehr gerne, macht ungerne Sport, und ist uebergewichtig. Sie wuerde allerdings gerne schlanker sein. Sie nimmt das Programm als Anlass, verschiedene Diaetenzu probieren, kann sich aber nicht daran halten. Versucht, zweimal woeentlich zum Sport zu gehen, kommt aber oft nicht dazu. Sie nimmt manchmal ab, ist aber nie auf Dauer erfolgreich, und bleibt uebergewichtig. In Einklang mit der Regelung bekommt sie keine € 100. Finden Sie das gerecht?
14e	Person E: isst sehr gerne, macht ungerne und selten Sport. Sie ist seit Jahren uebergewichtig, was ihr aber egal ist. In Einklang mit der Regelung bekommt sie keine € 100. Finden Sie das gerecht?
15	Gibt es sonst noch etwas, dass Ihnen während des Gesprächs doch den Kopf gegangen ist, und bei dem Sie dachten, das Sie mich noch darauf Aufmerksam machen sollten?



## **Appendix 5I: Initial code list**

Obtained from reviewing handwritten notes of salient points made during interviews. Individual codes were noted on pieces of paper, and then clustered as below:

### Unintended and bad consequences

Discrimination, discriminatory;  
 double whammy, they are punished enough;  
 no duty to think politics;  
 don't reward genetics;  
 rich people little money;  
 fraud, malpractice, mustn't lie, confidentiality;  
 you compromise evaluation data if you cheat;  
 counterproductive;  
 unnecessary;  
 too much responsibility: hypochondriacs;  
 mustn't force aggressive therapy;  
 patients go elsewhere if docs don't give in  
 doctor hopping: more specialist visits if copay waived  
 solidarity

### Alternatives and modifications

Broader focus needed, need to look at broader health system;  
 incentive envy, incentives for docs, where do docs fit in;  
 weight alone not such a determinant, don't measure just one, not just weight, don't  
 use single factor only;  
 do other things: gym;  
 don't use money;  
 need support (weight loss);  
 the effort counts (weight loss);  
 need to start earlier, too far in the future (cancer), early prevention needed;  
 reward what you do not what you weigh;  
 targeting;  
 do as token signal

### Exclamations

Menschen wille ist ein himmel reich;  
 makes my stomach turn (confidentiality);  
 that's difficult;  
 that's life;  
 help people, don't penalise them;  
 that's the way it is (that you lose out);  
 the beginning of the end;  
 means justify ends;  
 it's cheapest to die early;

### Drawing lines in policy

not fair re effort, but that's policy;  
 where do we draw lines; practically speaking (need to draw lines);

don't make patients pay all (incentive amount);  
need some play (wait);  
different league (colon v obesity cheating);  
watering can: too general;  
no sick pay if no cooperation, take holiday if ill because of sport;  
what about other risks, sport: skiing;

#### Why responsibility matters

Responsibility you can influence: rest not / less;  
might work for some;  
no therapy without engagement;  
solidarity

#### Evaluation and evidence

Evaluations, evidence;  
what's best evidence (cancer);  
art of medicine not science;

#### Framing

Carrot/stick is in eye of beholder: depends on what kind of patient you are  
gain sharing, not cost shifting

#### Other

Country of victims

**Appendix 5J: Final code list**

Type	Name	Memo Link	References
<b>A</b>	<b>Alternatives and modifications</b>		<b>46</b>
1		do as a token or signal	1
2		do as windfall not part of paycheck	2
3		do other things, such as...	3
4		don't link to insurance cost (but use money)	1
5		don't use money	2
6		incentive envy, what about docs	4
7		need look beyond HCS, need wider picture	5
8		need support also	3
9		need to start earlier, education	4
10		reward effort or improvement not weight	4
11		talk to patients	3
12		targeting	3
13		the effort counts	6
14		weight alone not enough	5
<b>B</b>	<b>Drawing lines in policy</b>		<b>36</b>
1		cancer and obesity in different league	1
2		don't make patients pay all, but some	2
3		economically this makes sense	3
4		even if no behavior change= ok as sends message	2
5		health should be focus, not money	1
6		need some play, tolerance re thresholds	3
7		no sick-pay if no cooperation	1
8		not fair re effort or SES but that's policy	3
9		relate amount to HC cost	2
10		this would be unique	3
11		use leave days if PR accident	2
12		watering can principle is too broad	2
13		we need to draw lines somewhere	6
14		what about other risks eg sport	6
<b>C</b>	<b>Evaluation and evidence</b>		<b>20</b>
1		ART of medicine not science	1
2		CC rarely caused by peoples action	3
3		cost-benefits data supports PRH	3
4		incentives don't work or only for some	4
5		Pts and docs don't always know best evidence	3
6		requires evaluation	4

		strongest evidence is needed to coerce	3
<b>D</b>	<b>Exclamations and strong reactions</b>		<b>21</b>
1		CC policy is offensive	1
2		crazy that rich people come for small incentives	1
3		ends justify means	1
4		help people don't penalize	3
5		if you money people do things	1
6		its cheapest to die early	3
7		makes my stomach turn	1
8		Menschen Wille ist ein Himmelreich	1
9		really unacceptable - stick obesity	1
10		that's difficult	3
11		that's life	2
12		that's problematic (so better don't do this)	1
13		the beginning of the end	1
		total nonsens _duty to think	1
<b>E</b>	<b>Framing</b>		<b>8</b>
1		gain sharing not cost-shifting	1
2		it's in the eye of beholder	2
3		use carrots not sticks	4
4		what framing is is perfectly clear	1
<b>F</b>	<b>Inconsistencies or not understood</b>		<b>2</b>
1		against right not to know but other against double whammy	1
2		pro right not to know but against penalizing CC generally	1
<b>G</b>	<b>Unintended and bad consequences</b>		<b>40</b>
1		can be barrier to care	2
2		can lead to too much PRH	3
3		cheating compromises evaluation-data	1
4		conflict w free will autonomy	2
5		could lead to higher cost in long run	1
6		creates inequality and tensions in society	1
7		discrimination	2
8		doctor hopping_ more specialist visits if copay waived	2

9		don't penalize innocent	1
10		don't reward genetics	2
11		double whammy, punished enough	3
12		fraud, malpractice, lying	4
13		harms confidentiality	1
14		is counterproductive	4
15		must respect existentialist situation	1
16		mustn't force aggressive therapy	1
17		no duty to think politics	2
18		patients go elsewhere if docs don't give in	2
19		solidarity	2
20		unnecessary, ineffective, wont make a difference	3
<b>H</b>	<b>why docs might lie</b>		<b>5</b>
1		know patient and family well	1
2		sympathy with refusing rough therapy	4
<b>I</b>	<b>why docs shouldn't lie</b>		<b>7</b>
1		institution of truth telling	2
2		should implement legitimate rules even if unfair or imperfect	1
3		to prevent patients health becoming worse	1
		you can be found out and penalized	3
<b>J</b>	<b>Q6b docs view on liars</b>		<b>4</b>
1	scenario is unrealistic		4
<b>K</b>	<b>Why responsibility matters</b>		<b>11</b>
1		behavior you docs can influence, environment not	2
2		health behavior has real clinical benefit	2
3		no therapy without engagement	1
4		solidarity	3
		some policies work for some people	3
<b>K</b>	<b>Other</b>		<b>2</b>
1		country of victims	1
2		some 'free' gym programs impossible to achieve	1
<b>L</b>	<b>Formal nodes</b>		<b>[20]</b>
1		01 factors	20

2	02 car insurance	20
3	03 obesity 150	20
4	04 obesity carrot-stick	20
5	05a obesity carrot how high	20
6	05b obesity stick how high	20
7	06a how many lie for obese	20
8	06b colleagues views on obesity liars	20
9	07 obese poor lose out	20
10	08 Colon cancer policy	20
11	09 CC carrot-stick	20
12	10 CC Three policy options	20
13	11 CC and right not to know	20
14	12 CC and right not to think	20
15	13a how many lie for CC	20
16	13b colleagues' views on CC liars	20
17	14a Five groups	20
18	14b Five groups	20
19	14c Five groups	20
20	14d Five groups	20
21	14e Five groups	20
22	General comments at end	12

## Appendix 6A: Changes in the survey instrument versus the guide used in physician interviews

In reviewing the instrument used for the physician surveys for use in a population level survey, substantial revisions to the form and content were made:

- Question 1 from the physician interviews asked respondents to decide which of three characterisations of the interplay of personal behaviour and other factors, and associated responsibility ascriptions they found most adequate. The double-barrelled structure was confusing to some and so was abandoned, with the responsibility element being removed.<sup>6</sup>
- Questions 5a and 5b from the physician interviews asked respondents to select a \$ or € amount for two weight-control incentive programmes: the first leading to a net reduction in health care cost ('carrot'), the second in increasing amounts for all plan members, with the incentive being to continue to pay previous levels ('false carrot'). To test experimentally framing issues, which emerged as a side-line in the physician interviews, and to make the scenarios more policy relevant in view of recent developments in the US,<sup>7</sup> a third scenario was added, in which those who do not meet the incentive requirement need to pay more ('stick'). Also, to allow for more fine-grained analysis, for all three scenarios respondents were asked to enter an amount between \$0€–2500 (or €0€–€1800), instead of selecting from a set range of amounts.
- Question 10 from the physician interviews asked respondents to compare the current German colon cancer policy that was described in Question 8 with

---

<sup>6</sup> The initial answer options are given below, with the modified version stated after →:

1. Personal choices are really the most important factor that determines health: we are largely responsible for how healthy we are. → How we behave is the most important influence on our health.

2. The role of personal choices is an important factor, and we always bear some responsibility for how healthy we are. But sometimes other factors can be more important: it's therefore not possible to say which factor is the most important one. → Personal behavior is important, but so are other things: it is impossible to say which is most important.

3. The importance of personal choices is overstated: many of the factors above such as the environment, how much money we have, and what genes we inherited influence quite strongly what kind of choice we can make, and they are practically always far more important. → Compared to other factors influencing health, personal behavior is less significant.

<sup>7</sup> While, in 2009, 36% of large employers used incentives, in 2011 54% used them, and 80% plan to do so in 2012. Eight per cent of large employers used incentives in the form of penalties in 2009. In 2011, this figure rose to 19%, and for 2012 38% are expected to impose higher costs on unhealthy employees. This increase represents a doubling over one year, and almost a fivefold increase over three years. National Business Group on Health, Towers Watson. 2011/2012 Staying@Work Report. New York: Towers Watson; 2011.

two alternative scenarios. Question 8 was somewhat complex and the alternatives added further complexity. In the interest of increasing completion rates Question 10 was therefore dropped.

- Questions 11 and 12 from the physician interviews concerned modifications of Question 8 and focused on one element only, viz attending counselling, or undergoing a colonoscopy, if one wanted to secure a lower co-payment threshold. To simplify the questions, and to bring forward the time at which the penalty would be imposed, a more straightforward incentive structure was chosen, so that the penalty was a simple surcharge (of half the value of the incentive amount in Question 8 of the physician interviews).
- New questions were added on respondents' past, current and intended future use of attainment- and participation-incentive programmes; health status; height and weight; use of colon cancer screening; cancer status; political views; health insurance status; and income.
- A welcoming and consent page was added, which informed participants about the focus of the survey, how results will be used and data protection arrangements, who was carrying out the survey, and how the platform works.



### Appendix 6B: Overview of the statistical significance of independent variables

- Tests were Pearson chi-square, ANOVA, and Wilcoxon/Kruskal-Wallis Tests, as appropriate, see Section 6.2.4.
- $p < .05$  is shaded light yellow,  $p < .01$  is light orange and  $p < .001$  is dark orange.
- “N/A” indicates that tests were not applicable as the variable was not used in the survey.
- “-” indicates that tests were not performed.
- “[ ]” indicates that in chi-square tests 20% of cells had an expected count of  $< 5$ , rendering the value suspect.

	Age	Gender	Income	Educ'n	Politics	Insur.	Health	BMI	cancer	scre'ng
1. Factors determining health										
US1	[.3481]	.0845	[.6089]	[.6933]	[.8475]	.2224	[.0069]	[.0445]	[.7724]	-
US2	.1195	.9999	.0009	<.0001	.0242	.1079	<.0001	N/A	N/A	N/A
Ger	[.2018]	.0002	[.5929]	[.5663]	[.0005]	-	[.0217]	.9060	-	-
2. Analogy: obese = risky drivers										
US1	.1997	.0013	.0003	.0052	.0625	.0075	<.0001	<.0001	-	-
US2	.5865	.0006	<.0001	<.0001	<.0001	.0001	<.0001	N/A	N/A	N/A
Ger	.0205	.0114	<.0001	.0011	<.0001	[<.0001]	<.0001	<.0001	-	-
3. \$150 for normal body weight										
US1	.1211	.6308	.0016	.3360	.0171	.0050	<.0001	.0018	-	-
US2	.2461	.2090	<.0001	<.0001	<.0001	.0013	<.0001	N/A	N/A	N/A
Ger	.1704	.7728	.0968	.0030	.0186	[.2262]	[.0015]	<.0001	-	-
4. Is \$150 weight incentive carrot or stick?										
US1	.4460	.0609	.4580	.9304	.1379	.2511	<.0001	.2161	-	-
Ger	.2996	.0158	.0834	.0015	.0010	.0142	.6197	.0002	-	-
5a How high should weight incentives be? Carrot, all amounts										
Mdn - US1	.0002	.2170	<.0001	-	-	-	<.0001	<.0001	-	-
Mdn - Ger	<.0001	.0059	.2751	-	-	-	<.0002	<.0001	-	-
M - US1	.0318	.1588	.2241	-	-	-	.0016	.1818	-	-
M - Ger	.0002	.0007	.4793	-	-	-	<.0001	.0123	-	-
5a How high should weight incentives be? Carrot, amounts >\$1										
Mdn - US1	.2595	.0865	.0038	-	-	-	.0511	.0148	-	-
Mdn - Ger	.0160	.0108	.8274	-	-	-	.0036	.1220	-	-
M - US1	.1449	.1291	.6851	-	-	-	.0500	.388	-	-
M - Ger	.0030	.0017	.6349	-	-	-	.0007	.4763	-	-
5b How high should weight incentives be? False carrot, all amounts										
Mdn - US1	<.0001	.7821	.0045	-	-	-	<.0001	<.0001	-	-
Mdn - Ger	<.0001	<.0001	.7720	-	-	-	<.0001	<.0001	-	-
M - US1	.0539	.3559	.0423	-	-	-	.0057	.0164	-	-
M - Ger	<.0001	<.0001	.3404	-	-	-	<.0001	.0058	-	-

	Age	Gender	Income	Educ'n	Politics	Insur.	Health	BMI	cancer	scree'ng
5b How high should weight incentives be? False carrot, amounts >\$1										
Mdn – US1	.1030	.2269	.0003	-	-	-	.0031	.0227	-	-
Mdn – Ger	.0017	<.0001	.5573	-	-	-	<.0001	.1435	-	-
M – US1	.5464	.2384	.1325	-	-	-	.1291	.2178	-	-
M – Ger	.0002	.0008	.3327	-	-	-	<.0001	.3766	-	-
5c How high should weight incentives be? Stick, all amounts										
Mdn – US1	.0029	.4986	.0022	-	-	-	<.0001	.1818	-	-
Mdn – Ger	<.0001	<.0001	.2777	-	-	-	<.0001	<.0001	-	-
M – US1	.2609	.6223	.0108	-	-	-	.0007	<.0001	-	-
M – Ger	<.0001	<.0001	.0317	-	-	-	<.0001	.0003	-	-
5c How high should weight incentives be? Stick, amounts >\$1										
Mdn – US1	.0814	.2107	<.0001	-	-	-	.1042	.0048	-	-
Mdn – Ger	.0030	.0217	.1147	-	-	-	.0039	.1202	-	-
M – US1	.7986	.6078	.0491	-	-	-	.1016	.03	-	-
M – Ger	<.0001	.0113	.0234	-	-	-	.0022	.3780	-	-
6 How many ask doctor to lie to get weight incentive?										
US1	.0014	.2898	.8076	.1236	.6063	.8423	.4869	.5607	-	-
Ger	.5370	.0049	.2305	.2121	.0119	.3632	.5020	.7769	-	-
7 A doctor who complies and lies acts [right/wrong]										
US1	<.0001	.3675	.1818	.6209	.0221	-	.0029	.9256	-	-
Ger	.0003	.0074	.3663	.0616	[.6007]	-	.0149	.6087	-	-
	Age	Gender	Income	Educ'n	Politics	Insur.	Health	BMI	cancer	scree'ng
8 Is it fair or unfair if poor benefit less from weight incentives?										
US1	.0003	.3887	<.0001	.9024	.0025	.5506	<.0001	<.0001	-	-
Ger	.0001	.0370	.0208	.0034	.0011	<.0001	<.0001	<.0001	-	-
9a Is it fair that 'lucky-virtuos' ones get weight incentives?										
US1	.6649	.5466	.0015	.7936	.1200	-	<.0001	.0046	-	-
Ger	.1123	.4498	.2483	.0004	.1210	.0516	.0008	<.0001	-	-
9b Is it fair that 'lucky-non-virtuous' ones get weight incentive?										
US1	.8951	<.0001	.8311	.1451	.3934	-	.0033	.6263	-	-
Ger	.0638	.0123	.1782	.0382	.0067	.1531	<.0001	<.0001	-	-
9c Is it fair that 'Yes I can' groups gets weight incentive?										
US1	.0279	.1018	.0602	.2521	.1204	-	[<.0001]	.0023	-	-
Ger	.1003	.0922	.2862	.2221	[.0130]	.0041	.0006	<.0001	-	-
9d Is it fair that 'I'll do it tomorrow' group does not get weight incentive?										
US1	.2986	.2650	.0118	.9011	.0012	-	<.0001	.0397	-	-
Ger	.1302	.0446	.3063	.1119	<.0001	.0097	<.0001	<.0001	-	-

	Age	Gender	Income	Educ'n	Politics	Insur.	Health	BMI	cancer	scre'ng
<b>9e Is it fair that 'leave me alone' group does not get weight incentive?</b>										
US1	.9649	.5018	.0083	.3932	.0358	-	<.0001	.5600	-	-
Ger	.8079	.0778	.1854	.0508	.0006	.2794	.0002	<.0001	-	-
<b>10 counsel+ comply</b>										
US1	.0384	.2235	.0039	.0646	.1639	.0103	.0009	-	[.0773]	.0281
US1 50-65	N/A	.6229	.8203	[.5401]	.9201	.2037	[.0411]	-	[.5375]	.5975
Ger	.0018	.0076	.0495	.0017	<.0001	.0139	.0001	-	[.0071]	.0214
Ger 50-65	N/A	.1005	.0140	.0010	[<.0001]	[.0156]	[.0089]	-	.0001	.0364
<b>11 Is CRC counsel&amp;comply policy carrot or stick?</b>										
US1	.9184	.8058	.2493	.1545	.5212	.6496	.0161	-	[.1332]	.0224
US1 50-65	N/A	.7176	.0839	[.8967]	.2753	.7210	.2705	-	[.2973]	.2836
Ger	.1663	.0022	.2797	.0011	.0277	.0697	.0639	-	[.2822]	[.1347]
Ger 50-65	N/A	.1005	.0064	.0003	[<.0001]	[.0011]	.0037	-	.0750	.0005
<b>12 CC: \$250 surcharge for not thinking about CRC</b>										
US1	.0693	.7431	.1674	.1892	.2472	.1044	.0395	-	[.0823]	.0010
US1 50-65	N/A	.3201	.3435	[.0317]	.5724	.3272	[.1245]	-	[.4967]	[.1408]
Ger	.0002	.0764	.0757	.2480	[.0001]	.0002	.0283	-	[.3653]	[.2478]
Ger 50-65	N/A	.1755	.0061	.0025	[<.0001]	[<.0001]	.0928	-	.0516	.0516
US2	.0056	.0070	<.0001	<.0001	.6843	.0004	.0001	N/A	N/A	N/A
US2 50-6	N/A	.0149	.0083	.0001	.7184	.0074	[.0218]	N/A	N/A	N/A
<b>13 CC: \$250 surcharge for not doing CRC</b>										
US1	.2757	.7660	.0020	.1169	.0720	.0346	.0046	-	[.0835]	.0049
US1 50-65	N/A	.7700	.1554	[.1731]	.3398	.2983	[.0543]	-	.2681	.1135
Ger	0.0003	.0976	.0723	.1121	<.0001	.1386	.0007	-	[.0089]	.0491
Ger 50-65	N/A	.0805	.0012	.0002	[<.0001]	[<.0001]	[.0011]	-	<0001	<0001
US 2	.0123	.1271	<.0001	<.0001	.0862	.0658	<.0001	N/A	N/A	N/A
US2 50-65	N/A	.0040	.3883	<.0001	[.3832]	.2193	[.0022]	N/A	N/A	N/A
<b>14 How many ask doctor to lie to get CRC incentive?</b>										
US1	.0001	.4045	.7963	.1281	.9694	.0971	.7158	-	.1717	.0029
US1 50-65	N/A	.0665	.8060	.4164	.2348	.0539	.0539	.4584	.2243	.0350
Ger	.5370	.0049	.2305	.2121	.0119	.3632	.5020	-	.5362	.7337
Ger 50-65	N/A	.2355	.0941	.9211	.0876	.7413	.8676	-	.8989	.7690
<b>15 A doctor who complies and lies acts [right/wrong]</b>										
US1	<.0001	.9177	.3330	.8683	.3735	.8760	.7378	-	[.0937]	.0002
US1 50-65	N/A	.3006	[.5185]	[.8697]	[.8743]	[.6406]	[.5552]	-	.7871	.3287
Ger	<.0001	.0005	.0051	0.0001	.0247	.4014	.0628	-	[.2586]	.0394
Ger 50-65	N/A	.0247	.0607	[<.0001]	[<.0001]	[.1194]	[<.0001]	-	.0141	.0184

	Age	Gender	Income	Educ'n	Politics	Insur.	Health	BMI	cancer	scree'ng
16 Past attainment-incentive use										
US1	.0814	.6676	.0279	.9780	.1269	.0500	.0985	.5926	-	-
Ger	.6456	.0148	.1157	.0701	.0012	.0725	.0192	.0581	-	-
17 Present attainment-incentive use										
US1	.7723	.9863	.0021	.3814	.5237	.0030	.0818	.2745	-	-
Ger	.9430	.0081	.5704	.0042	.2156	.4869	.0041	.6186	-	-
18 Future attainment-incentive use										
US1	.0028	.4669	.0018	.9150	.7842	.0033	.0042	.0048	-	-
Ger	.9878	.0877	.0565	.0076	.0424	.5722	.0131	.0247	-	-
19 Past participation -incentive use										
US1	.1033	.1288	.0659	.2899	.0765	.0009	.1350	.7241	-	-
Ger	.9137	.6468	.0651	.7769	.0665	.3052	.5880	.3516	-	-
20 Present participation -incentive use										
US1	.2020	.1875	.0177	.4796	.2552	<.0001	.0183	.5734	-	-
Ger	.0616	.8368	.1838	.1162	.0761	.2902	.0573	.0172	-	-
21 Future participation -incentive use										
US1	.0053	.4156	.0290	.0826	.1280	.0084	.0048	.0046	-	-
Ger	.2102	.4954	.1693	.0359	.0098	.1131	.0065	.0591	-	-

### **Appendix 6C: Notes on sample providers' methods of recruitment, use of incentives, participation and completion rates**

Samples for surveys US1 and Ger were provided by the company *Usamp*. *Usamp* is active in around 200 countries and the total panel size approaches 5 million. All panels constitute non-random samples: individuals may apply to be a panel member, and in some cases are actively recruited, for example through links to the *Usamp* recruitment page<sup>8</sup> being placed at the end of (non-*Usamp*-panel) surveys.

Respondents receive financial incentives for each completed surveys, and may choose between cash awards, gift cards, making charitable donations, or obtaining points for use with company partners (similar to using points collected through store reward cards or airlines that may then be used for goods or accommodation in hotels, etc). The average equivalent monetary amount is \$1 per completed survey of around 10 minutes. Panels are actively managed, which includes regular updating of demographic information held on panellists, controlling the number of survey invitation panellists receive (typically no more than one per week), preventing multiple submissions, and monitoring completion times in order to remove "speeders" who do not read the question text and simply click randomly to reach the end of a survey. Members who sign up receive notice of new surveys by emails, with the subject title typically reading "New exciting survey opportunity" or similar, and the body of the email detailing the broad topic (such as "health", "consumer electronics", etc.), estimated length of completion, and available monetary value, see the sample at Appendix 6C. More detail about *Usamp*'s approach to sampling, panel recruitment, management and data protection can be found in their responses to the ESOMAR 26, a set of questions generated by the European Society for Opinion and Market Research to assist in the evaluation of key parameter of online panels (USamp 2011).

The sample for survey US2 was provided by Knowledge Networks, which maintains the US' only nationally representative online access panel. Panellists are recruited randomly through Random Digit Dial (RDD) telephone and Address-Based Sampling (ABS). The company argues that this approach gives broader reach than traditional RDD used in telephone surveys alone. Qualifying households are provided with free internet access and laptops. Incentives similar to those provided

---

<sup>8</sup> See: <http://www.surveyhead.com/> accessed on 7 December 2011.

by *Usamp* are offered for surveys longer than 15 minute (not applicable in the case of survey US2). Sample management and distribution of notices of new surveys work in a similar fashion as described above, although respondents do not know about the survey topic when they receive an invitation, see Appendix 6D for a sample. More detail about the company's approach to sampling, panel recruitment, management and data protection can again be found in their responses to the ESOMAR 26 (Knowledge Networks 2011).

Surveys US1 and Ger were fielded between 25 September and 5 October 2011, and US 2 between 22 September and 3 October 2011. US2 was fielded as part of an Omnibus, and questions were combined with another preceding survey, which concerned two questions in which respondents were presented with terms ("book" and "fire" in the testing version) and asked to provide a specific personal event that each terms brings to mind. Two further questions asked respondents to list important world events occurring since the 1930s. While these questions were unlikely to have biased the responses to the following health survey, they can complicate response rate assessments, which, generally, are far from straightforward in online panels.

For traditional phone surveys it is natural to begin a survey by telling potential participants about the topic, and refusal to participate can itself be meaningful, as, for example, low response may indicate that only people with a strong interest have participated, calling into question the representativeness of the findings. The online situation combined with the access panel arrangement presents a different situation. Surveys can be distributed to vast numbers of people at the click of a mouse, and, as noted, the headers of emails do not usually mention the topic (in the case of the Knowledge Network, the body of the email also does not mention it). Therefore, non-response with regard to following up an email invitation (which are often sent to an extremely large number of people to speed up the gathering of responses) is unlikely to indicate lack of interest, or stronger reservations, with regard to the survey's specific subject matter. The meaningfulness of response rates that is conventionally expressed as the number of people completing a survey versus the number of people receiving an invitation differs, therefore, in the comparison of telephone surveys and online surveys such as those carried out here. Moreover, several studies have addressed the question of the role of response rates in predicting

accuracy of findings, and *Knowledge Networks* points to research which suggests that for probability samples higher rates do not predict notably increased accuracy (Groves 2006; Keeter et al. 2000).

In the case of US2, 52% of those who received an invitation to participate completed the survey (understood as clicking through all pages). Of those who started the survey, 91% per completed it. In the case of US1, of those who were invited, 9% completed it. Of those who started the survey, 96.9% completed it, with 1.5% declining to take part after reading the welcome page, and 1.6% dropping out at later stages. In the German sample, 5 % of all invitees completed, 2.7% declined to participate on reading the welcome page, and 1% dropped out later. Clearly, panel membership entails a high willingness to participate in surveys (of whatever kind). The far lower overall completion rates in the *Usamp* panels in comparison with the *Knowledge Networks* sample reflect differences in the business model and methodological approach between the companies, chiefly, because *Usamps'* method of so-called river sampling (trying to win everyone for panel membership who 'swims by') leads to far higher numbers of people joining the panel, but also higher numbers of people who, in fact, are less interested in taking surveys in general.

## Appendix 6C: Sample email from *Usamp* to survey panel members about new available surveys

Take Our Survey And Share Your Opinion  
Your Time And Opinion Are Important To Us

*Survey Topic:* Consumer

*Reward Amount:* \$1.50

*Survey Length:* 10 Minutes

[Start the Survey Now](#)

If you are unable to click on the survey link above, please copy and paste this link into your address bar to begin the survey: <http://surveyhead.com/ssal.php?e=@@@@.com&mi=11047142&gj=MjQ2Nzk=&bi=MTgzNzg1&sbi=MTgzNDYw&u=Y>

Important Message: Please provide thoughtful and honest responses at all times!

You will not receive your survey reward and your member account may be closed if you engage in any of the following activities:

- Register with [Surveyhead](#) by providing false or misleading information.
  - Provide false or misleading information in our surveys or profiles in an attempt to qualify.
  - Demonstrate inattentive behavior such as speeding through a survey or entering information that is ineligible.
  - Create multiple [Surveyhead](#) accounts.
  - Complete a survey more than one time.
  - Forward your invitation to someone else, unless directed to do so.
  - Complete a survey well under the average length of interview as recorded in our system.
  - Post our survey links on the Internet or advise others on how to qualify.
  - Register if you are under the age of 13 years old in the United States. Note: age restrictions vary by country.



**Appendix 6D: Sample email from *Knowledge Network* to survey panel members about new available surveys**

Your Latest KnowledgePanel Survey

Dear %%firstname%%,

Thanks for being an integral part of KnowledgePanel! Your latest survey can be accessed by clicking the following link:

<a href="%%Link%%" class="bold">Click Here to Start Survey</a>

(If it does not work automatically from email, please copy the below link to your internet browser).

%%Link%%

Our Panel Member Support Center is available if you have questions or comments. You can contact us at our toll free 1-800-782-6899 number, or simply reply to this email invitation. In order to better serve you, please be sure to include reference <span class="bold">#(%%ANO%%)</span> in the Subject of your message. We are always happy to hear from you!

Thanks,  
KnowledgePanel Support Team

**Appendix 6E: Overview of the effects of applying age-weights to sample Ger.**

	Unweighed sample	Weighed sample	Population	Difference weighed sample in relation to population
<b>Age<sup>9</sup></b>				
20–35	53	33	33	-
36–49	36	36	36	-
50–65	12	31	31	-
<b>Gender<sup>10</sup></b>				
Male	52	52	49	+3
female	48	48	51	-3
<b>Income</b>				
Low	30	29	25	+4
Middle	61	62	69	-7
high	9	9	6	+3
<b>Education<sup>11</sup></b>				
Less than HS	8	8	15	-7
HS	46	47	49	-2
Some college	24	22	10	+12
BA or higher	23	22	26	-4
<b>Politics<sup>12</sup></b>				
Left Party	8	10	5	+5
Green	15	15	13	+2
SPD	20	20	22	-2
FDP	2	2	7	-5
CDU/CSU	14	14	22	-8
Right wing	3	3	2	+1
Other/no pref.	38	37	29	-2
<b>Insurance<sup>13</sup></b>				
SF	83	83	90	-7
Private	15	15	9	+6
Other	2	2	1	+1
<b>Health<sup>14</sup></b>				
Poor	2	2	2	-
Fair	12	17	8	+9
Good	49	50	31	+19
Very good	32	27	46	-19
excellent	6	4	13	-9

<sup>9</sup> (Statistisches Bundesamt 2009)

<sup>10</sup> (Statistisches Bundesamt 2009)

<sup>11</sup> (Statistisches Bundesamt 2010)

<sup>12</sup> (Allensbach (Institut für Demoskopie) 2011)

<sup>13</sup> (Statistisches Bundesamt 2008)

<sup>14</sup> (Statistisches Bundesamt 2011c)

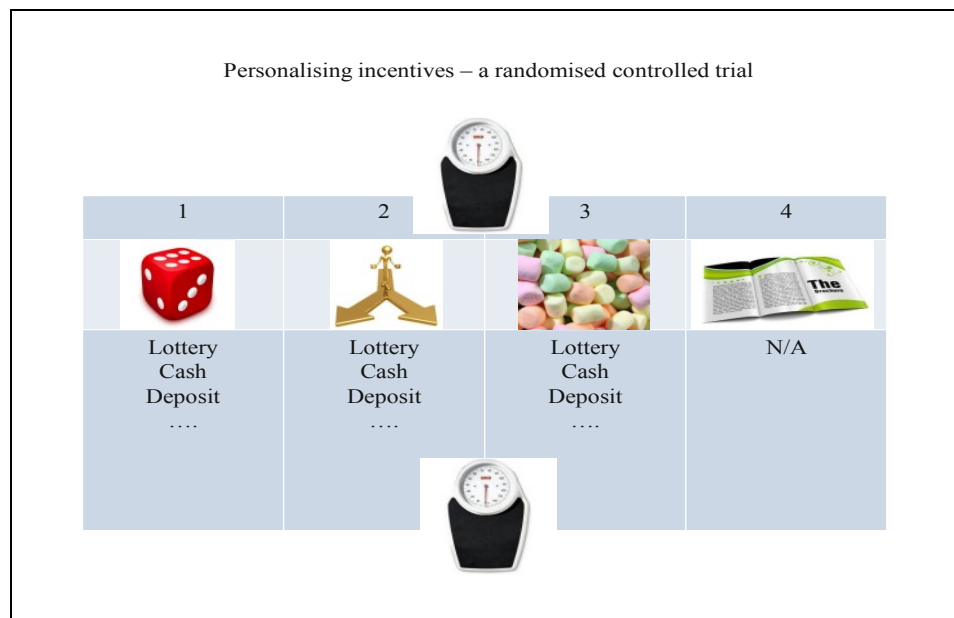
	Unweighed sample	Weighed sample	Population	Difference weighed sample in relation to population
<b>BMI<sup>15</sup></b>				
Underweight	5	4	2	+2
normal	51	45	50	+5
overweight	26	31	34	-3
obese	18	21	14	+7
<b>Used attainment</b>				
Yes	23	23	-	-
No	77	77	-	-
<b>Uses attainment</b>				
Yes	18	18	-	-
No	82	82	-	-
<b>Will use attainm.</b>				
Yes	57	57	-	-
No	43	43	-	-
<b>Used particip.</b>				
Yes	27	27	-	-
No	73	73	-	-
<b>Uses particip.</b>				
Yes	21	20	-	-
No	79	80	-	-
<b>Will use particip.</b>				
Yes	57	56	-	-
No	43	43	-	-

---

<sup>15</sup> (Statistisches Bundesamt 2011b)

### Appendix 8A: A randomized controlled trial to study personalized incentives

To gain a better understanding of the determinants of the effectiveness of incentives for particular types of people, a screening tool, perhaps in the form of a short questionnaire could be developed, to help identify which incentive structure is most likely to work for a given individual.



The study could take the following format. All participants agree reasonable weight loss goals with the study team of around a pound per week over a 16 week period; receive information and counselling on evidence based strategies; and weigh themselves at regular intervals, say every day, on scales that take their picture and transmit their weight to a central computer. In the first arm, participants are randomly assigned to one of three incentive structures that have been used with some success before (Volpp 2008): a lottery based one, in which participants who are on track with their weight loss trajectory have a 1:5 chance of winning \$10, and a 1:100 chance of winning \$100 (with the average daily amount of \$3); a monthly cash payment of the equivalent amount that is provided if the daily weight goal is met; and a deposit contract in which participants can put their own money on the line which is matched up to \$3 per day (but both the deposit and the matching amount are lost if the daily target is not met). In the second arm participants are not randomly allocated to these options, but may chose whatever intervention they find will work best for them. In the third arm they are directed to one of the three interventions, having taken a prior screening test that measures different motivational dimensions. In the fourth (control) arm participants are merely provided with information on weight loss (that is also given to all participants in the other arms). Comparing the weight loss between the different arms and the sub-options within them over the study period, and, importantly, at a suitable time after conclusion of the study (six and twelve months later) can help determine the extent to which the effectiveness of programmes can be maximised by personalising incentive programmes.