Health Inequality Measurement and Impacts

Methods and Applications

Christopher Lynderup Lübker

Ph.D.

University of York

Health Sciences

October 2022

Abstract

Monitoring health inequalities and assessing the impacts of policy interventions requires valid measurements of social characteristics and health outcomes, and the relationships between them. Despite recent progress, methodological challenges remain. First, standard bivariate health inequality measures (BHIMs) – linear associations between a health dependent variable and a socioeconomic independent variable – may contradict social justice because measured health inequality can increase when income inequality is reduced, and vice versa (the Nordic paradox critique). I generalise existing simulation studies to address the contribution of rank dependency and alternative underlying causal pathways, finding BHIMs only risk contradicting results when (a) using absolute, rather than ranked, income and/or (b) there is substantial income re-ranking. Second, BHIMs may reflect fair health differences associated with adult socioeconomic outcomes partly within a person's control (the inequality of opportunity critique). I propose parental gradients – bivariate associations between health in adulthood and fraction-ranked parental socioeconomic status in childhood, a determinant outside a person's control – as simple inequality of opportunity for health metrics. I find parental gradients in general self-assessed health are significant and substantial for UK adults, and 40-65% the magnitude of standard BHIMs. Third, an evaluation challenge is how to extend distributional cost-effectiveness analysis (DCEA) methods to quantify health inequality impacts of interventions outside the health sector. To illustrate a co-funding solution, I evaluate the lifetime population health and health inequality impacts of the Universal Infant Free School Meal (UIFSM) programme. I find UIFSMs are cost-effective up to a £92 million per year group health care co-fund and reduce health inequalities by disproportionately benefiting more deprived children. This thesis helps improve our understanding of strengths and limitations of BHIMs used for both research and policy monitoring purposes. It also advances the DCEA methodology by illustrating a cross-sectoral co-funding proposal, while providing new evidence about a topical childhood policy.

Contents

List of Tables

List of Figures

Acknowledgements

This work would not have been possible without the support of several people.

Firstly, I would like to thank my family and friends; I would not be where I am today without each of you. I would also like to thank my colleagues, old and new – at the Centre for Health Economics and Department of Health Sciences, University of York; Centre on Well-being, Inclusion, Sustainability and Equal Opportunity, Organisation for Economic Co-Operation and Development; and the Obesity Access Team, Novo Nordisk – for the stimulating conversations and inspiring work environments. Special thanks go to Matthew Robson and Ieva Skarda for their helpful comments throughout this undertaking, and to the members of my Thesis Advisory Panel: Karen Bloor, Susan Griffin, and Nils Gutacker.

Finally, I would like to thank my supervisors, Tim Doran and Richard Cookson. Your honest and patient guidance have helped me realise how much there is to learn in the field, while managing to make that finding endlessly fascinating, instead of demoralising. Thank you for making the journey thoroughly enjoyable, whatever the challenges and pandemics along the way.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as references.

Parts of this thesis have been presented and discussed at conferences. An earlier version of Chapter 4 was presented at Society for Social Medicine and Public Health in September 2021, titled *Comparing current and life course socioeconomic gradients in health: A simple inequality of opportunity metric*. An earlier version of Chapter 5 was presented at the Health Economics Study Group in January 2022, titled *Evaluating the Health Impacts of Universal Infant Free School Meals: A Distributional Cost-Effectiveness Analysis*.

This work was supported by Wellcome Trust (Grant No. 205427/Z/16/Z) within the Senior Investigator Awards in Humanities and Social Science, titled *Re-Engineering Health Policy Research for Fairer Decisions and Better Health* of Professors Tim Doran and Richard Cookson.

CHAPTER 1

Introduction

Since the landmark World Health Organisation (WHO) target in 1981 to reduce the difference in health between countries by 25% before 2000 (WHO 1981), policy-makers have increasingly recognised the urgent need to address health differences between and within countries by quantifying them and setting reduction targets accordingly (OECD 2015; UN 2015; Whitehead, Scott-Samuel, and Dahlgren 1998). Differences in health are commonly referred to as health inequities or inequalities if they are "unnecessary and avoidable", while also being "considered unfair and unjust" (Whitehead 1991, p. 5). For example, socially advantaged persons tend to live longer lives in better health than disadvantaged persons (Mackenbach et al. 2008), a consistent finding termed the social gradient (Dahlgren and Whitehead 1991; Marmot and Wilkinson 2005). The social gradient encompasses a wide range of health determinants, including individual, societal, and environmental factors and applies across the social distribution, not just in health differences between the worst- and best-off (Marmot et al. 2008).

Recently, the COVID-19 pandemic has shone a spotlight on vast social inequalities in health both between and within countries. In England and other high-income countries, social disadvantage is associated with higher likelihood of infection, along with greater risks of complications and mortality following infection (Wachtler et al. 2020). Studies from USA have also shown a rise in deaths of despair over the past two decades – deaths from suicide, drug overdose, alcohol poisoning, and liver cirrhosis – concentrated among populations with lower educational attainment (Case and Deaton 2021), a finding which has since been documented in several other high-income countries (Lübker and Murtin 2022).

Inequalities may be recorded using simple measures comparing social gaps in health, such as the difference in life expectancy between the highest and lowest occupational classes (Marmot et al. 2008). Since these measures do not account for differences in group sizes, or the middle of the distribution, bivariate measures of health inequalities have also been used to provide information on differences in life expectancy across the entire distribution, by estimating the linear association between all socioeconomic ranks and life expectancy (Mackenbach et al. 2008). These metrics have been instrumental in placing health inequality measurement on the policy agenda and establishing international monitoring programmes. However,

bivariate measures have also been criticised because they may be incompatible with social justice and may reflect fair health differences associated with socioeconomic outcomes in adulthood partly within a person's control, which are further explained in this thesis.

Quantifying health inequality impacts of interventions also informs health care and societal budget allocations in the context of scarce resources and opportunity costs (Drummond et al. 2015). Distributional cost-effectiveness analysis (DCEA) provides a framework for equity-informative health economic evaluation (Cookson et al. 2020) with the ability to inform decision-makers both about the magnitude of health inequality impacts and about potential trade-offs between inequality impacts and cost-effectiveness. In the global context of constrained health care budgets and the need to maximise health benefits of a costly health-related investment, economic evaluation and standard cost-effectiveness analysis (CEA) have become increasingly important. However, the rise in inequality considerations on the policy and societal agenda call for the quantification of the inequality impacts of resource allocations as well. Since DCEA was developed for health care evaluation with a fixed health sector budget, challenges remain in extending these methods to cross-sectoral policies with both health and non-health benefits and budget impacts.

Decision-makers increasingly recognise the urgent need to reduce health inequalities, and substantial progress has been made in developing methods of health inequality monitoring and evaluation to facilitate this. However, challenges remain in understanding and addressing the limitations of current methods to ensure policy makers have both better evidence about health inequalities and a better understanding of the strengths and limitations when applying these methods in the pursuit of social justice. This thesis addresses three such challenges: two about monitoring and one about evaluation. The monitoring challenges are that standard bivariate measures of socioeconomic inequalities in health (i) may sometimes contradict social justice, since measured health inequality can increase when income inequality is reduced, and vice versa (the Nordic paradox critique) and (ii) may partly reflect fair health differences associated with socioeconomic outcomes in adulthood partly within a person's control (the inequality of opportunity critique). An evaluation challenge is how to extend distributional cost-effectiveness analysis (DCEA) methods to quantify health inequality impacts of interventions outside the health sector.

In Chapter 2, I review methodological frameworks and empirical findings from the health inequality measurement literature to provide background for Chapters 3-5. First, I review bivariate health inequality measurement methods and the technical

concepts that underpin them, including the choice of level- or rank-dependent measures. Second, I summarise the literature on key causal pathways driving health and socioeconomic status over the lifetime and between generations. This informs frameworks on lifetime inequality of opportunity for health, which rely on the differentiation between socioeconomic choices and circumstances driving differences in health. Third, I introduce the distributional cost-effectiveness analysis and cross-sectoral economic evaluation frameworks as they relate to evaluating the health inequality impacts of health care co-funding of cross-sectoral interventions with important benefits and budget impacts across two or more policy sectors, such as school feeding programmes.

In Chapter 3, I investigate a critique of standard bivariate health inequality measures, henceforth referred to as the Nordic paradox critique. The critique arose following cross-country findings on education-related inequalities in mortality and morbidity, suggesting Nordic countries with more generous welfare arrangements and lower levels of economic inequality do not necessarily have lower levels of health inequality than the USA and other high-income countries with less generous social safety nets (Cavelaars et al. 1998; Doorslaer and Koolman 2004; Eikemo et al. 2008; Mackenbach and Kunst 1997; Mackenbach et al. 2008, 2018). This finding has been termed the "Nordic paradox" (Mackenbach 2012), though studies using more detailed data on income-related inequality in mortality reveal a more nuanced picture; Norway, for example, has lower levels of health inequality than the USA along the middle of the income distribution, if not in the lowest and highest percentiles (Kinge et al. 2019). However, until granular socioeconomic data become widely available, routine monitoring studies of socioeconomic health inequalities must rely on coarser data with fewer quantiles.

Previous simulation studies suggest bivariate health inequality measures may report greater socioeconomic health inequalities when income inequality increases, even if the health distribution remains unchanged (Brekke and Kverndokk 2012; Renard et al. 2019). If true, this critique of standard BHIMs would have major policy implications, meaning that widespread national and international systems of health inequality monitoring and evaluation established in recent decades may actually harm the cause of equality and justice. However, previous studies have not assessed the impact of alternative causal pathways and rank-preservation on the performance of level- and rank-dependent bivariate health inequality measures. I develop a simulation model to systematically address these shortcomings and explore the degree to which BHIMs are appropriate tools for the measurement and monitoring of health inequality.

In Chapter 4, I engage with the criticism of standard bivariate measures that they may partly reflect fair health differences associated with socioeconomic outcomes in adulthood partly within a person's control, henceforth referred to as the inequality of opportunity (IOp) critique. There is a commonly held belief in varying degrees of personal responsibility for socioeconomic success, such as job title or income, on the grounds that some of these factors may be improved through personal choices, such as time investments in education, employment, and physical activity. However, this belief is tempered by a recognition that not all choices are completely within a person's control and may be influenced by circumstances for which a person cannot be held responsible, such as the parents they were born to. The role of personal choice in driving health and social outcomes in adulthood means that associations between socioeconomic status and health in adulthood may potentially exaggerate the degree of social injustice in a population. Furthermore, it makes inequality target-setting more complex, since there may still be a social gradient in health even if equality of opportunity is achieved.

Distinguishing between choices and circumstances is not ethically uncontroversial. For example, lifestyle choices within a person's control may contribute to overweight and obesity, but these behaviours may also be driven by factors outside a person's control, such as learned norms from their upbringing. Distinguishing choices and circumstances is also empirically challenging, often including a wide array of variables – e.g., health, behavioural, and economic characteristics – across long time intervals, requiring careful deliberation about the relationships between them. This limits the scalability of sophisticated inequality of opportunity methods, since comprehensive longitudinal data from childhood are not commonly available and the statistical expertise required to conduct these studies is scarce. I present a simple inequality of opportunity metric that attempts to address these challenges, equipping decision-makers with a tool which may be routinely implemented for monitoring purposes.

In Chapter 5, I extend the distributional cost-effectiveness analysis (DCEA) framework to consider health care sector co-funding of education policies (Cookson et al. 2020). DCEA was developed for the health care sector, assuming a fixed health budget and challenges therefore remain in extending the framework to non-health policies. This is an important extension since non-health interventions may have greater lifetime health care and health inequality implications than health care and policy interventions (Marmot et al. 2008). This has previously been captured in the compensation approach, whereby the health care sector compensates non-health sectors for the health benefits accrued from non-health intervention, but has not been extended to capture health inequality impacts (Claxton et al. 2010).

As an illustrative example of a topical policy, I evaluate the cost-effectiveness and health inequality impacts of the Universal Infant Free School Meal (UIFSM) programme in during a child's first three years of school in England. The UIFSM programme increases access to free school meals (FSMs) beyond means-tested provision, whereby a child's household must be eligible for certain benefits to gain FSM access. While the Department for Education currently funds the UIFSM programme, the policy may also have substantial health benefits by improving the eating habits and educational outcomes of children early in life, potentially leading to greater health-related quality and length of life, while also yielding health sector savings. A health care co-fund of FSMs may be justified on these grounds.

The long-term inequality impacts of universal compared to means-tested FSMs are unknown: since universal provision incrementally increases FSM access more for less deprived populations, it may be possible that universal provision worsens health inequalities. Equally, it may be the case that universal provision ensures access to healthy food options for all those who need it, regardless of parental income levels, leading to more equal outcomes. These lifetime inequality impacts have not been examined and are key elements of health and educational policy in England. I develop a *de novo* lifetime Markov cohort model to estimate lifetime costs and effects of universal FSMs by deprivation quintile in England. I apply distributional cost-effectiveness analysis methods to assess lifetime population health and health inequality impacts, informing decision-making on the allocation of resources to a topical childhood policy with international relevance.

Chapter 6 concludes the thesis. First, I present the key findings and their implications as they relate to measuring health inequality and health inequality impacts. Second, I note the contributions of the thesis to the established literature, while explaining the impact of limitations on findings. Third, I suggest avenues for future research based to expand on the findings and address the limitations of the studies in this thesis.

This thesis aims to provide practical tools that inform two key elements of an effective health inequalities strategy; first, the measurement and monitoring of health inequalities, to set impactful targets and track progress towards them; and, second, the evaluation of the interventions implemented, to assess which policies have the greatest and fewest health equity benefits. In doing so, the methods and results from this thesis lay a groundwork to better understand and implement bivariate health inequality measures in research and population monitoring. Finally, this thesis advances the methodology of cross-sectoral distributional cost-effectiveness analysis by showing how to address questions

of co-funding, and provides new evidence about an important and topical childhood intervention.

The simulation study in Chapter 3 aims to investigate how bivariate health inequality measures, or health gradients, respond to rank-preserving changes in the income distribution under varying causal assumptions. This informs the extent to which changes in conventional health gradients may be spuriously driven by changes in the distribution of the socioeconomic variable, and therefore informs the extent to which these measures should be used to guide and monitor policies. The illustrative study in Chapter 4 aims to show how bivariate health inequality measures can be modified to provide insights on inequality of opportunity and the importance of childhood circumstances in predicting adult health. This provides a practical tool that allows for insights into an area of growing policy concern, while limiting the additional methodological complexity, which can otherwise be an obstacle to the uptake of tools outside research environments. The distributional cost-effectiveness analysis in Chapter 5 aims to show how equity-informative economic evaluations can be employed in a cross-sectoral intervention setting to quantify the maximum cost-effective co-fund and the health equity impacts of Universal Infant Free School Meals, which is funded by the Department for Education but has multi-sectoral benefits.

CHAPTER 2

Background

2.1 Inequality Measurement Concepts

2.1.1 Equity and Time

Most egalitarian philosophical literature is concerned with the question 'equality of what?' (Sen 1979). Less attention has been devoted to the question 'equality when?', or more precisely 'equality over which time period?'. Even within a single egalitarian theory, differing time perspectives can result in substantially different ethical implications, as we will see. Competing views can be classified under two broad categories: segment perspectives refer to a specific temporal segments, as defined by a time period (e.g. calendar dates, such as the year 1989) or age-groups (e.g. age 18-24), and lifetime perspectives refer to an individual's entire life course, from birth to death.

The distinction between lifetime and segment perspectives has been reflected in the conceptual differences between *age-groups* and *birth cohorts* (Bidadanure 2016, pp. 239-42). *Birth cohorts* are typically defined as individuals born in the same year. These cohorts then pass through *age-groups* together. The difference between these two concepts is elegantly explained the concept that birth cohorts age, while age-groups do not (Daniels 1988, p. 13).[1] We may therefore consider age-groups a segment perspective because they consider specific temporal segments through which individuals pass. The same essential definition holds true for time periods, such as 1920, which may therefore also be considered a segment view. Birth cohorts, on the other hand, may be considered a lifetime perspective because they consider individuals as they pass through temporal segments.

Lifetime and segment views result in distinct issues of justice (Daniels 1988, 2008). For example, if a hypothetical policy subsidised education for those over 25, this would result in different injustices than if the policy subsidised education for those

[1]This is to say, as a cohort passes through time their age changes, however the age of age-groups remains the same; 50-year-olds are 50 years old whether we consider them in 1920 or 2020.

born before 1994. The latter policy will never be applicable to those born from 1995 onwards, however the former policy will also be applicable to every person who turns 25 henceforth. Crucially, lifetime perspectives inherently justify intergenerational justice claims; if we pollute enough to harm future persons, this will generate an injustice in lifetime wellbeing between present and future persons (Holtug and Lippert-Rasmussen 2007, p. p.10). Segment views are intrinsically concerned with temporal segments, not individuals; they care about when inequality is occurring, and to what extent, but not between whom. Therefore, segment views do not require individuals to be separable, which is necessary to claims of intergenerational justice (see section 1.1).

Lifetimes were long assumed to be the only appropriate time perspective in philosophical egalitarian literature. However, Dennis McKerlie's defence of segment views sparked a discussion about the legitimacy of this assumption, which is summarised in the remainder of this section (McKerlie 1989). Distinctions between lifetime and segment views are important due to their influence on applied health research; should we care most about health equalising health now, or equalising health over lifetimes? In this section, I refer to wellbeing before turning to specific issues of health in later sections.

Time-Relevant Equity Concepts

Separability of Persons

When discussing lifetimes and segments, we must clarify whether an individual's wellbeing is separable. Separability here refers firstly to the extent to which one individual's wellbeing is separable from another's in morally important ways (Broome 2004; McMahan 2002). If individual wellbeing is not separable, the resulting theory of justice is unconcerned with distributions of wellbeing between persons. Utilitarianism is an example of such a distributive theory. However, separability here also refers to the extent to which an individual is separable in their own lifetime. For example, are we the same person at 10 years old as we are at 50 years old, despite our changes in personality?

The metaphysical issue of personal identity is dense and not always directly related to time perspectives. These issues will therefore not be resolved here since the focus in on applied ethics, not meta-ethics. It will suffice to say I assume it is plausible that individuals are separable from one another and consist of a continuing (if not unchanging) personality throughout their lifetime, such that past injustices are still unjust to the same person in the present and future. This is justified

by a connection of prior experiences and their relationship to current and future circumstance (Broome 2004; Parfit 1986; Rawls 1972).

These separateness of persons assumptions are necessary to consider lifetime egalitarian perspectives. If individuals are not separable from one another, the lifetime perspective is redundant, and if individual wellbeing is separable in lifetimes, it is impossible to consider lifetime injustices. All future subsections assume the plausibility of these assumptions, however this does not beg the question of whether we ought to endorse lifetime or segment views, it merely allows us to consider both time perspectives.

The Changing Places Problem

Dennis McKerlie (1989) criticised lifetimes as the gold standard egalitarian time perspective (McKerlie 1989). Much of the legwork was done by his caste-swap thought experiment considering a society of two classes, which swap places every 20 years (McKerlie 1989, p. 479). In each 20-year segment individuals in one class lead fantastic lives, full of wellbeing, at the expense of the other class, where individuals lead miserable lives. Even if lifetime wellbeing is equal for both classes, we may still be opposed to inequalities present in such a society. This example has become known as the changing places problem and implies lifetime perspectives do not fully capture our aversion to inequality (Bidadanure 2016, p. 241). To illustrate a simple example of the changing places problem, consider Case I in Figure 2.1.

Figure 2.1: Equity and time thought experiment: Case I.

Case I

	T_1	T_2	
A	7	3	= *10*
B	3	7	= *10*
	4	*4*	

In Case I, the lives of persons A and B are characterised by two stages, T_1 and T_2, each containing a wellbeing score. Assuming an additive relationship between wellbeing across life stages, lifetime *wellbeing* is calculated by summing each individual's wellbeing in T_1 and T_2, which is 10 for both A and B. Lifetime *inequality* is then calculated by taking the absolute value of subtracting Person B's lifetime wellbeing from Person A's, which is zero. Segment inequality is

calculated by taking the absolute difference between Person A and B at each time period – four for both T_1 and T_2 – and summing these values, which is eight. Segment inequality is higher because it is concerned with *any* inequality between *any* persons at *any* time, whereas lifetime perspectives are intrinsically unconcerned with temporary inequalities between separable individuals (Bidadanure 2016; McKerlie 1989; Temkin 1993).

Although there is no lifetime inequality in Case I, McKerlie refers to a situation such as in Case II, to illustrate how lifetime perspectives might not capture our aversion to inequality (see Figure 2.2). In Case II, wellbeing is equal across the lifetime and during all time segments; thus, on both segment and inequality views inequality is equal to zero. McKerlie claims few egalitarians would prefer a society where systematic oppression exists (Case I) to one where it does not (Case II), suggesting lifetime perspectives do not fully capture our stance on equality over time (McKerlie 1989, p. 479).

Figure 2.2: Equity and time thought experiment: Case II

Case II

	T_1	T_2	
A	5	5	= *10*
B	5	5	= *10*
	0	*0*	

Reciprocity

The concept of reciprocity – compensating those who have been unjustly harmed – is featured in most theories of justice (Rawls 1972). However, even within a single theory of justice, lifetime and segment perspectives generate diametrically opposing compensation demands. Lifetime perspectives make extreme reciprocity demands whereas segment perspectives make no reciprocity demands whatsoever.

Case III shows how segment egalitarians would allocate wellbeing in T_2, knowing the wellbeing distribution in T_1 (see Figure 2.3). Segment egalitarians distribute wellbeing equally in T_2, regardless of segment inequality T_1, on the grounds that any inequality at any time is undesirable and to impose further inequalities in T_2 would only serve to increase total inequality. Given the same choice, lifetime

egalitarians would allocate wellbeing in T_2 as in Case I, compensating Person B on the grounds that they were unjustly harmed in T_1.

Figure 2.3: Equity and time thought experiment: Case III

Case III

	T_1	T_2	
A	7	5	= *12*
B	3	5	= *8*
	4	*0*	

Neither of these extreme reciprocity demands seem ideal. Lifetime perspectives effectively require individuals to be punished for their high wellbeing in the past, regardless of how it has come about, and segment perspectives completely disregard any past inequalities, no matter how unjust. Our support of either lifetime or segment perspectives will therefore depend, firstly, on the extent to which we believe temporary inequality to be of intrinsic importance and, secondly, on the extent to which we believe the assumptions on the separateness of persons.

Separability of Time

Similar to the issue of separability of persons, separability of time is important to our choice of time perspective. Segment egalitarians claim the moral goodness or badness of a temporary inequalities is independent of any past or future inequalities (Hirose 2005). As Kasper Lippert-Rasmussen (2008) notes, this generates the conclusion that segment perspectives are concerned with inequalities at any instantaneous moment in time, since choosing a longer segment is subject to slippery slope arguments (Lippert-Rasmussen 2008, p. 34). This is a very strong requirement. The slippery slope argument and issue of time separability are clarified in Figure 2.4, which draws inspiration from Bidadanure (2016) (Bidadanure 2016, p. 244).

Case IV is identical to Case I, merely splitting each time period into subsections x and y. As this shows, the arbitrary choice of temporal segments can have a significant impact on the amount of inequality reported; Case I reports a total segment inequality of eight, whereas Case IV reports ten. This leaves segment views at the mercy of arbitrary choices of temporal segments, and generates a slippery slope argument requiring segment egalitarians to consider instantaneous inequalities.

Figure 2.4: Equity and time thought experiment: Case IV

Case IV

	T_1		T_2			
	x	y	x	y		
A	4	3	1	2	=	*10*
B	0	3	6	1	=	*10*
	4	*0*	*5*	*1*	=	*10*

Time separability presents significant theoretical obstacles to supporting segment perspectives due to the arbitrary selection of temporal segments. However, the Case IV thought experiment also has empirical applications. At the very least, the issue of time separability suggests we ought to deliberately consider and report the temporal segments we apply when measuring inequality, as this may have substantial impacts on how much inequality is reported.

Segment Displacement

Further to the arbitrary choice of temporal segment, the arbitrary displacement of lifetimes has also been presented as an objection to segment perspectives. As Temkin (1993) observes, if two individuals with identical sequential wellbeing patterns across their lifetimes are merely born at different times, substantial segment inequality is still reported (Temkin 1993, p. 241). Some may not find this arbitrary displacement of lives generates objectionable inequality. To use Lippert-Rasmussen's (2008) example, we might not see any reasonable moral difference between you suffering through a dentists appointment today, and me suffering through one tomorrow (Lippert-Rasmussen 2008).

Figure 2.5: Equity and time thought experiment: Case V.

Case V

	T_1	T_2	T_3		
A	7	3		=	*10*
B		7	3	=	*10*
		4		=	*4*

Consider Case V in Figure 2.5. Person A and B have identical sequential wellbeing (seven and three) but Person A is born in T_1 and Person B is born at T_2. Lifetime perspectives see no important difference between this arbitrary displacement of wellbeing since both individuals end up with a lifetime wellbeing score of ten. For segment egalitarians, however, the inequality between Person A and B in their overlapping time period, T2, is morally relevant, generating a total segment inequality of four. If we believe, as Daniels suggests, that there is no injustice in Case V since nobody is treated differently nor unfairly across their lifetime, then segment displacement presents considerable obstacles to supporting segment views of inequality (Daniels 1988, 2008).

Conclusions

Egalitarian philosophical literature has mostly supported the lifetime perspective because it allows for issues of intergenerational justice and does not hinge on the arbitrary selection or displacement of temporal segments. Nevertheless, extreme reciprocity demands and the changing places problem suggest that segment inequality also captures some of our aversion to inequality. As a result, pluralistic views incorporating segment perspectives or supplementary principles may be necessary for a coherent account of justice over time.[2]

The differing requirements of lifetime and segment perspectives rest on their underlying assumptions and value judgements. Segment perspectives consider any inequality at any time to be morally significant, whereas lifetime perspectives don't consider any temporary inequality intrinsically morally significant whatsoever. While examples in this section have been simplified for the sake of clarity, the underlying assumptions of segment and lifetime perspective have far-reaching implications on how we allocate health care resources equitably.

2.1.2 Time Perspectives in Health Inequality Measurement

The implicit value judgements we make when measuring health equity can have a decisive impact on the results obtained (Harper et al. 2010; Kjellsson, Gerdtham, and Petrie 2015). The literature on value judgements in health inequality measurement has largely been concerned with distinctions between individuals and

[2]See Bidadanure (2016) for an example of supplementation with segment sufficiency principles (Bidadanure 2016, pp. 252-254).

groups, relative and absolute inequality, and shortfalls and achievements. These topics have been extensively discussed elsewhere[3] but insofar as they do not pertain directly to issues of time perspectives, they are omitted here.

Time perspective issues have rarely been explicitly recognised in health inequality measurement. Interestingly, while the political philosophy literature generally assumes lifetime perspectives, health inequality measurements generally assume segment perspectives, assessing socioeconomic differences in health between individuals in the current time period, not across their lifetimes. This section reviews the literature on how differing time perspectives affect our measurement of health variables.

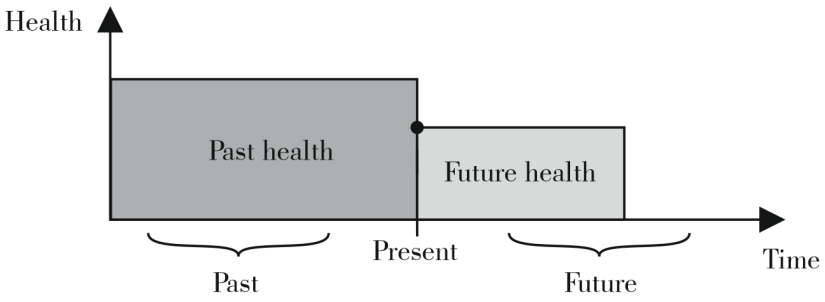
Past, Present, Future, or Lifetime Health?

On segment views, we can prioritise equality in past, present or future health, however lifetime perspectives consider all three. To illustrate the differences between past, present, future and lifetime health, an example is presented for each, inspired by (Olsen 2017). In these diagrams, past and future health are represented by the dark and light grey rectangles, respectively, and present health is measured as the height of the bold point where past and future health intersect (see Figure 2.6).

Past Health
Past health perspectives prioritise those with worse health states experienced until the present. This brings us back to the idea of reciprocity. While we cannot strictly speaking equalise past health because these health states have already been experienced, it can be compensated. If past health is the most equity-relevant time perspective for health inequalities, we ought to compensate those who have experienced the worst cumulative health, from birth until present. While the incorporation of reciprocity is important to an equitable conception of health, the lack of concern for future and present health generates problematic decision-rules.

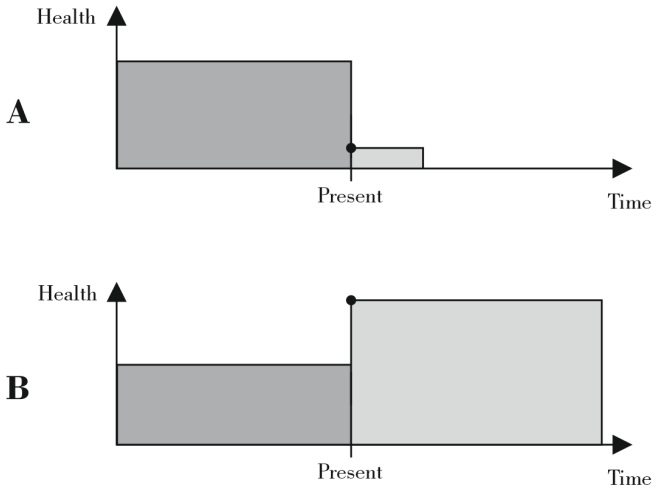
[3]For an overview of implicit value judgements in health inequality measurement, see Kjellsson et al. (2015) and Harper et al. (2010) (Harper et al. 2010; Kjellsson, Gerdtham, and Petrie 2015).
For a discussion regarding absolute and relative inequality, see Harper et al. (2010) and Asada (2010) (Asada 2010; Harper et al. 2010). To see how this influences the empirical measurement of health inequality, see Erreygers and Van Ourti (2011) (Erreygers and Van Ourti 2011).
For a discussion regarding achievement and shortfall inequality, see Sen (1995) (Sen 1995). To see how this influences the empirical measurement of health inequalities, see Lambert and Zheng (2011)(Lambert and Zheng 2011).
For a discussion regarding the distinction between group and individual inequality, see Asada (2013) and Eyal (2018) (Asada 2013; Eyal 2018).

Figure 2.6: Past, present, and lifetime health measurement.



In Figure 2.7, for example, we are forced to favour B over A, despite A's low future, current and lifetime health. Past health segment views therefore completely disregard future or current circumstances, which some may find objectionable.

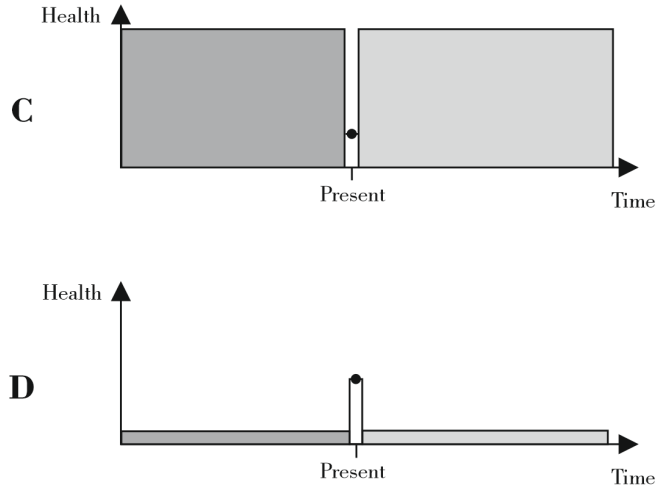
Figure 2.7: Past, present, and lifetime health measurement: scenarios A and B.



Present Health

Present health views coincide with need principles, which were especially popular in early health equity policy literature (Culyer and Wagstaff 1993; Whitehead 1991). On this view, the current severity of health states matters most, which is often justified by Rawls' difference principle (Rawls 1972). As might be expected, this time perspective also leads to problematic decision-rules. In Figure 2.8, the present health perspective suggests prioritising D over C, even though D is only temporarily in a better health state and has otherwise experienced much worse health in the past and will continue to do so in the future. This approach therefore embodies a recency bias few decision-makers would endorse.

Figure 2.8: Past, present, and lifetime health measurement: scenarios C and D.

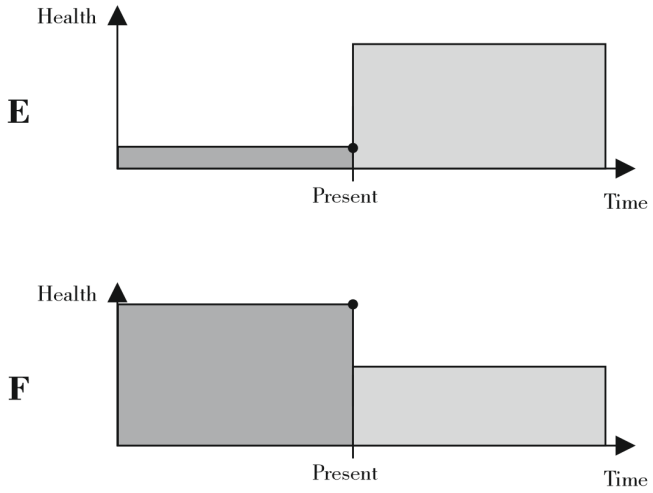


Future Health
Future health views place the primary importance on how much health we can expect to achieve (or lose) in the future. An example of such a view is NICE's end-of-life premium in the UK, in which those expected to have fewer than 2 years left to live are entitled to a larger share of health care resources than other individuals (Cookson 2013). The shortcomings of future health views can be seen in Figure 2.9. Even though F has perfect health currently and have experienced for their entire lifetime until the present, they will be chosen over E due to their poor future health prospects. This may seem unfair due to the very poor health experienced by E in the past and at present, suggesting future health views alone do not provide a comprehensive account of justice.

Lifetime Health
Lifetime health perspectives consider the culmination of all health experiences: past, present and future. As a result, in each of the former examples, the lifetime perspective is not forced to accept the unfavourable decision rules otherwise mandated. Take Figure 2.7, for example. While past health perspectives are forced to prioritise B over A, on the lifetime perspective A is worse off than B and is therefore given priority. The same applies to each of Figures 2.8 and 2.9, and the lifetime perspective therefore seems to provide a better metric to evaluate health.

The lifetime perspective does not inherently consider quality of life intrinsically important. A long miserable life is of the same value as one half as long, but lived in twice the health, as in Figure 2.10. It is not obvious that G and H are equally well off, especially if good health is valued more during certain life stages.

Figure 2.9: Past, present, and lifetime health measurement: scenarios E and F.

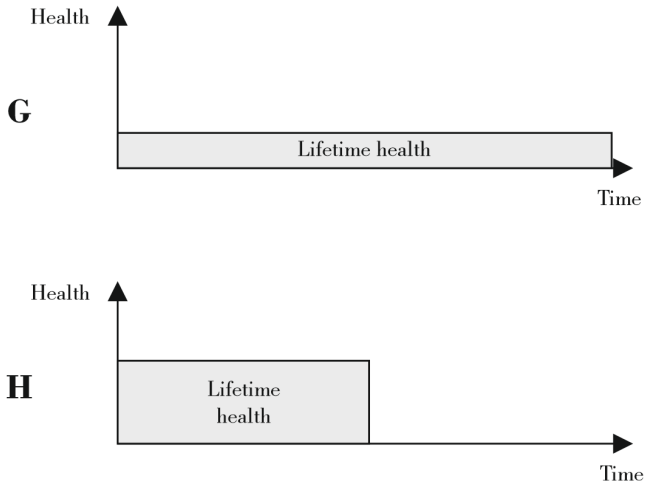


While this trade-off between quality and time is worth mentioning, is not exclusive to the lifetime perspective; all examples in Figures 2.6-2.8 encounter the same issue. Once more, the issue stems from the fact that segment approaches (i.e. past, present and future) consider only inequalities in their segment, regardless of context and lifetime approaches do not consider any temporary inequalities significant whatsoever. However, we might plausibly value good health more at 20 than at 80, or at least consider health shortfalls at 20 more unjust. Neither lifetime nor segment perspectives are inherently sensitive to these differential time preferences for health.

Conclusions

There is no consensus on which time perspective is preferable and for which reasons. Researchers are often forced to use whichever health data are available to them and therefore rarely acknowledge underlying value judgements implicit in their health variable. In general, lifetime and pluralistic views – combining present health with either future or past health – have been proposed in health care decision-making. Whichever time perspective is taken, acknowledging and reporting implicit value judgements inherent in health variables is important to measure the health inequality we are most interested in.

Figure 2.10: Past, present, and lifetime health measurement: scenarios G and H.



2.2 Bivariate Health Inequality Measurement

Bivariate health inequality measures are a mainstay of health inequality research and monitoring (Mackenbach et al. 2008). These measures estimate the association between a dependent variable that relates to health status and an independent variable that relates to socioeconomic status (SES), on the grounds that health should not be systematically stratified along socioeconomic lines (Braveman and Gruskin 2003). SES is a complex construct of social advantage often proxied by education, occupation, income, wealth, and area-level deprivation depending on the inequalities of research interest (Galobardes et al. 2006a,b).[4]

The implicit value judgements in the choice of socioeconomic health inequality measure can have a decisive impact on the results obtained (Harper et al. 2010; Kjellsson, Gerdtham, and Petrie 2015). This section introduces the key concepts and instruments of bivariate health inequality measurement, with a focus on topics relevant to Chapter 3.[5]

[4]Composite SES measures have also been operationalised in health inequality research (Robson, Doran, and Cookson 2019).

[5]For a discussion regarding achievement and shortfall inequality, see Sen (1995). To see how this influences the empirical measurement of health inequalities, see Lambert and Zheng (2011).

2.2.1 Concepts

The health inequality measurement literature draws substantially from the income inequality literature (Cowell 2000), however there are key differences between health and socioeconomic variables which influences inequality measurement (Griffin et al. 2012). First, equality is more easily attainable in socioeconomic variables than in health. Second, health can be seen as both intrinsically and instrumentally important – that is, important for its own sake, and important to something else with intrinsic value – whereas socioeconomic variables may only have instrumental importance. Perhaps as a result of this, socioeconomic status is generally assumed to exhibit diminishing marginal returns, unlike health. Third, health variables often have both upper and lower bounds, as is the case by definition for the quality-adjusted life year (QALY) and by biological restrictions for life expectancy. These differences between income and health may generate different normative requirements for inequality measures.

Univariate and Bivariate Inequality

Bivariate measures of health inequality (BHIMs) measure associations between a socioeconomic and health variable, such that greater associations are interpreted as more unjust. However, health inequality may also be measured univariately by considering the distributional characteristics of a health variable. Univariate health inequality measures (UHIMs) are often derived based on the magnitude of distributional variance in a variable. It is assumed that greater variance represents greater levels of inequality because it implies greater disparities in the variable of interest, such as life expectancy (Erreygers and Kessels 2017).

The choice between BHIMs and UHIMs rests on the underlying value judgements on which inequalities are deemed unfair. While BHIMs are most popular in empirical research and monitoring, there has been substantial debate between the two approaches.[6] Arguments generally centre on: (i) the assumed intrinsic or instrumental moral importance of health inequalities; (ii) whether the primary concern is for group or individual inequalities; (iii) the importance of comparability across populations and time; and (iv) the extent to which we wish to provide either an overall perspective on health inequality or a direct policy target (Asada 2013).

If health is intrinsically important, UHIMs are justified because health differences

[6]See, for example, the back-and-forth discussion between Murray, Gakidou, and Frenk (2000) and Murray, Gakidou, and Frenk (1999) and Braveman, Krieger, and Lynch (2000)

are a cause for concern in their own right, and not due to relationships to any other variables. However, if health is instrumentally important – valuable only insofar as it aids in other, intrinsically important, pursuits – univariate inequality assessments do not carry the same moral weight; differences in observations are only morally worrisome to the extent they occur systematically by unfair means. Decision-makers might, for example, be more opposed to health inequalities associated with geographical location, than those associated with health behaviours, such as smoking.

UHIMs and BHIMs also differ in their unit of measurement. BHIMs tend to focus on group inequalities, whereas UHIMs tend to focus on individual inequalities (Asada 2013, pp. 38-40). Measuring group inequalities requires averaging across observations, thereby washing out distributional information (Asada 2013; Temkin 1993). Individual-level UHIMs provide a clear comparison of inequality measurements over time and across populations, since individuals are maintained as the central unit of analysis regardless of context (Erreygers and Kessels 2017). BHIMs, however, encounter more obstacles comparing measurements across populations. Firstly, there are pragmatic obstacles since BHIMs require more data, and secondly, the distribution of stratifying variables, such as income, is likely to change over time and across populations. BHIMs often avoid this problem by ranking the stratifying variable (more on this in the 'Ranks and Levels' subsection).

Finally, UHIMs and BHIMs differ in the direct policy relevance of their output. If, for example, BHIMs find life expectancy differs substantially between education levels, this provides a clear policy directive: initiatives to reduce health inequalities may benefit from examining the education-health relationship and potentially exploring the impact of educational policies on life expectancy. However, it may be argued that BHIMs ignore the overall picture of health inequalities, which UHIMs provide (Asada 2013, p. 40). Simply stated, UHIMs tend to be general and informative, while BHIMs are specific and prescriptive.

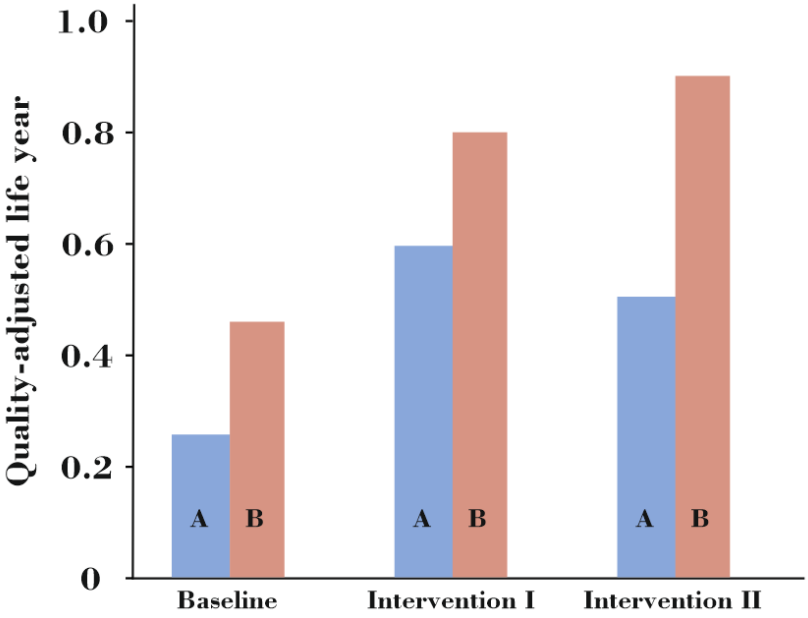
The debate between univariate and bivariate inequality has not been settled. BHIMs remain a mainstay of health inequality measurement, while arguments for univariate assessments garner significant theoretical support, particularly if health is considered intrinsically valuable. Eyal (2018) suggests a remarriage of the two stances. Eyal claims that while only avoidable univariate individual inequalities are of intrinsic importance, group inequalities, as measured by BHIMs, are still useful to measure because they capture determinants of individual inequalities [pp. 160-163]. This is to say, both UHIMs and BHIMs are important and complementary methods to informing health inequality research; using one should not preclude the

use of the other.

Relative and Absolute Inequality

Health inequalities may be measures in terms of relative or absolute health inequalities. Figure 2.11 adapts an example from Asada (2010), considering two populations, A and B (see Figure 2.11) with a quality-adjusted life-year (QALY) score of 0.25 and 0.45, respectively.[7] There is an absolute gap of 0.2, and a relative ratio of 1:1.8. Suppose we had a choice between two interventions, in which we attempt to improve outcomes while keeping inequality the same. Intervention I increases QALYs of both A and B by 0.35, resulting in 0.6 QALYs for A and 0.8 QALYs for B. Intervention II increases QALYs proportionally by 100% instead, resulting in a QALY score of 0.5 for A and 0.9 for B.

Figure 2.11: Quality-adjusted life years of a baseline population and two hypothetical interventions.



Intervention I maintains the 0.2 QALY absolute difference while decreasing the relative ratio to approximately 1:1.33, whereas Intervention II maintains the relative

[7]The quality-adjusted life year is a generic measure of length and health-related quality of life (Kaplan and Hays 2022). A QALY varies between zero and one, with zero indicating death and one indicating perfect health.

ratio of 1:1.8 while increasing the absolute difference from 0.2 to 0.4 QALYs. The choice between relative and absolute measures has been contentiously debated in the philosophical literature (Scanlon 2018; Sen 1979; Temkin 1993) and elicitations of public preferences have not been unanimous (Asada 2010; Gakidou, Murray, and Frenk 2001). It is therefore generally recommended that studies report both absolute and relative inequalities (Mackenbach and Kunst 1997).

Broadly, absolute measures are more appropriate if a specific threshold value is considered important, or perhaps a human right (Sen 1979). A decision-maker might, for example, be less concerned with inequalities following Intervention I, because they believe experiencing less than 0.6 QALYs is unjust. Relative inequality measures, on the other hand, are scale-consistent across distributions and more highly comparable across populations and time. Whether BHIMs should reflect absolute or relative inequality therefore rests on our implicit values with regards to the health variable in question and the characteristics of the inequality measure.

Ranks and Levels

The use of rank-dependent measures has bled over from income inequality measurement into health inequality measurement, providing a pragmatic solution to the comparability problem. However, several authors have questioned whether we ought to apply these assumptions to health inequality measurement (Asada 2010; Erreygers 2009). Ordinal ranking converts a continuous variable to an ordinal variable wherein the distance between between observations is constant such that the variable remains scale-consistent across populations and over time (Cowell 2000). Fraction-ranking may also be applied to account for differing distributions of socioeconomic variables across populations, particularly when the number of groups being compared is low and weighting has greater impact.

Although (fraction-)ranking variables addresses the comparability issue for BHIMs, the transformation results in distinct value judgements. By using socioeconomic rank, measures imply the moral importance of relative socioeconomic position, rather than absolute level, and therefore focus on relative socioeconomic inequalities, which is rarely explicit. Furthermore, ranking a variable also renders it insensitive to rank-preserving changes to the underlying distribution which may be misleading in some cases (Erreygers and Kessels 2017).

2.2.2 Inequality Measures

Pearson's Correlation Coefficient

Pearson's Correlation Coefficient (PCC) is a simple statistical measure of the association between two variables used in early studies of socioeconomic inequalities in health (Wilkinson 1992). Pearson's correlation coefficient (PCC) is used in this thesis because it illustrates a level-dependent BHIM, comparing the levels of health and income. The PCC, ρ, is calculated as:

$$\rho_{Y,H} = \frac{cov(Y,H)}{\sigma_Y \sigma_H} \tag{2.1}$$

Where Y is the outcome level of the income variable, H is the outcome level of the health variable, *cov* is the covariance, and σ is the standard deviation. This will yield results ranging from -1 to 1, with values greater than (less than) zero indicating a positive (negative) linear relationship, and a value of 0 indicating no linear relationship. The larger the PCC, the more strongly the two variables are correlated. The PCC is sensitive to distributional changes in either variable, especially given a large sample size, and has therefore not been consistently implemented in studies of socioeconomic inequalities in health (Regidor 2004, p. 900).

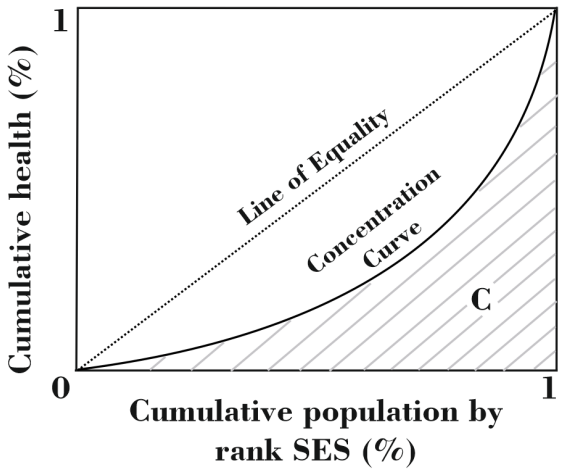
Concentration Indices

The concentration index (CI) is a rank-dependent BHIM based on the concentration curve, which takes after the Lorenz curve (Lorenz 1905), except it considers both a health and socioeconomic variable (Kakwani 1977). The CI plots the cumulative population proportion by socioeconomic rank against the cumulative proportion of health (see Figure 2.12). Similar intuitions apply as with the Lorenz curve: if the concentration curve lies below the diagonal line of equality, then the distribution of health is regressive , with poor health concentrated among low-SES observations and good health concentrated among high-SES observations.

The concentration index (CI) is then defined in terms of the concentration curve $C(x)$:

$$CI = 1 - 2\int_0^1 C(x)dx \tag{2.2}$$

Figure 2.12: Health concentration curve



Note: SES, socioeconomic status.

Or, empirically:

$$CI(h|y) = \frac{2cov(h_i, R_i)}{\bar{h}} \tag{2.3}$$

Where h_i is the health level of individual i, R_i is the rank of individual i in socioeconomic variable y, and $\bar{h}$ is the average population health level. Thus, when examining health achievement variables, the more regressive (progressive) the concentration curve, with health concentrated among higher (lower) socioeconomic ranks, the more the concentration index tends towards 1 (-1), and perfect equality is represented by a concentration index of 0. The generalised, or absolute, transformation of the CI is then:

$$CI(h|y) = 2cov(h_i, R_i) \tag{2.4}$$

As with most BHIMs, it should be noted the same CI may reflect different distributions of health by socioeconomic rank. Suppose one concentration curve follows the line of equality exactly, resulting in $CI = 0$. Another concentration index also reports a value of zero, however in this sample, health is concentrated around the mean of observations. This could result in a concentration curve which at first seems regressive, and then progressive crossing the line of equality and also resulting in $CI = 0$.

Slope and Relative Index of Inequality

The SII subtracts the predicted health of the highest socioeconomic rank from the predicted health of the lowest socioeconomic rank according to a linear regression coefficient between an ordinal socioeconomic independent variable (scaled 0-1) and a health level dependent variable (Mackenbach et al. 1997). By using the linear approximation values, rather than the gap between the health of the highest and lowest socioeconomic groups, the SII utilises information from the full socioeconomic distribution unlike simple gap metrics. The SII is calculated in a regression framework using an ordinary least squares (OLS) regression with robust standard errors (Moreno-Betancur et al. 2015):

$$y = g_\alpha(x) = y_0 + \alpha x \tag{2.5}$$

When SES is fraction-ranked resulting in a maximum of 1 and minimum of 0:

$$SII = \alpha = g_\alpha(1) - g_\alpha(0) \tag{2.6}$$

Where α is the magnitude and sign of the linear association between x and y, the socioeconomic and health variables, respectively, when the socioeconomic variable is fraction-ranked between 0 and 1. SII values greater than zero in a health achievement variable indicate the socioeconomically best-off tend to have better health and vice versa. If the SII is zero, or statistically indistinguishable from zero, this indicates no association between socioeconomic fraction-rank and the health variable.

The SII captures the socioeconomic dimension in health inequalities and is able to use data from the entire distribution. The measure is an absolute index of inequality since it uses levels of health (Wagstaff, Paci, and Doorslaer 1991). Thus, if all health levels were to rise proportionally, so would the SII, since it is sensitive to the mean.

The relative index of inequality (RII) is the relative transformation of the slope index of inequality. Since the SII is sensitive to the mean, the RII is easily calculated by dividing the SII by the mean health of the sample (Pamuk 1988; Wagstaff, Paci, and Doorslaer 1991). As such, it is possible for the entire population to become more or less healthy, however this measure of inequality may still remain the same. Furthermore, the SII and RII could even report opposing trends (Pamuk 1988). In this thesis I follow a regression framework and use a GLM function with Gaussian

family, log link, and robust standard errors to calculate the RII (Moreno-Betancur et al. 2015):

$$y = f_\beta(x) = y_0 exp(\beta x) \tag{2.7}$$

Thus, when SES is fraction-ranked, exhibiting a maximum of 1 and minimum of 0:

$$RII = exp(\beta) = \frac{f_\beta(1)}{f_\beta(0)} \tag{2.8}$$

Where $exp(\beta)$ is the magnitude of the linear association between x and y, the socioeconomic and health variables, respectively, when the socioeconomic variable is fraction-ranked between 0 and 1. An RII estimate equal to one indicates no relative advantage between predicted lowest and highest socioeconomic ranks. If the RII is greater than one for a health achievement (shortfall) variable, higher socioeconomic rank is associated with relatively better (worse) health and when RII is less than one, higher socioeconomic rank is associated with worse (better) health.

Inequality Measures in Practice

Health inequality measures have been operationalised in national and international monitoring efforts. The English health inequalities strategy between 1997 and 2010 aimed to reduce the life expectancy gap between the most and least deprived local authority quintiles by 10% (Mackenbach 2011). The initiative implied a focus on relative health inequalities and monitored these using a simple gap metric. In 2021, the Office for Health Improvement and Disparities was established in the UK with the aim of improving the health and wellbeing of the most disadvantaged, thereby reducing health disparities (OHID 2022). Monitoring dashboards compile health indicators on key national public health priorities, such as cardiovascular, kidney, and respiratory disease, by deprivation and includes a bespoke health inequalities dashboard with a focus on systematic health differences between deprivation groups. Three inequality priority areas now also use the SII: suicide rates and mortality rates in under-75s for cardiovascular disease and cancer. At the international level, the World Health Organisation (WHO) commonly reports the gap in a health variable (such as malaria or HIV incidence) by the richest and poorest quintile (WHO 2022) while the Organisation for Economic Co-operation and Development

has also begun routinely report inequalities in life expectancy by education among its member countries regression-based inequality measures, such as the SII and RII, to complement simple gap-based metrics Lübker and Murtin 2022; Murtin et al. 2022.

2.3 Inequality of Opportunity Measurement

In the economics literature, there has been a general shift in health inequality measurement towards inequality of opportunity (Arneson 2002; Rawls 1972), resting on the assumption that there are "fair" and "unfair" determinants of health. In health economics and epidemiology, equality of opportunity most often refers to reducing health inequalities attributed to circumstances beyond a person's control, while allowing health inequalities attributed to choices to exist (Roemer 1998). There are generally two approaches to measuring inequality of opportunity for health: ex ante and ex post. The ex ante method considers differences in opportunity sets, or circumstances, before the health outcomes are achieved (Jusot and Tubeuf 2019). The ex post method, on the other hand, considers health outcomes stratified by efforts to identify the influence of circumstances.

Whether estimating ex ante or ex post inequality of opportunity, circumstances may influence choices. For example, while smoking is a choice, the parents we are born to and environment we grow up in may influence our likelihood of smoking (Leonardi-Bee, Jere, and Britton 2011). Disentangling the causal network of circumstances and choices driving health inequality is therefore not ethically uncontroversial or empirically straightforward.

2.3.1 Establishing Causality

As mentioned in the previous sections, the primary competing causal hypotheses for the association between health and socioeconomic status (SES) are social causation and health selection (Hoffmann, Kröger, and Pakpahan 2018). Under social causation, SES causes health, e.g. higher income leads to better living conditions and therefore better length and health-related quality of life. Under health selection, health causes SES, e.g. a sudden health shock renders a person unable to participate in the labour market.[8] Health and SES could also be jointly

[8]Often referred to as reverse causation in economics because it reverses the causal pathway between income and health in the widely popularised human capital model, which assumes income

determined by other factors, which would also explain the SES-health gradient, e.g. unobservable heterogeneity in individual genetics, risk or time preferences, and non-cognitive skills.

In practice, the relationship between health and SES is likely to be multi-directional. Robust research designs must therefore be employed to allow for consistent and unbiased effect estimation. Randomised control trials (RCTs) represent the gold standard of experimental evidence, generating causal evidence provided the experimental design is robust. However, RCTs are require substantial time and monetary investments, so quasi-experimental methods are often employed in research to estimate causal effects (Cookson et al. 2021a). Quasi-experimental methods such as difference-in-difference or regression discontinuity can exploit natural experiments, such as a change in taxation or education laws, to establish causality. Even in the absence of natural experiments, researchers often use statistical methods such as instrumental variables, propensity score matching, and twin fixed-effects models to make causal claims. Causal inference using quasi- and non-experimental methods may require rich datasets to provide significant results and often encounter method-specific complications. Nonetheless, quasi-, non-, and experimental methods have all made essential contributions to causal research regarding the health gradient.

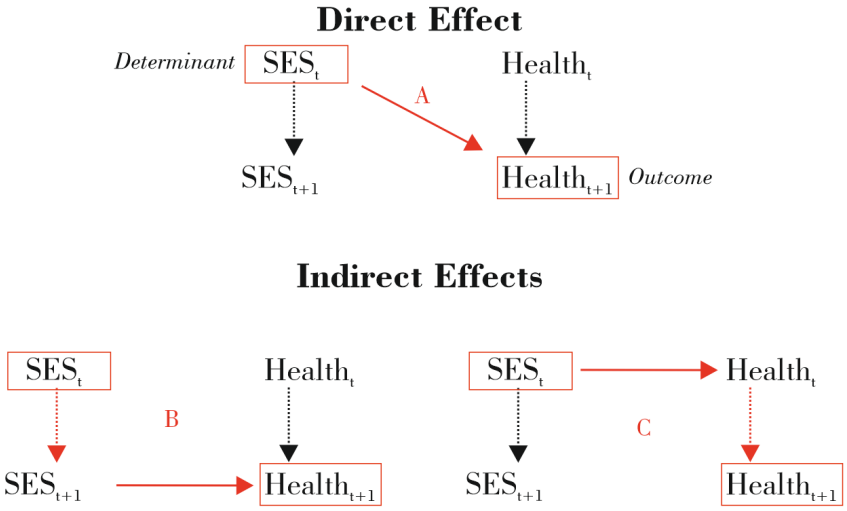
When examining causal effects over time, it is useful to distinguish between total, direct and indirect long-term effects (Pearl 2009). Figure 2.13 illustrates this distinction, drawing on the field of causal inference through the use of directed acyclic graphs (DAGs) (Pearl 2000). As the figure shows, it could be the case that current income directly impacts mortality (Path A: direct effect). However, it is also possible current income influences retirement income, which then predicts mortality (Path B: indirect effect). Or, perhaps current income influences current health which then predicts mortality (Path C: indirect effect). In all likelihood, all three effects are simultaneously influential, which empirical research tends to cumulatively estimate; the total effect.

Isolating direct and indirect effects is empirically challenging (Pearl 2009). Direct effects reflect the sensitivity of an outcome to its determinant while all mediators remain fixed . Direct effects can therefore be identified by holding mediators – in this case indirect effects – constant, but not by selecting out, adjusting for, or conditioning on mediators.[9] Total indirect effects can then be calculated by

determines health (Grossman 1972).

[9]Using the latter methods can create misleading estimations and/or introduce spurious associations between variables (Pearl 2009, p.132).

Figure 2.13: Total, direct, and indirect effects of socioeconomic status on health.



subtracting the direct effect from the total effect. Identifying individual indirect effects is more troublesome. Given the necessity of direct effects for a DAG to exist, indirect effects cannot be estimated by the same means of holding the direct effect constant. Instead, Pearl (2001) identifies path-specific indirect effects by calculating changes in total indirect effects when successively forcing each indirect pathway to be inactive (Pearl 2001). Calculating path-specific effects allows decision-makers to identify the strongest individual pathways and deliver more informed policy recommendations.

The reasons for separating of direct and indirect effects can also shed light on how SES-health gradients develop over the lifetime. Using Figure 2.13 once more, it is possible to distinguish between path A being a direct long-term effect – because SES and health interact directly across time periods – and paths B and C being indirect *short-term* effects - because the interaction between SES and health occurs in a single time period. This distinction is explored referring to empirical evidence in the following section.

Separating age, period and cohort effects is also key to correctly identifying causal effects driving the gradient (Heathcote et al. 2018; Lauderdale 2001). Yang and Land (2013, pp. 1-2) provide a succinct description of these concepts. Period effects occur at specific moments in time such as a natural disaster or the implementation of a vaccination programme (see Figure 2.14). Age effects correspond to Daniels' age-groups, encountered by all individuals through their biological and social aging (see Figure 2.15). Cohort effects refer to the differing experiences of individuals

according to their time of birth, sharing age and period effects to a large extent when situated in the same population (see Figure 2.16). If age, period and cohort effects are not separated analyses run risks of endogeneity, due to the likelihood of each time effect confounding the others. For example, if a natural disaster occurred in 1980, it is difficult to distinguish the age effects of this disaster as a result of being 60 or 80 years old from the cohort effects of being born in 1920 or 1900.

Figure 2.14: Age, period, and cohort effects: period effect.

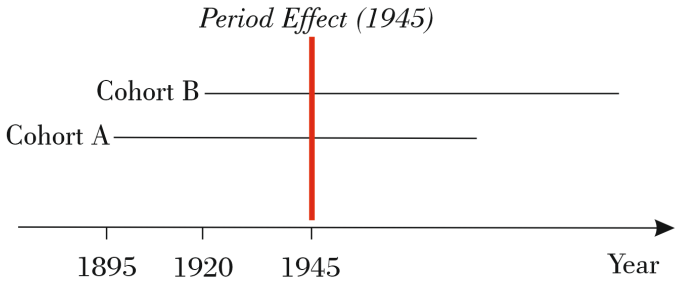


Figure 2.15: Age, period, and cohort effects: age effect.

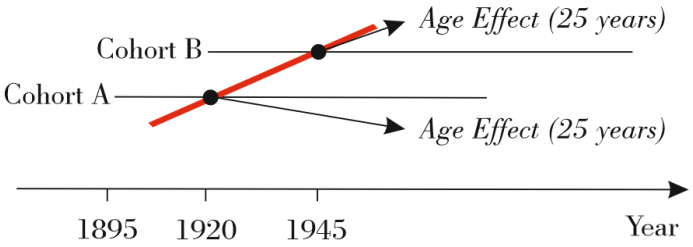
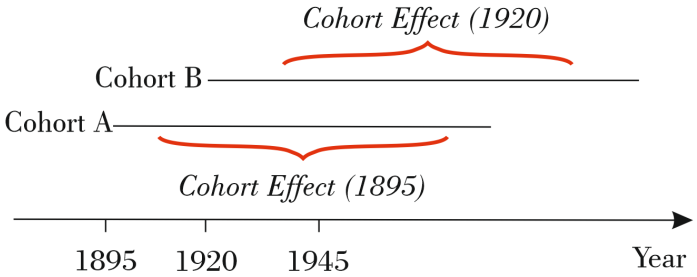


Figure 2.16: Age, period, and cohort effects: cohort effect.



In the absence of exceptional research designs, the estimation of age, period and cohort effects is subject to some strong assumptions, such as constant period effects. In a study of health inequalities over the life cycle, Kippersluis et al. (2009) exclude period effects on the basis of (a) their commonality among all individuals, implying

a smaller effect than cohort and age effects, and (b) a pragmatic acknowledgement that unrestricted age, period and cohort effects cannot be simultaneously estimated. To identify cohort effects, the authors group observations by birth year to compare different cohorts. To identify age effects, the authors control for individuals' aging within a cohort. As a robustness check they include period effects, finding the addition "reduces the precision of estimates while leaving age and cohort profiles much the same" (p. 827). The authors also emphasise the benefits of longitudinal data in separating age, period and cohort effects. By following individuals over time, unobserved individual heterogeneity is more easily tackled, thereby overcoming obstacles cross-sectional data struggle with on this matter.

Thus, three key factors influence the ability to manage confounding between age, period and cohort effects. First, the research design must be robust and aimed at identifying two of the three time effects at most. Secondly, it must be plausible that one of the effects, most likely period effects, is constant. Thirdly, having longitudinal data eases the separation of age, cohort and period effects and facilitates the former two factors.

2.3.2 Health Inequalities Between and Within Lifetimes

There is a large body of evidence documenting associations between socioeconomic status (SES) and health, often referred to as gradients. Several frameworks have been proposed to conceptualise the development of health inequalities within lifetimes, across multiple disciplines. The Dahlgren and Whitehead (1991) 'rainbow' framework places the individual at the centre, who is influenced by – in order of proximity to the individual – individual lifestyle factors (e.g., drinking, smoking, and other risky behaviours), societal and communal influences (e.g., crime and availability of local services), living and working conditions (e.g., educational institutions, workplaces, and housing), the macro-level environment (e.g., political, environmental, and economic conditions). Diderichsen, Evans, and Whitehead (2009) describe a health inequality framework whereby initial social stratification – influenced by macro-level factors – is amplified through differential exposure and vulnerability to ill health, followed by differential consequences thereof, ultimately leading to further stratification. This framework provides clear policy directives, suggesting effective interventions reduce social stratification, differential exposure, differential vulnerability, and differential outcomes in ill-health.

Fewer frameworks have been proposed explaining the development of health inequalities both within and between lifetimes. The growing body of research

indicating health inequalities later in life may originate in childhood. For example, life course epidemiology points to the first years of life as a critical period whereby initial social and health circumstances dictate later life health and social outcomes (Kuh et al. 2003). Some hardships experienced early in life may compound over the life course, through risk chains, and others may have latent effects, only arising later in life.

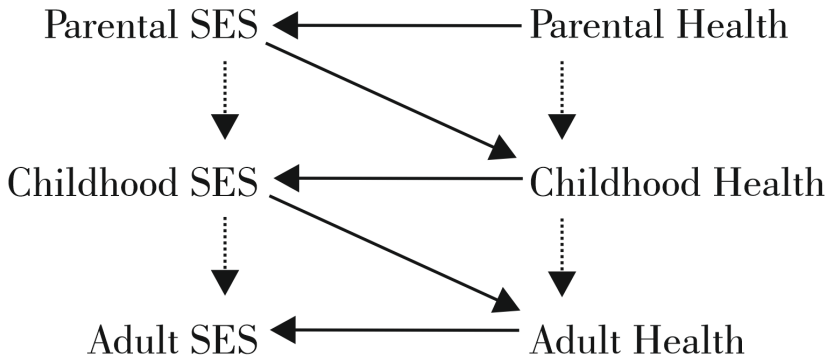
Pearce et al. (2019) adapt the Dahlgren and Whitehead (1991) and Diderichsen, Evans, and Whitehead (2009) frameworks to consider their applications to inequalities in child health. In the Dahlgren and Whitehead (1991) model, the most proximal factors influencing child health are the parenting and child health behaviours, followed by the parent and carer characteristics. In the Diderichsen, Evans, and Whitehead (2009) model, the differential exposure according to household socioeconomic position, differential vulnerability according to parental health, and differential consequences according to child health. As in the original model, social stratification is the determinant of health inequalities, and is further worsened through differential exposure, vulnerability, and outcomes. Through these adaptations, the conventional models consider the material, psychosocial, behavioural, and structural pathways to inequalities in child health.

This section reviews key empirical causal literature on the health-socioeconomic status relationship over the lifetime and between generations. In doing so, the review informs one example of a coherent explanation for the gradient across the lifetime and between generations, as seen in Figure 2.17.[10] Disentangling the causal mechanisms underlying social gradients in health is key to understanding these inequalities and designing policies to address them.

This model of lifetime gradient mechanisms (Figure 2.17) makes two simplifying assumptions: first, each arrow assumes exogeneity – so causality only runs in the direction indicated and not the opposite – and second, social causation is entirely captured by direct long-term effects and health selection is entirely captured by indirect short-run effects. Both assumptions are unlikely to be true, but are nonetheless useful in separating time effects and building simulation models in later chapters.

[10]Intra- and intergenerational transmission are assumed in this figure. Intragenerational transmission describes health (SES) in one life stage as predictive of health (SES) in the next life stage(s). Intergenerational transmission describes the same prediction, but between generations, which is more contentious but nonetheless extensively studied (Black and Devereux 2011; Currie and Moretti 2007).

Figure 2.17: Framework for health and socioeconomic status effects across generations and lifetimes.



Note: SES, socioeconomic status. Dotted lines indicate the transmission of the same variable from one life stage or generation to the next. Filled lines indicate gradient mechanisms; lines connecting SES to health represent social causation mechanisms and lines connecting health to SES represent health selection mechanisms.

Short-Term Causal Mechanisms

Childhood Health → Childhood Education

Most evidence on the relationship between health and SES in childhood supports the health selection hypothesis. Many studies examine the impact of early childhood health, and even prenatal environments, on late childhood educational attainment. This calls into question whether effects truly occur in the long or short term. The few studies on this topic find short-term effects are stronger than long-term effects. Evidence suggests the short-term gradient in childhood is driven by early childhood health predicting late childhood health, which then influences late childhood education (Case, Fertig, and Paxson 2005; Currie and Stabile 2003, 2006; Currie et al. 2010).

Mechanism: Mental Health

One of the key health selection mechanisms in childhood is the impact of mental health on education (Reiss 2013). Currie and Stabile (2006, 2009) use sibling fixed effects model with rich Canadian childhood administrative data to assess the impact of conduct disorders on educational attainment. Having a conduct disorder or higher hyperactivity and aggression scores had a threefold negative educational effect: decreased attainment, increased chances of repeating a year, and increased dropout rates. Similar results have also been obtained in the USA (Ding et al. 2009; Kessler et al. 1995), New Zealand (McLeod and Kaiser 2004; Miech

et al. 1999), Australia (Webbink et al. 2012), and Germany (Salm and Schunk 2012). In a later paper, Currie et al. (2010) study the effects ADHD diagnoses or conduct disorders on attending 12^{th} grade by age 17, finding large and substantial decreases in educational attainment caused by contemporaneous conduct disorders and ADHD. Overall, evidence suggests mental health issues during schooling years have significant impacts on educational attainment and retention.

Mechanism: Birth Weight
Lower birth weight has consistently been shown to predict lower educational attainment. Drawing on earlier non-causal evidence, Behrman and Rosenzweig (2004) use a monozygotic twin fixed-effects model to estimate the causal impact of fetal growth and birth weight on education. The authors find heavier infants stay in school longer. These results are consistent with other studies since, also using twin fixed-effects in a variety of contexts (Black, Devereux, and Salvanes 2007; Lin and Liu 2009; Royer 2009). Case, Paxson, and Bryan (2010) use greater height in early childhood as an instrumental variable for birth weight, finding taller children stay in school longer. Evidence therefore suggests the continued influence of birth weight on educational attainment later in childhood. It is not clear whether this effect is driven by indirect or direct effects, however if findings from the mental health literature are transferable, effects are likely to be driven by indirect effects through the prediction of later childhood health (Case, Fertig, and Paxson 2005; Currie and Stabile 2003, 2006; Currie et al. 2010).

Mechanism: Health Shocks
Physical health shocks during schooling years also impact educational attainment; de-worming and clean water policies have consistently demonstrated positive spillover effects on educational attainment (Baird et al. 2016; Bleakley 2007). However, the strongest evidence reporting the influence of health shocks on SES originates from lower- and middle-income countries (LMICs). Consequently, it may be questioned whether results are generalisable to countries with higher baseline levels of education. In a study of US children, Currie and Stabile (2003) report physical health shocks exhibit a gradient in the short-term but not in the long-term; lower-SES children do take longer to recover from health shocks but their long-term health outcomes are not significantly different from their higher-SES peers. Currie and Stabile's analysis therefore indicates the differing sensitivity to and recovery from health shocks are important determinants of the socioeconomic gradient in health in high-income countries as well, although only in the short term.

Mechanism: Chronic Health Conditions
The effect of chronic health conditions on education in childhood is contested. Salm

and Schunk (2012) find asthma reduces cognitive ability, however Currie et al. (2010) find no effect of asthma after controlling for health in adolescence. The effect of chronic health conditions on education appears to be larger for lower SES children, which aligns with human capital theory: higher SES parents manage their children's health better due to better health knowledge, widening inequalities (Galama and Kapteyn 2011; Grossman 1972; Kenkel 1991).

Summary
Most empirical evidence suggests health selection is the key short-term gradient determinant in childhood. These effects appear to be primarily driven by birth weight and contemporaneous mental health. However, recent evidence consistent with joint determination suggests the importance of non-cognitive skills, particularly in childhood (Chiteji 2010; Conti et al. 2010; Cunha, Heckman, and Schennach 2010). Non-cognitive skills may therefore explain some of the effect of childhood health on later educational achievement. However, when considering effect sizes, health selection appears to be the dominant causal pathway.

Adult Health → Adult Labour Market Outcomes
Labour markets play a critical role in the establishment, mobility, and retention of SES, capturing many of its constituents. In an early review of the literature, Currie and Hyson (1999) find substantial but varying impacts of health on labour market outcomes. Smith (2009) predicted this effect to be enhanced in early adulthood at labour market entry due to higher social mobility than later in the working life. These early works have since been supported by strong causal evidence suggesting health selection in labour force participation accounts for a large proportions of the gradient in adulthood.

Mechanism: Mental Health
Much of the health selection effect in adulthood may be explained by mental health, especially when considering smaller scale mental health impacts are rarely captured. Frijters, Haisken-DeNew, and Shields (2005) use an instrumental variable identification strategy on longitudinal Australian data to find large and significant effects of mental health on unemployment. The authors report one standard deviation reduction in mental health reduces employment by 30 percentage points. García-Gómez (2011) likewise reports large causal effects in various European countries but notes effect sizes are likely to vary by differences in social security arrangements.

Garcia-Gomez and Lopez-Nicolas (2006) study the Spanish population, finding health selection effects operate mainly through withdrawal from the labour market

(5% lower chance of being employed), however income is also directly affected. In a later paper, García-Gómez, Jones, and Rice (2010) note effects of health on labour are likely also heterogeneous among individuals, because work may be psychologically distressing or restorative, depending on the individual, their job, and the health shock sustained.

Thus far, studies have mostly been limited to large mental health shocks, while lower-level recurrent negative mental health shocks have not been causally interpreted. More research is needed to distinguish between large and recurrent mental health impacts on SES.

Mechanism: Physical Health
Among adults, physical health selection also drives the gradient. In the Netherlands, García-Gómez et al. (2013) find acute hospitalisations permanently reduce income and chances of employment by 5% and 7%, respectively across all workers. Worse still, the authors find health shocks have ripple effects onto dependants or carers; household income decreases by 50% more than individual income. Health shocks may therefore be even more damaging to household than individual SES.

Physical health shocks are particularly impactful among manual and older workers. Case and Deaton (2005) find evidence for health selection in the US among manual workers, whose health depreciates more rapidly with age than non-manual workers, despite upward health-based selection bias of those forced to withdraw from the manual labour market. Among adult men, Bound, Stinebrickner, and Waidmann (2010) find poor physical health outcomes are more likely to force early labour market withdrawal as they near retirement, causing a significant impact on earnings and thereby SES. An earlier study by French (2005) supports this assessment. French finds no effects of poor physical health on labour market outcomes for those under 38 years, but substantial effects for older adults, which peak around 62 years. The effects of physical health on SES begin to decline among those older than 66, even though individuals are more likely to suffer from poor physical health as they age. Evidence therefore suggests the importance of the labour market as a major channel through which the gradient operates.

Summary
Throughout early adulthood the dominant health selection mechanism appears to occur through mental health. However, as individuals near retirement, physical health becomes increasingly influential. Both mechanisms highlight the pivotal role of labour markets in determining the gradient. Evidence for health selection is further strengthened by the paucity of evidence suggesting income has a substantial

causal effect on health in adulthood (O'Donnell, Van Doorslaer, and Van Ourti 2015). Where evidence does exist for social causation in adulthood, it suggests long-term, but not short-term, changes in SES may influence health (Deaton 1999; Deaton and Paxson 1994; Kuehnle 2014).

Long-Term Causal Mechanisms

Parental Socioeconomic Status → Childhood Health

Long-term determinants of the gradient are worrying due to their ability to accumulate over time. This also pertains to the widely documented intergenerational transmission of inequalities; children with low-SES parents are more likely become low-SES adults themselves, and to have poor health (Almond, Currie, and Duque 2018). By uncovering some of the key causal mechanisms driving the intergenerational socioeconomic gradient in health, policy might be better informed to tackle this issue.

Mechanism: Maternal Education

The main evidence of social causation in early childhood has examined the effect of maternal education on birth weight. Currie and Moretti (2003) use US data to assess the impact of maternal education on birth weight. Using higher education availability during the future mother's seventeenth year as their instrumental variable (IV), the authors find higher maternal education increases birth weight and gestational age. The authors also focus on mechanisms through which this effect might occur, including smoking, marriage, fertility and use of prenatal care. Their results suggest higher maternal education increases the use of prenatal care, reduces parity, and renders mothers more likely to be married and less likely to smoke. These secondary results indicate prenatal health behaviours and environments are important mechanisms in explaining the effect of maternal SES on childhood health. Evidence presented here is consistent with later contributions to the literature, highlighting the importance of intergenerational social causation transmitted through parental health behaviours (Almond, Chay, and Lee 2005; Lien and Evans 2005; McCrary and Royer 2011; Royer 2009). Since health behaviours often take time to accrue before their effects are observed, it would suggest social causation through health behaviours occurs through direct long-term effects.

Mechanism: Parental Income

Income also accounts for some of the causality between parental SES and birth weight, however the evidence suggests effect magnitudes are generally small. Currie and Moretti (2007) found poverty raises the chances of transmitting poor health

outcomes, such as low birth weight, between generations. Increased parental earnings have also been shown to increase birth weight and decrease the probability of low birth weight slightly in multiple settings and using various identification strategies (Hoynes, Miller, and Simon 2015; Johnson and Schoeni 2011; Mocan and Altindag 2014). This small but causal effect of parental income on child health likely persists after birth and into childhood (Burgess, Propper, and Rigg 2004).

Kuehnle (2014) used exogenous variation in the labour market in an instrumental variable identification strategy, finding family income has a small but significant effect on self-perceived childhood health. Interestingly, Kuenhle's results indicate families make consumption decisions – including medical care and insurance – based on average income in the long run, not in the short run. This aligns with early economic theories proposed by Friedman (1957) and Modigliani (1966), as well as recent empirical evidence (Carlson 2015; Deaton and Paxson 1994; Golberstein, Gonzales, and Meara 2019). Evidence therefore suggests the impact of parental income on their children's future health is facilitated by long-term effects.

Mechanism: Parental Investment
Studies assessing parental time and resource investments in their child(ren) have often examined the impact of parental leave and maternal employment on child health. Overall, it seems to be a case of diminishing returns; short maternity leaves are very beneficial to child health when compared to little or no leave, however longer maternity rarely generate statistically significant additional benefits (Almond, Currie, and Duque 2018, pp 1406-1429).

Other research has focused on the effect of parental health knowledge on childhood health. Higher SES parents mayacquire better health knowledge, which is supported by human capital theory (Becker 2009; Grossman 1972) and empirical findings showing fewer children from high SES backgrounds become deaf as a result of common ear infections (O'Neill 1999). Furthermore, parental investments have also been shown to be influential in a child's development of non-cognitive skills, a key mechanism of the joint determination gradient hypothesis (Cunha and Heckman 2008). If parental investment in their child's development is predicted by parental SES, due to time and resource availability, it would suggest strong evidence for the influence of intergenerational social causation on the gradient and the joint determination hypothesis.

Mechanism: Childhood Residential Environment
The final key mechanism driving social causation in early life is the childhood environment. Parents who are better off are more likely to live in better areas, and

leave areas in the event of health hazards, thereby subjecting their children to fewer health risks (Currie 2011). Chay and Greenstone (2003), and others since, report pollution concentration increases infant mortality (Arceo, Hanna, and Oliva 2016). This is worrying because poorer households are more likely to be concentrated in areas of high pollution due traffic and congestion, for example (Currie and Walker 2011; Knittel, Miller, and Sanders 2016). Even after birth, pollution has been shown to cause poor health outcomes and chronic health conditions (Currie and Neidell 2005; Currie, Neidell, and Schmieder 2009). These health effects in early childhood are worse among low SES families (Neidell 2004). Evidence therefore suggests childhood residential environment, largely dictated by parental SES, explains some of the long-term social causation driving the gradient in childhood.

Summary
Once more, it is plausible that joint determination drives relationship between parental SES and childhood health. (Cunha and Heckman 2008; Cunha, Heckman, and Schennach 2010) find parental investments time and material resource investments are key to the development of non-cognitive skills in late childhood. However, evidence indicates higher SES parents invest more time and resources in their children, justifying the intergenerational long-term social causation pathway because childhood non-cognitive skill development may actually be driven by parental SES. It also seems likely that intergenerational effects occur if a child is very ill, forcing their parent(s) to withdraw from the labour market but no such evidence was found in this review of the literature.

Childhood SES → Adult Health
Perhaps the most widely documented long-term determinant of the gradient is education. Cutler and Lleras-Muney (2010) found education effects on health behaviours accounts for 60-80% of this effect. This supports the direct causal effect from childhood SES to adult health in Figure 2.17, since lifestyle effects take time to accrue and are rarely significant in the short-term.

Mechanism: Education
Most studies assessing whether the effect of SES on health is causal use educational reforms in their design, providing a strong identification strategy. Lleras-Muney (2005) used a change in compulsory education laws in the US and a regression discontinuity design on a synthetic cohort, finding one additional year of education lowers the risk of mortality by up to 3.6 percentage points and increases the life expectancy of 35 year-olds in 1960 by nearly two years. This effect is much larger than studies have found since, which may be explained by the context of the data: baseline education levels were low and the study took place in the US whereas

most studies have taken place in Europe.

Albouy and Lequien (2009) exploit two French educational reforms extending the compulsory school attendance age from 13 to 14 years (Zay reform in 1923) and then from 14 to 16 years (Berthoin reform in 1953). Using a regression discontinuity design, the authors find the survival rates are higher the more years are spent in education, however they are not able to prove causality. This study highlights some drawbacks of using mortality as a health outcome; age of death is a crude measure of health and may ignore other aspects of health affecting quality of life. This explains some of the controversy surrounding the causality of the education returns to health, since studies finding small or no causal health returns to education often use mortality (Clark and Royer 2016).[11] Another explanation is diminishing returns of education on health. This concept is reflected in wider evidence, suggesting later childhood interventions are often less effective than those implemented in early childhood (Cunha and Heckman 2007; Cunha et al. 2006). Despite these obstacles, the causal effect of education on health and it's importance in determining the gradient is still one of the most widely documented phenomena in the inequality literature (Cutler, Lleras-Muney, and Vogl 2008; Galama, Lleras-Muney, and Kippersluis 2018).

Health behaviours and lifestyle are key to explaining long-term effects of education on health. In one of the earliest works on the topic, Fuchs (1986) claims lifestyle is the key determinant of health variation once essential health standards have been attained. Recent evidence supports Fuchs' claim. Using smoking, drinking, exercising and body mass index as proxies for health behaviour, Brunello et al. (2016) find health behaviours account for approximately one quarter of the short-term effects of education on health and one third of the long-term effects. This may appear to suggest the comparable importance of health behaviours in the short- and long-term, however since long-term outcomes are generally more uncertain, this long-term predictive power of health behaviours is exceptional.

The evidence is consistent with the wider literature, claiming health differences can in large part be explained by health behaviours (Cawley and Ruhm 2011; Cowell 2006), particularly as individuals age (Mocan and Altindag 2014). Contoyannis and Jones (2004) note the importance of accounting for unobserved heterogeneity when estimating the influence of health behaviours and lifestyle. Their results imply many of the studies not properly accounting for endogeneity will underestimate the effect of lifestyle on the SES gradient by over 75%, explaining some of the controversy in the literature.

[11] A notable exception is the 2011 study by Kippersluis, O'Donnell, and Doorslaer (2011).

Summary
There is a wealth of evidence documenting long-term effects of education on future health. This effect largely operates through health behaviours, but as Cutler and Lleras-Muney (2010) note, the gradient is multifaceted and some of it still remains unexplained. While it is plausible non-cognitive skills mediate this pathway, at this point of writing the strength and quantity suggests the dominant intragenerational long-term direct effect runs from childhood SES to adult health.

Conclusions

This review of casual gradient evidence has uncovered some key mechanisms driving the life course SES-health relationship, informing the circumstance, choice, and outcome variables often selected in inequality of opportunity analyses. First, the evidence supports the influence of parental SES, particularly education, on prenatal environment and birth weight. These early childhood health conditions then predict later childhood health. This, in turn, influences educational attainment, which predicts adult SES. Educational attainment also influences a young person's future health, much of which can be explained by the positive effect of education on health behaviours and lifestyle. Finally, health in adulthood influences SES through the ability to participate in the labour market. This interaction of SES and health creates one example of a coherent explanation for the gradient across the lifetime and between generations as seen in Figure 2.17.

Long-term mechanisms of the gradient are particularly important due to their capability to widen disparities over lifetimes and generations. Short-term mechanisms may be equally influential within lifetimes if they occur early on, since their effects may compound. Understanding the time periods over which the gradient is determined and the key mechanisms through which it operates allows decision-makers to better target and prioritise policies aiming to reduce the socioeconomic gradient in health.

This review has revealed education and labour markets to be key mechanisms through which the SES-health gradient is generated, which are both influenced by short- and long-term mechanisms. However it should be noted that these findings may be influenced by publication bias, as well as the availability of data and natural experiments, and most of the socioeconomic variation in health remains unexplained.

2.4 Distributional Cost-Effectiveness Analysis

Distributional cost-effectiveness analysis (DCEA) is a framework for equity-informative health economic evaluation (Cookson et al. 2020). While other frameworks are also available, such as extended cost-effectiveness analysis (ECEA; Verguet, Kim, and Jamison (2016)) and multi-criteria decision analysis (MCDA; Marsh et al. (2016) and Thokala et al. (2016)), this thesis focuses on DCEA due to the ability to inform decision-makers on the trade-off between equity and efficiency, or cost-effectiveness.[12] Standard economic evaluation informs decision-makers on the efficiency of a given resource allocation; the costs , benefits, and opportunity costs of the allocation (Drummond et al. 2015). In the global context of constrained health care budgets and the need to maximise health benefits of a costly health-related investment, economic evaluation and standard cost-effectiveness analysis (CEA) have become increasingly important. However, inequality considerations have also risen on the policy and societal agenda, calling for the quantification of the inequality impacts of resource allocations.

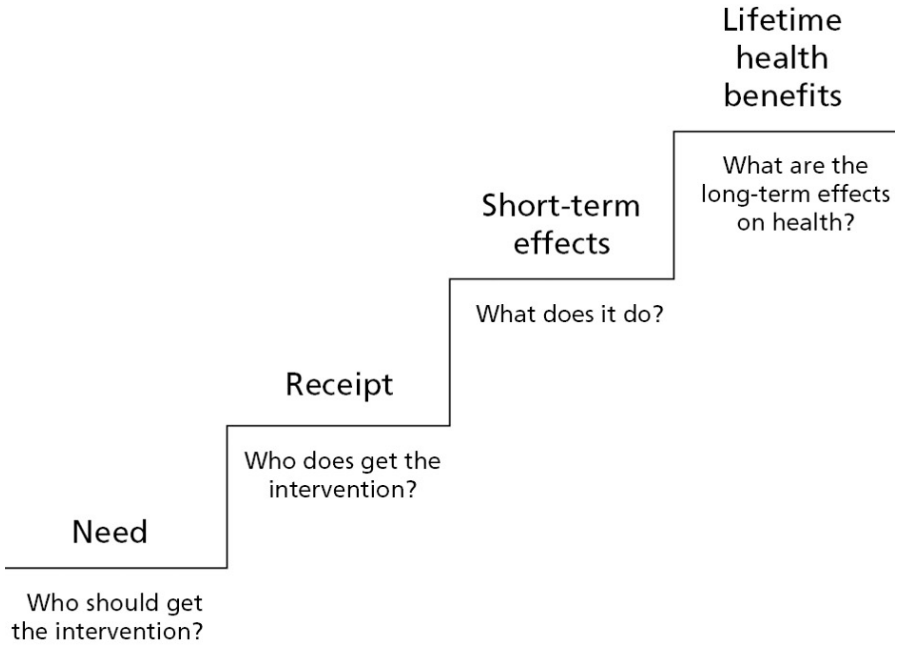
2.4.1 DCEA Design and Data Requirements

DCEA extends the data requirements beyond traditional CEAs. In theory, each step in designing a standard health economic model may be conditioned on a socioeconomic variable, provided data are available. In practice, the staircase of inequality provides a useful framework to consider the key stages during DCEA design in which distributional impacts might be considered (Figure 2.18).

The illustration facilitates discussion on the distributional inputs required for key model inputs. The first step considers whether health needs are unequally distributed in the analysis population. As an illustrative example, in designing a DCEA for a smoking cessation intervention, the model may consider prevalence of daily smokers (need), likelihood of seeking care (uptake), and intervention efficacy and adherence (short- and long-term treatment effects) across income quantiles. While the staircase captures key aspects of the model conceptualisation, which may be adapted to account for inequality, others such as treatment and health care costs may also be considered.

[12] For a comprehensive introduction to the underlying principles of DCEA and implementation guidance, see Cookson et al. (2020).

Figure 2.18: Staircase of inequality



Source: Cookson et al. (2020).

To model the relative incremental equity impacts of an intervention, it is necessary to estimate the baseline distribution of health by the stratifying variable of interest, such as deprivation (Love-Koh et al. 2015). To account for both quality and length of life, the quality-adjusted life expectancy (QALE) may be used to inform this distribution. QALE is a continuous ratio scale variable based on the quality-adjusted life year (QALY) as a generic measure of overall health, estimated through preference-based health related quality of life (HRQoL) questionnaires.

Likewise, while opportunity costs are often reflected in the cost-effectiveness thresholds in standard economic evaluations, the distribution thereof may also be a key factor in considering the equity impact of a given resource allocation. Opportunity costs estimations inform decision-makers on the benefits forgone by choosing a given resource allocation, which may not be equally distributed among deprivation groups. For example, evidence from the UK National Health Service suggests that for every investment in the health care system approximately 24% of the benefits are accrued by the worst off and 14% by the best off, indicating a pro-poor gradient (Love-Koh et al. 2020). However, this does not account for evidence that health care outcomes are often worse in more disadvantaged populations and so may exaggerate the extent to which more disadvantaged groups disproportionately benefit from additional health expenditure (Cookson et al.

2021c). Forthcoming evidence accounting for these factors suggests opportunity costs between deprivation groups are not statistically distinguishable from each other, and therefore do not follow a gradient.

2.4.2 Equity Impacts

Net health benefits may be used to visualise and calculate equity impacts, presenting the raw data without imposing value judgements on the distribution of health gains or losses across equity-relevant subgroups. Net health benefits for each group j in the variable of interest is calculated as:

$$NHB_j = \Delta QALYs_j - \Delta HOC_j \tag{2.9}$$

$$\Delta HOC_j = \frac{\Delta Costs}{MarginalProductivity} * P_j \tag{2.10}$$

Net health benefits, NHB, are measured in QALYs, equal to the incremental benefits, $\Delta QALYs$, less the health opportunity costs, Δ HOC, by SES quintile, $_j$. Health opportunity costs comprise the incremental cost of the intervention $\Delta Cost$ divided by the sector-specific marginal productivity rate, in GBP per QALY, $MarginalProductivity$, multiplied by the proportion of opportunity costs, P, falling on a given SES quintile, $_j$. This marginal productivity rate is informed by empirical estimations of the average health returns per monetary investment in the health care system, expressed in the UK and GBP per QALY. The resulting distribution of net health benefits across the stratifying variable may also be used to calculate bivariate inequality measures described in Chapter 2.1, providing an inequality impact summary measure.

2.4.3 Equity Trade-offs

To inform trade-offs between efficiency and equity impacts, DCEA uses an equity-regarding health-related social welfare function (SWF) by combining population health and health inequality into a single index (Asaria, Griffin, and Cookson 2016). An inequality aversion parameter applies implicit weights to socioeconomic groups according to their health relative to the total population average. There is considerable uncertainty around the value of inequality aversion

parameters. Inequality aversion parameters may be elicited using experiments in a representative sample, asking participants to trade between population health or health inequality improvements (Robson et al. 2017). There are several factors which may bias such experiments, including framing effects or selection bias, affecting the representativeness of the sample. To allow for decision-maker heterogeneity on this, an inequality aversion sensitivity analysis is often conducted to estimate the degree to which the magnitude of the inequality aversion parameter influences the equity-weighted cost-effectiveness of a resource allocation. Consequently, sensitivity analyses on the inequality aversion parameter are recommended (Asaria, Griffin, and Cookson 2016).

There are several possible SWFs to use in the DCEA framework, but in practice most empirical applications have used the Atkinson index, which measures relative inequality:

$$A_\epsilon = 1 - \left[\frac{1}{N} \sum_{i=1}^{N} \frac{H_i}{\bar{H}}^{(1-\epsilon)} \right]^{(1-\epsilon)} \tag{2.11}$$

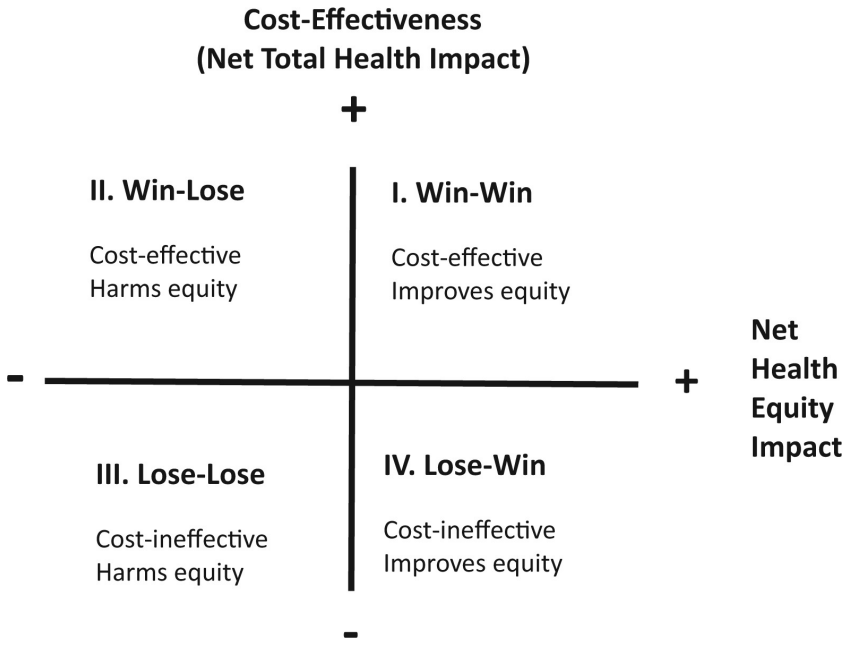
Where ϵ is the inequality aversion parameter, N is size of the total population, Q_i is the health estimate for a given individual ith, and $\bar{H}$ is the mean population health.

The equity impact plane uses a SWF to apply equity weights to the net health benefits of each group, thereby calculating a population equity-weighted net health benefit and illustrating the trade-off between equity and efficiency, reporting net health benefits on the y-axis, while reporting equity-weighted net health benefits on the x-axis (Figure 5.8).

2.4.4 Cross-Sectoral Economic Evaluation

Health economic evaluations typically take a single, health sector, perspective. While this is appropriate when benefits and opportunity costs of a resource allocation fall within a single sector, it may underestimate benefits for interventions which have spill-over benefits on non-health sectors. The free school meal (FSM) programme in England represents one such intervention. Historically, FSMs in England have been provided on a means-tested basis, whereby children living in households eligible for

Figure 2.19: Equity impact plane



Source: Cookson et al. (2017)

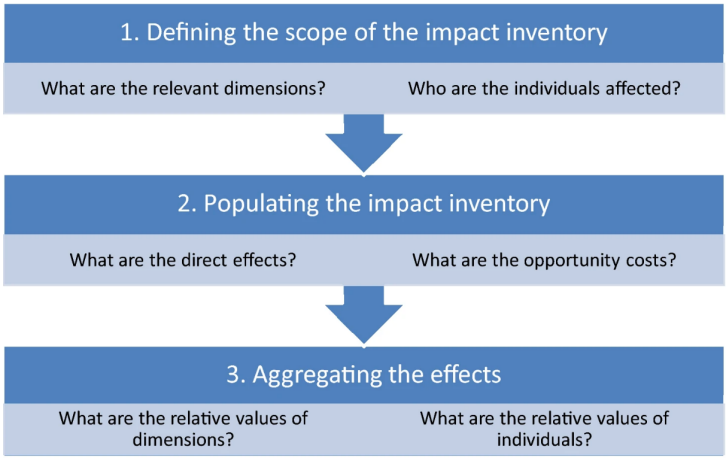
certain benefits may claim one FSM token for lunch every school day of the year. This is still the case for children in Year 3 and above in English state-funded schools, however since 2014, all children in English state-funded schools have access to FSMs their first three years of school (ages 4/5 to 7/8). This is a widely publicised scheme, especially upon the onset of COVID-19 pandemic, where lockdowns caused approximately half of the FSM-eligible children to lose access to five meals per week (Parnham et al. 2020).

Providing free school meals (FSMs) to all primary school children is often advocated on humanitarian and educational grounds, given the positive documented educational and social intervention outcomes (Cohen et al. 2021). However, FSMs may also have long-term benefits on health (Lleras-Muney 2005), facilitated by early changes to the childhood food environment and the multidirectional relationship between health and socioeconomic factors and health throughout the life course (See Chapter 2.2). The FSM scheme is currently funded by the Department for Education, concentrating the financial burden on a resource-constrained government arm. Given the benefits and opportunity costs are multisectoral, cross-sectoral frameworks for cost-effectiveness analysis are warranted.

A sophisticated cross-sectoral economic evaluation framework has been developed

by Walker et al. (2019), building on previous literature by Claxton et al. (2010), Claxton, Sculpher, and Culyer (2007), and Sculpher et al. (2014). The framework conceptualises an impact inventory capturing where the effects and opportunity costs of an intervention fall across all sectors considered in the analysis (See Figure A), which can inform compensation provisions to sectors with disproportionately high costs or low benefits. Similar to the health sector, cross sectoral opportunity costs are calculated according to marginal productivity, but also require estimates for non-health outcomes and consumption (Woods et al. 2016). The framework is therefore consistent with current resource allocation and decision-making approaches.

Figure 2.20: Impact inventory approach to conducting a cross-sectoral economic evaluation



Source: Walker et al. (2019)

Equity considerations may be especially important in the types of interventions that warrant cross-sectoral evaluations, such as public health interventions where the reduction of inequalities may be a key intervention outcome of interest (Weatherly et al. 2009). To incorporate equity considerations into the cross-sectoral economic evaluation framework through DCEA, analyses must estimate the distribution of benefits and opportunity costs across the relevant sectors, which is a substantial. Alternative approaches may also be considered based on a wellbeing QALY or "WELBY" as a generic summary measure of wellbeing (Cookson et al. 2021b). These methods may also be applied to the DCEA framework (Skarda, Asaria, and Cookson 2022). Similar concepts have also been recommended by NICE to capture non-health effects, including those to caregivers, based on the EuroQol Health and Wellbeing instrument (Brazier et al. 2022).

CHAPTER 3

Simulating the Impact of Income Inequality and Growth on Bivariate Health Inequality Measures

Background: Simulation studies have shown it is possible for bivariate health inequality measures (BHIMs) of income-health associations to increase when univariate income inequality decreases. However, these studies have not systematically examined the conditions under which health inequality measures may yield misleading results of this kind.
Aim: This simulation study investigates how different types of BHIMs respond to rank-preserving changes in the income distribution under varying causal assumptions.
Methods: The Pearson's correlation coefficient, slope index of inequality, relative index of inequality, concentration index, and generalised concentration index are estimated following rank-preserving changes in income inequality, absolute income growth, and relative income growth in three simulation models representing different plausible causal pathway assumptions: (i) income causes health; (ii) health causes income; (iii) income and health are both caused by soft skills.
Results: Rank-preserving changes to the income distribution that leave the health distribution unchanged do not influence rank-dependent BHIMs, but decrease level-dependent BHIMs. Assuming income causes health, rank-preserving increases in income inequality and growth increase rank-dependent BHIMs, such as the slope index of inequality, but decrease level-dependent BHIMs, such as Pearson's correlation coefficient.
Discussion: BHIMs only risk contradicting results when (a) using absolute, rather than ranked, income and/or (b) there is substantial income re-ranking. Guidelines should be established with suitable checks for whether health inequality trends could be misleading due to factors such as rank-altering changes in the income distribution.

3.1 Introduction

International comparisons of socioeconomic inequalities in mortality and morbidity suggest Nordic countries do not necessarily report lower levels of health inequality than non-welfare states (Eikemo et al. 2008; Mackenbach et al. 1997; Mackenbach et al. 2008, 2018; Murtin et al. 2022), despite having more generous social safety nets, which might be expected to improve health outcomes for the most socioeconomically disadvantaged populations (Beckfield and Krieger 2009; Muntaner et al. 2011). This finding has been termed the Nordic paradox (Mackenbach 2012), though studies using more detailed data on income-related inequality in mortality reveal a more nuanced picture; Norway, for example, has lower levels of health inequality than USA along the middle of the income distribution, if not in the lowest and highest percentiles (Kinge et al. 2019). However, until granular socioeconomic data become widely available, routine monitoring studies of socioeconomic health inequalities must rely on coarser data with fewer quantiles.

Previous studies have suggested standard bivariate health inequality measures (BHIMs) – linear associations between a health variable and a socioeconomic variable – are counterproductive because they respond in the opposite direction of social justice; when income inequality decreases, measured health inequality increases (the Nordic paradox critique). If true, this critique of standard BHIMs would have major policy implications, meaning that widespread national and international systems of health inequality monitoring and evaluation established in recent decades, with the support of international and governmental organisations, may actually harm the cause of equality and justice.

Brekke and Kverndokk (2012) use mathematical proofs and two simple simulations to show correlation coefficients, odds ratios, and concentration curves all report increased bivariate inequality when income inequality is decreased. However, it is unclear whether income ranks changed as a result of the variance change, and if so, the extent to which this may have influenced results. In testing these indices, the authors also assume causality runs from health to income. Consequently, changes to the income distribution do not change the health distribution. Before concluding their paper, the authors return to this causal assumption and claim the inadequacies of bivariate health inequality measures are in fact even worse under the assumption of social causation, however no systematic evidence is presented to support this conclusion. Renard et al. (2019) use simulated data to estimate the population attributable risk (PAR) and the slope and relative indices of inequality (SII and RII, respectively) of education inequalities in mortality rates. When allowing population

sizes of the educational attainment groups to fluctuate, the authors find the SII and RII could either increase or decrease.

Previous simulation studies have not systematically addressed the issues of rank-preserving changes to the income distribution and the effect of differing underlying causal pathways. This study simulates three plausible causal structures driving an income-health association to systematically assess the impact of rank-preserving income inequality and growth changes on bivariate health inequality measures. Bivariate inequality measures are assessed against the normative criterion that they should be insensitive to changes in income that leave health unchanged.

The chapter is structured as follows. Section 2 provides a conceptual framework for bivariate health inequality measurement, describing background concepts and deriving a normative criterion to assess measures against. Section 3 presents the simulation models and analyses. Section 4 presents the simulation results from changes in income inequality, level, and growth. Section 5 summarises findings, compares them with previous research, states strengths and limitations, draws policy implications, and makes recommendations for future research.

3.2 Conceptual Framework

Bivariate health inequality measures (BHIMs) estimate the association between a health variable (e.g. mortality rates or life expectancy) and a socioeconomic variable (e.g. deprivation or income). BHIMs therefore aim to document unfair inequalities on the grounds that health should not be systematically associated with socioeconomic variables, such as income, occupation, and education (Braveman and Gruskin 2003; Marmot et al. 2008). For example, it may be considered unfair if having low income causes ill health, and equally it may be unfair if having poor health causes someone to be poor.

The socioeconomic variable in BHIMs may either be a rank or level variable according to the BHIM (see Chapter 2.2). Popular BHIMs such as the slope index of inequality (SII) or the concentration index (CI) use a weighted fraction-rank of the socioeconomic variable to facilitate comparisons between populations and over time where the distribution of the socioeconomic variable may differ.

3.2.1 Normative Criterion

Previous simulation studies have criticised BHIMs, claiming they are inadequate policy tools due to potential confounding by distributional changes in the stratifying socioeconomic variable (Brekke and Kverndokk 2012; Renard et al. 2019). However, these assertions presuppose a normative judgement about what BHIMs aim to measure. To make claims regarding their adequacy, it is necessary to identify the normative criteria for BHIMs. Establishing these criteria will likely involve an iterative process between researchers and decision-makers, however the following criterion seeks to initiate this discussion and form the basis for analysis in this chapter.

Criterion: Bivariate health inequality measures should be insensitive to changes in income which leave health unchanged.

The criterion captures the main premise of Brekke and Kverndokk (2012) and Renard et al. (2019); insofar as BHIMs are primarily concerned with health inequalities, distributional changes in the socioeconomic variable (such as income) which leave the distribution of health unchanged should not influence BHIMs. Otherwise it is possible for BHIMs to respond in the opposite direction of social justice; when income inequality decreases, measured health inequality increases. This threatens the national and international evidence base built upon these measures.

So, it is plausible policy-makers would generally not want distributional changes in the socioeconomic variable to influence BHIMs, unless these also lead to changes in the health distribution. This extends to several changes in the income distribution, including increases in income inequality, absolute income growth, and relative income growth.

3.2.2 Impact of Income Distribution Changes

In this section, I explain the intuition behind the Nordic paradox critique of BHIMs, using a diagrammatic approach. I illustrate BHIM responses to changes in income inequality, absolute income growth, and relative income growth. Income inequality is assumed to be reflected in the variance of a continuous income variable. Absolute

income growth refers to a context where the income of all individuals changes by the same absolute quantity, such as 250 GBP higher net income per month. Relative income growth, on the other hand, refers to a context where income changes for all individuals by the same relative amount, a 5% increase in annual net income, for example. When relative income growth occurs, both the mean and variance of income changes, therefore combining changes in income inequality – where the variance changes but the mean does not – and absolute income growth – where the mean changes but the variance does not.

Income Inequality

Figure 3.1 illustrates a bivariate inequality metric under a scenario in which the income variance increases but income does not cause health, where a greater gradient indicates greater inequality. Moving from income distribution A to B, the income distribution widens, thereby decreasing the BHIM, despite no changes occurring in the health distribution. These BHIM changes may be considered undesirable, since it is not clear that greater income inequality should be represented as lesser income-related health inequality, when monitoring populations.

Figure 3.1: Impact of income inequality on level-dependent bivariate health inequality measures when income does not cause health.

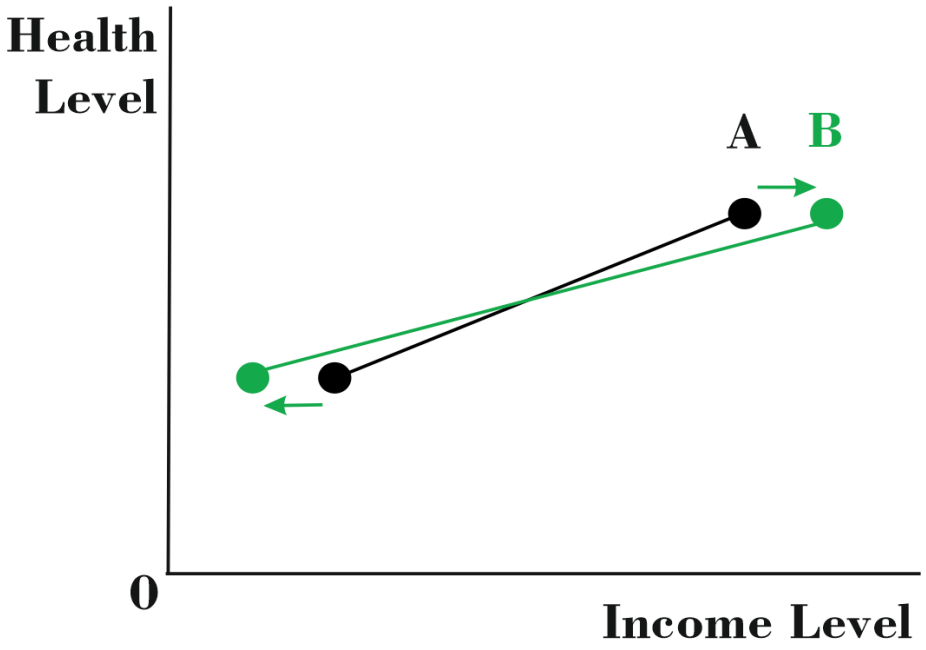


Figure 3.2 illustrates the same scenario as Figure 3.1, however in a context where

income causes health, hence the increase in income variance is reflected in an increased variance of the health distribution. Consequently, the BHIM change is ambiguous, it may increase or decrease depending on the magnitude of changes in the income and health distributions.

Figure 3.2: Impact of income inequality on level-dependent bivariate health inequality measures when income causes health.

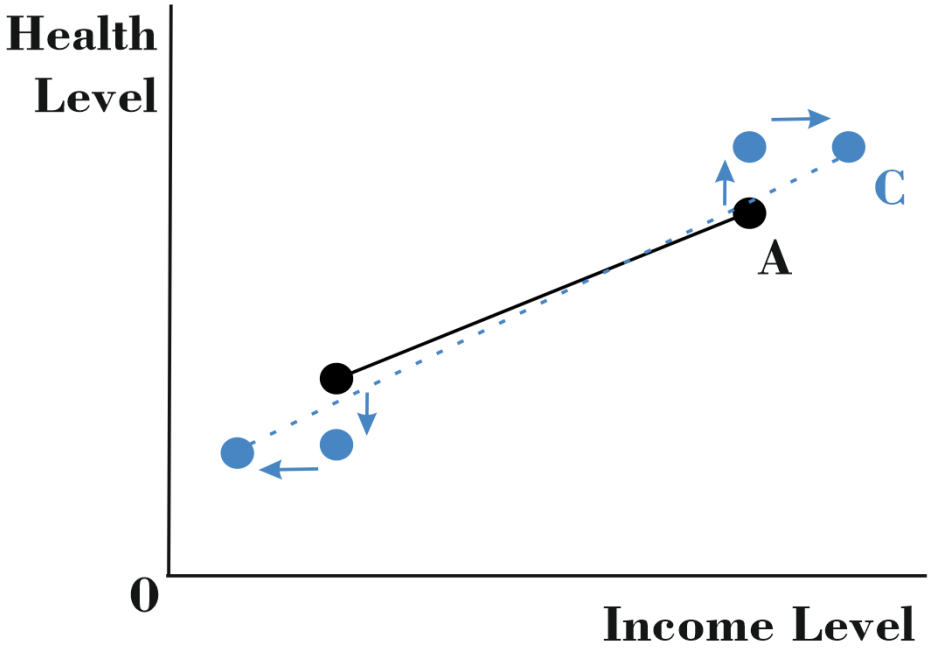
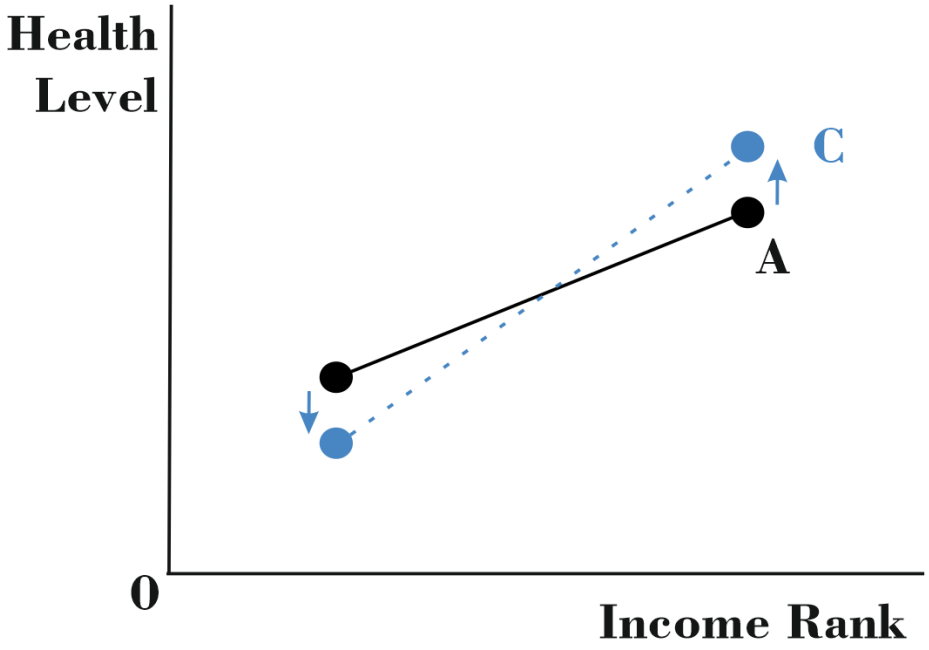


Figure 3.3 illustrates the same scenario as Figure 2, but uses a rank-dependent bivariate index. Consequently, increased income variance does not change the income ranking, however the increased variance of the health distribution is reflected, increasing the BHIM.

Figure 3.3: Impact of income inequality on rank-dependent bivariate health inequality measures when income causes health.



Absolute Income Growth

Using the intuition from Figures 3.1-3.3, income level changes are not expected to change absolute BHIMs, whether they are rank- or level-dependent, because the change does not influence the slope of the gradient in the illustrative figures. However, relative BHIMs are expected to decline with increasing income level if income causes health, to reflect the resulting greater mean health.

Relative Income Growth

Relative income growth changes are expected to have ambiguous impacts on level-dependent BHIMs. If income does not cause health, relative income growth increases are expected to decrease level-dependent BHIMs, and vice versa, due to the resulting greater income inequality. When income does cause health, positive income growth may either increase or decrease level-dependent BHIMs because, for example, and increase in both income inequality and level would be reflected by a complementary increase in health inequality and level. In the same scenario, rank-dependent BHIMs are expected to remain unchanged when income does not

cause health and are expected to increase when income causes health, due to the subsequent changes to the health distribution, where the health of the most advantaged improves more than that of the least advantaged.

3.3 Methods

3.3.1 Simulation Models

This study uses three investigational simulation models based on causal pathways driving the income-health relationship: social causation (income causes health); health selection (health causes income); and joint determination (income and health caused by a third variable, e.g. noncognitive skills), as described in Chapter 2.2. Variables in the models are selected for illustrative purposes. I use ratio scale variables, such that both rank- and level-dependent BHIMs can be used. Beyond this, the precise distributional characteristics and resulting inequality magnitudes are not the focus of this study, which is instead on patterns of the changes in inequality metrics.

Health is represented by health-adjusted life expectancy (HALE) at birth, with imposed lower and upper bounds at 0 and 120 years, respectively. I assume HALE at birth is normally distributed, with a mean roughly equal to Office for National Statistics (ONS) findings from the UK in 2017-19 of approximately 63 years (ONS 2021). Income is represented by ONS estimates of equivalised annual household net income in 2018, with an average of approximately 34 000 GBP and a right-skewed distribution (ONS 2019). I impose a lower bound at zero and no upper bound.

For the joint determination variable, I use noncognitive, or soft, skills due to the evidence on their causal effect on both income and health (Chiteji 2010; Heckman, Stixrud, and Urzua 2006). While there are many methods for measuring noncognitive skills, I simulate data informed by empirical measures using a seven-point scale (1-7) across the Big Five personality traits: extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience (Cobb-Clark and Schurer 2012). The resulting variable is normally distributed on a scale 5-35, with a mean of 25 and imposed upper and lower bounds.

Each simulation model follows a two-step process (see equations 3.1-7)

Social Causation

$$Income_i = LogNormal(Mean, StDev) \tag{3.1}$$

$$Health_i = \beta * LogIncome_i + HealthCovariates_i \tag{3.2}$$

Health Selection:

$$Health_i = Normal(Mean, StDev) \tag{3.3}$$

$$Income_i = exp(\gamma * Health_i + IncomeCovariates_i) \tag{3.4}$$

Joint Determination

$$NoncognitiveSkills_i = Normal(Mean, StDev) \tag{3.5}$$

$$Health_i = \delta * NoncognitiveSkills_i + HealthCovariates_i \tag{3.6}$$

$$Income_i = exp(\rho * NoncognitiveSkills_i + IncomeCovariates_i) \tag{3.7}$$

To simulate social causation, annual equivalised household net income is first simulated for 10 000 individuals according to a log-normal distribution, with a mean and standard deviation chosen to approximate the UK distribution in 2018 (ONS 2019) (Equation 3.1). The log-normal distribution is chosen to reflect the right-skewed distribution of income. Second, health-adjusted life expectancy is simulated for each individual by: (i) multiplying the log income by a coefficient, β, reflecting the assumed causal effect of current household income on health-adjusted life expectancy; and (ii) adding a randomly sampled value from a normal distribution that reflects a vector of health covariates, assuming complete observability (Equation 3.2). Log income is used to normalise the right-skewed income distribution.

To simulate health selection, health-adjusted life expectancy is first simulated for 10 000 individuals according to a normal distribution, with a mean and standard deviation chosen to approximate the UK distribution in 2017-19 (ONS 2021) (Equation 3.3). Second, the health-adjusted life expectancy of each individual is multiplied by a coefficient, γ, and added to a random sample of income covariates from a normal distribution, the sum of which is exponentiated to calculate the

annual equivalised household net income for each individual (Equation 3.4). The sum is exponentiated to reflect the relationship between health and income, and the right-skewed distribution of income.

To simulate joint determination of health and income, noncognitive skills are first simulated for 10 000 individuals according to a normal distribution, with a mean and standard deviation chosen to approximate published empirical evidence (Cobb-Clark and Schurer 2012) (Equation 3.5). The health and income of each individual is then simulated simultaneously in equations 3.6 and 3.7, respectively. The health-adjusted life expectancy is simulated for each individual by: multiplying noncognitive skills by a coefficient, δ and adding a randomly sampled value from a normal distribution that reflects a vector of health covariates, assuming complete observability. The annual equivalised net household income of each individual is calculated by multiplying noncognitive skills by a coefficient, ρ, adding a randomly sampled value from a normal distribution to reflect observed income covariates, and exponentiating the sum.

3.3.2 Bivariate Health Inequality Measures

I measure income-related inequalities in health-adjusted life expectancy using five bivariate health inequality measures, as described in Chapter 2.2: the Pearson's correlation coefficient (PCC), slope index of inequality (SII), and relative index of inequality (RII). The concentration index (CI) and generalised concentration index (GCI) are reported in Appendix A.2 for completeness. In this study, the PCC reflects absolute level-dependent BHIM, whereas the SII and RII reflect absolute and relative rank-dependent BHIMs, respectively. This allows for coverage across popular monitoring indices and key properties thereof.

The Pearson correlation coefficient (PCC) yields results ranging from -1 to 1, with values greater than (less than) zero indicating a positive (negative) relationship, and a value of 0 indicating no relationship. The greater the absolute value of the PCC, the greater the income-related inequalities in health-adjusted life expectancy. SII values greater than zero indicates higher income is associated with greater health-adjusted life expectancy tend to have better health and vice versa. For example, an income-life expectancy SII of 10 indicates the person with the highest predicted income has a life expectancy 10 years greater than the person with the predicted lowest income rank. An RII estimate equal to one indicates no relative advantage between predicted lowest and highest socioeconomic ranks. If the RII is greater (less) than one, higher income is associated with relatively longer (shorter)

health-adjusted life expectancy. For example, an income-life expectancy RII of 1.1 indicates the person with the highest predicted income has a 10%, or 1.1 times, greater life expectancy than the person with the predicted lowest income rank.

3.3.3 Analyses

I systematically assess changes to BHIMs in response to changes in the income inequality, absolute income growth, and relative income growth by applying change vectors the income standard deviation, mean, and a multiplicative factor, respectively. Changes in the income distribution are reported in relative terms – as a percentage change in the income standard deviation, mean, or growth. Changes in the bivariate index of inequality are also represented in relative terms to ensure comparability across simulation models and inequality indices. For presentation purposes, main graphs include the PCC, SII, and RII, while results with CI and GCI are included in Appendix A.2.

Models are simulated in *R v4.1.2* and ranks are fixed using R's *set.seed()* function.[1] The models are assumed to be exogenous and the β, γ, δ, and ρ parameters in Equations 3.1-3.7 can therefore be assumed to reflect causal relationships. It is also assumed the models can be estimated with linear measures, since BHIMs often rely on this assumption. Histograms of the income and health distributions are available in Appendix A.1.

3.4 Results

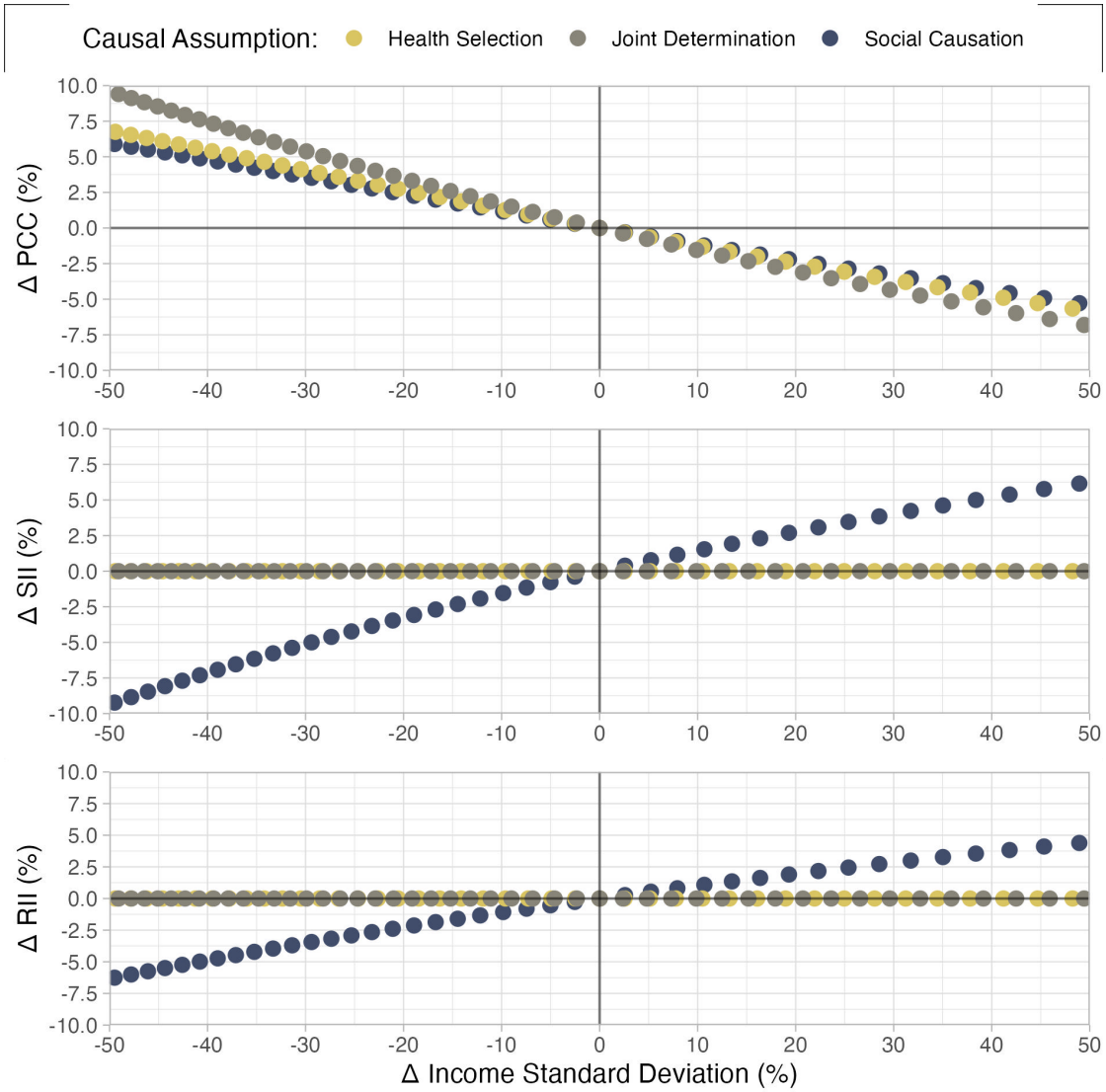
3.4.1 Income Inequality

Figure 3.4 shows the impact of changing the standard deviation of the income distribution, which may be considered a proxy for income inequality, on the Pearson's correlation coefficient (PCC), slope index of inequality (SII), and relative index of inequality (RII) under social causation, health selection, and joint determination assumptions. Increasing income inequality decreases the PCC, a level-dependent absolute bivariate health inequality measure (BHIM), under all

[1] Results for rank-altering changes in the income distribution can be found in Appendix A.2., replicating the results of Brekke and Kverndokk (2012).

three causal assumptions. Variations in magnitude of changes are assumed to be a feature of the simulation models, rather than an indication of empirical findings. For the SII and RII, rank-dependent BHIMs, increasing income inequality increases bivariate inequality under the social causation pathway but are unchanged under the health selection and joint determination pathways. This occurs due to widening health distributions from widening income distribution. Due to the log relationship between income and health, the SII and RII respond non-linearly to changes in income inequality, increasing at a decreasing rate indicating SII and RII changes are greater (lesser) when income inequality is lesser (greater).

Figure 3.4: Impact of rank-preserving income inequality changes on Pearson's correlation coefficient, slope index of inequality, and relative index of inequality under social causation, health selection, and joint determination causal pathway assumptions.

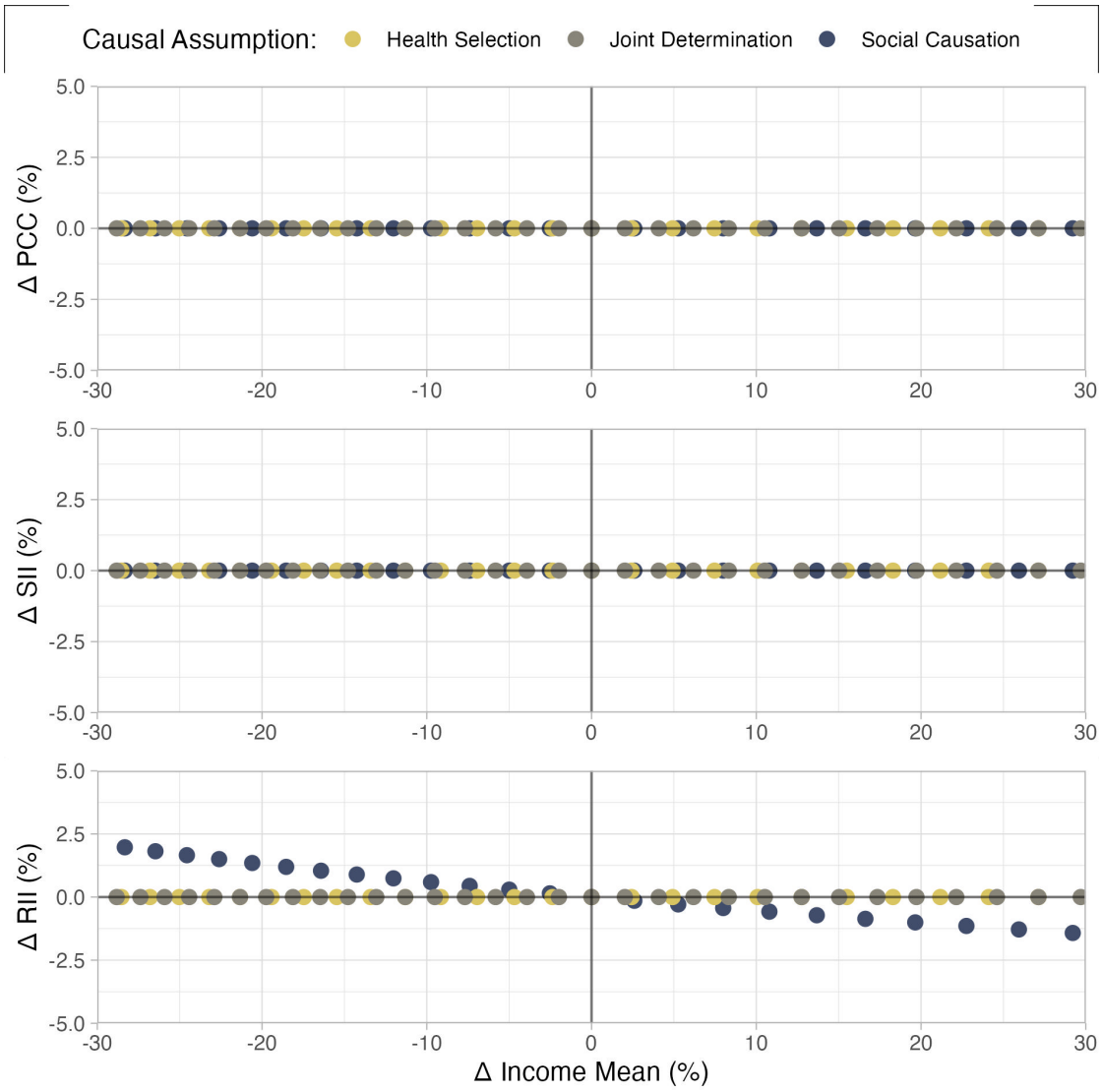


Note: Income standard deviation is assumed to represent univariate income inequality. Social causation assumes income causes health, health selection assumes health causes income, and joint determination assumes income and health are both caused by noncognitive skills. Abbreviations: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality.

3.4.2 Absolute Income Growth

Figure 3.5 shows the impact of changing the income distribution mean, while retaining the same income variance, on the PCC, SII, and RII under the social causation, health selection, and joint determination causal assumptions. Rank-preserving changes to income mean do not impact the PCC or SII under any causal pathway included in the simulation. The RII is also insensitive to income mean changes under the health selection and joint determination pathways. However, increasing income mean changes decreases the RII when income causes health to reflect higher health levels. The magnitude of RII changes are small and nonlinear, such that RII decreases are greater (lesser) when income growth is lesser (greater). Increasing (decreasing) income growth by 30% leads to an approximately 1.5% decrease (2.5% increase) in RII.

Figure 3.5: Impact of rank-preserving income level changes on Pearson's correlation coefficient, slope index of inequality, and relative index of inequality under social causation, health selection, and joint determination causal pathway assumptions.



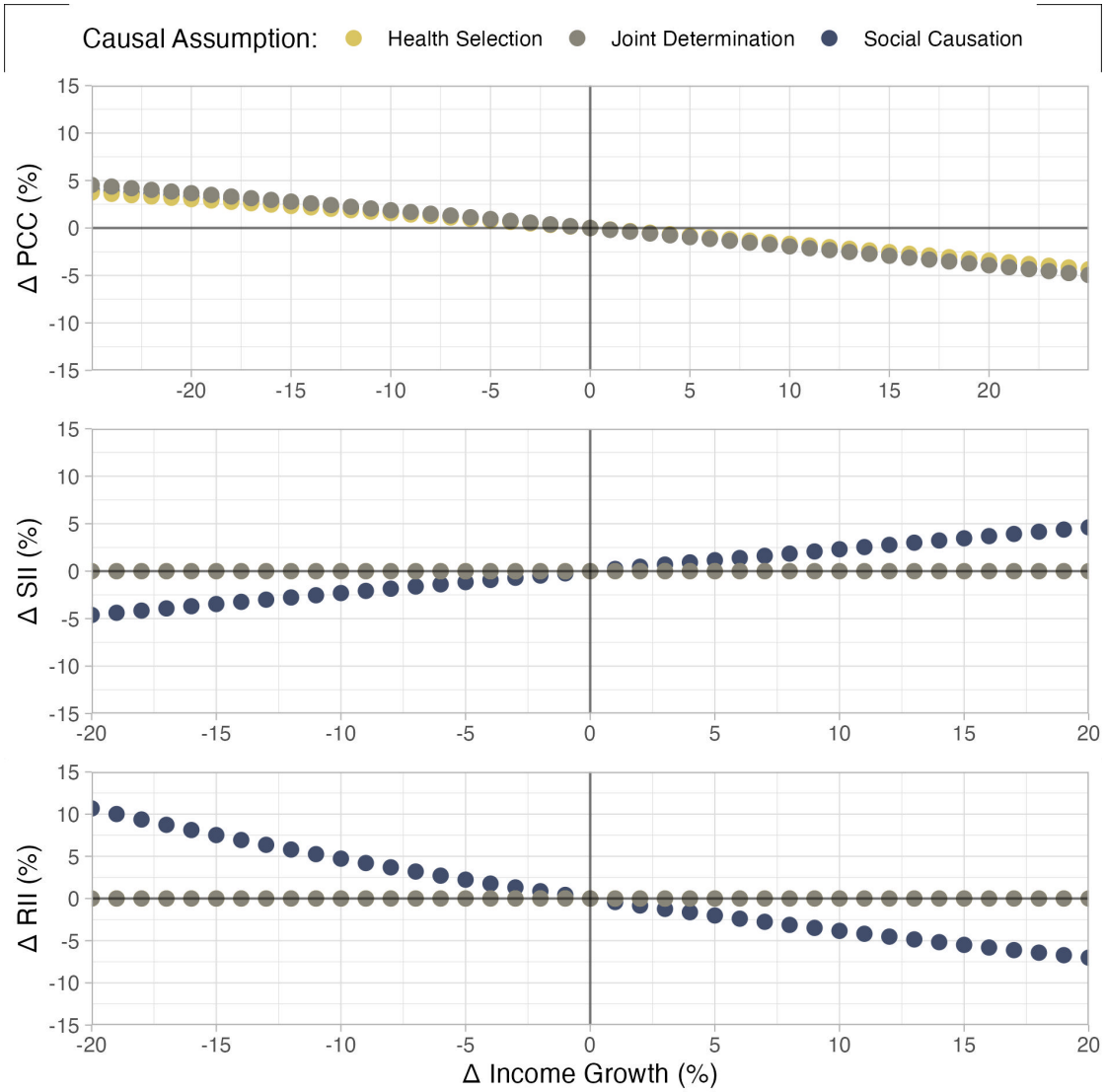
Note: Social causation assumes income causes health, health selection assumes health causes income, and joint determination assumes income and health are both caused by noncognitive skills. Abbreviations: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality.

3.4.3 Relative Income Growth

Figure 3.6 shows the impact of rank-preserving relative income growth – increasing both income inequality and levels – on the PCC, SII, and RII under social causation, health selection, and joint determination causal pathway assumptions. Under the health selection and joint determination pathways, income growth changes do not affect the SII or RII, but lead to decreases in the PCC, reflecting changes in income inequality. Under social causation assumptions, increasing income growth leads to lower PCC and RII, while the SII increases.

SII changes are linear, and a 20% increase (decrease) in income growth results in an approximately 5%, reflecting the strength of the causal coefficient dictating the income-health relationship. RII changes are non-linear; increases in income growth lead to greater (lesser) changes in RII at lower (higher) income growth levels, such that a 20% decrease (increase) in income growth is reflected by an approximately 10% increase (7.5% decrease) in RII.

Figure 3.6: Impact of rank-preserving income growth changes on Pearson's correlation coefficient, slope index of inequality, and relative index of inequality under social causation, health selection, and joint determination causal pathway assumptions.



Note: Income growth is simulated using a multiplicative factor on the income distribution, thereby changing both income mean and variance. Social causation assumes income causes health, health selection assumes health causes income, and joint determination assumes income and health are both caused by noncognitive skills. Abbreviations: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality.

3.5 Discussion

3.5.1 Primary Findings

This study assesses the impact of rank-preserving changes to income inequality, level, and growth on bivariate health inequality measures (BHIMs) across three causal assumptions: social causation, where income causes health; health selection, where health causes income; and joint determination, where income and health are both caused by a third variable, which is proxied by noncognitive skills in this study. The results show BHIMs only risk results that contradict social justice when (a) using absolute, rather than ranked, income and/or (b) there is substantial income re-ranking.

Rank-preserving changes to the income distribution that leave the health distribution unchanged do not influence rank-dependent BHIMs, but decrease level-dependent BHIMs. Assuming income causes health, rank-preserving increases in income inequality and growth increase rank-dependent BHIMs, such as the slope index of inequality (SII), but decrease level-dependent BHIMs, such as Pearson's correlation coefficient (PCC). Consequently, a country may report greater income-related health inequalities, solely due to a narrower income distribution, when using level-dependent BHIMs. Increases in income levels may also decrease relative rank-dependent BHIMs when income causes health due to increases in health levels.

3.5.2 Comparison with Previous Findings

Brekke and Kverndokk (2012) and Renard et al. (2019) find BHIMs are inadequate policy monitoring tools. Both studies use simulated data to show some BHIMs may increase when inequality in the socioeconomic variable decreases, even when there is no change in the health variable. However, these studies were subject to some limitations, explaining the differences in findings from this study.

Brekke and Kverndokk (2012) use mathematical proofs and two simulations to show correlation coefficients, odds ratios, and concentration curves all report increased bivariate inequality when income inequality is decreased under health selection causal assumptions. Under the same conditions, results align with those in this study. However, Brekke and Kverndokk do not systematically address the issues

of rank-preserving income changes and differing causal assumptions. While their findings hold true for level-dependent BHIMs, which may respond undesirably to changes in the income distribution, this is not the case for rank-dependent BHIMs following rank-preserving income distributional changes.

Renard et al. (2019) use simulated data to estimate educational inequalities in mortality using population attributable risk (PAR) and the slope and relative indices of inequality of education inequalities in mortality rates. The authors find increasing mean educational attainment increases the reported bivariate inequality by the SII and RII, due to changes in the socioeconomic distribution when using a categorical socioeconomic variable, such as education. This provides useful insights to the implications of changing distributions of categorical socioeconomic variables on socioeconomic health inequalities.

3.5.3 Strengths and Limitations

This study systematically simulates the impact of income inequality and growth changes on BHIMs, expanding on previous literature by addressing three plausible causal assumptions driving the income-health association: social causation, where income causes health; health selection, where health causes income; and joint determination, where income and health are both caused by a third variable, which is proxied by noncognitive skills in this study. The study also adds to the literature by illustrating the impact of rank-preserving income changes on BHIMs.

This study is subject to limitations. First, this is a simulation study, which simplifies the complex empirical income-health relationship. For example, all three causal structures may influence the income-health relationship simultaneously. Second, while data were used to inform simulation model inputs, these may differ substantially between empirical contexts. The distributional properties of and relationships between income and health may therefore not be representative of all populations, such as those in lower- and middle-income countries. Finally, this simulation study has not considered the impact of socioeconomic rank mobility on BHIMs. Previous literature has shown income mobility may account for up to half of the income gradient in life expectancy in Denmark (Kreiner, Nielsen, and Serena 2018). It may therefore be expected that changes to mobility in the income distribution would reduce the magnitude of changes in BHIMs.

3.5.4 Policy Implications

This study indicates rank-dependent BHIMs may be appropriate policy instruments for monitoring socioeconomic health inequalities to the extent that income ranks are stable. For example, findings from the updated Marmot Review suggest socioeconomic health inequalities have increased in the UK while income inequality has remained mostly unchanged (Marmot et al. 2020). Consequently, the income distribution may not have altered substantially and conclusions on the development of socioeconomic health inequalities are more plausible.

To assess the appropriateness of BHIMs, estimates should also be accompanied by sensitivity analyses on income re-ranking and the univariate health and income distributions. Income re-ranking may be assessed using metrics from the income inequality literature on horizontal equity (Atkinson 1980; Kakwani 1984; Plotnick 1981). When income re-ranking is found to be substantial, more advanced methods that account for re-ranking in the weighted fraction-rank of the socioeconomic variable may be considered (Allanson, Gerdtham, and Petrie 2010; Coveney et al. 2020; Van Ourti, Doorslaer, and Koolman 2009). Univariate analyses of the income and health mean and variance provide insights into the underlying distributional changes and the income-health relationship. While causal analyses are not feasible in routine health inequality measurement, univariate distributional checks of income and health provide insights into possible drivers of changes in rank-dependent BHIMs. Furthermore, univariate distributional checks are able to capture spurious changes in BHIMs due to bounds in income (e.g. zero lower bound) and health (e.g. zero lower bound and biological upper bound on life expectancy). This may be especially important in developing countries with higher prevalence of zero income or high infant mortality. In these scenarios, increases in mean health levels – whether caused by income or otherwise – may increase reported bivariate health inequality.

3.5.5 Future Research

Future research should explore the impact of income mobility on the consistency of BHIM responses to changes in the income distribution in inequality monitoring environments. This sheds light the impact of relaxing the rank-preservation assumption in this study. For example: if income inequality and mean remained constant but income mobility increased, how would this impact BHIMs?

There are at least two key avenues to ensure income re-ranking is considered for

the routine monitoring of health inequalities. First, by using more complex health inequality metrics developed to account for income re-ranking, drawing from the literature on measuring health-related income mobility (Jones and López Nicolás 2004) and income-related health mobility (Allanson, Gerdtham, and Petrie 2010). However, these methods may exhibit barriers to entry for national-level statistical offices. A second method would therefore be to examine how simple estimates of income mobility assessments, such as the rank-rank slope may be applied to assess the degree of income rank comparability between populations, especially in international comparisons (Chetty et al. 2014). A future research study may investigate both approaches, applying and comparing them in a case study using previous measurements of socioeconomic health inequality. An impactful output of this study would be clear guidance to decision-makers on inequality measurement contexts in which income re-ranking should be accounted for, and how to do so, using practical tools already in common use by national statistical offices where possible.

Future research may also explore the possibility of allowing multiple causal models to dictate the income-health relationship, which may be more reflective of the empirical environment. This simulation would also allow insights into the impact of changing causal parameter strength and relative contribution of causal pathways in BHIMs, shedding further light on methodological groundings for explanations of the Nordic paradox. In doing so, future research should contribute to establishing guidelines for BHIM implementation.

CHAPTER 4

Comparing Parental and Current Socioeconomic Gradients in Health: A Simple Inequality of Opportunity Metric Applied to a UK Birth Cohort

Background: Sophisticated lifetime inequality of opportunity (IOp) measures require disentangling causal networks between lifetime circumstances and choices throughout a person's life. Bivariate associations between childhood circumstances and adult health might provide a simple indicator of societal lifetime IOp in health.
Aim: This study compares conventional current social gradients in health with parental social gradients: bivariate associations between fraction-ranked parental socioeconomic status in childhood and adult health.
Methods: I calculate gradients using slope and relative indices of inequality in 36-item Short Form Health Survey (SF-36) scores at age 50 from the 1958 National Child Development Study, applying multiple imputation methods to account for attrition. Socioeconomic variables include income, education, and occupational class, measured across childhood, ages 0-16, and in adulthood at age 50.
Results: Parental socioeconomic gradients in adult health at age 50 are significant and substantial; growing up in the most socioeconomically advantaged household is associated with 7-15 points, or 12-50%, higher SF-36 scores at age 50, compared to growing up in the least socioeconomically advantaged household. These parental gradients are roughly 40-65% the magnitude of current socioeconomic gradients and remain significant after controlling for socioeconomic variables in adulthood.
Discussion: Parental gradients are accessible summary measures of "unfair" inequalities, which are easy to implement, interpret, and compare. Results indicate current socioeconomic gradients in adult health may be substantially determined during childhood and adolescence for people born in the UK in 1958, especially for static socioeconomic variables such as education.

4.1 Introduction

Conventional bivariate measures of health inequality estimate "current" socioeconomic gradients in health: associations between health and socioeconomic status (SES) measured at the same point in time. These bivariate inequality measures have consistently found higher current SES is associated with better current health (Mackenbach et al. 1997; Mackenbach et al. 2008, 2018; Marmot et al. 2008). However, conventional current bivariate gradients have been criticised because they may partly reflect "fair" outcome differences from choices for which people can be held responsible. Furthermore, research has increasingly shown adult health may also trace back to parental SES in childhood (Almond and Currie 2011; Almond, Currie, and Duque 2018; Conti et al. 2010). This raises serious concerns about inequality of opportunity (IOp), since it is reasonable to assume childhood circumstances are beyond personal choice or responsibility (Jusot and Tubeuf 2019; Roemer and Trannoy 2015).

Sophisticated measures of lifetime inequality of opportunity for health require undertaking the empirically challenging and ethically controversial task of disentangling the causal influence of circumstances and choices throughout an individual's life, including unfortunate circumstances beyond individual control arising after childhood (Rosa Dias 2009).[1] This raises the question about whether bivariate associations between childhood socioeconomic circumstances and adult health might be useful as a simple indicator of lifetime inequality of opportunity (IOp) for health.

In this chapter, I propose the parental gradient as a simple tool to measure and monitor inequality of opportunity for health. Parental gradients refer to bivariate associations between parental socioeconomic fraction-rank in childhood and adult health, while standard "current" gradients refer to bivariate associations between socioeconomic fraction-rank and health contemporaneously in adulthood. I use the 1958 National Child Development Study (NCDS) as an illustrative example to estimate parental gradients in respondent the 36-item Short Form Health Survey (SF-36) scores at age 50 by parental education, income, and occupation during respondents' childhood. I conduct robustness checks to investigate the potential impacts of: (i) intergenerational socioeconomic rank mobility; (ii) controlling for current socioeconomic fraction-rank; (iii) different timings of socioeconomic variable measurement during childhood; (iv) missing data methodology.

[1] For a comprehensive review of IOp frameworks and empirical literature, see Jusot and Tubeuf (2019).

The chapter is structured as follows. Section 2 introduces the 1958 cohort, variables used, and missing data strategy. Sections 3 and 4 discuss analysis methods and their results, respectively. Section 5 discusses key results, while considering strengths and limitations of the study, suggesting implications of the results, and proposing avenues for future research.

4.2 Data

Health inequality studies using parental socioeconomic variables in childhood commonly obtain data from birth cohort studies or surveys of older adults. The former directly surveys parents of respondents during childhood and the latter asks respondents to recollect their socioeconomic conditions growing up. Both study designs have limitations, which must be considered in their respective analyses.

Birth cohort studies often experience attrition; fewer cohort members respond in each follow-up sweep, leading to (potentially systematic) missing (Mostafa and Wiggins 2015) and unrepresentative (Mostafa et al. 2020) data. If childhood socioeconomic rank predicts attrition, as has been shown by Silverwood et al. (2020), analyses are likely to underestimate associations between childhood socioeconomic rank and adult health by excluding the worst-off. Longitudinal surveys of older adults, on the other hand, are subject to recall inaccuracy and bias. Recall inaccuracy can arise because studies rely on older people remembering the socioeconomic characteristics of their parents several decades ago during their childhood. Recall bias can arise because a respondent's memory may be influenced by their subjective assessment of how their life turned out (Hardt and Rutter 2004). Data from surveys of older adults may be further biased by survivorship. For example, it is not uncommon for surveys of older adults to exclude institutionalised individuals and individuals who have already died (Börsch-Supan et al. 2013). This could lead to the underestimation of gradients, especially if socioeconomically disadvantaged individuals are more likely to be institutionalised or dead by late adulthood.

Systematic approaches to handling missing data due to attrition have become widely accessible through modern statistical software, whereas recall bias is still difficult to account for. Consequently, the birth cohort study design is preferred for this analysis and the United Kingdom (UK) context is selected due to data availability. This study uses the 1958 National Child Development Study (NCDS) because, at this time of writing, it strikes the best compromise between data quality

and a sufficient age achieved by cohort members for health disparities to begin emerging. Older cohorts may be better placed to identify parental socioeconomic determinants of health inequalities, since poor health is less common in younger cohorts, where lagged or accumulative effects have not yet developed (Galama, Lleras-Muney, and Kippersluis 2018).

4.2.1 1958 National Child Development Study

The 1958 NCDS initially consisted of 17 415 children born 3rd-9thMarch 1958 in the UK, capturing key information through their life course (Power and Elliott 2006). Since then, nine follow-up sweeps have been conducted, most recently at age 55. Childhood follow-ups at age 7, 11, and 16 included cohort members, their parents, and teachers. These three childhood sweeps also supplemented the sample with immigrants born in the same week of March 1958, in response to emigration from the UK in the original sample, increasing the total cohort to 18 558 individuals. Although the most recent available sweep is at age 55, I use primary outcomes from the sweep at age 50, due to the availability of the 36-Item Short Form Health Survey (SF-36) questionnaire.

Missing Data

Missing data are prevalent in the 1958 NCDS, as in many cohort studies; response rates – the percentage of eligible respondents participating in the survey – decrease in almost every successive sweep (Table 4.1). The largest drops in response rates occur at the age 23 sweep (the first sweep in adulthood; 13% lower than the preceding sweed at age 16) the age 46 sweep (the first sweep to include a biomedical survey; 11% lower than the age 42 sweep). At age 50 the study had 9 789 participants, a 60% response rate among eligible participants.

Non-response poses a risk to the representativeness of the sample if attrition is non-random. Previous work suggests data in the 1958 NCDS are not Missing Completely at Random (MCAR), such that the likelihood of non-reponse is independent of any observed or unobserved variables (Mostafa et al. 2020). Instead, data are likely to be Missing at Random (MAR); several variables are associated with non-response, such as socioeconomic background, mental health, cognitive ability, and voting behaviour (Mostafa et al. 2021). Consequently, available or complete case analyses (ACA; CCA) may be biased and a systematic missing data

Table 4.1: Missing data across sweeps in the 1958 NCDS.

Year (Age)	Sample Size			Response Rate	
	Total	Eligible[a]	Respondents	% Total	% Eligible
1958 (Age: 0)	17 638	17 638	17 415	99	99
1965 (Age: 7)	18 016[b]	16 720	15 425	86	92
1969 (Age: 11)	18 287[b]	16 746	15 291	84	91
1974 (Age: 16)	18 558[b]	16 886	14 642	79	87
1981 (Age: 23)	18 558	16 886	12 536	68	74
1991 (Age: 33)	18 558	16 402	11 407	62	70
2000 (Age: 42)	18 558	16 174	11 387	61	70
2004 (Age: 46)	18 558	16 091	9 534	51	59
2008 (Age: 50)	18 558	15 963	9 789	53	61

[a] Ineligible participants include cohort members who have died or emigrated.
[b] Original sample supplemented with immigrants born in the same week of March 1958 due to emigration in the original sample.

approach is preferable (Briggs et al. 2003). Data may also be Missing Not At Random (MNAR) – whereby observed variables cannot account for the differences in the likelihood of missingness – however this hypothesis is not testable (Carpenter and Kenward 2013).

Following guidance from the Centre for Longitudinal Studies at University College London, which currently manages the 1958 NCDS, I apply multiple imputation (MI) methods to restore sample representativeness (Silverwood et al. 2020). Multiple imputation is an iterative stochastic simulation approach to handling missing data, whereby missing data points are approximated with plausible values (Briggs et al. 2003).

MI consists of three phases: imputation, analysis, and pooling (Sterne et al. 2009). In the imputation phase, completed datasets are created by estimating the variance

about missing data points, drawing from the distribution of the observed data for the variable.[2] In the analysis phase, each imputed dataset is analysed using appropriate statistical methods. Since some variables in this MI model are non-normal or binary, I employ Multiple Imputation using Chained Equations (MICE) (White, Royston, and Wood 2011). MICE uses a sequence of conditional regressions appropriate to the distribution of the variables; in this case logistic for binary variables, multinomial logistic for unordered categorical variables, ordered logistic for ordered categorical variables, and Ordinary Least Squares (OLS) regressions for normally distributed continuous variables.

Finally, in the pooling phase, estimates from each dataset in the analysis phase are combined to construct the MI estimates. Hence, the more datasets included, the more accurate the reflection of uncertainty about the missing data point. Theoretically, as few as five imputed datasets have been suggested as sufficient, however empirically as many as 50 or more have been suggested to reduce sampling variability (Horton and Lipsitz 2001). Following the suggestions of Silverwood et al. (2020) and White, Royston, and Wood (2011), the MI estimates in this chapter consist of 70 imputations of 35 iterations each .[3]

Health Outcome: 36-Item Short Form Health Survey (SF-36)

The SF-36 general health component assesses health through a range of questions, each with 5-point Likert scale answers, including self-assessed health in general, respondents' current health compared to the health of their peers and their own health 12 months ago, how frequently they are ill compared to their peers, and expectations of how their health will develop in the future (McDowell 2006). The resulting general health perception score ranges from 0-100, where 0 (100) indicates the worst (best) possible general health.

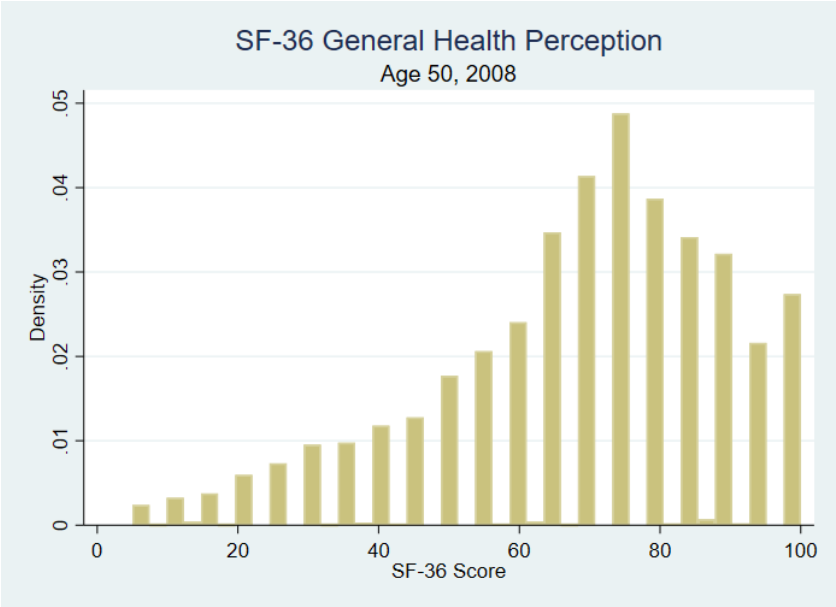
I use the SF-36 score as the primary health variable due to a more granular scale in SF-36 scores (0-100) than Likert-scale self-assessed general health (1-4 or 1-5) and therefore greater identifiable variation. This is important for this analysis because the sample population is relatively young. Consequently, less granular health measures, such as a Likert-scale self-assessed health, may be more insensitive

[2]For more information on multiple imputation and a full list of imputation variables, see Appendix B.2.

[3]Iterations refer to i, the number of times the conditional regressions are performed to calculate the variance about the missing data point. Imputations refer to the total number of m datasets (each consisting of i iterations) used for the MI estimation. For information on MI model checks, see Appendix B.3.

to changes in the health distributions across socioeconomic groups. Figure 4.1 shows the distribution of the SF-36 score in the total sample, which is left-skewed, likely due to the youth and therefore good health of the cohort.

Figure 4.1: Distribution of SF-36 general scores at age 50.



Source: 1958 National Child Development Study (Power and Elliott 2006). SF-36, 36-Item Short Form Health Survey.

Childhood Socioeconomic Variables

Occupation
I use household occupational class of the highest-ranked household member when the respondent is aged 16, combined with father's occupational class when respondents are 0, 7, and 11 years old, to create a childhood indicator based on the Office for National Statistics (ONS) 1990 Standard Occupational Classification (SOC 90) (ONS 2020b). Ranks across the four sweeps are summed to create a scale between 4-23, where lower (higher) values reflect lower (higher) parental occupational class across childhood.

The 1958 NCDS measures father's, but not mother's, occupational class consistently across all childhood sweeps.[4] Due to the collinearity between father's occupational class and household occupational class at age 16 (Pearson's Correlation Coefficient:

[4]Question formulation changes slightly between sweep's between "father" and "mother's

0.87), both variables cannot be included in the multiple imputation model (see 'Missing Data' section). I therefore include household occupational class to account for households where the father does not have the highest occupational class. This difference in variable composition across may influence results, however, given the high correlation, household occupational class is assumed to be a reliable proxy for father's occupational class.

Father's occupational class at ages 7 and 11 is measured using the same SOC-90 six-point scale as household occupational class at age 16, however at age 0, father's occupational class is only measured on a five-point scale. Since variables are fraction-ranked for analyses (see Section 3), variables are not transformed, instead opting to retain all possible information (Figure B.7). Robustness checks compare gradients using father's occupational class at 0, 7, and 11 with gradients using household occupational class at age 16.

Education
When respondents are 16, their mother's and father's education is measured by the age they left full-time education (Figure B.3). The resulting variable ranges from 12-24 years, with 12 indicating the parent left education at age 12 or younger, and 24 indicating a parent left at age 24 or older. Where the questionnaire option contains two ages, for example "13-14 years", the variable is given the midpoint value, 14 years in this example. For primary analyses, I use the education level of the parent with the highest educational attainment for consistency with the occupation variable.

Parental education variables are not captured consistently in childhood sweeps at ages 0, 7, and 11. However, parental education at age 16 may be considered a reliable proxy for parental education throughout childhood, since it is plausible that educational attainment is typically fixed from early adulthood. I include robustness checks using mother's educational attainment when respondents are aged zero and father's educational attainment when respondents are aged seven.

A binary survey question at the age zero sweep asks whether the mother stayed in school beyond the minimum school leaving age is also available (Figure B.6). Given the binary nature of the variable it is included as a robustness check, but interpretations are cautiously compared with results using other independent variables. I calculate father's school leaving age when respondents are seven years

husband/partner". In the 1958 NCDS, divorce rates are low and most children live with the same parents throughout their childhood (Power and Elliott 2006). As such, "father" is used here for brevity.

old by combining two variables. The first asks whether the respondents' father stayed in school after the minimum leaving age and the second asks the age at which he left school, if he remained in school after the minimum limit (see Figure B.6). I use the father's birth year to allocate a continuous age data point if the dummy variable indicates he left at the minimum leaving age, which varies between ages 13-15 according to the father's year of birth in the sample between 1880 and 1939.

Income
I use parental financial hardship measured at ages 7, 11, and 16, as a proxy variable for household income in childhood (Figure B.5). This variable has also been commonly utilised as an income proxy in studies using the 1958 NCDS due to non-response in housheold income measured continuously (Lindeboom, Llena-Nozal, and Klaauw 2009).

The financial hardship variables are binary, where a value of (zero) one represents (not) experiencing financial hardship in the past 12 months. Since the variable is measured on a negative scale, where a higher value indicate worse socioeconomic rank, variables are transformed to aid comparability with results using other socioeconomic variables. I construct the average childhood variable by summing the values at ages 7, 11, and 16, producing a score ranging from 0, indicating no financial hardship, to 3, indicating financial hardships reported in all three sweeps.

I include a continuous household income variable measured at age 16, as a robustness check. Household income at age 16 is calculated by summing the mother's, father's, and other annual net pay (Figure B.1), banded into 12 groups. I allocate the midpoint value to all respondents in any given income band and convert values to 2008 GBP to aid comparability with income at age 50 (Bank of England 2020). I use the square-root scale to equivalise income (OECD 2013). This variable is not included in the primary analyses because it is not measured before respondents are 16 years old and may therefore not be representative of income circumstances across childhood ages.

Adulthood Socioeconomic Variables

Occupation
I use the ONS SOC-90 to categorise occupations into six occupational classes at age 50, ranging from 'I', professional occupations, to 'V', unskilled labour (ONS 2020b). As for the childhood variable, I record household occupational as the

occupational class of the highest-ranked household member. To aid comparability with education and income variables, I reverse the ranking order such that a higher rank represents higher occupational class (see Figure B.2).

Education
I record education level at age 50 as the highest qualification attained by the respondent, according to the National Vocational Qualification scale (levels 0-5, with level 0 being equivalent to less than GCSE or O-Level grades D-G, or lower secondary education, and 5 being equivalent to an undergraduate degree or higher, or tertiary education).

Income
I record household income at age 50 on a continuous scale by summing mother's, father's, and other annual net pay. As with parental continuous income at age 16, I use the square-root scale to equivalise household income for different household sizes, by dividing annual net income by the square root of the number of household members (OECD 2013).

4.2.2 Sample

Table 2 shows summary statistics from the total sample prior to the MI procedure. At age 50, the average SF-36 general score is approximately 69/100, while 82% of the survey question respondents reported 'Good' or better self-assessed health (SAH), for example. The SF-36 had a more missing data than SAH at around 11%, compared to the almost 0% missing for SAH of the participants in the age 50 sweep.

Meanwhile, missing data proportions for income are high: 29% for grouped parental equivalised income (mean: 4.5) when respondents are 16 and 22% for log household equivalised income (mean: 9.6) when respondents are 50. The occupational class variables show an increase in household occupational class from 3.6 to 4.2 between the parental variable when respondents are 16 in 1974 and the respondent's own household occupational class at age 50 in 2008. The proportion of missing data is high at age 16, 24%, but low at age 50, 5%. For visual summaries of the degree of missing data for key analysis variables, see Figures B.17 and B.18.

Table 3 shows effects of MI on analysis variables. For SF-36 scores and SAH, averages decrease for both men and women; 3-4% for SF-36 and 5-7% for SAH.

This indicates those with worse general health are less likely to respond in follow-up sweeps. Standard deviations also decrease for both health variables 12-15%, with slightly larger decreases for men than women. This is likely a product of the increased number of observations, centring about the decreased mean, therefore decreasing the standard deviation.

For the socioeconomic variables at age 50 – income, occupational class, and education – means and standard deviations all decrease as a result of the MI procedure. The largest change in mean occurs in education, decreasing the mean by 17 and 20% for women and men, respectively, while standard deviations only decrease about 5% for both men and women. This mean-standard deviation trend is reversed in income, where the mean only decreases about 1-2% however standard deviation decreases by 16-17%. Social class mean and standard deviation both decreases by 6-9%. Overall, mean decreases are to be expected since it has been shown low socioeconomic rank is a predictor for non-response and decreases in standard deviation may be explained by the higher number of observations, converging on the decreased mean (Mostafa et al. 2020; Silverwood et al. 2020).

For socioeconomic variables at age 16, there is no observable trend in MI data versus non-MI data and mean and standard deviation changes are slight. The lack of observable trends between MI and non-MI data is likely due to the low missing data incidence at earlier sweeps, since it is earlier on in the study where less systematic attrition has taken place.

Histograms for analysis variables are available in Appendix B.1.

Table 4.2: Descriptive statistics for health and socioeconomic analysis variables following multiple imputation.

Imputed 1958 NCDS Dataset	**Female**					**Male**				
	Mean	**St.Dev.**	**Min.**	**Max.**	**Obs.**	**Mean**	**St.Dev.**	**Min.**	**Max.**	**Obs.**
Health										
SF-36 (Age: 50)	66.828	19.094	5	102	8,963	65.184	17.736	5	100	9,599
Income										
Current (Age: 50)	9.393	0.677	1.791	12.598	8,963	9.515	0.635	-0.507	13.718	9,599
Parental (Age: 16)	8,826	4,066	-280	22,449	8,963	8,857	4,005	130	21,403	9,599
Financial Hardship (Age 16)	0.202	0.328	0	1	8,963	0.202	0.327	0	1	9,599
Financial Hardship (Age 11)	0.140	0.299	0	1	8,963	0.139	0.295	0	1	9,599
Financial Hardship (Age: 7)	0.092	0.253	0	1	8,963	0.094	0.255	0	1	9,599
Social Class										
Current (age 50)	3.919	1.217	1	6	8,963	3.903	1.182	1	6	9,599
Parental (age 16)	4.011	1.338	1	6	8,963	4.004	1.360	1	6	9,599
Father's (Age: 11)	3.370	1.216	1	6	8,963	3.362	1.224	1	6	9,599
Father's (Age: 7)	3.231	1.190	1	6	8,963	3.234	1.196	1	6	9,599
Father's (Age: 0)	2.896	0.880	1	5	8,963	2.899	0.887	1	5	9,599
Education										
Current (Age: 50)	2.133	1.341	0	5	8,963	2.105	1.336	0	5	9,599
Mother's (Age: 16)	15.895	1.377	13	24	8,963	15.829	1.335	13	24	9,599
Father's (Age: 16)	16.237	1.803	12.993	30.116	8,963	16.209	1.799	13	30.142	9,599
Father's (Age: 7)	15.107	1.979	11.190	39.000	8,963	15.088	1.944	11.466	35	9,599
Mother's (Age: 0)	0.253	0.427	0	1	8,963	0.250	0.425	0	1	9,599

Table 4.3: Raw descriptive statistics for health and socioeconomic analysis variables.

Original 1958 NCDS Dataset	**Female**					**Male**				
	Mean	**St.Dev.**	**Min.**	**Max.**	**Obs.**	**Mean**	**St.Dev.**	**Min.**	**Max.**	**Obs.**
Health										
SF-36 (Age: 50)	69.233	21.986	5	100	4,530	68.086	20.895	5	100	4,211
Income										
Current (Age: 50)	9.504	0.812	1.791	12.598	3,992	9.636	0.775	-0.507	13.718	3,683
Parental (Age: 16)	13,255	6,103	301	35,738	5,058	13,276	5,999	266	36,430	5,343
Financial Hardship (Age 16)	0.106	0.308	0	1	5,547	0.104	0.306	0	1	5,819
Financial Hardship (Age 11)	0.116	0.320	0	1	6,506	0.111	0.315	0	1	6,857
Financial Hardship (Age: 7)	0.084	0.277	0	1	6,388	0.085	0.279	0	1	6,675
Occupation										
Current (age 50)	4.180	1.333	1	6	4,768	4.219	1.284	1	6	4,564
Parental (age 16)	3.645	1.260	1	6	5,374	3.612	1.263	1	6	5,726
Father's (Age: 11)	3.328	1.274	1	6	6,498	3.321	1.289	1	6	6,814
Father's (Age: 7)	3.243	1.244	1	6	6,887	3.242	1.247	1	6	7,227
Father's (Age: 0)	2.903	0.900	1	5	7,950	2.902	0.906	1	5	8,524
Education										
Current (Age: 50)	2.564	1.415	0	5	4,966	2.607	1.419	0	5	4,817
Mother's (Age: 16)	15.978	1.504	13	24	5,579	15.906	1.457	13	24	5,852
Father's (Age: 16)	15.987	1.857	13	24	5,404	15.944	1.852	13	24	5,687
Father's (Age: 7)	15.086	2.084	13	39	6,802	15.053	2.029	13	35	7,192
Mother's (Age: 0)	0.249	0.433	0	1	8,354	0.246	0.431	0	1	8,922

4.3 Methods

4.3.1 Research Strategy

I calculate current and parental socioeconomic gradients in SF-36 scores at age 50, using both the slope and relative indices of inequality (SII; RII), since their results may differ in some instances (Pamuk 1988). Socioeconomic variables are fraction-ranked, with a 0-1 scale and mean of 0.5, aiding comparisons across socioeconomic variables and time. Fraction-ranking also aids interpretation of results, indicating the linear approximation of the difference or ratio between the highest and lowest socioeconomic advantage. I control for age through the cohort study design and for sex by conducting analyses on separate samples (Mackenbach et al. 1997; Moreno-Betancur et al. 2015).

When measuring health inequalities, socioeconomic variables are commonly grouped into income quintiles or deciles. This approach may wash out potentially informative distributional variation through averaging and the grouping of socioeconomic variables is therefore minimised where possible. Instead, priority is given to maximising the use of all available data points from each variable. All analyses are performed using *Stata* version 17.0. Selected illustrative and representative results are presented in the Results section, while all results can be found in Appendix B.2.

4.3.2 Inequality Measurement

I use the slope index of inequality and relative index of inequality, as described in Chapter 2.2, to estimate parental and current gradients in SF-36 scores at age 50.

The SII is calculated using a generalised linear model (GLM) with Gaussian family, and identity link with robust standard errors to ensure homoskedastic error terms. SII values greater than zero indicate the socioeconomically best-off tend to have greater SF-36 scores than the socioeconomically worst-off. Conversely, if the SII is less than zero, the less advantaged have higher SF-36 scores. If the SII is zero, or statistically indistinguishable from zero, this indicates no association between socioeconomic fraction-rank and SF-36 score.

The relative index of inequality (RII) is calculated using a GLM function with Gaussian family, log link, and robust standard errors. A RII estimate equal to one

indicates no relative advantage between predicted lowest and highest socioeconomic ranks. If the RII is greater than one, higher socioeconomic rank is associated with relatively higher relative SF-36 scores and when RII is less than one, higher socioeconomic rank is associated with lower relative SF-36 scores. For example, a RII of 1.2 indicates the predicted highest socioeconomic rank has a 20%, or 1.2 times, higher SF-36 score than the predicted lowest socioeconomic rank.

4.3.3 Robustness checks

I conduct four robustness checks to assess the sensitivity of results. First, current adult socioeconomic status may be a key explanatory variable in the association between parental socioeconomic status in childhood and general health at age 50. I therefore include current socioeconomic status as a control in robustness checks of the parental gradients, to examine changes in the magnitude of coefficients and standard errors.

Second, I estimate the rank-rank slope index as a relative measure of intergenerational socioeconomic mobility which regresses current fraction ranked socioeconomic status on fraction-ranked parental socioeconomic status in childhood (Chetty et al. 2014). I use the rank-rank slope rather than the intergenerational income elasticity due to greater comparability between socioeconomic measures and between mobility estimates over time (Black and Devereux 2011). The resulting coefficients estimate the association between parental socioeconomic rank in childhood and current socioeconomic rank in adulthood. Greater coefficients indicate greater associations and therefore lesser socioeconomic mobility for a given variable. Thus, an increase in parental income rank in childhood is associated with a greater increase in current income rank in adulthood. The rank-rank slope, $\rho_R RS$, is calculated by regressing the cohort member's socioeconomic rank at age 50, C_i, on the socioeconomic rank of their household when growing up, P_i:

$$\rho_{RRS} = Corr(P_i, C_i) \tag{4.1}$$

Third, I assess the impact of differential timing of parental socioeconomic variable measurement during childhood. The primary parental socioeconomic variable for this analysis utilises the maximum possible data points by summing socioeconomic variables across sweeps at ages 0, 7, 11, and 16 to provide an indication of parental socioeconomic rank throughout childhood. However, previous research

has suggested the presence of "critical periods" in development (Almond, Currie, and Duque 2018; Currie 2009).

Finally, multiple imputation methods are generally recommended in cohort studies to account for attrition, especially the longer the study runs for. However, multiple imputation can be computationally demanding and statistically complex to implement if many variables are required to inform the imputation model. Consequently, I conduct available and complete cases analyses as robustness checks; the former making use of all non-missing data points for each regression model and the latter restriction analysis samples to data points which are non-missing for all regression models.

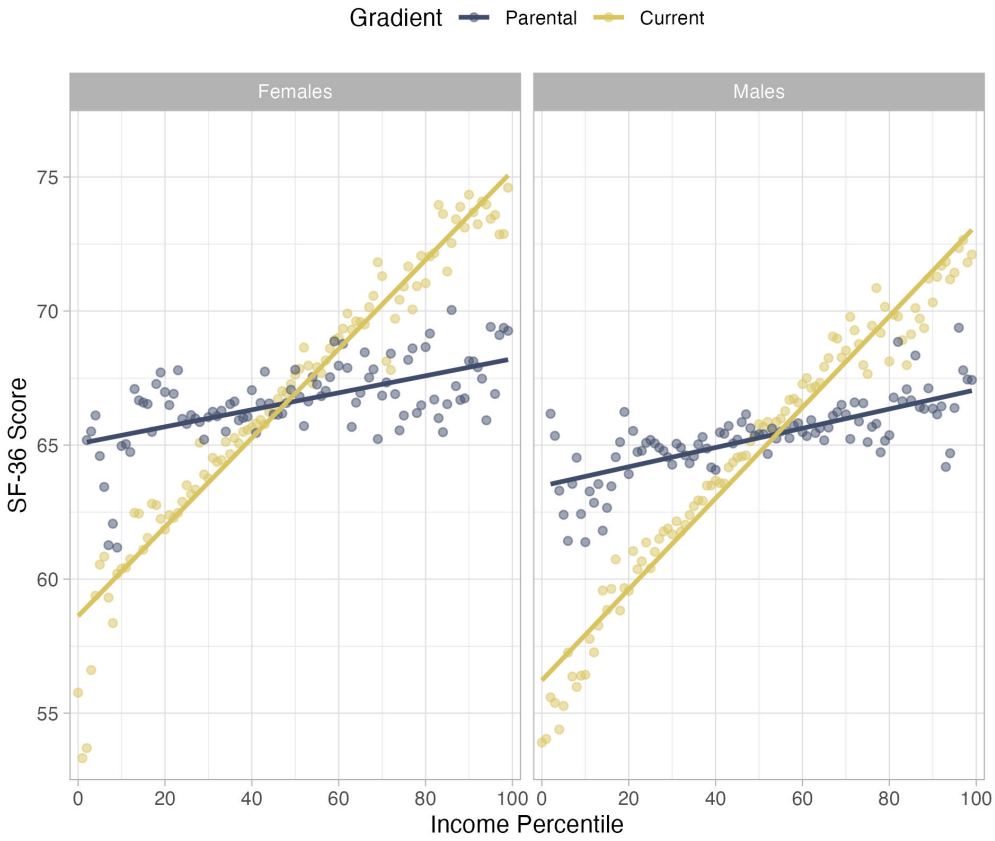
4.4 Results

4.4.1 Parental and Current Socioeconomic Gradients in Adult Health

Results suggest current gradients are steeper than parental gradients, as illustrated in Figure 4.2, showing a scatter plot of parental and current income against SF-36 score at age 50 for women and men. The scatter plot also reveals current gradients follow a more linear and predictable pattern across income percentiles, whereas the dispersion is greater when using parental income, particularly at the highest and lowest income ranks. Figures B.8 and B.9 in Appendix C.2 show the scatter plots using education and occupation variables, where differences between current and parental gradients are smaller but still apparent.

Figure 4.3 shows absolute current and parental gradients in SF-36 scores at age 50 using socioeconomic status – education, income, and occupation – in adulthood (age 50) and childhood (age 7, 11, and 16) as indicated by the slope index of inequality (SII) for women and men. All SII results are positive and significant, indicating the presence of parental and current absolute health inequalities favouring the socioeconomically better-off. Female gradients mean estimates are greater than male gradients, except for current income gradients, however confidence intervals are overlapping across all socioeconomic variables. Parental gradients are lesser than current gradients across all socioeconomic variables with the greatest difference for occupation for both men and women and income for men.

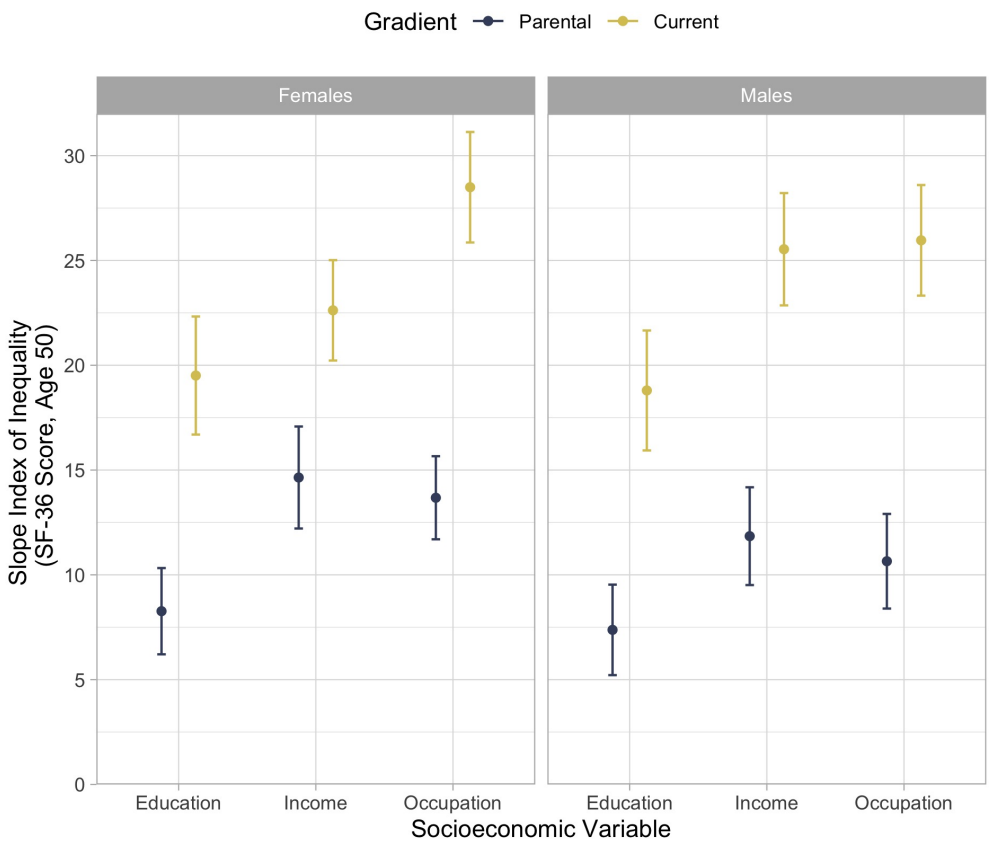
Figure 4.2: Scatter plot of parental and current income against SF-36 score at age 50 for females and males.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

Mean absolute parental socioeconomic gradients vary between 7-15 points on the SF-36 score. Parental gradients are greatest for income, where growing up in a household with the least recurrent financial hardship is associated with SF-36 scores 15 points higher for women and 12 points higher for men at age 50, compared to children growing in in a household with the greatest recurrent financial hardship (Women: 14.6 points [95% CI: 12.2, 17.1]; 11.8 points [95% CI: 9.5, 14.2]). This is closely followed by occupational class, where parental gradients are one point lower on the SF-36 age 50 than for parental gradients using income for both women and men (Women: 13.7 points [95% CI: 11.7, 15.7]; Men: 10.6 points [95% CI: 8.4, 12.9]). Parental education gradients have the smallest magnitude, but are still significant and substantial; growing up with the most educated parent is associated with SF-36 scores 8 points higher for women and 7 points higher for men at age

Figure 4.3: Current and parental slope indices of inequality in SF-36 score by education, income, and occupation.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

50, compared to growing up with the least educated parent (Women: 8.3 points [95%CI: 6.2, 10.3]; Men: 7.4 points [95% CI: 5.2, 9.5]).

Mean absolute parental socioeconomic gradients are 39-65% the magnitude of current socioeconomic gradients in SF-36 scores at age 50; higher (lower) for women and when income (education) is the proxy for socioeconomic status. In adulthood, occupational gradients are greatest, living in a household with the highest occupational class at age 50 is associated with SF-36 scores 29 and 26 points higher than living in a household with the lowest occupational class, for women and men, respectively (Women: 28.5 points [95% CI: 25.9, 31.1]; Men: 26.0 points [95% CI: 23.3, 28.6]). Current absolute income gradients closely reflect occupational gradients but 3 points lower for women and 1 point lower for men

(Women: 22.6 points [95% CI: 20.2, 25.0]; Men : 25.5 points [95% CI: 22.9, 28.2]). As with parental socioeconomic gradients, educational gradients have the smallest magnitude; having the highest educational attainment is associated with 20 and 18 points higher SF-36 scores at age 50, for women and men, respectively, compared to having the lowest educational attainment (Women: 19.5 points [95% CI: 16.7, 22.3]; Men: 18.8 points [95% CI: 15.9, 21.7]).

Relative parental and current gradients reflect absolute gradients; parental gradients are 38-60% the magnitude of current gradients with similar sex and socioeconomic proxy trends (Figure 4.4). Compared to growing up in the least advantaged households, growing up the most advantaged household is associated with 12-24% higher SF-36 scores at age 50. Being in the most advantaged household at age 50 is associated with 32-51% higher SF-36 scores than being in the most advantaged household at age 50.

4.4.2 Robustness Checks

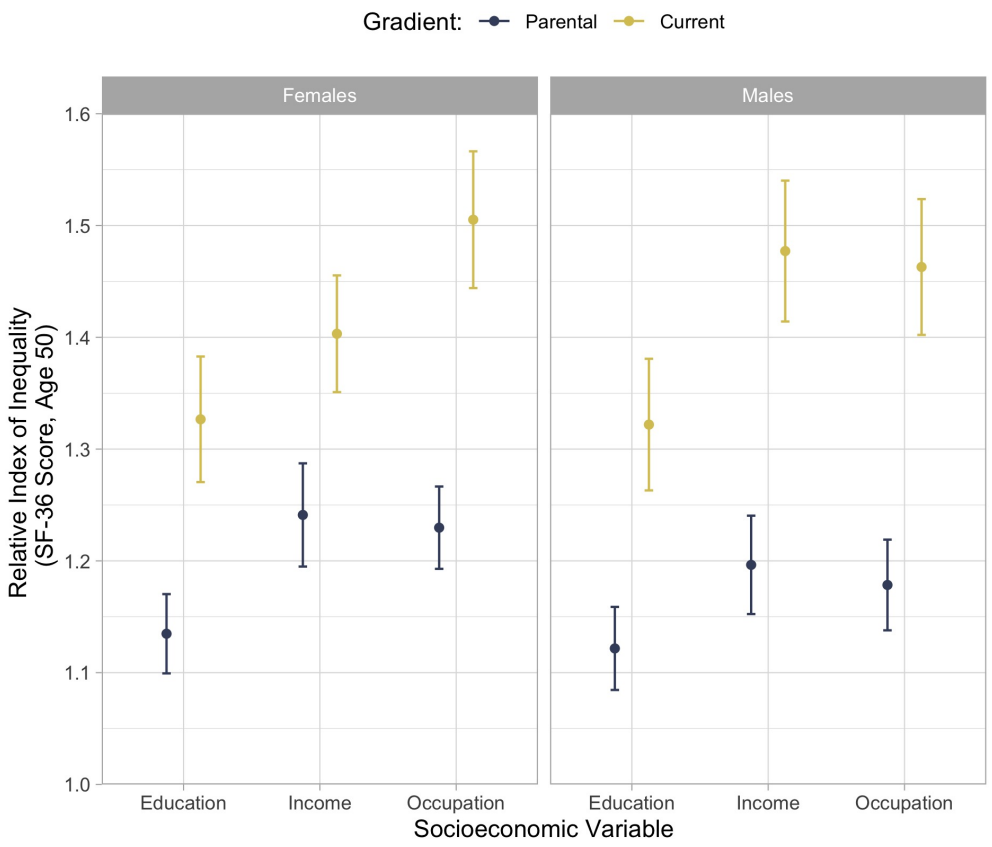
Intergenerational Socioeconomic Rank Mobility

Rank-rank slope index estimates indicate significant intergenerational mobility gradients; increasing parental socioeconomic rank in childhood by one quintile, or 20 percentiles, is associated with a 3-6 percentile increase in socioeconomic rank in adulthood (Figure 4.5). Estimates are higher for men than women in the 1958 NCDS, indicating greater intergenerational socioeconomic mobility for women, however differences are small in magnitude. Sex-differences are largest in occupation, with smaller differences in income and education.

Controlling for Current Socioeconomic Status

Controlling for respondents' socioeconomic fraction rank at age 50 in parental gradients decreases absolute parental gradients in SF-36 scores at age 50 by 20-62%, with the largest (smallest) decreases in women (men) and in occupational class (income) (Figure 4.6). After controlling for current socioeconomic fraction-rank, parental gradients remain significant and substantial at approximately 22-40% the magnitude of current gradients. Relative index of inequality results for parental gradients with and without controlling for current socioeconomic fraction rank are available in Figure B.10.

Figure 4.4: Current and parental relative indices of inequality in SF-36 score by education, income, and occupation.

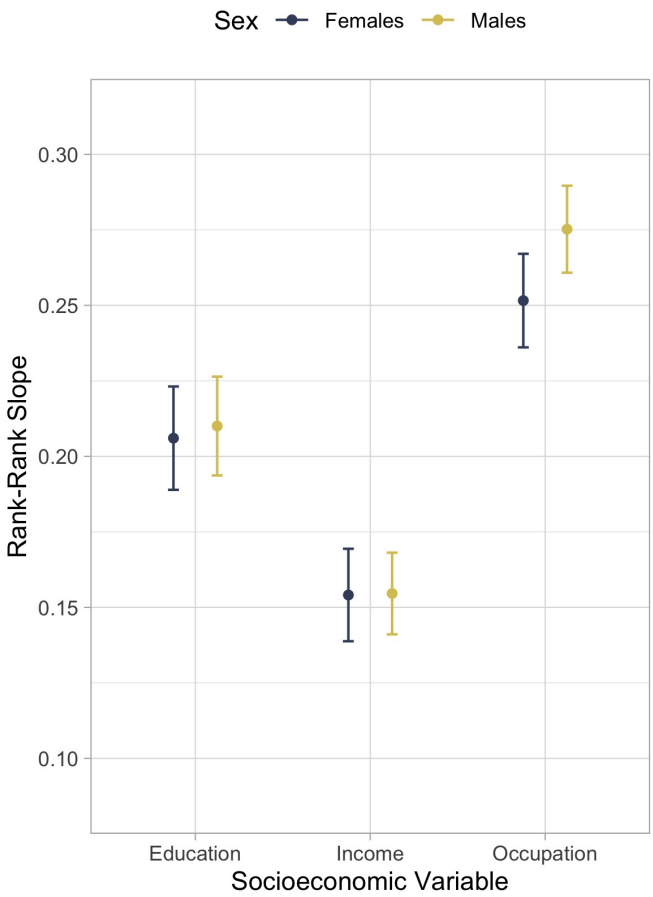


Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

Differential Timing in Parental Socioeconomic Variable Measurement During Childhood

Robustness checks do not provide a consensus on the impact of different timings of parental socioeconomic variable measurement on the magnitude of parental gradients; with increasing respondent age at measurement, parental education gradients tend to increase, while parental income gradients tend to decrease, and parental occupational class gradients are largely unchanged. Parental education gradients tend to increase with increasing respondent age at measurement (Figure B.12 and B.12). Compared to using the educational attainment of the highest ranked household member at age 16, parental gradients are greater when using

Figure 4.5: Rank-rank slope socioeconomic mobility for education, income and occupation for women and men.

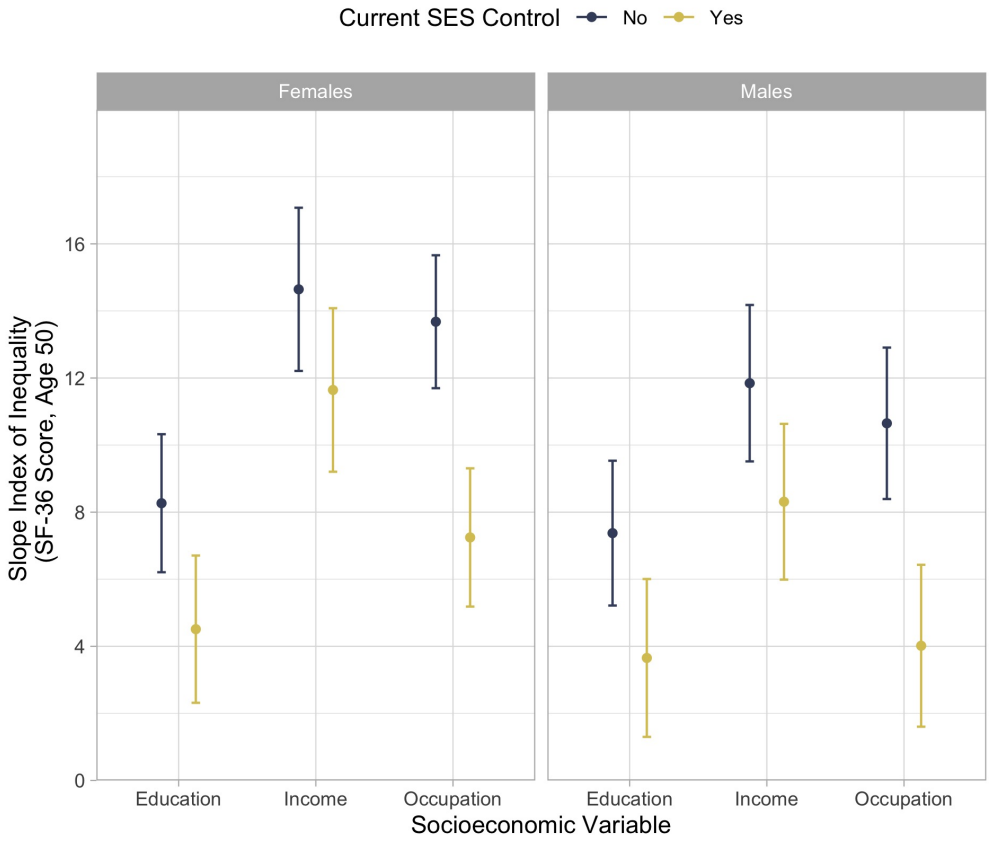


Data: 1958 National Child Development Study.

father's education measured when the respondent is 7 and 16 years old, and when using mother's education when the respondent is 16 years old. Parental educational gradients are smallest when using mother's education at age zero. Parental occupational gradients are similar across measurements at ages 0, 7, 11, and 16, or the childhood average, and across mother's and father's occupation (Figures B.14 and B.11). However, using the maternal grandfather's occupation results in lower gradients.

Parental income gradient trends indicate the inverse trend of parental educational gradients; the absence of financial hardship earlier in childhood is associated with greater SF-36 score increases than the absence of financial hardship later in

Figure 4.6: Parental slope indices of inequality in SF-36 score by education, income, and occupation, including current SES control.



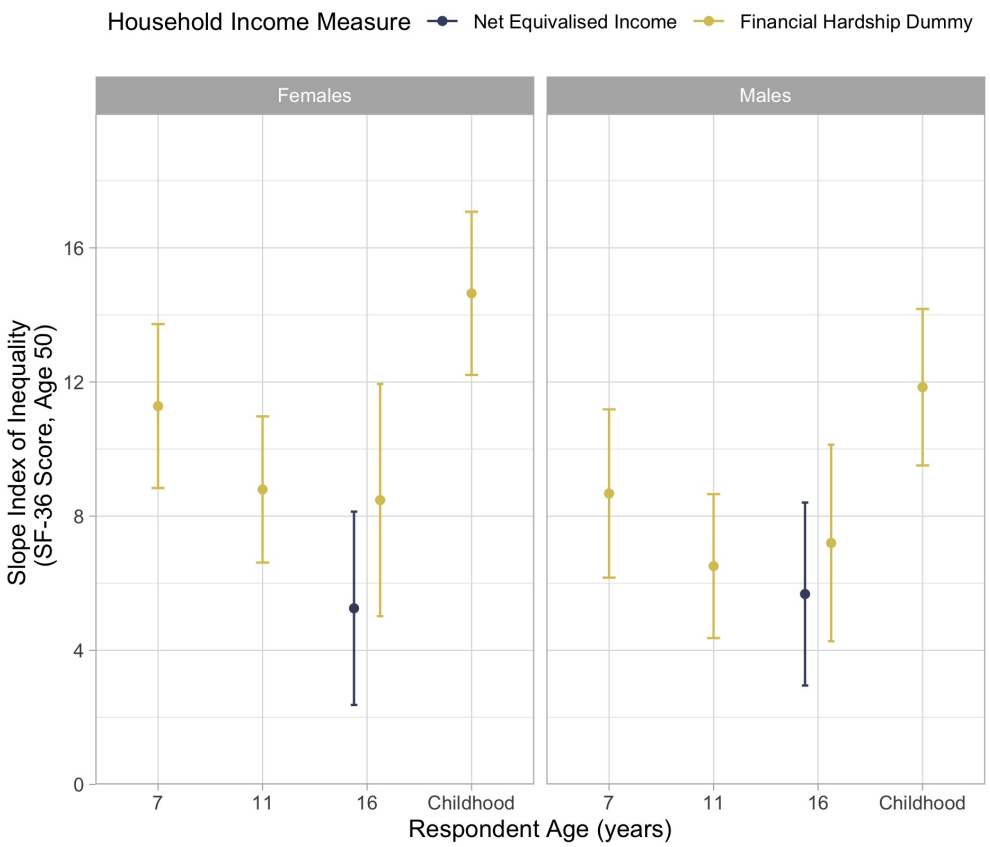
Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey; SES, socioeconomic status.

childhood (Figures 4.7 and B.11). The fraction-ranked sum of financial hardship incidences across sweeps at ages 7, 11, and 16 indicates the greatest parental gradient, perhaps indicating the harm of repeated financial hardship across childhood. Parental gradients using household equivalised income measured continuously at age 16 are lowest among income variables.

Available and Complete Case Analyses

Available and complete case results tend to underestimate current gradients compared to multiple imputation estimates, and magnitudes are larger for absolute inequality measures and current gradients. The magnitude of differences in parental

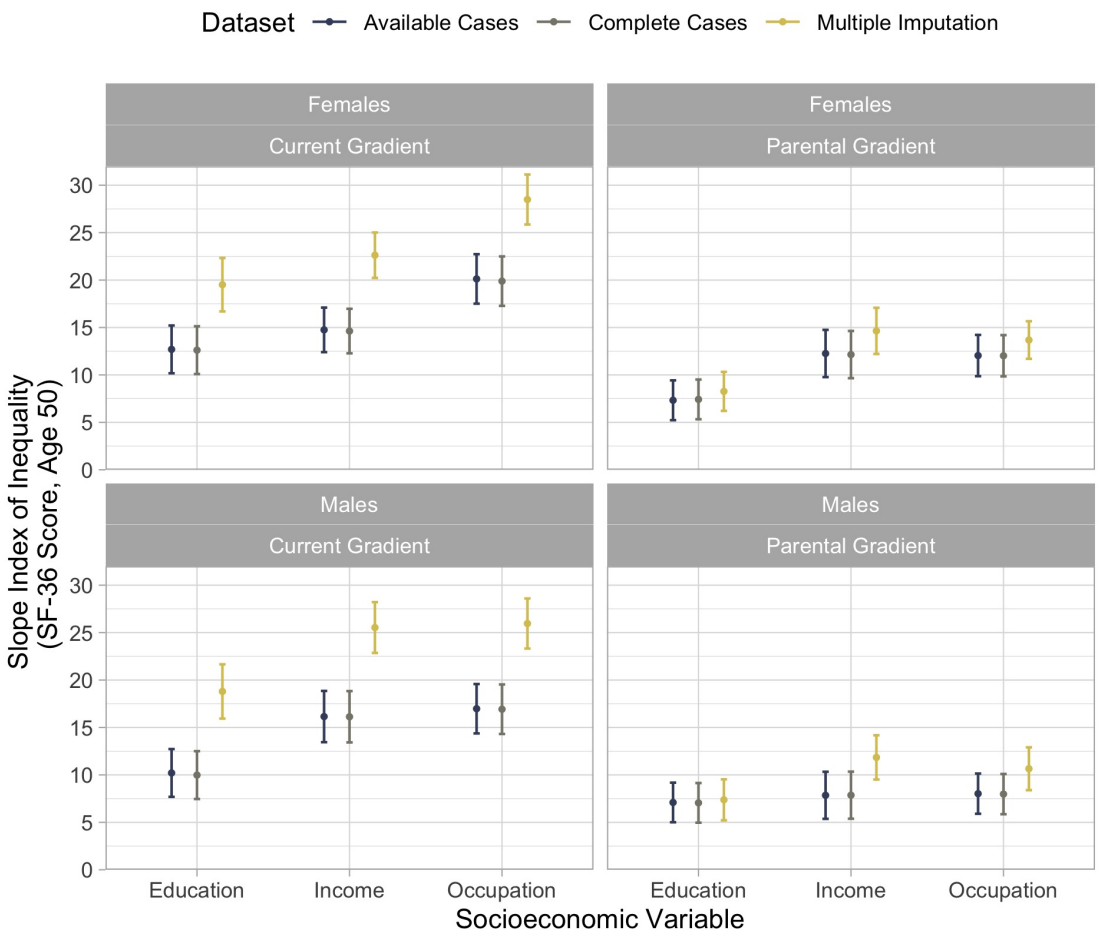
Figure 4.7: Impact of measurement timing on the slope index of inequality in SF-36 score at age 50 by parental income in childhood.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

gradients are generally small, especially when using education as the socioeconomic proxy. Results for the relative parental gradients are available in Figure B.16.

Figure 4.8: Parental gradient estimates under multiple imputation, available case, and complete case analyses using the slope index of inequality



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey. "Childhood" is the sum of binary financial hardship variables at ages 7, 11, and 16.

4.5 Discussion

4.5.1 Principal Findings

Absolute parental socioeconomic gradients in SF-36 scores at age 50 are significant and substantial in the 1958 National Child Development Study, but smaller than current socioeconomic gradients. Mean absolute parental socioeconomic gradients

vary between 7-15 points on the SF-36 score; higher when measuring average household income childhood and lower when measuring parental education during childhood. These parental gradients are roughly 40-65% the magnitude of current socioeconomic gradients in SF-36 scores at age 50, which are higher (lower) when income or occupation (education) is the proxy for socioeconomic status.

Comparing parental and current gradients indicate socioeconomic health inequalities may be substantially determined during adolescence, especially for more static socioeconomic variables such as education, and occupational class for women. Parental and current gradients are consistently slightly greater for women indicating socioeconomic inequality of opportunity for health and socioeconomic health inequalities may be greater among women born in the UK in 1958 than their male peers.

Robustness checks reveal several supporting findings. (1) Increasing parental socioeconomic rank in childhood by one quintile, or 20 percentiles, is associated with a 3-6 percentile increase in socioeconomic rank in adulthood. This indicates significant intergenerational transmission of relative advantage, however results also indicate income mobility may be greater in the UK than in USA and more in line with findings from Denmark and Canada (Chetty et al. 2014). (2) Introducing adult socioeconomic fraction rank decreases absolute parental gradients in SF-36 scores at age 50 by 20-62%. Socioeconomic inequality of opportunity for health may therefore be substantially driven by the association between parental and current socioeconomic fraction-rank. (3) Robustness checks do not provide a consensus on the impact of different measurement timings on the magnitude of parental gradients across socioeconomic measures. However, repeated financial hardship appears to be substantially more harmful than any individual instance of financial hardship. (4) Failing to account for attrition using multiple imputation methods is likely to lead to substantially underestimates of current socioeconomic gradients in health. However, available and complete case results do not alter parental gradients magnitudes substantially, particularly when using education as the socioeconomic proxy.

Comparison with Previous Findings

Previous studies have found childhood circumstances account for roughly 20-25% of current health inequalities, which may be an overestimate due to the use of CCA (Jones, Roemer, and Dias 2014; Rosa Dias 2009, 2010). The difference between previous results and those of this study may have multiple causes. First, previous

studies have employed more sophisticated statistical models, which control for multiple explanatory factors, which is notable absent when measuring parental gradients. Second, the smaller IOp finding may be due to the younger age of the populations studied. If childhood socioeconomic rank has an accumulative or lagged effect, studies assessing health earlier in the parental may identify fewer health disparities, potentially decreasing the proportion of inequality associated with childhood circumstances.

4.5.2 Strengths and Limitations

This study presents some key advantages over previous literature. First, this study provides an accessible IOp metric, which increases comparability between contexts as well as the ease of implementation and interpretation. Parental gradients also contextualise the growing parental literature by providing a direct comparison with the established literature on current socioeconomic gradients in adult health. Second, this study uses representative data by applying multiple imputation methods to a cohort study design. Third, several robustness checks provide new insights into contextual factors when estimating socioeconomic inequalities in health originating in childhood.

This study is also subject to limitations. First, methodologies applied in this study are associative. Multiple causal theories are therefore plausible (see Chapters 2 and 3), especially given the length of time elapsed between the measurement of socioeconomic and health variables. Second, some socioeconomic variables encounter comparability issues across sweeps. For example, measurement of education changes between childhood and adulthood (parental education is measured in years but respondents' own education is measured in qualification levels). Fraction-ranking variables addresses some problems but cannot guarantee direct comparability. If the parental and current variables are incomparable, implications of results may be misleading. Lastly, multiple imputation results rely on the MICE procedure and therefore the missing-at-random assumption. Although multiple imputation methods are widely accepted – and even endorsed for the 1958 NCDS dataset specifically – the model should be scrutinised appropriately (Silverwood et al. 2020). Multiple checks for the MI model are therefore presented in Appendix C.3, which support the plausibility of the MI model in this study.

4.5.3 Policy Implications

Parental gradients are accessible summary measures of "unfair" health inequalities, which are easy to implement, interpret, and compare. This simple metric provides an indication of the presence and magnitude of inequality of opportunity, which may be monitored or used to justify the need for further research using more sophisticated methods. Parental gradients may be especially well-paced to inform decision-makers because they use well-known inequality metrics, which are comparable to standard current gradients.

Parental gradients may also be a supplementary tool to monitor long-term term changes following childhood polices. This may be especially important with the increasing focus on early life health interventions and ill-health prevention (Almond, Currie, and Duque 2018). When measuring parental gradients, the impacts of intergenerational socioeconomic rank mobility and current socioeconomic rank controls should be investigated where possible to account for the associations between parental and adult socioeconomic rank. Evidence from this study suggests multiple imputation methods may not be required to provide an accurate indication of inequality of opportunity for health using parental gradients.

4.5.4 Future Research

This study presents several avenues for future research. First, parental gradients could be used to further examine how inequality of opportunity for health has have changed over time, using studies such as the 1946 National Birth Cohort from 1946, British Cohort Study from 1970, Next Steps from 1989-90, and Millennium Cohort Study from 2000 (Bann et al. 2018; Fluharty et al. 2021).

Second, parental gradients may be applied using objective health outcomes and health behaviours to explore differences from self-assessed general health. It may be plausible steeper parental gradients exist in health behaviours than in objective or general health earlier in life, since health behaviours have been shown to be a key pathway for the development of health inequalities later in life (Brunello et al. 2016).

Third, future research could explore parental gradients using more readily available data, such as retrospective socioeconomic status and Likert-scale self-assessed health or mortality. This would provide an indication of how parental gradients

could be applied at scale, both for continuous national monitoring and international comparative purposes. Such studies would have to overcome challenges in the collection of reliable parental socioeconomic data from childhood, either using methods to adequately address recall bias or leveraging rich registry data (such as in Scandinavian countries).

CHAPTER 5

Evaluating the Lifetime Distributional Health Impacts of Universal Infant Free School Meals

Background: Providing free school meals (FSMs) to children is often advocated on humanitarian and educational grounds, however long-term health impacts of universal compared to means-tested FSM provision for different social groups are unknown.
Aim: This study evaluates the population health and health inequality impacts of universal, compared to means-tested, FSM delivery in the UK among 4/5- to 7/8-year-olds from a health sector perspective.
Methods: I extrapolate short-term causal conditional average distributional bodyweight treatment effects (Holford and Rabe 2022) using a *de novo* lifetime Markov cohort model to estimate lifetime costs and effects of universal FSMs by socioeconomic quintile group in England. The model is informed by age-sex-bodyweight-income specific transition probabilities derived from UK longitudinal studies and I use distributional cost-effectiveness analysis methods to assess population health, health inequality impacts, and equity-efficiency trade-offs.
Results: Universal FSMs generate 5 323 additional discounted lifetime QALYs [95% CI: 214, 10 521] and a lifetime discounted health care cost savings of £12 million [95% CI: -2, 22] per year group of 654 473 children, justifying a health care co-fund of £92 million [95% CI: 5, 180] per year group, or 19% [95% CI: 1, 37] of incremental programme costs, assuming bodyweight treatment effects fade out linearly to zero at age 100. Net health benefits are positive for all but the best-off socioeconomic quintile, resulting in a pro-poor overall distribution. Consequently, at a 10% co-fund from the health care budget, universal FSMs improve both population health and health equity.
Discussion: The cost-effectiveness of universal, compared to means-tested, FSMs from a long-term health sector perspective depends largely on assumptions about the fadeout of bodyweight treatment effects over the lifetime. Universal free school meals may reduce health inequality beyond means-tested provision, despite strong pro-rich incremental uptake gradients.

5.1 Introduction

Providing free school meals (FSMs) to children is often advocated on humanitarian and educational grounds (European Union 2021), with 388 million children receiving school meals across 161 countries in 2020 (UN 2021). FSMs are associated with a wide range of short- and long-run health and economic outcomes (Cohen et al. 2021; Lundborg, Rooth, and Alex-Petersen 2022) and may address early-life social inequalities (Bardin, Washburn, and Gearan 2020; Spence et al. 2014). School feeding policy has also risen sharply on the policy agenda following the onset of the COVID-19 pandemic (Dunn et al. 2020; Kenney et al. 2022; Parnham et al. 2020).

One potential FSM benefit is improved childhood bodyweight outcomes (Holford and Rabe 2022), through reduced caloric intake (Heymsfield and Wadden 2017). Childhood obesity is strongly predictive of living with obesity in adulthood (Simmonds et al. 2016), which can cause metabolic conditions (Neeland, Poirier, and Després 2018) associated with comorbidities, such as cardiovascular disease, type 2 diabetes and non-alcoholic fatty liver disease (Guh et al. 2009). Consequently, obesity is associated with lower health-related quality of life (Kolotkin and Andersen 2017), greater global burden of disease (GBD 2015 Obesity Collaborators 2017), and greater mortality risk (MacMahon et al. 2009). More recently, obesity has been identified as a key risk factor for complications of COVID-19 infection (Popkin et al. 2020; Sanchis-Gomar et al. 2020; Yang, Hu, and Zhu 2021).

Obesity is widely recognised to be costly to health care systems (Withrow and Alter 2011), however obesity is also associated with substantial economic costs (Tremmel et al. 2017), leading to greater societal burden of disease beyond the health care sector from factors such as productivity losses (Goettler, Grosse, and Sonntag 2017) and premature mortality (Okunogbe et al. 2022). Socioeconomically disadvantaged groups suffer higher obesity rates than their advantaged peers (Bann et al. 2018; OECD 2019). Interventions targeting obesity therefore have the potential to substantially improve health and economic outcomes while also reducing inequalities.

Since 2014, the Universal Infant Free School Meal (UIFSM) programme has offered one free lunch per school day to all children in English state-funded schools during their first three years of school (Sellen et al. 2018). Previously, FSMs were only available on a means-tested basis to children whose parents claimed certain benefits, such as Universal, Child Tax, and Working Tax Credit. While short-term incremental costs and effects of UIFSMs have been evaluated from

a public sector perspective (Sellen et al. 2018), long-term health outcomes are unknown. Furthermore, there is considerable uncertainty about the socioeconomic distribution of health impacts under universal compared to targeted FSM delivery, since universal delivery may benefit better-off children more, since they were less likely to have access to FSMs under means-tested provision.

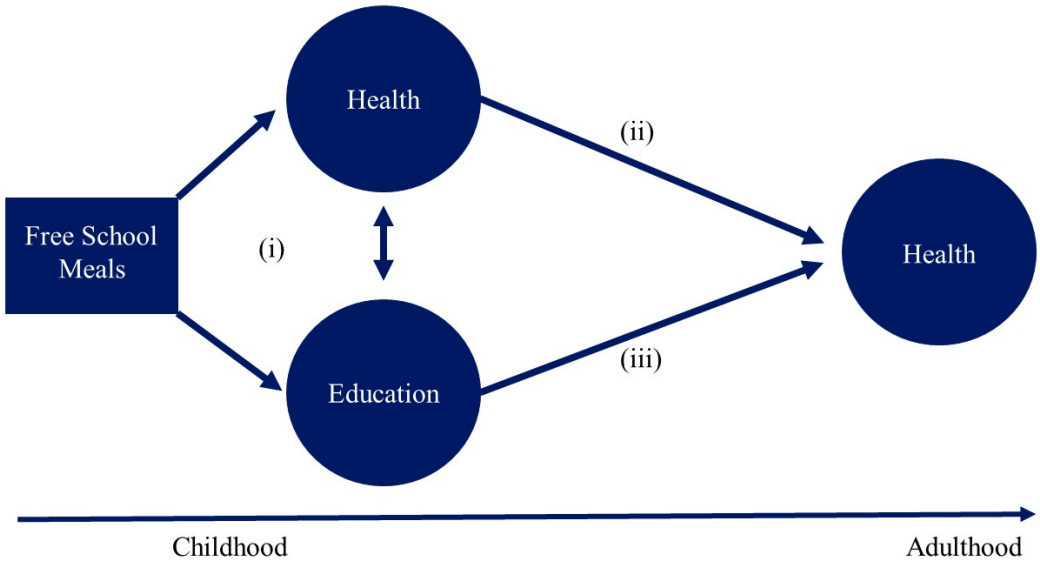
This study evaluates the (i) cost-effectiveness and (ii) distributional equity impacts and trade-offs of bodyweight changes from universal, compared to means-tested, FSM delivery among 4/5- to 7/8-year-olds in England from a long-term public health perspective. I estimate (distributional) lifetime costs and effects of universal FSM provision using empirical estimates of causal UIFSM distributional bodyweight treatment effects from Holford and Rabe (2022) and a *de novo* lifetime Markov cohort model informed by age-sex-bodyweight-deprivation specific transition probabilities derived from four UK longitudinal studies. I use distributional cost-effectiveness analysis methods to assess the cost-effectiveness and health inequality impacts of the programme.

The chapter is structured as follows. Section 2 provides background on school meals and UIFSM impacts. Section 3 outlines the methodological approach. Section 4 presents cost-effectiveness and health inequality impacts of UIFSM. Section 5 outlines primary findings, strengths and limitations, policy implications, and future research recommendations. Section 6 concludes the chapter.

5.2 Conceptual Framework

In this chapter, I focus on the protective health effects of school meals. This protective effect on health is driven by multiple potential mechanisms, including: (i) direct short-term effects on bodyweight and education in childhood (Holford and Rabe 2020), which likely have a reinforcing relationship; (ii) direct long-term bodyweight effects extrapolated from direct short-term bodyweight outcomes (Simmonds et al. 2016); and (iii) indirect long-term effects via other factors, such as education (Lleras-Muney 2005). Figure 5.1 illustrates how free school meals may impact lifetime health.

Figure 5.1: Conceptual framework for free school meal impacts on lifetime health.



5.2.1 Short-term Mechanisms

School meals have direct protective short-term bodyweight effects by altering caloric intake, such that more children attain, or remain at, a healthy bodyweight (Heymsfield and Wadden 2017). In high-income countries reducing caloric intake has great impact potential, since fewer children are underweight and a growing proportion of children are overweight or obese (NCMP 2020). However, a childhood food environment encouraging dietary variety has also been linked to decreased incidence of eating disorders often associated with low bodyweight in adolescence, such as bulimia nervosa and anorexia nervosa (Marchi and Cohen 1990), which may have long-term health consequences (Kotler et al. 2001; O'Brien et al. 2017).

Following the introduction of nutritional standards in England for FSMs in 2008/9, packed lunches were found to contain 18% more calories than school lunches on average (Evans et al. 2010). Free school meals have also been found to have better nutritional content than packed lunches, such as lower non-milk extrinsic sugar intake (Spence et al. 2019), lower saturated fat and sodium intake (Spence et al. 2014), while increasing vitamic C and folate intake (Spence et al. 2013). Furthermore, the UIFSM programme was found to reduce school absences due to illness or medical appointments by 0.7 days per year on average, among children eligible for means-tested FSMs (Holford and Rabe 2020). It is therefore plausible

that increasing the proportion of students eating school lunches could lower population caloric intake, improve average bodyweight outcomes, and improve population health, all other things being equal.

5.2.2 Long-Term Mechanisms

Bodyweight outcomes in childhood are significantly associated with adulthood bodyweight (Simmonds et al. 2015). Improving childhood bodyweight may therefore lead to long-term improvements in lifetime adult bodyweight, in turn potentially increasing health-related quality of life (Maheswaran et al. 2013), decreasing incidence rates of ill-health outcomes associated with obesity and overweight (GBD 2015 Obesity Collaborators 2017), and driving down health care costs (Hamilton, Dee, and Perry 2018).

School meals may also improve lifetime health indirectly, via education and income, for example. Improved educational outcomes have been shown to substantially improve health in the long run and lifetime income is also associated with better health (See Chapter 2.2). It may therefore be plausible that FSM provision improves lifetime health indirectly via non-health benefits.

Swedish reforms between 1959-69 to provide school meals free-of-charge increased lifetime income by 3% and university attendance by 5% (Lundborg, Rooth, and Alex-Petersen 2022). Difference-in difference estimates from the UK show UIFSMs improved educational outcomes; difference-in-difference estimates suggest children aged 7 who ineligible for means-tested FSMs performed better on average test scores across reading, writing, maths, and science, than would be expected if universal FSMs were not introduced outcomes (Holford and Rabe 2020). Household savings may also provide a small indirect effect, since providing school meals free-of-charge relieves a financial burden for parents (Holford and Rabe 2020), potentially leading to improvements in children's health (See Chapter 2.2).

In this study, I build a model to estimate the direct long-term effects of FSMs from extrapolated empirical estimates of short-term effects. The potential impact of indirect effects is reflected in optimistic treatment effects fadeout scenario sensitivity analyses.

5.3 Methods

This study takes a health care perspective on the long-term incremental costs and benefits of universal, compared to means-tested, FSM provision attributable to reductions in obesity and overweight by deprivation quintile. This study does not explicitly account for potential non-health benefits – such as impacts on educational attainment and lifetime earnings – or potential non-health sector cost savings – such as savings to education and social welfare budgets, however this effect may be approximated by making more generous assumptions on the long-term fadeout after the UIFSM programme stops at age 7/8 (see subsection "Bodyweight Treatment Effects").

5.3.1 Short-Term Treatment Effects

Holford and Rabe (2022) study the causal effect of UIFSMs on bodyweight outcomes using a difference-in-difference design and anthropometric data from the National Child Measurement Programme for all children ages 4-5 in England in academic years 2008/09–17/18. Primary treatment effects are reported by school-level FSM-eligibility (FSMe) quintiles and Index of Deprivation Affecting Children Index (IDACI[1]) are included for comparative purposes (see Section 3.2 "Bodyweight Treatment Effects").

Bodyweight treatment effects are recorded using the change in body mass index (BMI) Z-Score (BMIz) and the percentage point change in healthy and obese weight category prevalence rate according to the BMIz. The BMIz measures standardises a child's BMI via external growth curves to define bodyweight categories according to their standard deviations from the mean growth curve (Cole, Freeman, and Preece 1995). For this analysis, BMIz scores are considered the primary outcome and category treatment effects are included as a robustness check.

BMIz treatment effects follow a pro-rich gradient where treatment effects increase with lower school-level FSM eligibility proportions, except for the best-off, where treatment effects were small (see Figure C.1.). The null hypothesis of no treatment effect cannot be ruled out for the least and most advantaged subgroup. The finding

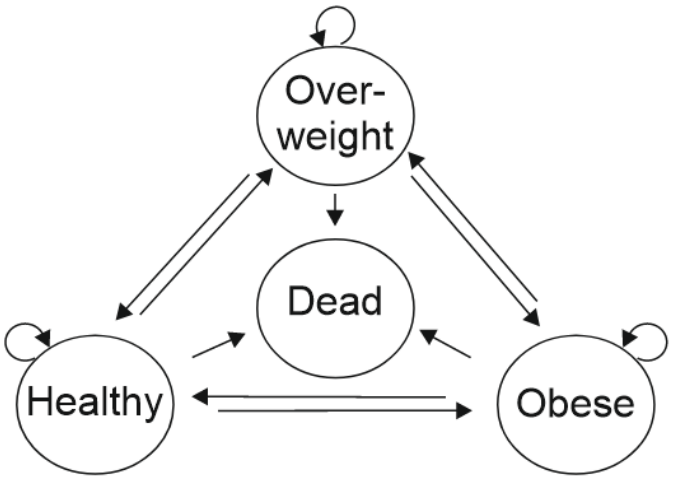
[1]The IDACI measures the fraction of children ages 0-15 living in low-income households, defined as households with no employed adults and households with a combined income sufficiently low to satisfy means-testing for benefits (DLUHC 2022).

of no significant treatment effects among the best-off corroborates previous findings that the most advantaged children benefit least from free school lunches (Lundborg, Rooth, and Alex-Petersen 2022). The finding of no significant treatment effects among the worst-off may be explained by lower incremental take-up among pupils in the least advantaged schools, since these pupils were more likely to be receiving means-tested FSMs.

5.3.2 Extrapolating Short-Term Effects

To extrapolate short-term conditional average treatment effects by socioeconomic quintile, I use a *de novo* lifetime Markov cohort model (MCM) with 10 cohorts – one for each deprivation group across males and females – which run from age 5 to 120, with one cycle every year. In each cycle, possible transitions occur between three bodyweight categories (healthy, overweight, and obese), remaining in the same bodyweight category, or dying (see Figure 5.2). Underweight is assumed to be encompassed in the healthy weight state. While health care costs and health-related quality of life may differ substantially between underweight and healthy weight populations, overall prevalence rates of underweight were low during the study period, so this simplification is unlikely to substantially alter results (NCMP 2020).

Figure 5.2: Markov cohort model design.



Note: Healthy weight includes underweight and bodyweight is categorised according to body mass index (BMI) growth curves until 17 years (Cole, Freeman, and Preece 1995) and standard BMI classifications from 18 years.

The 10 cohorts reflect the whole English population of five-year-olds at baseline, with healthy, overweight, and obesity rates conditioned on deprivation quintile and

sex (see "Study Population"). Deprivation-specific treatment effects are applied. For the first three years in each cohort, after which treatment effects fade out over the lifetime. Transitions between states across the lifetime are conditioned on age, sex, and deprivation, with a half-cycle correction. Lifetime incremental costs and effects are calculated by comparing across a UIFSM-treated and an untreated population. Model input parameters are summarised in 5.3 and detailed in the following sections.

Figure 5.3: Model input parameters

Parameter	Socioeconomic Measure	Source
Bodyweight prevalence at baseline	IMD	National Child Measurement Programme (2020)
BMI Short-Term Treatment Effects	FSMe & IDACI	Holford and Rabe (2020)
FSM Uptake	FSMe	Department for Education (2022)
Study population	IMD	Office for National Statistics (2020)
Transition Probabilities		
Age 5-17	Parental household equivalised income	Millennium Cohort Study (Connelly et al. 2014)
Age 18-42	Household equivalised income	1970 British Cohort Study (Elliott et al. 2006)
Age 42-50	Household equivalised income	National Child Development Study (Power et al. 2006)
Age 50+	Household wealth	English Longitudinal Study of Ageing (Steptoe et al. 2013)
Baseline distribution of health	IMD	Love Koh et al. (2015)
EQ-5D-3L Utility estimates	IMD	Health Survey for England (Mindell et al. 2012)
Bodyweight relative risk of mortality	N/A	DYNAMO-HIA (Lhachimi et al. 2012)
UIFSM Costs	N/A	Sellen et al. (2018)
Health Care Costs		
Adults	N/A	le Roux et al. (2018)
Children	N/A	Trasande et al. (2017)
Discount Rate		
Health Care Budget	N/A	National Institute for Health and Care Excellence (2013)
Public Health Budget	N/A	National Institute for Health and Care Excellence (2020)
Marginal Productivity		
Health Care Budget	N/A	Lomas et al. (2019)
Public Health Budget	N/A	Owen et al. (2018) & Martin et al. (2020)
Health Opportunity Costs Distribution		
Pro-Deprived	IMD	Love-Koh et al. (2020)
Equally Distributed	IMD	Cookson et al. (2021)

5.3.3 Study Population

The cohort size is estimated by assuming a year group of Reception-age children is one-fifth the size of the 5- to 9-year-old population group from 2014 Office for National Statistics (ONS) data by Index of Multiple Deprivation (IMD) quintile, age, and sex (ONS 2020a). The study population is comprised of 654 473 individuals in total, stratified by sex and deprivation group, representing one school year group in England.

Within each sex-deprivation cohort, baseline overweight and obesity prevalence rates are assigned according to National Child Measurement Programme rates from 2013/4 by a weighted average of prevalence rates in IMD deciles (NCMP 2020). I assign baseline healthy weight prevalence rates as the remaining population not in the obese or overweight states. Overweight rates are not differentiated by sex within deprivation groups, so I apply an overall observed male-to-female sex ratio in 4- to 5-year-olds of 1.1 male pupils for every female pupil. See Table 5.1 for an overview the baseline study population in each cohort at age 5.

Table 5.1: Sample observations and bodyweight distribution by socioeconomic status and sex.

SES Quintile	Sex	Obs.	% Obese	% Overweight	% Healthy
1	Female	77 902	10.7	12.7	76.6
	Male	81 650	11.7	12.7	75.5
	All	159 552	11.2	12.7	76.1
2	Female	65 513	9.9	13.4	76.7
	Male	68 595	10.9	13.4	75.7
	All	134 108	10.4	13.4	83.4
3	Female	58 976	9.2	13.3	77.5
	Male	61 714	10.1	13.3	76.5
	All	120 689	9.7	13.3	84.6
4	Female	57 372	8.4	13.3	78.3
	Male	60 189	9.2	13.3	77.4
	All	117 561	8.8	13.3	77.9
5	Female	59 681	7.5	12.5	79.9
	Male	62 881	8.3	12.5	79.2
	All	122 562	7.9	12.5	79.6

5.3.4 Model Parameters

Transition Probabilities

I calculate transition probabilities from UK cohort data, using marginal effects from an ordered logistic regression model to generate separate probabilities of transitioning between weight states by age, sex, and household square-root equivalised income (Mead et al. 2016). I calculate bodyweight transition probabilities up to age 50 using the Millennium Cohort Study (MCS) sweeps at ages 5, 11, and 17 (Connelly and Platt 2014); the 1970 British Cohort Study sweeps at ages 16, 30, and 42 (Elliott and Shepherd 2006); and the 1958 National Child Development Study sweeps at ages 42 and 50 (Power and Elliott 2006). For older adults, grouped into 50-69 years and 70+ years, I use the English Longitudinal Study of Ageing sweeps in year 2008 and 2012 (Steptoe et al. 2013). I calculate one-year transition probabilities using standard rate-probability conversion methods (Briggs, Claxton, and Sculpher 2006).

Bodyweight categories are classified according to the UK90 external reference for children under 18 years and standard BMI categories for adults aged 18 and over (Cole, Freeman, and Preece 1995). In MCS childhood sweeps, parental income is used, and in late adulthood sweeps from ELSA, wealth is used since it may be a better indicator of socioeconomic advantage in later life (Pollack et al. 2007; Zaninotto and Lassale 2019). Household equivalised income is used as a proxy for deprivation due to limited availability of IMD collection across UK surveys. Likewise, the UK population is assumed to proxy the English population. While the English population represents the majority of the sample in most cases, regional differences may influence results.

Mortality and Health-Related Quality of Life

I use previous analyses of baseline length and quality of life by SES calculated from 2011 UK life tables to establish baseline mortality rates by IMD quintile for 5-year age groups (Love-Koh et al. 2015). I adjust mortality probabilities by bodyweight using (i) age-, sex-, and bodyweight-specific relative risks of mortality from the DYNAMO-HIA UK reference dataset (Lhachimi et al. 2012) and (ii) Health Survey for England 2009-2013 observed prevalence rates of obesity, overweight, and healthy weight by age and IMD quintile (Mindell et al. 2012). These adjustments incorporate differential relative risks of mortality by bodyweight using observed prevalence rate data, while maintaining the overall UK age-sex-deprivation mortality rate.

Health-related quality of life is measured by the three-level, five-dimension EuroQoL (EQ-5D-3L) questionnaire scores (Brooks and De Charro 1996). The EQ-5D-3L is a general health questionnaire, where a respondent's health is assessed across five dimensions – mobility, self-care, usual activities, pain, and anxiety – with three levels indicating (ill) health. Patients are also asked to rate their health on a vertical visual analogue scale from best to worst health imaginable. For this analysis EQ-5D-3L scores are regressed on 10-year age groups, sex, IMD quintile, and three-weight BMI category using Health Survey for England data from years 2012 and 2014. Following literature recommendations, EQ-5D-3L estimates from the 16- to 24-year-old reference group are used for individuals aged 15 and under, since no child-specific HRQoL with approved value sets are yet available (Kreimeier and Greiner 2019).

Bodyweight Treatment Effects

Deprivation Measure

I regard FSMe quintiles as the primary deprivation measure due to some weaknesses of the IDACI's unit of measurement. Both the IDACI and FSMe quintiles are based on school-level data and therefore potentially overlook important within-school individual-level variation. However, while the FSMe measure aggregates individual-level eligibility data within each school, the IDACI is based on school postcode and carries a greater risk of misclassifying the deprivation levels of pupils attending the school.

BMI Measure

I use BMI Z-Score treatment effects from Holford and Rabe (2022) in line with previous studies of obesity interventions in children (Levy et al. 2011). The effect of BMIz changes on obesity, overweight, and healthy weight prevalence rates are calculated using BMI data from 3 683 4- to 5-year-olds in Health Survey for England (HSE) data from 2005-13. Since obesity and overweight rates are substantially lower in the Health Survey for England than rates observed in the National Child Measurement Programme, relative prevalence rate changes are applied in the model. I also use BMI category prevalence rate treatment effects as a robustness check. See Figure C.1 for changes obesity and healthy weight prevalence rates derived from both BMIz and BMI category treatment effects.

I apply treatment effects as percentage changes in prevalence rates of obesity and healthy weight bodyweight categories. Underweight is assumed to be encompassed in the healthy weight category and overweight treatment effects are inferred from

changes in obesity and healthy weight prevalence rates. Upper confidence intervals of the healthy weight category treatment effects represent a "positive" outcome – in the sense that a greater TE would make the children better off – whereas the upper confidence interval of obesity treatment effects represent a "negative" outcome – in the sense that a greater treatment effect would make the children worse off than a lesser treatment effect. I will therefore use the term "maximum" ("minimum") treatment effects, combining the upper (lower) 95% confidence interval healthy weight treatment effect and lower (upper) 95% confidence interval obesity treatment effect, to reflect the range of possible effects rather than the strict definition of upper and lower confidence intervals.

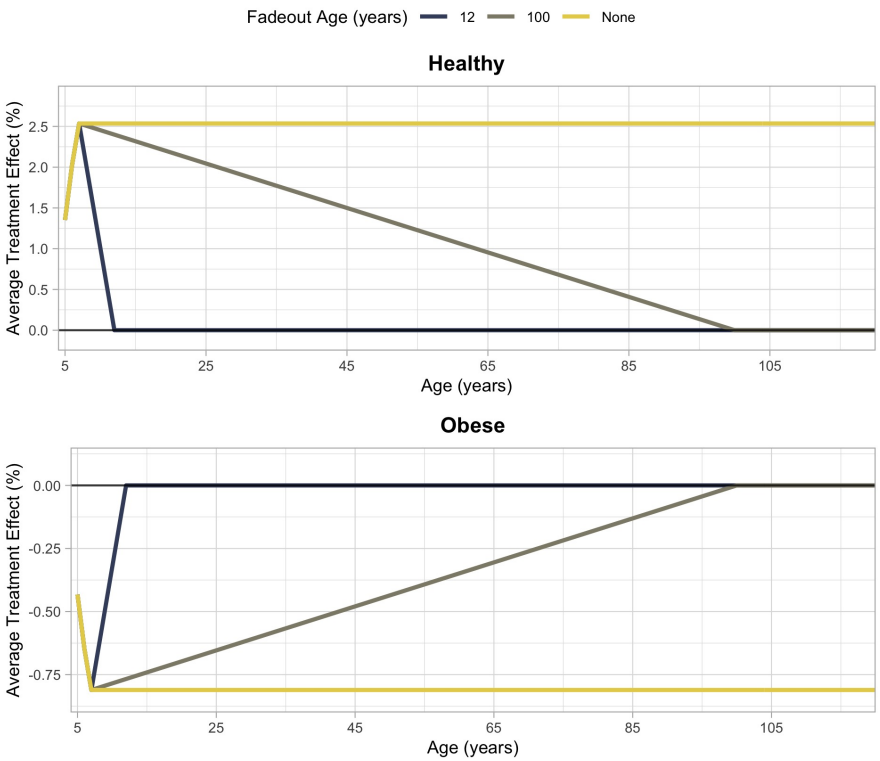
Treatment Effect Fadeout

Treatment effects from Holford and Rabe (2022) are based on Reception-age children only. No further evidence is available on UIFSM treatment effects in the second and third year of the programme, so they are assumed. Treatment effects trends throughout the first year of the policy suggest a consistent decrease in BMIz, using six data points between September and June Holford and Rabe 2020, p. 14. Based on this trend and childhood educational intervention fadeouts (Bailey et al. 2016), I assume incremental treatment effects are halved every year of the policy.

There are also no follow-up data available after the cessation of the intervention available for UIFSM – or sufficiently similar policies (Mead et al. 2017) – so lifetime treatment effect fadeouts are also assumed. There are two main reasons UIFSM treatment effects may last longer than other obesity interventions. First, the policy lasts three years, which is over three times longer than the average childhood obesity intervention (Mead et al. 2017). The length of the policy may lead to greater and longer-lasting effects on bodyweight by influencing long-term behaviours. Second, although the UIFSM stops after Key Stage 1, school meals are still available following the intervention, so treatment effects are more likely to be maintained.

I assume three lifetime treatment effect fadeout scenarios (Figure 5.3). The central scenario assumes treatment effects fade out linearly over the entire cohort lifetime to age 100, where the treatment effect is zero. This long fadeout scenario is also assumed to capture the health benefits of improved educational outcomes discussed in Section 2. The pessimistic scenario assumes treatment effects linearly fade out to zero at age 12. The optimistic scenario assumes the treatment effects are maintained throughout the lifetime, which has previously been the central assumption in cost effectiveness studies of childhood bodyweight interventions (Bates et al. 2020).

Figure 5.4: Average population-level Universal Infant Free School Meal treatment effect on healthy weight and obesity rates by fadeout assumption and age.



Note: Healthy weight includes underweight and bodyweight is categorised according to body mass index (BMI) growth curves until 17 years (Cole, Freeman, and Preece 1995) and standard BMI classifications from 18 years.

Costs: Universal Free School Meals

A detailed evaluation of UIFSM costs has indicated the total public sector cost of UIFSMs amount to £556 million per year on average (Sellen et al. 2018). This estimate includes government costs and net school financial impact. Government costs consists of revenue funding of £494.3 million per year on average (at a FSM cost of £2.30/meal) and average capital funding of £19 million per year (from the initial £190 million investment to support UIFSM implementation). Net school costs average £44.3 million per year, including additional direct costs of delivering school meals and staff costs.

I estimate incremental uptake of FSMs pre- and post-UIFSM - in 2012/3-13/14 and 2014/15-15/16, respectively - from the Department for Education dataset

"Schools, Pupils, and their Characteristics" (DfE 2022) (see Figure C.2 in Appendix C.1). This comprehensive dataset includes FSM uptake and eligibility data from all primary schools in England. I use FSM eligibility proportions in the pupil population to stratify schools in to FSMe quintles, to proxy individual-level SES. Since pre-UIFSM data do not distinguish pupils by age – only including "infants" and "non-infants" – I assume take-up rates are the same for infants as other primary school years. Applying UIFSM public sector cost data to these incremental take-up data, incremental costs ranging from an average of £193-307 per person per year, with higher incremental costs for pupils in schools with lower proportions of FSM-eligible pupils.

All costs are reported in 2015 GBP.

Costs: Health Care

The incremental costs of overweight and obesity are populated using UK estimates from the Clinical Practice Research Datalink and Hospital Episodes Statistics (Roux et al. 2018). The authors identify 396 091 eligible individuals[2] and track their health care expenditure in the five-year period between 1st January 2011 and 31st December 2015. Controlling for age and sex, overweight individuals had £102.00 higher total health care costs per year than their normal weight peers, and obese individuals had £363.15 higher health care costs per year.[3] Total costs consisted of contacts with a general practitioner (GP), hospitalisations, and prescriptions.

Direct childhood health care costs attributable to obesity and overweight are not available in the UK. However, figures are available from the USA in the 2005 Nationwide Inpatient Sample and 2001-05 Medical Expenditure Panel Survey (MEPS) (Trasande 2017). Since UK and USA health care cost levels are not directly comparable, I use ratios between childhood and adulthood incremental costs attributable to obesity and overweight. The health care cost ratio for obese 6- to 11- (12- to 19-) year-olds is 0.115 (0.281) compared to obese adults.

All costs are reported in 2015 GBP.

[2]Defined as persons with acceptable data by CPRD standards, aged 18 years and above, without type 1 diabetes, and with a BMI greater than 18.5 kg/m^2).

[3]Baseline health care utilisation costs for healthy weight adult is £1827.40 per year (Roux et al. 2018).

Budget Assumptions: Discount Rates and Marginal Productivity

Costs and effects (QALY gains/losses) of UIFSM are discounted to reflect lower valuations for costs and effects occurring in the future than in the present. Lower future valuations are justified by catastrophic risks, pure time preferences, opportunity costs, and consumption growth (Attema, Brouwer, and Claxton 2018). Marginal productivity rates inform the health opportunity cost estimates (see subsection "Health Opportunity Costs").

Costs and effects in UK health care interventions are both discounted at a recommended rate of 3.5% (NICE 2013). According to current NICE guidance, public health interventions use differential discount rates of 3.5% per year for costs and 1.5% per year for effects (NICE 2020).The differential discount rate for public health interventions reflects conditions where costs are incurred up-front and benefits are gained in the future. The marginal productivity of the health care sector has been estimated between £5 000 - £15 000 per QALY in 2003-12 (Lomas, Martin, and Claxton 2019), while the marginal productivity of the public health sector has been estimated at £3 800 per QALY in 2013/14 (Martin, Lomas, and Claxton 2020) and £7 843 per QALY in 2011-16 (Owen et al. 2018).

Consequently, I assume two budget scenarios with differing discount rates for QALYs gained and marginal productivity rates (the amount needed to invest elsewhere in to gain one QALY). First, a health care budget where costs and QALYs are discounted at 3.5% per year, while assuming a marginal productivity of £15 000 per QALY. Second, a public health sector budget, discounting costs at 3.5% per year and QALYs at 1.5% per year, while assuming a marginal productivity of £5 000 per QALY.

Health Opportunity Costs

I express the opportunity cost of net universal FSM monetary costs in terms of health, i.e. how many QALYs might have been gained by investing the funds elsewhere in the health care or public health systems. Net costs are equal to the incremental costs of policy implementation less the change in lifetime health care utilisation costs. I calculate incremental health opportunity costs as the incremental monetary costs of universal FSM provision divided by the marginal productivity of the associated budget. Consequently, every £15 000 spent from the health care sector budget is assumed to displace one QALY elsewhere in the health care sector and every £5 000 spent from the public health budget is assumed to displace one QALY elsewhere in the public health sector.

Based on forthcoming findings, I assume the distribution of health opportunity costs for the health care budget is equal for each SES group as the base case. Previous research has also found a pro-poor health opportunity cost gradient, whereby the worst-off SES groups benefit proportionally more from health care investments than the best-off SES groups (Love-Koh et al. 2020). However, this does not account for evidence that health care outcomes are often worse in more disadvantaged populations and so may exaggerate the extent to which more disadvantaged groups disproportionately benefit from additional health expenditure (Cookson et al. 2021c). I therefore conduct sensitivity analyses on this assumption. In the absence of evidence, I assume the same health care opportunity cost distributions apply to a public health budget.

5.3.5 Analyses

The Markov cohort model is designed and implemented in *Microsoft Excel*. I use *Stata v17.0* to compute cohort transition probabilities, EQ-5D estimates, and to transform BMI Z-Score treatment effects to BMI category treatment effects. Results are visualised using *R v4.1.2*.

Maximum Cost-Effective Co-Fund

The first analysis informs decision-makers on the maximum level up to which a co-fund of the UIFSM programme is cost-effective, depending on the budget assumption. The maximum co-fund is calculated as the lifetime incremental QALYs, ΔQALYs, divided by the marginal productivity rate, plus the health care cost savings.

$$Co - Fund = (\Delta QALYs * MarginalProductivity) - \Delta Cost_{HealthCare} \quad (5.1)$$

I express the maximum cost-effective co-fund in 2015 GBP and as a percentage of the total lifetime discounted costs of UIFSMs per year group.

Distributional Equity Impacts

Distributional health equity impacts are described in net health benefits, NHB, as incremental QALYs, $\Delta QALYs$, less the health opportunity costs, Δ HOC, by deprivation quintile, $_j$. Health opportunity costs comprise the incremental cost of universal infant FSM provision, $\Delta Cost$, divided by the sector-specific marginal productivity rate, in GBP per QALY, $MarginalProductivity$, multiplied by the proportion of opportunity costs, P, falling on a given SES quintile, $_j$.

$$NHB_j = \Delta QALYs_j - \Delta HOC_j \tag{5.2}$$

$$\Delta HOC_j = \frac{\Delta Costs}{MarginalProductivity} * P_j \tag{5.3}$$

Distributional Equity Trade-Offs

I use the Atkinson social welfare function (SWF) to assess trade-offs between total health and health inequality, where ε represents inequality aversion; the extent to which a decision-maker is willing to sacrifice total health for the sake of greater relative health equality. ε can vary between zero and infinity: a value of zero represents the utilitarian perspective where a decision-maker is not willing to sacrifice any health level for a more equal health distribution, and a value of infinity represents a decision-maker will sacrifice any level of health for a more equal health distribution. In the base case, I assume $\varepsilon = 10.95$ according to empirical estimates from the UK (Robson et al. 2017). I conduct sensitivity analyses, varying ε between 0 and 20, to assess the robustness of base case results to differing levels of inequality aversion.

The health equity impact plane illustrates the trade-off between equity and efficiency, reporting net health benefits on the y-axis, while reporting equity-weighted net health benefits on the x-axis. Equity-efficiency trade-offs are also represented using equally distributed equivalent (EDE) levels of health (Cookson et al. 2020). The EDE for health represents the level of population health generating the same level of social welfare as that of the observed health distribution, if equally distributed (Asaria, Griffin, and Cookson 2016). For a given level of inequality aversion, the difference between the EDE and mean levels of health represents the health a decision-maker would be willing to forgo to achieve health equality. The Atkinson SWF EDE is calculated as follows:

$$h_{EDE} = \left[\frac{1}{n} \sum_{i}^{n} (h_i)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \tag{5.4}$$

5.4 Results

Henceforth, the base case assumes treatment effects fade out linearly to zero at age 100; UIFSMs are co-funded by 10% by the the health care budget (with a 3.5% QALY discount rate per year and a marginal productivity of £15 000 per QALY), with health opportunity costs distributed equally between deprivation quintiles, using BMI Z-score treatment effects by FSMe.

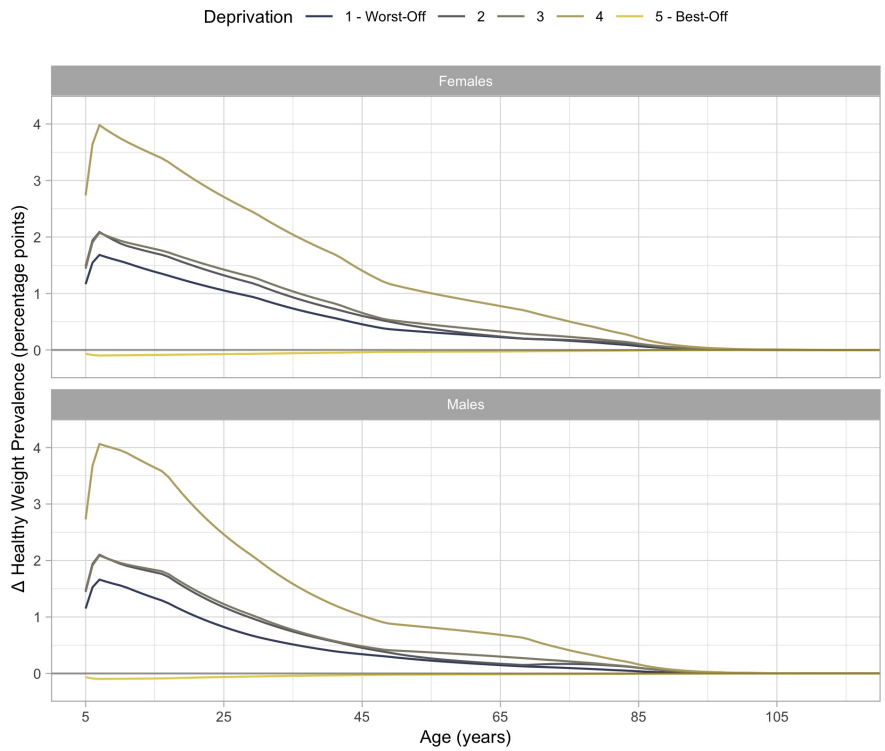
5.4.1 Lifetime Treatment Effects

Short-term treatment effects are extrapolated for each cohort in the Markov model: Figure 5.5 illustrates the base case treatment by age and deprivation. Treatment effects peak at age 7 for all deprivation groups, when the UIFSM programme stops, after which they decrease over the lifetime to zero at age 100. Although treatment effects decline linearly, the effect on health state proportions may be non-linear according to the altered mortality risk of changed population weight status, which is increases exponentially with age. As with short-term treatment effects, extrapolated long-term treatment effects tend to increase with lower deprivation – with very similar treatment effects for the middle and second most deprived quintiles – however the least deprived experience a small decrease in healthy weight rates.

5.4.2 Cost-Effectiveness of Co-Funding

In the base case, UIFSMs generate 5 323 [95% CI: 215, 10 521] lifetime population-level QALYs per year group, and saves the health care system £12 million [95% CI: 2, 22], potentially justifying a total cost-effective health care spend of £92 million per year group [95% CI: 5, 180], or 19 % of UIFSM costs [95% CI: 1, 37] (Table 5.6, see Section 1.4 and Appendix C.2 for sensitivity analyses). Depending on treatment effect fadeout, UIFSMs justify a mean health care co-fund of £15-123 million, or 3-25% of programme costs, per year group.

Figure 5.5: Lifetime treatment effects of Universal Infant Free School Meals on healthy weight proportions (percentage points) by age, sex, and deprivation.



Note: Healthy weight includes underweight and bodyweight is categorised according to body mass index (BMI) growth curves until 17 years (Cole, Freeman, and Preece 1995) and standard BMI classifications from 18 years.

UIFSMs justify a higher level of co-funding from the health care budget than the public health budget due to greater marginal productivity of the public health sector. Nonetheless, UIFSMs potentially justify a mean cost-effective public health spend of £54 million per year group, or 11% of programme costs, given a lifetime cost-saving of £12 million and 8 433 incremental population-level QALYs, assuming treatment effects fade out to zero at age 100.

5.4.3 Distributional Equity Impacts

The distribution of lifetime net health benefits of UIFSM are generally pro-poor, except for the second least deprived quintile, reflecting distributional trends in BMI Z-score treatment effects (see Figure C.1) and budget-specific discount rates and

Figure 5.6: Maximum cost-effective co-fund of Universal Infant Free School Meals per year group (2015 GBP millions and proportion of programme costs) by budget and treatment effect fadeout assumption.

Fadeout Age (Years)	Budget	Maximum Co-Fund											
		£, millions						% Cohort UIFSM Costs					
		Mean TE	(	Min	-	Max	)	Mean TE	(	Min	-	Max	)
12	Public Health	5.6	(	0.4	-	10.9	)	12.0	(	0.8	-	23.2	)
12	Health Care	15.2	(	1.0	-	29.6	)	12.0	(	2.1	-	63.0	)
25	Public Health	14.6	(	1.1	-	28.3	)	31.0	(	2.2	-	60.1	)
25	Health Care	35.0	(	2.3	-	68.2	)	74.4	(	4.8	-	144.9	)
50	Public Health	31.2	(	2.1	-	60.8	)	66.3	(	4.5	-	129.2	)
50	Health Care	63.0	(	3.8	-	123	)	134.0	(	8.0	-	261.9	)
100	Public Health	53.9	(	2.9	-	106	)	114.7	(	6.1	-	225.7	)
100	Health Care	91.6	(	4.8	-	180	)	194.8	(	10.1	-	383.0	)
Never	Public Health	83.6	(	3.7	-	165	)	177.7	(	7.9	-	351.7	)
Never	Health Care	123.3	(	5.7	-	243	)	262.2	(	12.2	-	517.6	)

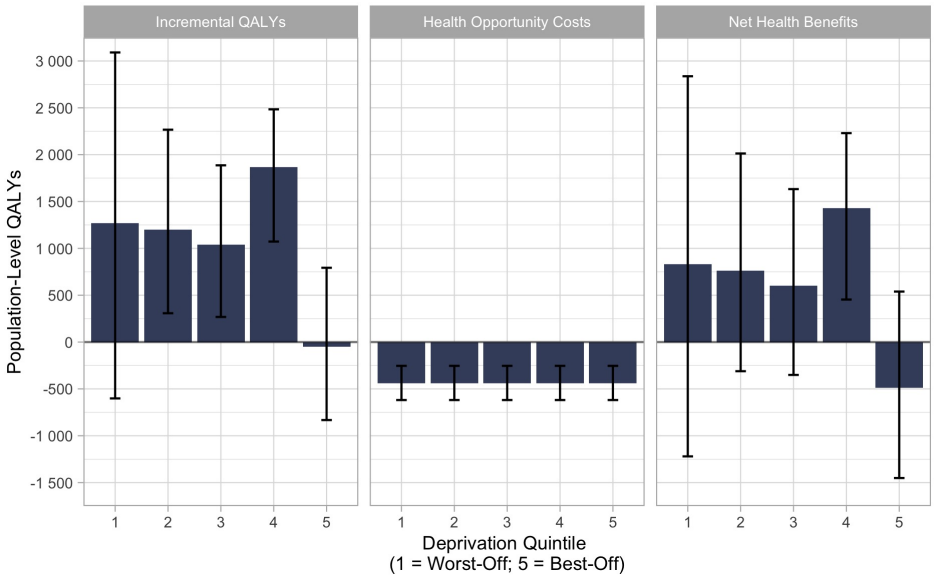
Notes: Costs are reported in 2015 GBP. Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: UIFSM, Universal Infant Free School Meals.

opportunity costs. Figure 5.7 shows distributional lifetime incremental QALYs, health opportunity costs, and net health benefits from a health care perspective, assuming a 10% co-fund of the UIFSM programme and treatment effect fadeout at 100 years. See Section 1.4 and Appendix C.3 for sensitivity analyses.

In the base case, UIFSMs benefit lower deprivation groups more at over 1 200 QALYs for the most (1 267 [95% CI: -601, 3 091]) and second most (1 201 [95% CI: 308, 2 267] deprived quintiles. The middle deprivation quintiles experience approximately 1 000 incremental QALYs (1 038 QALYs [95% CI: 268, 1 886]) while the second least deprived quintile benefit the most: 1 866 QALYs [95% CI: 1 072, 2 484]. UIFSMs slightly harm the least deprived, generating a loss of -49 QALYs [95% CI: -832, 793] over the cohort lifetime. Health opportunity costs are identical across deprivation quintiles at 438 QALYs [95% CI: 254, 619]. Confidence intervals around health opportunity cost estimates reflect differences in cost savings to the health care system from upper and lower 95% confidence interval treatment effect estimates.

The resulting lifetime net health benefits per year group are positive for all but the least deprived quintile, however the null hypothesis of zero net health benefit cannot be excluded for all but the second least deprived quintile. The least deprived

Figure 5.7: Incremental quality-adjusted life years, health opportunity costs, and net health benefits by deprivation quintile.



Notes: Costs are reported in 2015 GBP. Assumes a QALY discount rate of 3.5% per year and a marginal productivity rate of £15 000 per QALY. Abbreviations: QALY, quality-adjusted life years.

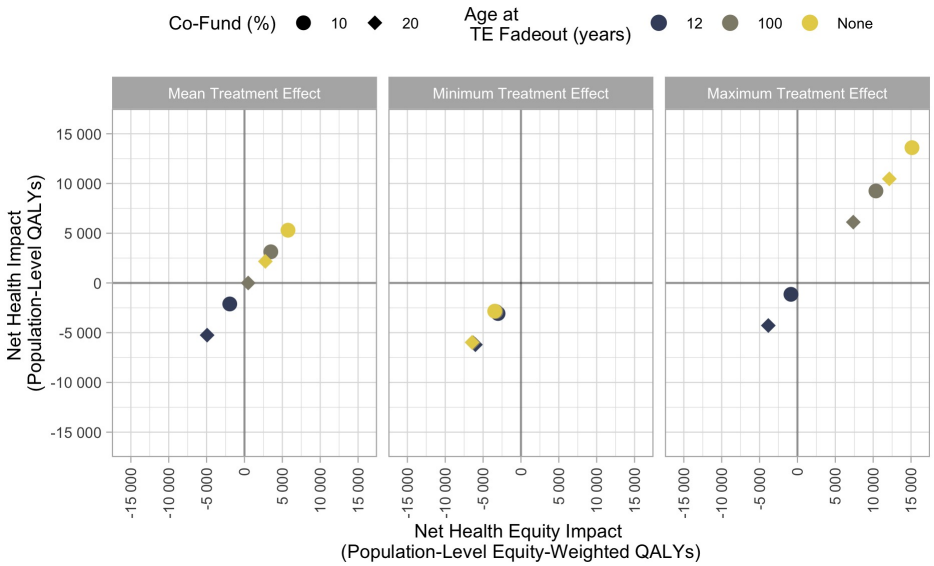
quintile experience the lowest net health benefits of -487 QALYs [95% CI: -1 904, 973] while the second least deprived quintile experience experience the largest net health benefit of 1 428 QALYs [95% CI: -604, 2 568]. For the three most deprived quintiles, net health benefits increasing with increasing deprivation at 829 QALYs [95% CI: -1 220, 2 837] for the most deprived, 763 QALYs [95% CI: -311, 2 013] for the second most deprived, and 600 QALYs [95% CI: -531, 1 632] for the middle quintile.

5.4.4 Distributional Equity Trade-Offs

Figure 5.8 shows the population-level trade-off between efficiency (net health impact) and equity (net health equity impact) assuming $\varepsilon = 10.95$, for each treatment effect fadeout scenario and a co-fund of 10% and 20% of the total UIFSM cohort costs on the equity plane. Estimates in the North-West quadrant both improve health equity and are cost-effective, while estimates in the South-East quadrant harm equity and are not cost-effective. In the base case, where treatment effect fadeout

occurs at age 100, a 10% UIFSM co-fund from the health care sector generates positive net health and health equity impacts; therefore falling into the North-West quadrant of the equity impact plane and improving equity while being cost-effective. Due to negative net health benefits for the least deprived, equity weighted QALYs (3 479 QALYs [95% CI: -3 307, 10 365]) are slightly higher than non-weighted QALYs (3 133 QALYs [95% CI: -2 879, 9 251]). Sensitivity analyses are available in Appendix C.4.

Figure 5.8: Incremental quality-adjusted life years, health opportunity costs, and net health benefits by deprivation quintile.

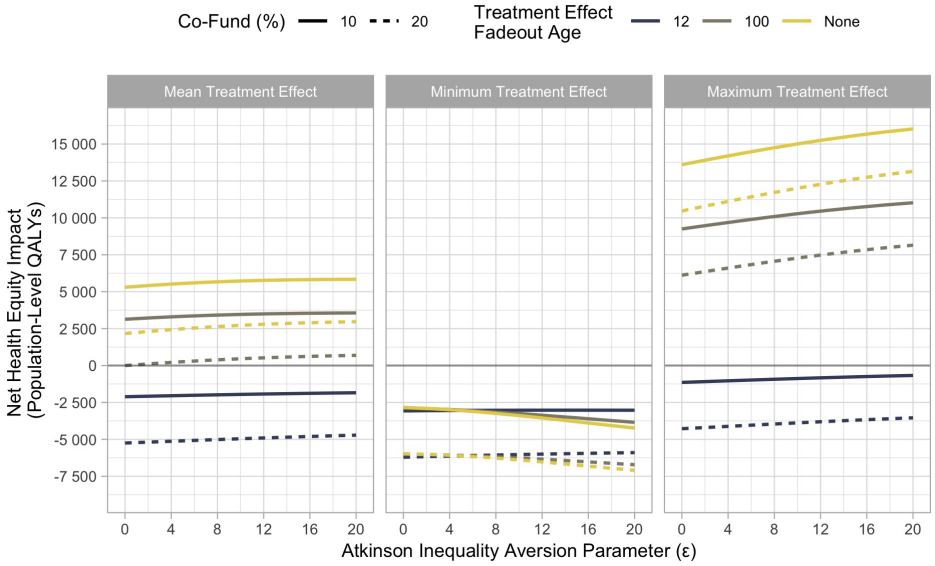


Notes: Costs are reported in 2015 GBP. Assumes a QALY discount rate of 3.5% per year and a marginal productivity rate of £15 000 per QALY. Abbreviations: QALY, quality-adjusted life years.

Assuming a 20% co-fund, UIFSMs generate small net health loss but a health equity benefit: 490 equity-weighted QALYs [95% CI: -6 295, 7 376] and -2 QALYs [95% CI: -6 015, 6 166] when treatment effects fade out at age 100. If treatment effects are maintained throughout the lifetime, net health and net health equity impacts are positive, even at a 20% health care co-fund; 2 759 equity-weighted QALYs [95% CI: -6 455, 12 135] and 2 167 QALYs [95% CI: -5 972, 10 467]. On the other hand, if treatment effects fade out to zero at age 12, net health and net health equity impacts are both negative, a "lose-lose" scenario falling into the South-Eastern quadrant. Treatment effect confidence intervals are large and cannot exclude the null hypothesis of zero treatment effect.

Figure 5.9 shows the sensitivity of net equity-weighted net health benefits to the magnitude of the Atkinson inequality aversion parameter for each treatment effect fadeout scenario and a co-fund of 10% and 20% of the total UIFSM cohort costs. Positive estimates indicate an improvement in health equity, while negative estimates indicate health equity is harmed by UIFSMs. Where estimates cross zero, equity impacts are neutral, indicating the break-even inequality aversion parameter; if a decision-maker is more (less) inequality averse than this parameter value, the intervention is considered to improve (harm) health equity. Under base case assumptions and using mean treatment effects, higher levels of aversion to inequality lead to small increases in equity-weighted net health benefits of UIFSMs, indicating that the policy is modestly inequality improving. However, the magnitudes of these changes are small and do not change the interpretation of results.

Figure 5.9: Incremental quality-adjusted life years, health opportunity costs, and net health benefits by deprivation quintile.



Notes: Costs are reported in 2015 GBP. Assumes a QALY discount rate of 3.5% per year and a marginal productivity rate of £15 000 per QALY. Abbreviations: QALY, quality-adjusted life years.

Using the lower 95% confidence interval treatment effects equity-weighted QALYs are negative at all inequality aversion levels analysed. Greater inequality aversion leads to lower equity-weighted net health benefits, indicating that, contrary to the mean treatment effects, the lower bound treatment effects exhibit a pro-rich distribution. Considering the upper 95% confidence interval of treatment effects, increasing the level of inequality aversion increases equity-weighted net health

benefits, therefore indicating a pro-poor distribution, greater than when considering mean treatment effects.

5.4.5 Sensitivity Analyses

Full sensitivity analyses tables and figures can be found in appendices D.2 for maximum cost-effective co-fund (Table C.1), D.3 for equity impacts (Tables 2-13), and D.4 for equity trade-offs (equity-efficiency impact planes in Figures D.3-10 and inequality aversion sensitivity analyses in Figures D.11-18).

Results are strongly influenced by the treatment effect fadeout assumption. If treatment effects fade out linearly to zero at age 12, rather than age 100, UIFSMs generate 19% of the population-level lifetime QALYs and 4% of the population-level lifetime cost savings, per year group. Net health benefits are therefore negative for every quintile. If, however, treatment effects do not fade out, UIFSMs generate 34% more QALYs and 35% more cost savings over the lifetime per year group. The resulting population-level total net health benefits are approximately 72% greater than the assumption of treatment effect fadeout at 100 years.

Funding UIFSMs from the public health, rather than health care, budget leads to less favourable cost-effectiveness estimates, despite the lower QALY discount rate, due to higher marginal productivity rates in the public health sector. A higher co-fund level of 20%, compared to 10%, reduces net health benefits due to higher health opportunity costs and only results in positive population-level net health benefits under favourable treatment effect fadeout assumption, such as no treatment effect fadeout.

Using BMI category, rather than Z-Score, treatment effects result in lower overall QALY gains and wider 95% confidence intervals, which often cross zero; therefore the null hypothesis of no treatment effect cannot be rejected. According to treatment effects from Holford and Rabe (2020), BMI category bodyweight-improving treatment effects are concentrated among the middle three SES quintiles with harmful effects to the socioeconomically worst-off and small improvements for the socioeconomically best-off (see Figure C.1). Consequently, using the net health benefit distribution is pro-rich and therefore harms equity.

Using IDACI, rather than FSMe, quintiles as the deprivation measure for treatment effects generates approximately 5% fewer total incremental QALYs and a more

homogenous and pro-rich distribution of treatment effects across quintiles. Using a pro-poor, rather than equal, distribution of health opportunity costs results in a modestly more pro-rich distribution of net health benefits, and therefore a more equity-harming outcome.

5.5 Discussion

5.5.1 Primary Findings

Universal Infant Free School Meals (UIFSMs) generate 5 323 incremental lifetime QALYs [95% CI: 215, 10 521] and a lifetime health care cost savings of £12 million [95% CI: 2, 22] per year group of 654 473 children assuming bodyweight treatment effects fade out linearly to zero at age 100 and a discount rate of 3.5% per year for costs and QALYs. If the health care sector co-funded up to £92 million [95% CI: 5, 180] per year group, or 19% [95% CI: 1, 37] of incremental programme costs, the investment would be considered cost-effective at £15 000 per QALY. Assumptions on the time to treatment effect fadeout substantially alter the incremental QALY gains, health care costs saved, and therefore the cost-effectiveness of universal compared to means-tested FSMs.

Assuming the health care sector co-funded £49 million of UIFSM incremental costs per year group, 10% of incremental programme costs, net health benefits are positive for all but the least deprived quintile (-487 QALYs [95% CI: -1 904, 973]). Net health benefits are generally higher for more deprived quintiles, resulting in a pro-deprived overall distribution. However, the pro-deprived gradient is modest, largely explained by the disproportionate net health benefits accrued by the second least deprived quintile (1 428 QALYs [95% CI: -604, 2 568]), which is over two-thirds greater than the next highest net health benefit.

Using Atkinson's social welfare function and an inequality aversion parameter of 10.95, equity-weighted QALYs are approximately 11% greater than unweighted QALYs in the base case, indicating an equity improvement. These results are insensitive to inequality aversion levels due to small gradient magnitudes in the distribution of net health benefits. Short-term treatment effects have wide confidence intervals, which is compounded by uncertainty in long-term extrapolations due to a paucity of evidence on bodyweight treatment effect fadeouts, especially among children.

Literature Comparison

This is the first study of childhood bodyweight interventions to consider lifetime cost-effectiveness and health inequality impacts. Previous UK studies (Brown et al. 2013; Hollingworth et al. 2012) have used the National Heart Foundation model developed by the Foresight Group (McPherson and Brown 2007) to assess average cost-effectiveness of targeted childhood obesity interventions. Due to previous analyses assessing targeted interventions, average effects per person are greater than the UIFSM programme.

One study finds BMIz score reductions of 0.03 and 0.13 resulted in incremental life year gains of 0.02 and 0.19 on average per person, respectively, assuming no discounting and treatment effect maintenance over the lifetime (Hollingworth et al. 2012). Under the same assumptions, UIFSMs generate an additional 0.020 life years per person on average (0.021 for SES 1-3, 0.034 for SES 4, and -0.001 for SES 5). Another study finds a 5% BMI reduction in overweight 12- to 17-year-olds leads to an increase of 0.16 and 0.14 QALYs in overweight males and females, respectively, assuming weight is regained at a rate of 2% per year and a QALY discount rate of 3.5% per year (Brown et al. 2013). Under the same assumptions, UIFSMs generate an additional 0.009 QALYs per person on average (0.009 for SES 1 and 3, 0.010 for SES 2, 0.017 for SES 4 and 0.000 for SES 5).

5.5.2 Strengths and Limitations

This study synthesises available evidence to provide novel evaluations of the lifetime cost-effectiveness and health inequality impacts of UIFSMs. Short term effects are based on robust quasi-experimental causal evidence with previously unknown differential impacts. These short-term effects are extrapolated using lifetime Markov cohort models based on best available evidence about longitudinal obesity trends from UK cohort studies and designed to accommodate distributional costs and effects. Cost-effectiveness estimates shed new light on the extent to which the Department for Health and Social Care could justify co-funding UIFSMs. The DCEA framework provides a summary measure of lifetime health inequality impacts and allows for the explicit analysis of equity-efficiency trade-offs.

This study is subject to limitations. First, this study relies on assumptions on the fadeout of UIFSM treatment effects across (i) the second and third years of the intervention and (ii) the lifetime following the intervention. Second, UIFSM treatment effects are not available at the individual level, so school-level estimates

are used instead. Third, household equivalised income quintile and IMD quintile are used interchangeably where unavoidable. The interpretation of results therefore relies on the degree of comparability between IMD and household equivalised income quintiles (Crawford and Greaves 2013). Likewise, free school meal eligibility (FSMe) is also used as a proxy for socioeconomic status in childhood when estimating treatment effect estimate. While this is not a perfect correlation, research has shown the FSMe binary variable is a reliable proxy for pupils in Wales (Taylor 2018). Fourth, some BMI data from the UK cohort studies, ELSA, and HSE, rely on self-reported BMI, which may be subject to reporting error (Spencer et al. 2002). Fifth, the Markov design cannot incorporate the individual accumulation of risk and complex interactions between SES and bodyweight over the lifetime. Sixth, this study does not systematically address uncertainty via probabilistic sensitivity analysis due to inconsistency in the reporting of standard errors across data sources. Lastly, the study does not consider the full range of societal impacts of UIFSMs. Consequently, it is likely to underestimate the benefits of universal FSM provision.

5.5.3 Policy Implications

In England, 10% of children in Reception, ages 5-6, were obese or severely obese in 2019/2020, rising to 21% in Year 6, age 10-11 (NCMP 2020). These prevalence rates exhibit a social gradient; in 2015/16 – 2019/20 the obesity prevalence rate among the most deprived children in Reception (12.4%) was nearly twice the rate among their least deprived peers (6.3%), as measured by the Index of Multiple Deprivation (IMD). This gap was even wider in Year 6, where obesity incidence rates among the most/least deprived children is 26.1/12.9%.

Rising incidence rates and steepening socioeconomic gradients in obesity are a public health concern because children cannot be held fully responsible for their health-related behaviours (Lobstein et al. 2015), but these behaviours may nonetheless contribute substantially to lifetime health outcomes (Reilly and Kelly 2010; Simmonds et al. 2016). The Long-Term Plan of the UK National Health Service (NHS) calls for "stronger NHS action on health inequalities" (NHS 2019, pp. 39-43) and "a strong start in life for children and young people" [pp. 45-55], while highlighting the growing issue of obesity [pp. 36-37]. Free school meals represent one such policy.

This study informs FSM policy on at least three fronts. First, the study shows co-funding 10% of the UIFSM programme from the health care or public health budgets may be cost-effective under many circumstances. Second, this study

indicates universal FSM provision may improve health equity beyond means-tested provision of FSMs. Considering the comparator of means-tested FSMs, which presumably exhibits a strong pro-poor health gradient, equity-improvements through universal provision are an encouraging result in the context of a strong pro-rich incremental FSM uptake gradient following universal provision (see Figure C.2). Previous research has shown reducing FSM stigma reduces FSM non-participation among eligible pupils (Holford 2015), which may explain positive incremental treatment effects in socioeconomically disadvantaged groups, which were already more likely to be FSM-eligible under the means-tested scheme.

Third, this study informs UK and international FSM policy regarding differences between targeted and universal implementation strategies. By avoiding deadweight loss – funds spent on FSMs for healthy-weight children and the best-off SES quintile (Lundborg, Rooth, and Alex-Petersen 2022) – targeted interventions have higher potential for cost-effectiveness (Dodge 2020). However, this is not to say FSMs should be targeted, particularly when considering results from a narrow health care perspective alone. The universal implementation of FSMs may be justified on humanitarian grounds and by spill-over effects on other sectors, which future research should explore.

5.5.4 Future Research

Future research could build on this study by (i) empirically informing the treatment effect fadeout interval following childhood bodyweight interventions and (ii) conducting a cross-sectoral analysis to incorporate the full range of benefits of universal FSM provision.

The effectiveness of UIFSMs depends on assumptions around the fadeout of treatment effects over the cohort lifetime. Future research would benefit from empirical evidence on UIFSM treatment effect fadeouts. The NCMP gathers data from England in Year 6, which allows for a 4-year follow-up study after the cessation of UIFSMs. This will allow for evidence-based model inputs on long-term effects of FSMs and inform the public health and bodyweight intervention literature, where there is currently a paucity of follow-up data. Similar international evidence may be analysed to assess differences between FSM policies high-income countries. Variations in policy implementation age-groups could describe whether critical periods exist for childhood FSM provision. Where possible, these future studies should include individual-level SES data to explore the distribution of health effects.

The societal costs and effects of obesity and overweight substantially outweigh the direct health impacts (Hamilton, Dee, and Perry 2018; OECD 2019). Future FSM evaluations would therefore benefit from a cross-sectoral approach capturing the full range of benefits from UIFSMs. If these cross-sectoral studies also aim to conduct an equity-informative economic evaluation, they may encounter methodological obstacles, since this will require estimates on the marginal productivity and distribution of opportunity costs for the considered sectors.

CHAPTER 6

Conclusion

This thesis has aimed to contribute practical tools to monitor health inequalities and evaluate the equity impacts of interventions implemented. Ensuring equity monitoring instruments behave desirably by using rank-dependent measures of bivariate health inequality and conducting univariate checks of the health and socioeconomic distributions. Second, illustrating a simple modification that addresses the common criticism that conventional measures of social inequalities in health may partly reflect fair differences in health. Finally, showing how the health equity impacts of non/health interventions may be considered in a simple co-funding scenario using distributional cost-effectiveness analysis. In doing so, this thesis has addressed key elements of the design, implementation, and evaluation of an effective health inequalities strategy.

In this concluding chapter, I first summarise the main findings and implications of Chapters 3–5, then discuss their strengths and limitations, before finally turning to unanswered questions and potential future research directions.

6.1 Summary and Implication of Findings

Health inequalities are a long-standing and growing policy concern. Quantitative evidence and analysis can help decision makers reduce health inequalities by monitoring progress towards appropriate inequality reduction targets and by evaluating the health inequality impacts of policy interventions. However, methodological challenges remain in understanding and addressing the limitations of current methods of health inequality monitoring and evaluation. This thesis has sought to provide decision-makers with a better understanding of the strengths and limitations of current methods, practical tools for addressing those limitations, and new evidence relevant to reducing health inequalities in one illustrative policy area, childhood nutrition.

In Chapter 3, I investigate the policy relevance of a potentially disruptive critique of standard bivariate health inequality measures (BHIMs), by systematically analysing the conditions under which this critique is and is not valid. BHIMs

are linear associations between a current health dependent variable and a current socioeconomic independent variable. They have been criticised because they may theoretically yield misleading conclusions about social justice through their response to changes in the income distribution. In theory, if income inequality reduces while health remains unchanged, measured health inequality may increase, and vice versa. Under certain conditions, therefore, trends in BHIMs may suggest social justice is deteriorating when it may in fact be improving. I have referred to this as the Nordic paradox critique, because it has been raised as one possible explanation for the Nordic paradox (Mackenbach 2012); the finding that countries with more generous welfare arrangements do not necessarily record lower levels of health inequality than countries with less generous social safety nets (Cavelaars et al. 1998; Doorslaer and Koolman 2004; Eikemo et al. 2008; Mackenbach and Kunst 1997; Mackenbach et al. 2008, 2018).

Previous simulation studies have found that both level- and rank-dependent BHIMs may decrease when income inequality increases (Brekke and Kverndokk 2012; Renard et al. 2019). If this is the case, it threatens the validity of the vast international evidence base built upon these measures because they may imply health equity is harmed when income inequality decreases. However, these studies do not systematically consider rank-preservation in the income distribution and multiple plausible causal pathways when interpreting BHIM results. I find rank-preserving increases in income inequality may decrease level-dependent BHIMs, regardless of the underlying causal pathway driving the income-health relationship. This is not the case for rank-dependent BHIMs, which remain unchanged if income does not cause health, when subjected to the same change.

These findings suggest BHIMs are appropriate monitoring instruments so long as they are (i) based on correlations with socioeconomic rank rather than absolute income, which is generally recommended (Mackenbach et al. 2015; Moreno-Betancur et al. 2015), and (ii) there is little change in income ranks. Income ranks are plausibly stable when evaluating within-country health care policies and other social policies with limited impacts on the income distribution, as is ordinarily the case. However, in instances where income ranks may be substantially altered, such as major societal and/or political changes, BHIMs may risk encountering the Nordic paradox critique. This may also be the case for cross-country comparisons where the income distributions differ substantially. In such cases, univariate checks of the income distribution may be conducted and, if appropriate, decision-makers may choose to employ more complex methodologies capable of capturing income re-ranking.

In Chapter 4, I tackle the criticism of BHIMs that they may reflect health outcome differences associated with social determinants at least partly within a person's control. I illustrate how commonly used bivariate health inequality measures may be adapted to provide a simple indication of socioeconomic inequality of opportunity for health by using a socioeconomic determinant outside a person's control. This simple inequality of opportunity metric, a parental socioeconomic gradient in health, estimates the association between a person's parental socioeconomic fraction-rank during childhood and a person's own health in adulthood. Parental gradients offer a practical alternative to more complex inequality of opportunity frameworks that require the ethically controversial task of disentangling causal pathways between many different determinants of health for which a person can and cannot be held responsible.

I find significant parental socioeconomic gradients in general health at age 50 for a representative UK cohort, indicating substantial inequality of opportunity for health whereby health is determined by socioeconomic factors beyond a person's control. For children born in the UK in 1958, growing up in the most advantaged household was associated with 12-24% higher self-reported health (as measured by the 36-Item Short Form Health Survey) at age 50, compared to growing up in the least advantaged household. Parental gradients are approximately 40-65% the magnitude of current socioeconomic gradients in self-reported general health at age 50. This may imply that roughly half the social gradient in health can be attributed to childhood circumstances beyond a person's control. Robustness checks indicate parental gradients may be mediated by intergenerational socioeconomic mobility and the association between current socioeconomic characteristics and health. However, parental gradients remain significant and substantial, even after controlling for current socioeconomic variables.

Decision-makers may use parental gradients to routinely monitor the development of national- or local-level inequality of opportunity, and conduct cross-country studies, provided robust data including parental SES are available. Parental gradient findings may then be used to justify the need for further research using more sophisticated inequality of opportunity methods. Parental gradients are easy to compute and interpret, provided valid data are available, because they leverage decision-makers' familiarity with standard current gradients. Parental gradients may also be directly compared to current gradients, using a person's own socioeconomic variables in adulthood. Consequently, parental gradients equip policy makers with a practical tool to quantify a stricter subset of total health inequalities, which may be raised by critics of conventional current gradients.

Finally, in Chapter 5, I apply distributional cost-effectiveness analysis (DCEA) to quantify the health inequality impacts of interventions outside the health sector. I find the Universal Infant Free School Meal (UIFSM) programme generates an additional 5 323 population-level discounted incremental lifetime QALYs and a lifetime health care cost savings of £12 million per year group of 654 473 children. Consequently, the health care sector could co-fund 19% of incremental programme costs and the investment would be considered cost-effective, assuming bodyweight treatment effects fade out linearly to zero at age 100.

Assuming the health care sector co-funded £49 million of UIFSM incremental costs per year group, net health benefits are positive for all but the least deprived quintile. Net health benefits are generally higher for more deprived quintiles, resulting in a pro-deprived overall distribution. Hence, incremental equity-weighted QALYs are over 10% greater than unweighted QALYs. These findings suggest universal FSM delivery may have equity benefits beyond means-tested FSM delivery; health benefits are concentrated among children in the most deprived households, despite a strong pro-rich UIFSM incremental uptake gradient.

Results and methods from Chapter 5 illustrate how DCEA methods may be applied to quantify the health inequality impacts and trade-offs of intervention benefits falling across sectors. This is especially important in the context of constrained governmental and educational budgets where investments, and their opportunity costs, must be carefully considered when allocating resources. Decision-makers may use this approach to spread the financial burden of policies which have benefits across sectors. For example, findings from this chapter show that up to one-fifth of UIFSM incremental costs could be funded by the Department for Health and Social Care and the investment would be considered cost-effective and equity-improving from a health care perspective.

6.2 Strengths and Limitations

The studies in this thesis aim to provide practical methodological improvements on the methods for measuring health inequality levels and impacts, with policy-makers in mind. To do so requires a balance of empirical rigour and applicability to organisations which must use these methods routinely and may be both resource and time constrained when doing so. The methods presented in this thesis have attempted to strike this balance. While this is a strength of the studies in this thesis, it also presents some limitations. Where limitations have been encountered,

they have been transparently documented and accompanied by sensitivity analyses to test the validity of simplifying assumptions.

The simulation model in Chapter 3 is the first study to systematically consider the conditions under which the Nordic paradox critique is relevant to decision-makers relying on BHIMs to monitor the development of health inequality. Unlike previous studies, the simulation models in Chapter 3 include three plausible causal assumptions driving the income-health association: income causing health; health causing income; and income and health caused by a third factor, proxied by non-cognitive skills. The study also contributes to the literature by illustrating the impact of rank-preserving income changes on both level- and rank-dependent BHIMs. Lastly, the simulation models are applied to five of the most commonly used BHIMs: the Pearson's correlation coefficient, slope index of inequality, relative index of inequality, concentration index, and generalised concentration index. This ensures results are relevant to decision-makers who rely on BHIMs to provide indicators for the development of health inequality.

While results from Chapter 3 lend confidence to within-country routine monitoring when the income ranks are largely stable – such as findings that socioeconomic health inequalities have increased in the UK (Marmot et al. 2020) – it may cast doubt on cross-country studies where income distributions differ substantially and within-country studies where income ranks are volatile. However, the simulation study is subject to limitations due to the simplicity of the simulation approach, which cannot reflect the expected complexity of multiple causal pathways in empirical settings. It may therefore be possible that findings from theoretical investigations are difficult to extrapolate to empirical settings. Furthermore, the simulation models do not consider the impact of socioeconomic rank mobility on BHIMs due to programming constraints, which may nonetheless play an important role in explaining the magnitude of socioeconomic gradients in health (Kreiner, Nielsen, and Serena 2018).

In Chapter 4, parental gradients illustrate inequality of opportunity metrics which are easy to compare, estimate and interpret, while addressing the growing literature and political focus on childhood origins of later life outcomes. I use rich birth cohort longitudinal data from 1958 National Child Development Study and account for attrition using multiple imputation by chained equations to restore sample representativeness, which is likely to increase the accuracy of socioeconomic health inequality estimates. Parental gradients provide decision-makers with an accessible inequality of opportunity measurement tool which can be scaled up for monitoring purposes. The study also conducts several robustness checks to contextualise

existing literature and inform future research on: (i) intergenerational socioeconomic rank mobility; (ii) the impact of controlling for current socioeconomic variables on parental gradient magnitudes; (iii) the impact of socioeconomic measurement timing during childhood on parental gradients; and (iv) how missing data may bias parental gradient estimation using longitudinal cohort evidence. These analyses illustrate practical checks that may be conducted alongside parental gradient estimations to validate and contextualise results.

The parental gradient approach to inequality of opportunity measurement is limited by data availability, requiring longitudinal data from childhood or cross-sectional survey data that accounts for recall and survivorship bias. Comparisons with parental gradients calculated from longitudinal data from childhood also require both missing data methods to account for attrition and comparable socioeconomic variables between parental measurement in childhood and respondents' own measurement in adulthood. However, parental gradients are less data-intensive than sophisticated inequality of opportunity econometric approaches. Parental gradients are also limited because they are associative, hence multiple causal pathways may plausibly be driving the relationship between childhood socioeconomic circumstances and health outcomes in adulthood. Parental gradients in adult health may encounter pragmatic difficulties in informing policy design because interventions are likely to experience long lags between childhood socioeconomic status and adult health outcomes. Finally, parental gradient measures systematically underestimate inequality of opportunity by selecting a single variable to proxy childhood circumstances and thereby ignoring other determinants of the social gradient in health beyond a person's control.

In Chapter 5, the DCEA study sheds new light on the previously unknown long-term cost-effectiveness and health inequality impacts of universal FSMs from a health care perspective to inform national and international governmental decision-making on a topical and influential policy. The DCEA is based on a *de novo* Markov cohort model that builds distributional considerations directly into the economic model by leveraging rich longitudinal data from the UK to extrapolate robust quasi-experimental empirical evidence on short-term UIFSM treatment effects. In doing so, the study illustrates how DCEA can be used to inform the extent to which the Department for Health and Social Care could co-fund non-health interventions with health benefits, while accounting for both cost-effectiveness and equity considerations.

This study is subject to methodological limitations due to a lack of data. First, the DCEA assumes the development of treatment effects with the UIFSM treatment

period and the fadeout of treatment effects after treatment cessation. This is the greatest source of uncertainty in results. Second, data used to inform the Markov model parameters were not all disaggregated by socioeconomic quintile, and those that were sometimes used different measures of socioeconomic advantage. The analysis is therefore limited to the extent that (a) data points without distributional information may influence the lifetime distribution of net health benefits and (b) different proxies for socioeconomic advantage are comparable. Finally, the study does not conduct a comprehensive cross-sectoral economic evaluation, due to a lack of data on opportunity costs in non-health sectors and their distribution across deprivation quintiles. A narrow health care focus is likely to underestimate the total benefits of UIFSMs, including the total health benefits, by not accounting for other benefits of FSMs, such as improved educational outcomes. These multi-sectoral impacts on health are approximated in the model by allowing favourable assumptions on UIFSM treatment effect fadeouts, however more systematic analyses within a cross-sectoral should be conducted where possible.

6.3 Future Research and Policy

Findings in this thesis present opportunities for future research in measuring health inequality and health inequality impacts. Across studies, data limitations have also been encountered, which future studies should address.

To further validate BHIMs, future research should explore the impact of income mobility on the consistency of BHIM responses to income distribution changes in inequality monitoring environments. This study may encounter programming challenges in attempting to systematically vary the degree of rank mobility. Likewise, a simulation study allowing for multiple co-existing causal pathways driving the income-health relationship may encounter programming challenges, but would allow for simulation analyses that more closely replicate the empirical environment. This research may benefit from estimating both current and parental rank-dependent bivariate health inequality measures as well as socioeconomic mobility.

Future studies may provide further insights to the Nordic paradox. The results could be used to generate a checklist to inform analysts and decision-makers on the circumstances in which certain bivariate health inequality measures may be appropriate, and tests to conduct to assess this appropriateness. In contexts where BHIMs may risk misleading results, more advanced methods that account for substantial re-ranking in the weighted fraction-rank of the socioeconomic variable

may be considered (Allanson, Gerdtham, and Petrie 2010; Coveney et al. 2020; Van Ourti, Doorslaer, and Koolman 2009).

Parental gradients illustrated in Chapter 4 allow for the routine monitoring of inequality of opportunity for health within and between countries. In the UK, future research would benefit from assessing how parental gradients have developed for cohorts born since the mid-20th century, using readily available both cohort data (Connelly and Platt 2014; Elliott and Shepherd 2006; Power and Elliott 2006; Wadsworth et al. 2006). In doing so, it may be possible to indicate how changes in social, political, and economic circumstances during childhood have influenced lifetime inequality of opportunity for health using comparable measures. Evidence from this study suggests multiple imputation methods may not be required to provide an accurate indication of inequality of opportunity for health using parental gradients provided longitudinal socioeconomic data from childhood are used. Parental gradients may be applied to study inequality of opportunity for health internationally; for example, monitoring the parental education gradients in mortality. International comparisons may be especially impactful if comparable and robust socioeconomic and health measures can be regularly collected. Where these studies rely on retrospective self-reported childhood socioeconomic circumstances in survey data, the issue of recall and survivorship bias must be addressed.

Questions also remain how parental gradients develop across life course. The study in Chapter 4 assessed health differences at age 50 to capture health differences that tend to emerge later in life (Galama and Kippersluis 2018). Studies assessing the dynamics of parental gradients over the life course may inform decision-makers on critical periods for intervention. Where prevention is the desired policy aim, parental gradients in early life health behaviours may be well-placed to capture inequality of opportunity for health before they manifest in later life outcomes. The limitations of self-reported health used in Chapter 4 also present opportunities for future research; self-reported general health measures may be subject to biases or have limited comparability across populations. Parental gradients using health data from clinical practices, biobanks, or mortality registers would therefore inform decision-makers on the degree to which parental gradients persist when using objective health outcomes.

Future studies on the long-term health inequality impacts of FSMs would benefit from leveraging follow-up data from the National Child Measurement Programme in Year 6 to address data paucities on long-term fadeout of childhood interventions. Multiple data points collected at shorter intervals would also provide novel data on the degree to which fadeout effects following a childhood intervention are linear or

non-linear. This would have implications beyond the FSM programme, as there is a paucity of follow-up data from childhood interventions to childhood behavioural and environmental change treatment effect fadeouts. Where possible, this follow-up data should be disaggregated by deprivation since inequality concerns are embedded in the policy objectives and therefore a key outcome of interest. Disadvantaged children are more likely to maintain access to FSMs via the means-tested scheme after they age out of the UIFSM scheme, so it may therefore be plausible that treatment effects are slower to fade out among more deprived children, but this remains to be tested.

Future studies on FSMs, and other non-health interventions with health benefits, could also extend the study in Chapter 5 by incorporating a societal perspective in an established framework. An impact inventory approach would include the distribution of intervention costs, effects, and opportunity costs across both the health and education sectors, when considering the childhood impacts FSMs (Walker et al. 2019). Lifetime cross-sectoral analyses could broaden the scope further to include labour market outcomes, thereby capturing potential productivity gains and reductions in employment-related health inequalities associated with improved childhood educational and bodyweight outcomes. Impact inventory approaches will rely on extensive data collection and analyses, such as estimations of the marginal productivity and distribution of opportunity costs for the sectors considered, which may prove to be a methodological obstacle for this approach.

Future research may complement and expand upon the objectives of this thesis by further adapting and implementing methods to address a decision-maker's most pressing concerns. This will likely include practical instruments for health inequality measurement that may be used routinely to provide relevant and actionable insights on health and social variables, as well as the relationship between them. Limitations encountered in this thesis also call for the routine collection of health data disaggregated by social variables. This is essential to ensure future research is capable of informing decision-makers on health inequality impacts alongside other metrics of interest, such as efficiency and quality of care.

Findings from this thesis may strengthen national health inequality strategies rooted in the quantification of health inequalities for measurement and monitoring purposes, as well as in the design, implementation, and evaluation of interventions related thereto. Findings from this thesis show that bivariate health inequality measures used by the Office for Health Improvement and Disparities (OHID) could be strengthened by supplementing with: (i) assessments of the distributional changes and rank mobility in both the health and socioeconomic variables, and

(ii) parental socioeconomic gradients in adult health. This would strengthen the national health inequality monitoring programmes on two fronts. First, checking the income and health variables would lend more credibility to the estimates currently being monitored, such that decision makers can be more confident changes in health gradients are not spuriously driven by changes in the distribution of the socioeconomic variable. Second, parental gradients address a monitoring gap by providing quantifiable and interpretable evidence on the otherwise unexplored importance of equality of opportunity and childhood circumstances in explaining current health gradients. This is especially important since it aligns the monitoring programme more closely with the increasing policy and research focus on childhood. However, both extensions to the current monitoring programme require collaborations across the Office for National Statistics, Department for Health, and the Treasury to provide the necessary data. Lastly, interventions and policies should be evaluated in terms of both cost-effectiveness and equity impacts, and their trade-offs considered, through the use of equity-informative economic evaluation methods, such as distributional cost-effectiveness analysis. As with reporting parental gradients, equity-informative cost-effectiveness analysis imposes greater data requirements, some elements of which may be difficult to capture in childhood interventions. Nonetheless, the informational infrastructure in Nordic countries is an actionable example of how high-quality registry data may be collected and linked, providing a strong foundation for an effective health inequalities strategy.

Appendix

A Appendix to Chapter 3

A.1 Descriptive Statistics

Figure A.1: Distribution of health-adjusted life expectancy and household equivalised annual income in the social causation simulation model

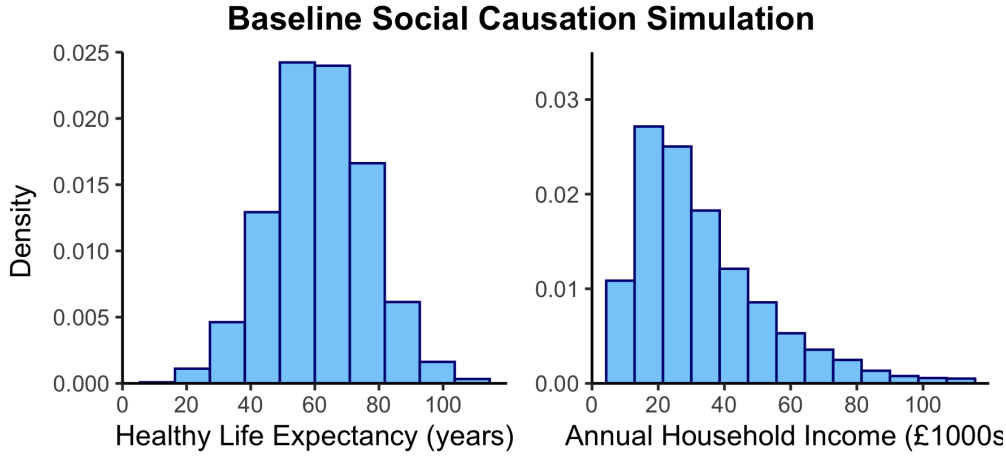


Figure A.2: Distribution of health-adjusted life expectancy and household equivalised annual income in the health selection simulation model

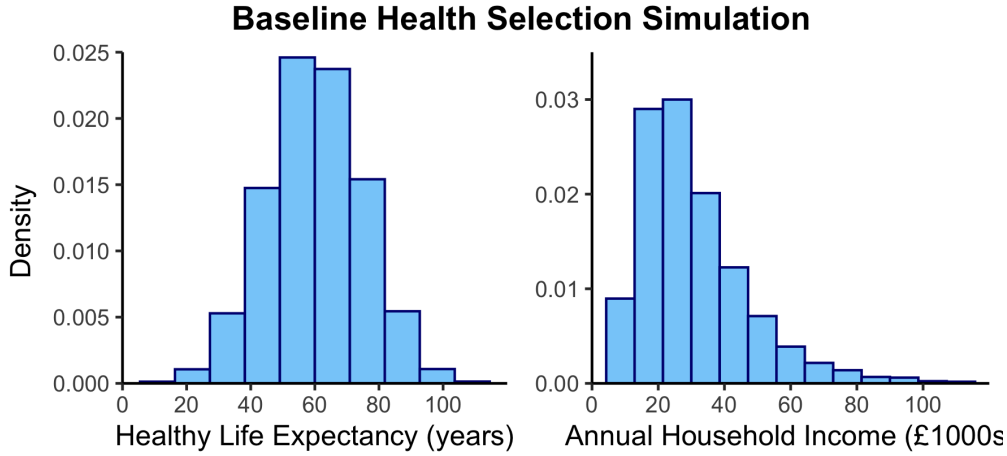
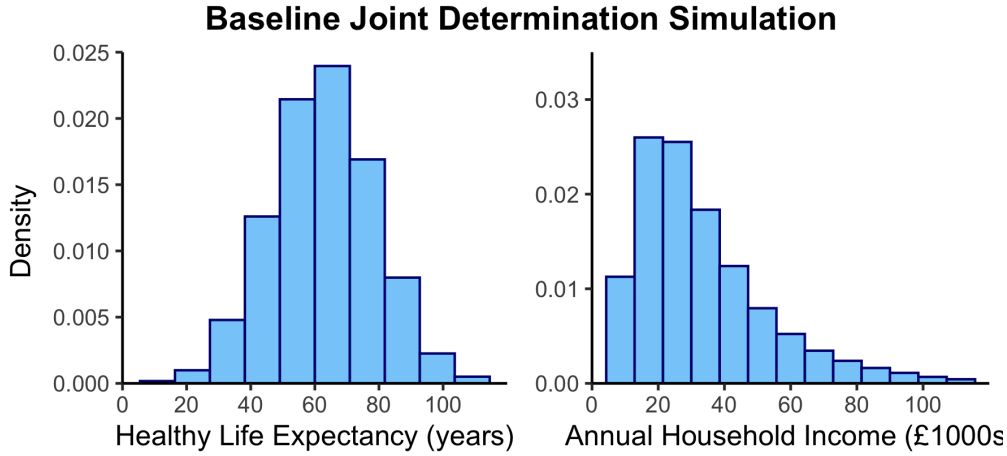


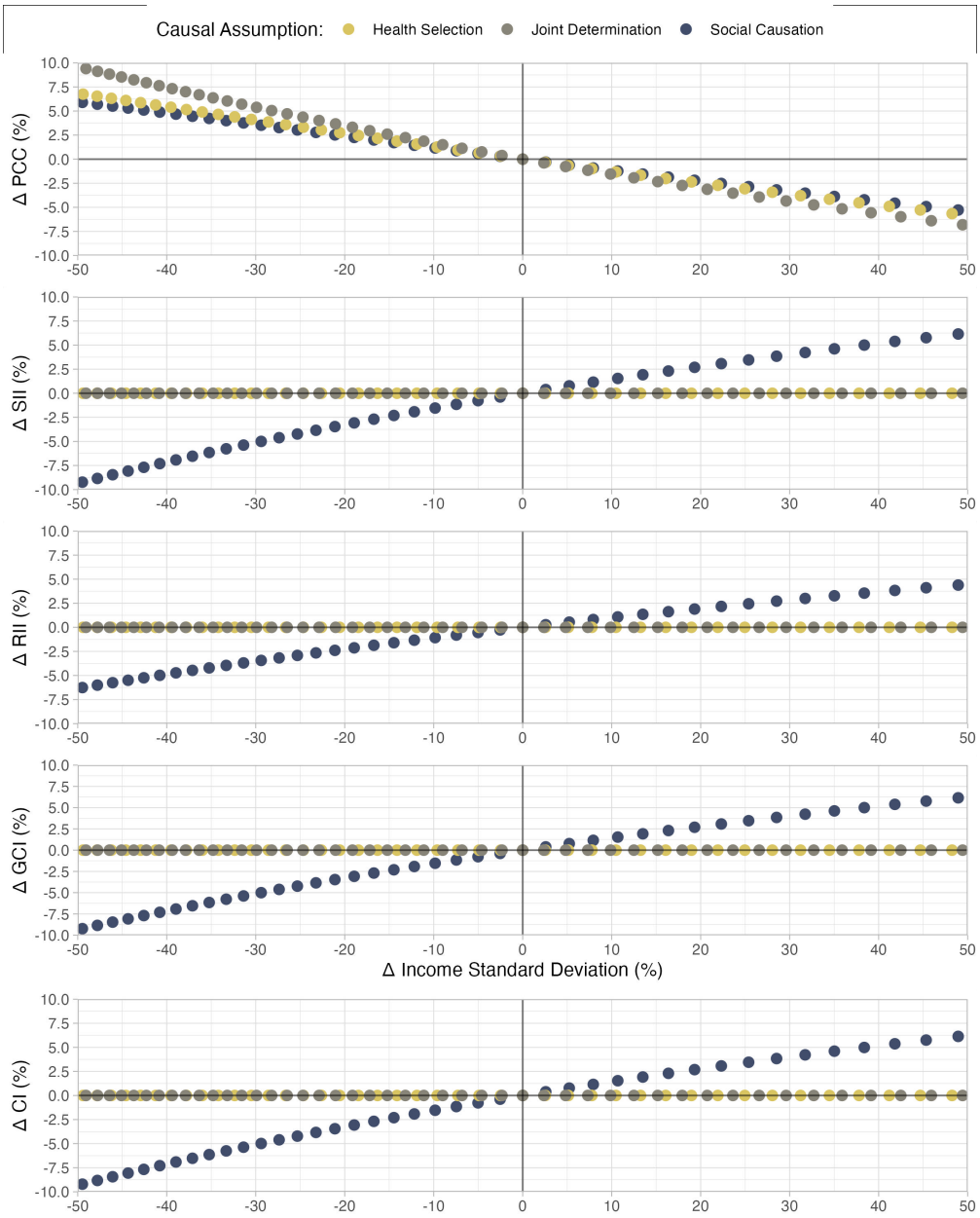
Figure A.3: Distribution of health-adjusted life expectancy and household equivalised annual income in the joint determination simulation model



A.2 All Simulation Output Graphs

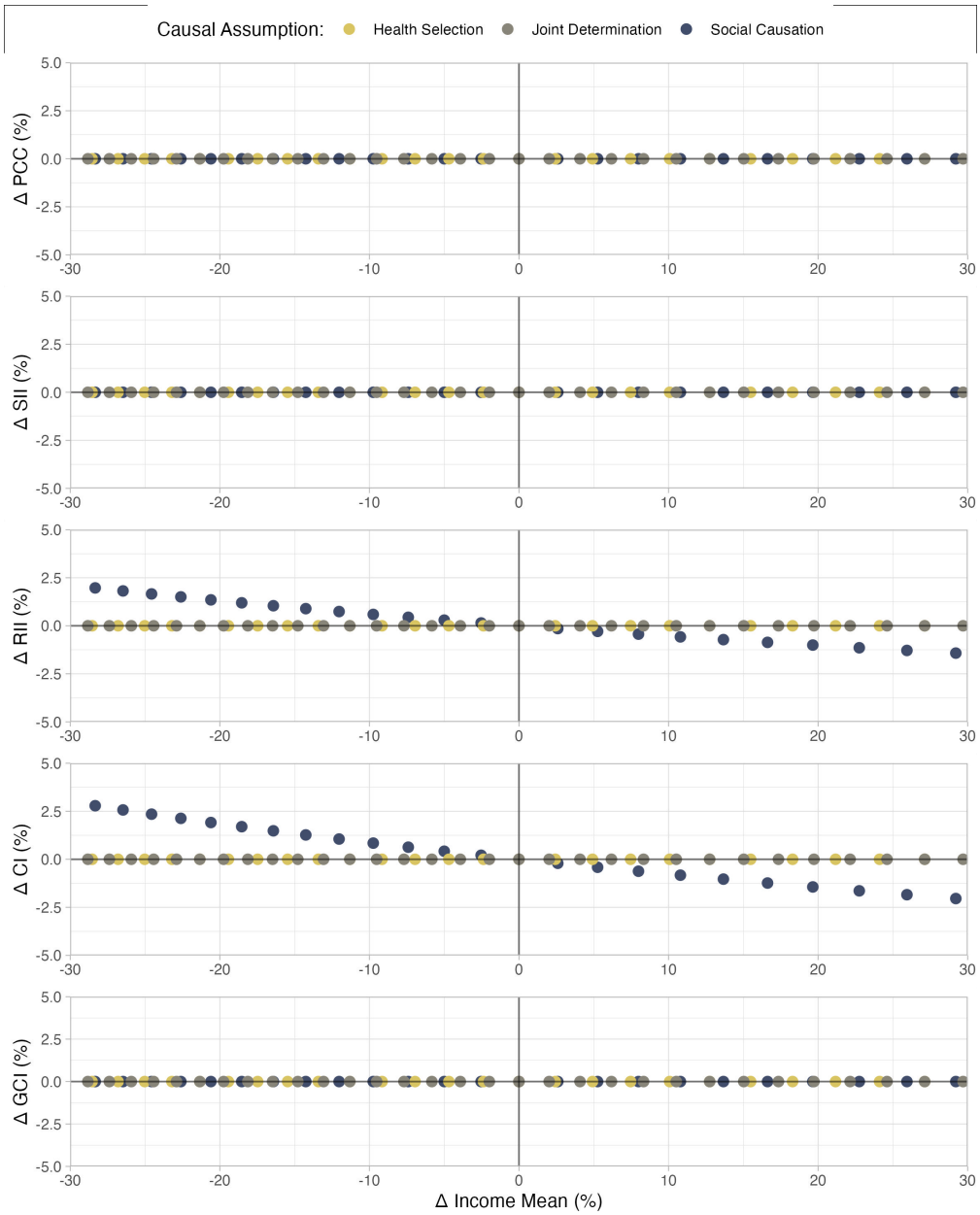
Rank-Preserving Income Changes

Figure A.4: Impact of rank-preserving income inequality changes on bivariate health inequality measures by causal assumption.



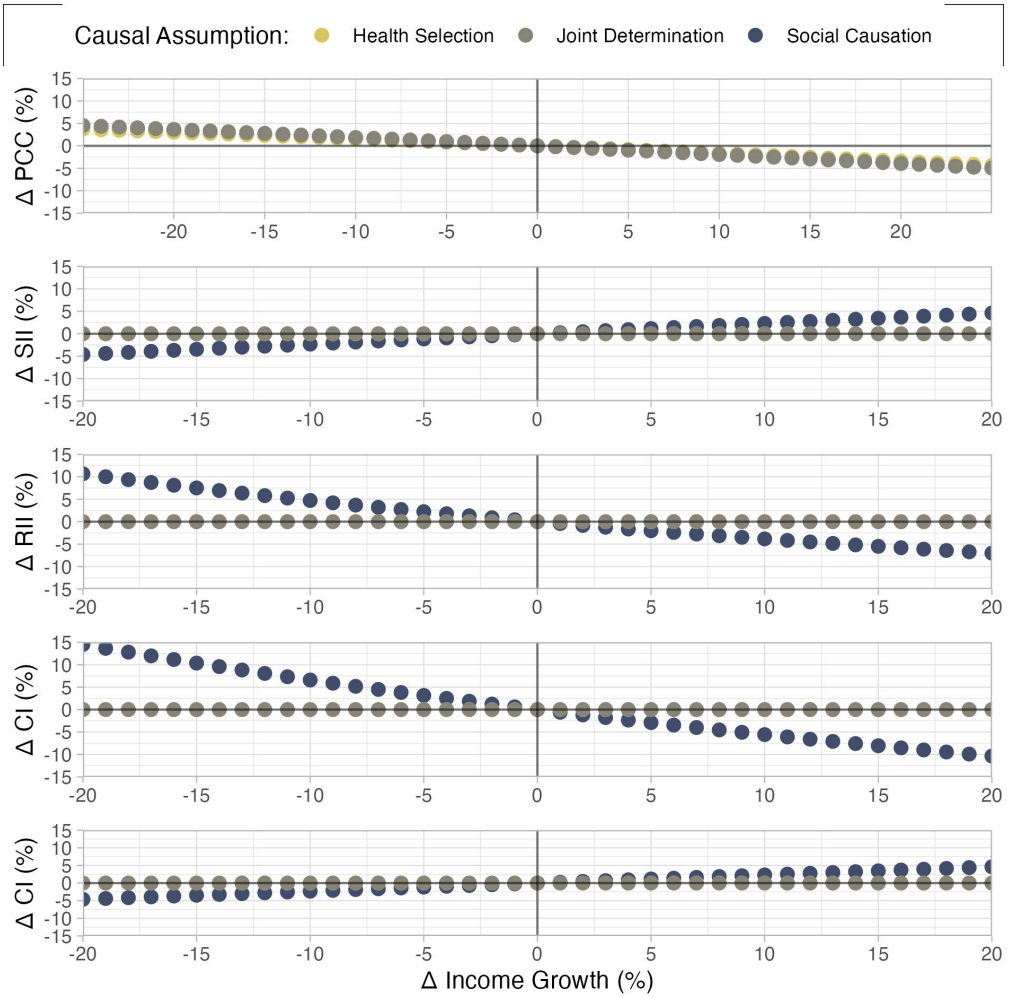
Note: Social causation assumed income causes health, health selection assumes health causes income, and joint determination assumes income and health are both caused by noncognitive skills. Abbreviations: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality; GCI, generalised concentration index; CI, concentration index.

Figure A.5: Impact of rank-preserving income level changes on bivariate health inequality measures by causal assumption.



Note: Social causation assumed income causes health, health selection assumes health causes income, and joint determination assumes income and health are both caused by noncognitive skills. Abbreviations: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality; GCI, generalised concentration index; CI, concentration index.

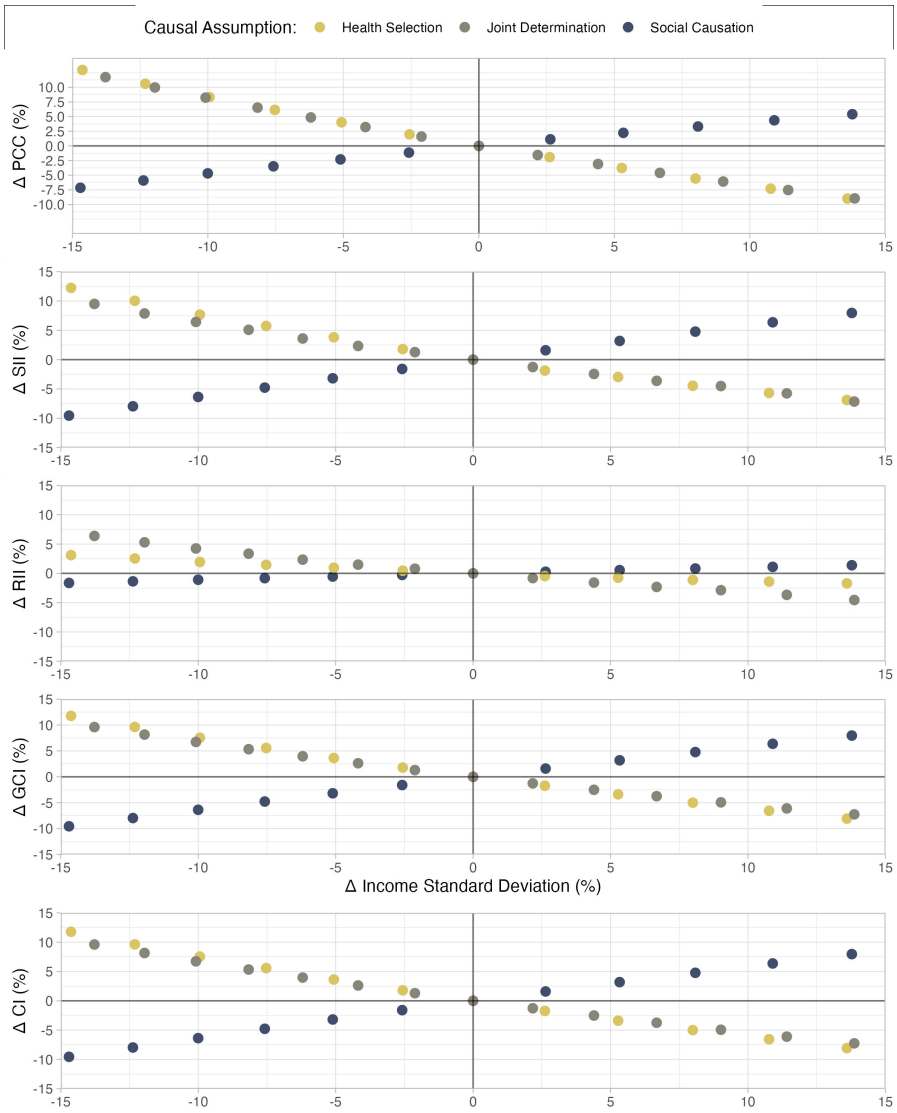
Figure A.6: Impact of rank-preserving income growth changes on bivariate health inequality measures by causal assumption.



Note: Social causation assumed income causes health, health selection assumes health causes income, and joint determination assumes income and health are both caused by noncognitive skills. Abbreviations: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality; GCI, generalised concentration index; CI, concentration index.

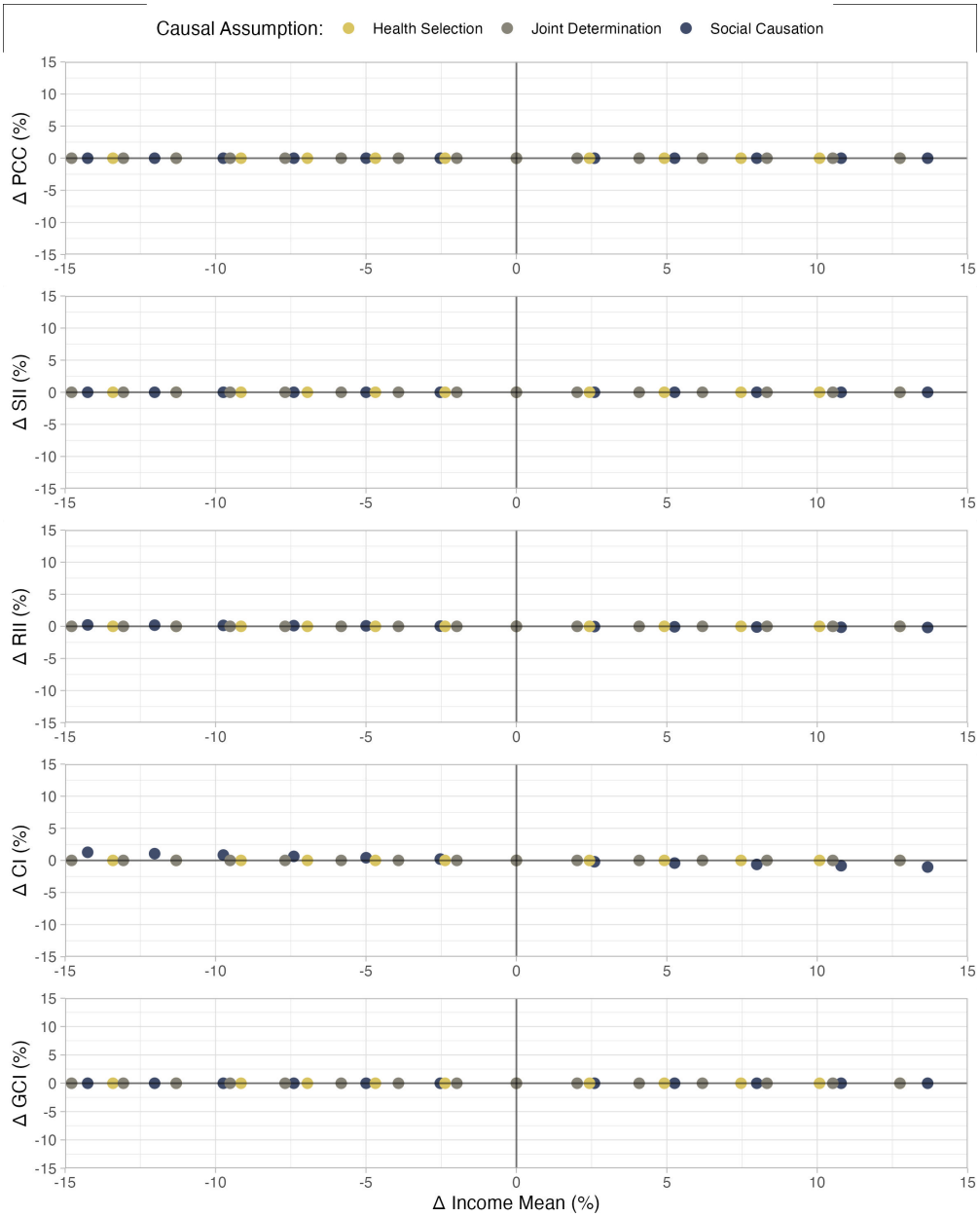
Rank-Altering Income Changes

Figure A.7: Impact of rank-altering income inequality changes on Pearson's correlation coefficient, slope index of inequality, relative index of inequality, concentration index, and generalised concentration index under social causation, health selection, and joint determination causal pathway assumptions.



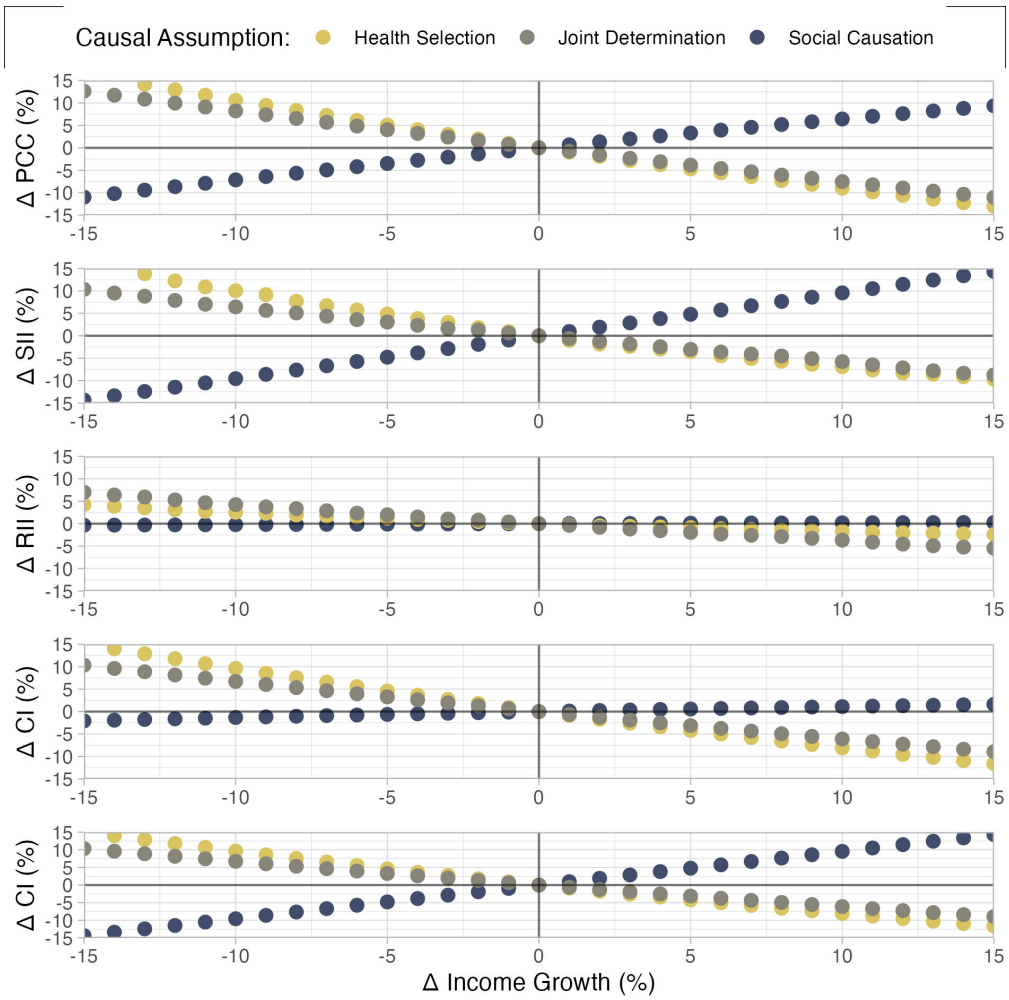
Note: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality; GCI, generalised concentration index; CI, concentration index.

Figure A.8: Impact of rank-altering income level changes on Pearson's correlation coefficient, slope index of inequality, relative index of inequality, concentration index, and generalised concentration index under social causation, health selection, and joint determination causal pathway assumptions.



Note: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality; GCI, generalised concentration index; CI, concentration index.

Figure A.9: Impact of rank-altering income growth changes on Pearson's correlation coefficient, slope index of inequality, relative index of inequality, concentration index, and generalised concentration index under social causation, health selection, and joint determination causal pathway assumptions.



Note: PCC, Pearson's correlation coefficient; SII, slope index of inequality; RII, relative index of inequality; GCI, generalised concentration index; CI, concentration index.

A.3 Simulation Reporting

Table A.1: Simulation reporting checklist according to Cheng et al. (2016).

Item	**CONSORT Item number**	**Extension for SBR**	**Page No.**	**Relevant text from manuscript**
Title and abstract	1	In abstract or key terms, the MESH or searchable keyword term must have the word "simulation" or "simulated."	55	"Simulating" – Title "Simulation" – Abstract
Background	2	Clarify whether simulation is subject of research or investigational method for research.	62	"investigational" – Methods
Interventions	5	Describe the theoretical and/or conceptual rationale for the design of each intervention. Clearly describe all simulation-specific exposures, potential confounders, and effect modifiers.	62	"This study uses three simulation models based on causal pathways driving the income-health relationship: social causation (income causes health); health selection (health causes income); and joint determination (income and health caused by a third variable, e.g. noncognitive skills), as described in Chapter 2.2." – Methods
Outcomes	6	In describing the details of methods of assessment, include (when applicable) the setting, instrument, simulator type, timing in relation to the intervention, along with any methods used to enhance the quality of measurements.	62	"Variables in the models are selected for illustrative purposes. I use ratio scale variables, such that both rank- and level-dependent BHIMs can be used. Beyond this, the precise distributional characteristics and resulting inequality magnitudes are not the focus of this study, which is instead on patterns of the changes in inequality metrics." – Methods

		Provide evidence to support the validity and reliability of assessment tools in this context (if available).		
Blinding (masking)	11	Describe strategies to decrease risk of bias, when blinding is not possible.	N/A	See CONSORT item 20 SBR extension.
Statistical methods	12	Clearly indicate the unit of analysis (e.g., individual, team, system), identify repeated measures on subjects, and describe how these issues were addressed.	63	"individuals" – Methods
Baseline data	15	In describing characteristics of study participants, include their previous experience with simulation and other relevant features as related to the intervention(s).	N/A	N/A
Numbers analyzed	16		63	"10 000 individuals"
Outcomes and estimation	17	For assessments involving >1 rater, interrater reliability should be reported.	N/A	N/A
Limitations	20	Specifically discuss the limitations of SBR.	75	" First, this is a simulation study, which simplifies the complex empirical income-health relationship. For example, all three causal structures may influence the income-health relationship simultaneously. Second, while data were used to inform simulation model inputs, these may differ substantially between empirical contexts." – Discussion

External Validity	21	Describe generalizability of simulation-based outcomes to patient-based outcomes (if applicable).	75	“ The distributional properties of and relationships between income and health may therefore not be representative of all populations, such as those in lower- and middle-income countries.” – Discussion
Funding	25	List simulator brand and if conflict of interest for intellectual property exists.	67	Models are simulated in R v4.1.2 and ranks are fixed using R’s set.seed() function.

B Appendix to Chapter 4

B.1 Descriptive Statistics

Histograms

Figure B.1: Distribution of household equivalised income, age 16 and age 50.

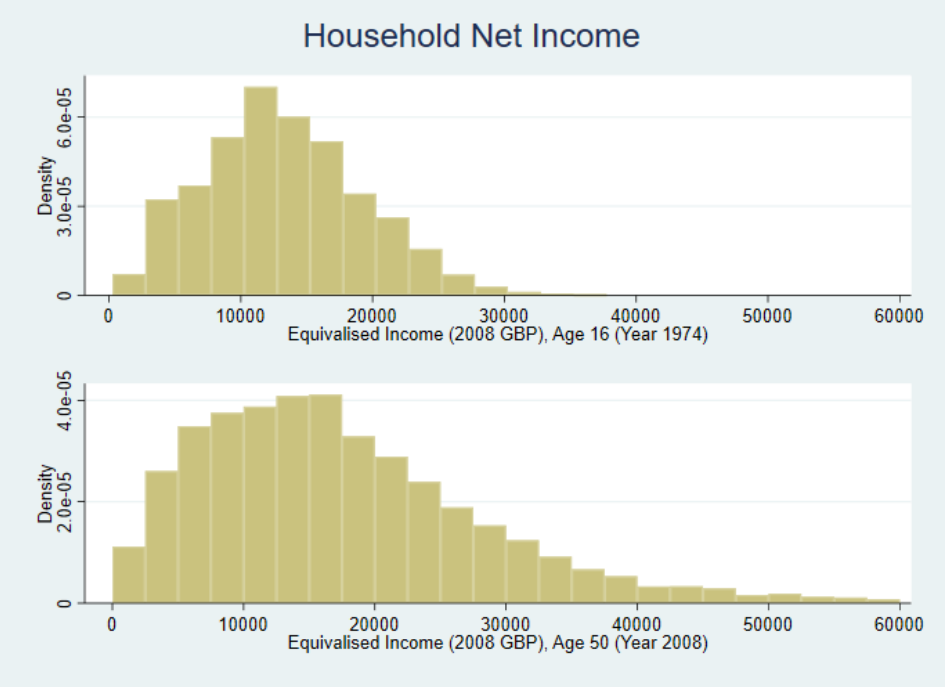


Figure B.2: Distribution of household social class by occupation (SOC 90), age 16 and age 50.

Figure B.3: Distribution of parental education by school leaving age (respondent age 16).

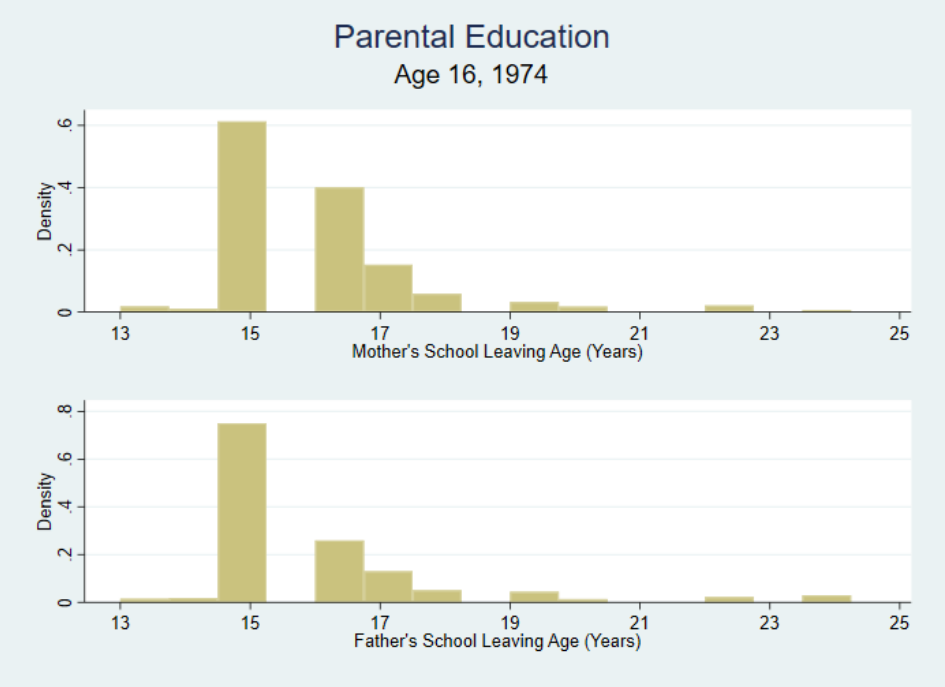


Figure B.4: Distribution of respondents' educational attainment (NVQ Level), age 50.

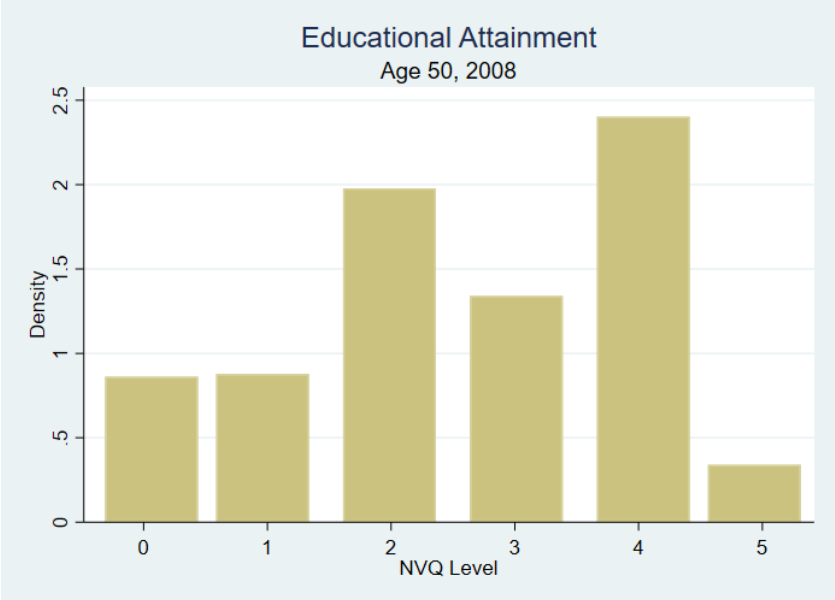


Figure B.5: Distribution of financial hardship, ages 7, 11, and 16.

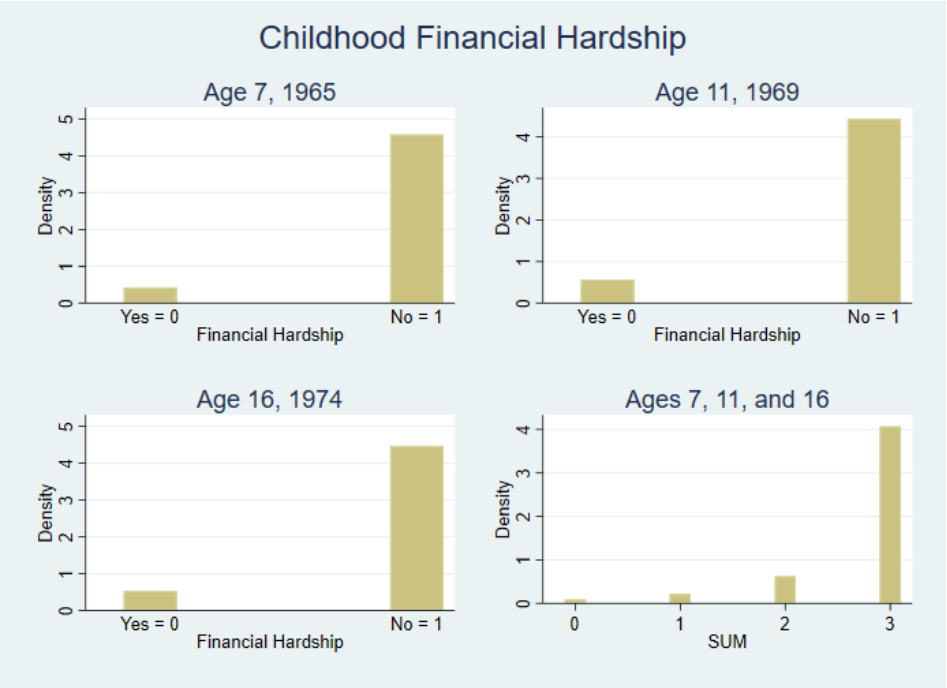


Figure B.6: Distribution of mother's and father's education, respondent age: 0 and 7.

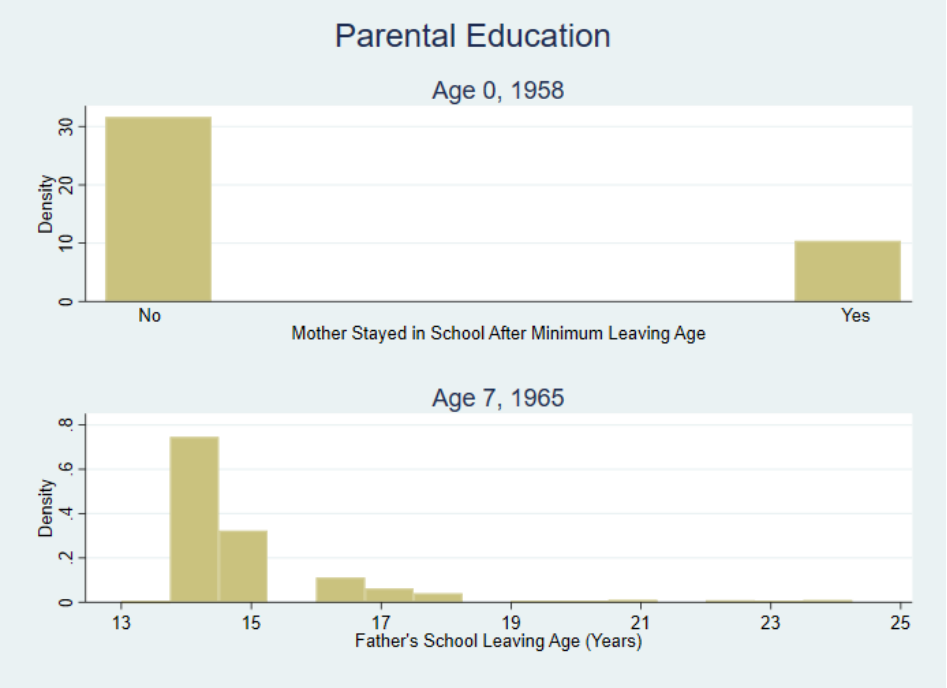
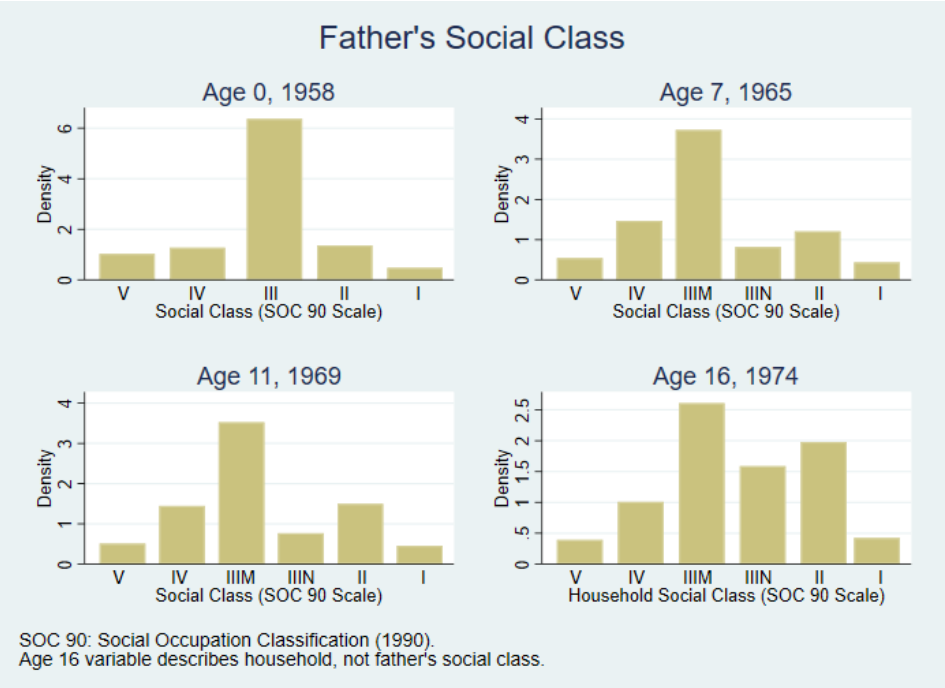


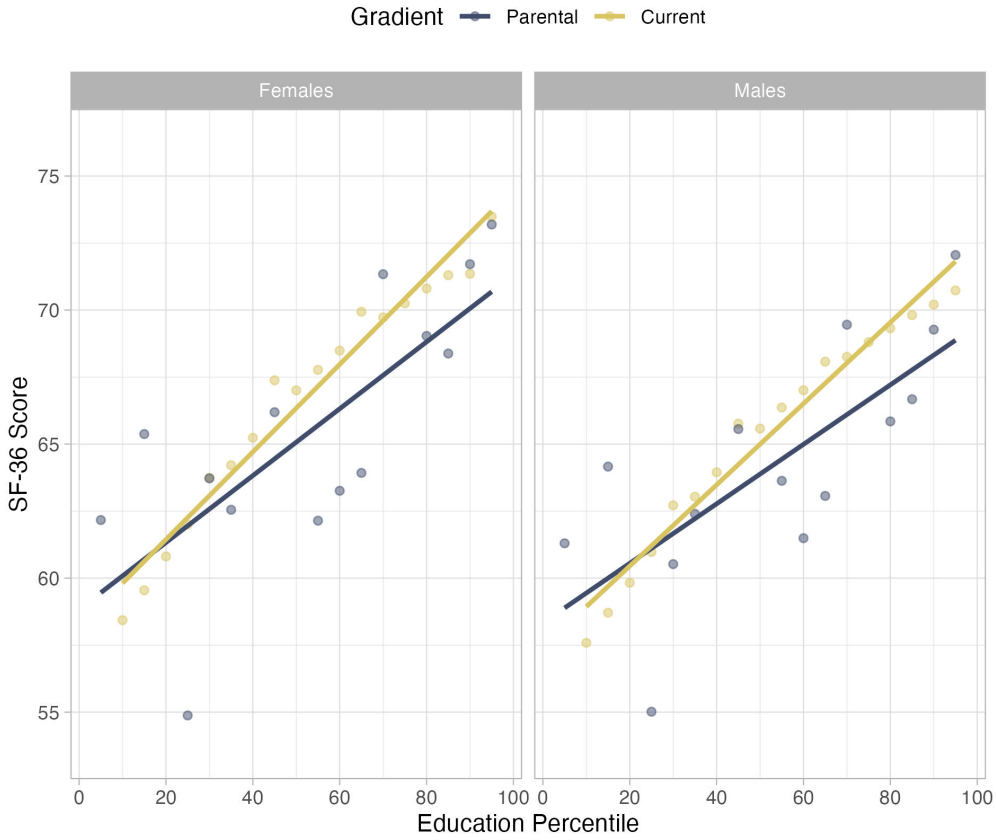
Figure B.7: Distribution of father's social class, ages 0, 7, 11, and 16.



B.2 Supplementary Results

Scatter plots

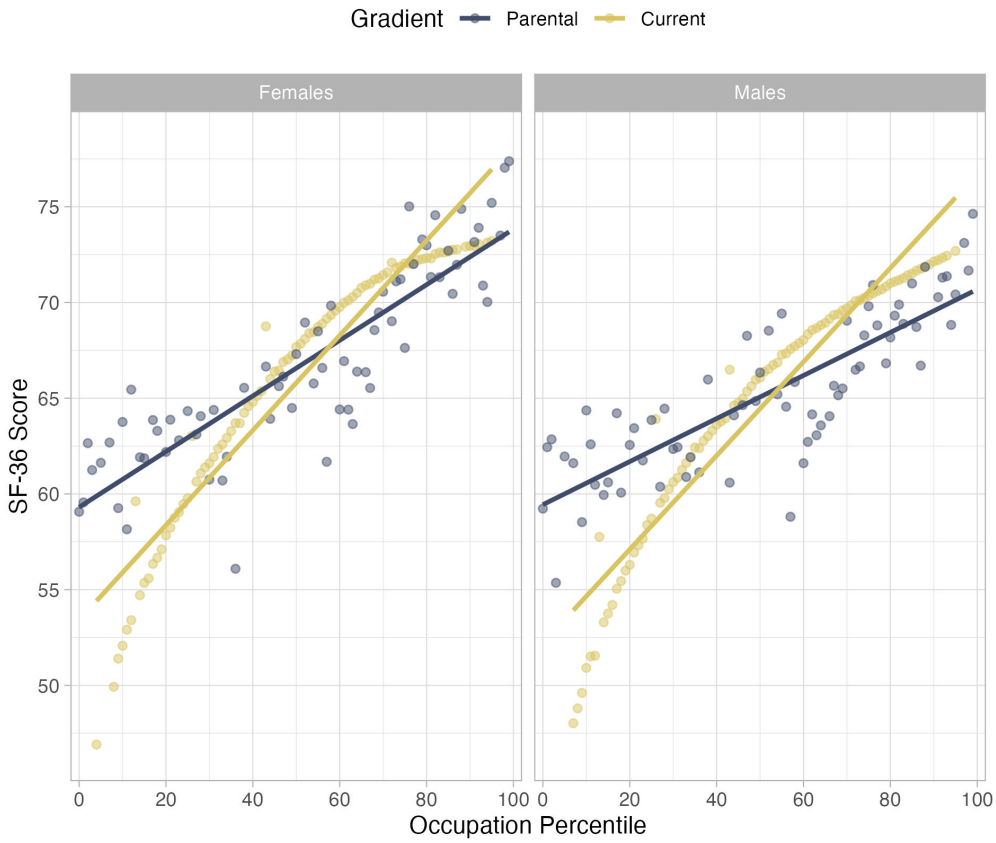
Figure B.8: Scatter plot of parental and current education against SF-36 score at age 50 for females and males.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

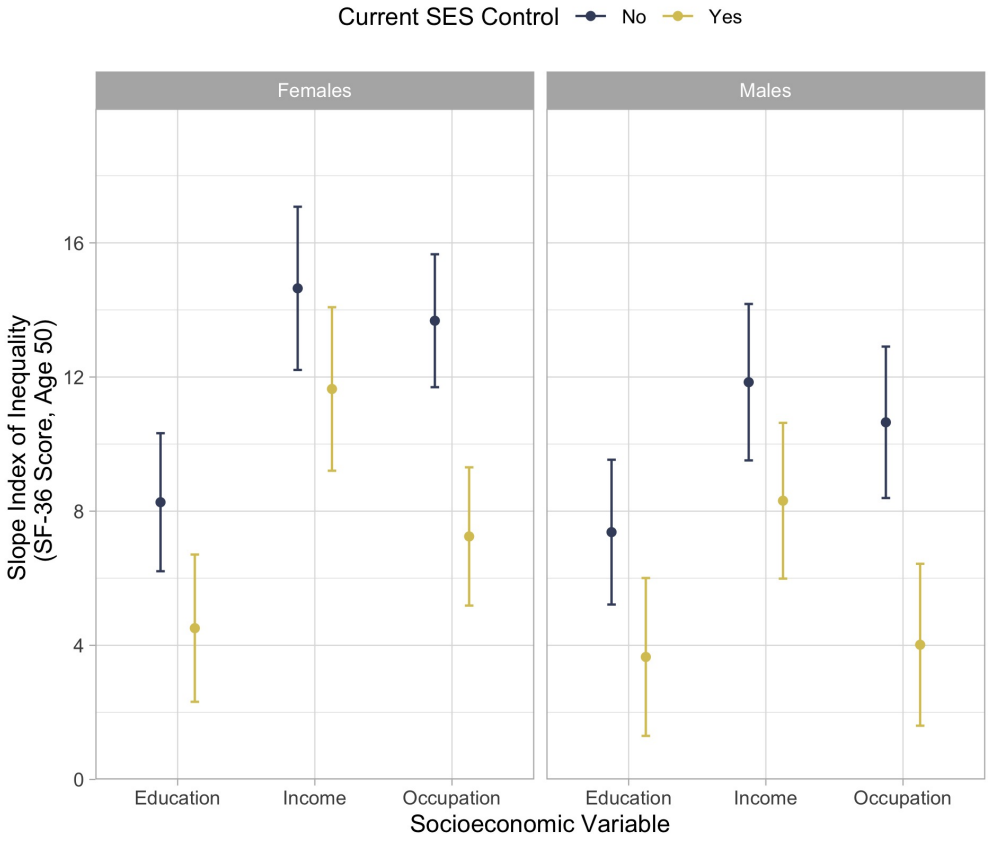
Controlling for Current Socioeconomic Status in Parental Gradients

Figure B.9: Scatter plot of parental and current occupation against SF-36 score at age 50 for females and males.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

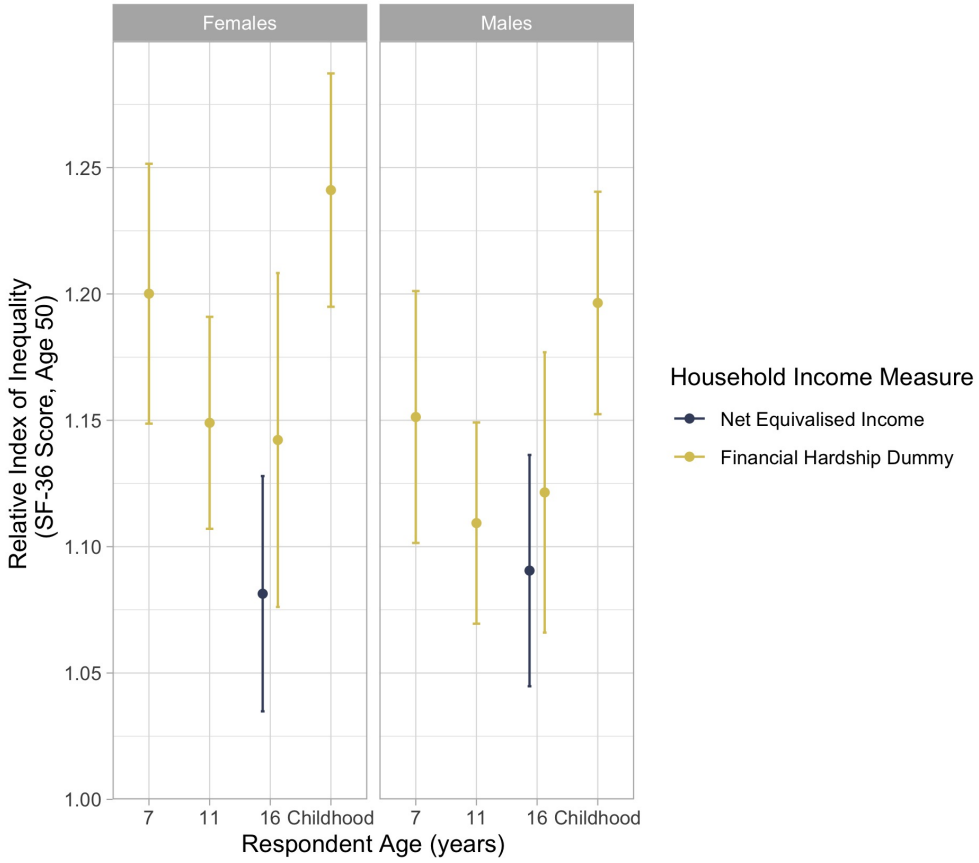
Figure B.10: Current and parental relative indices of inequality in SF-36 score by education, income, and occupation, including current SES control.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey; SES, socioeconomic status.

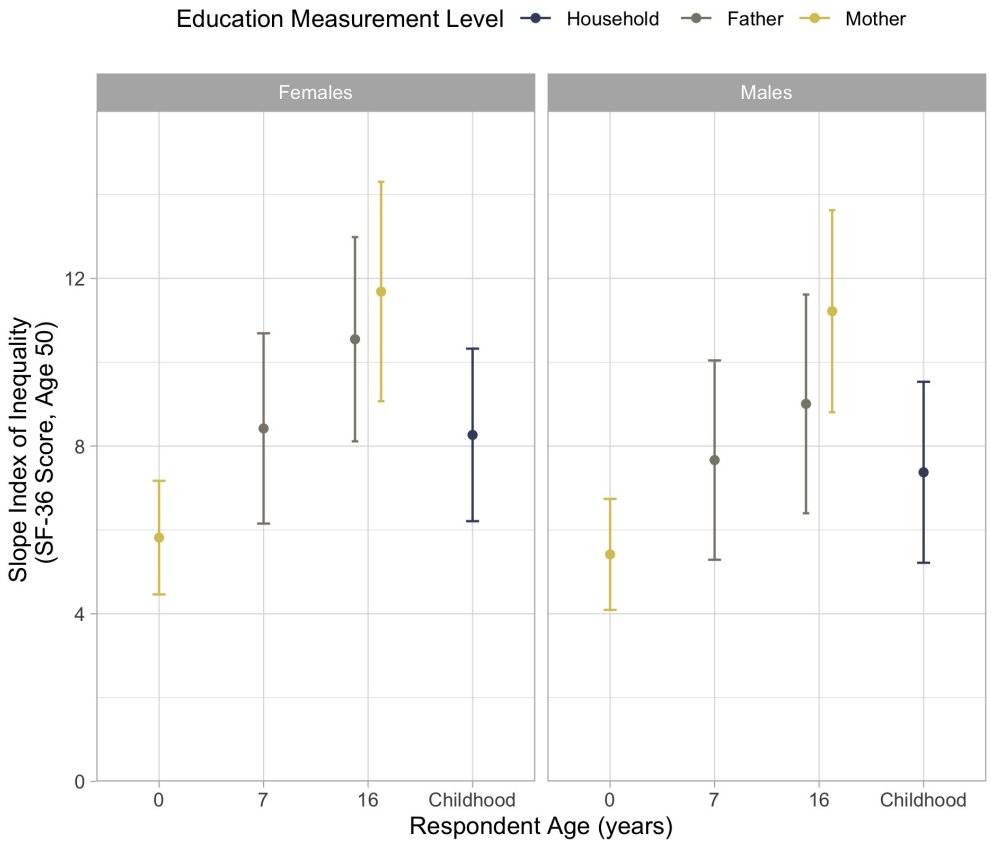
Differential Timing in Parental Socioeconomic Variable Measurement During Childhood

Figure B.11: Impact of measurement timing on the relative index of inequality in SF-36 score at age 50 by parental income in childhood.



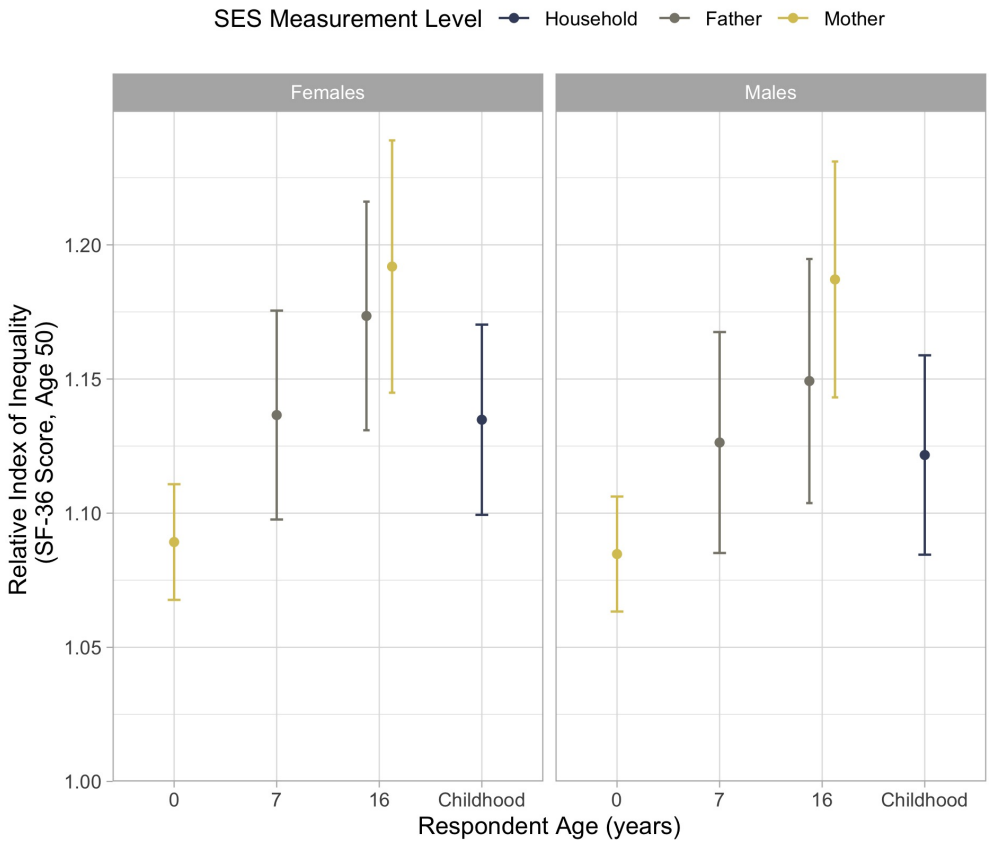
Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

Figure B.12: Impact of measurement timing on the slope index of inequality in SF-36 score at age 50 by parental education in childhood.



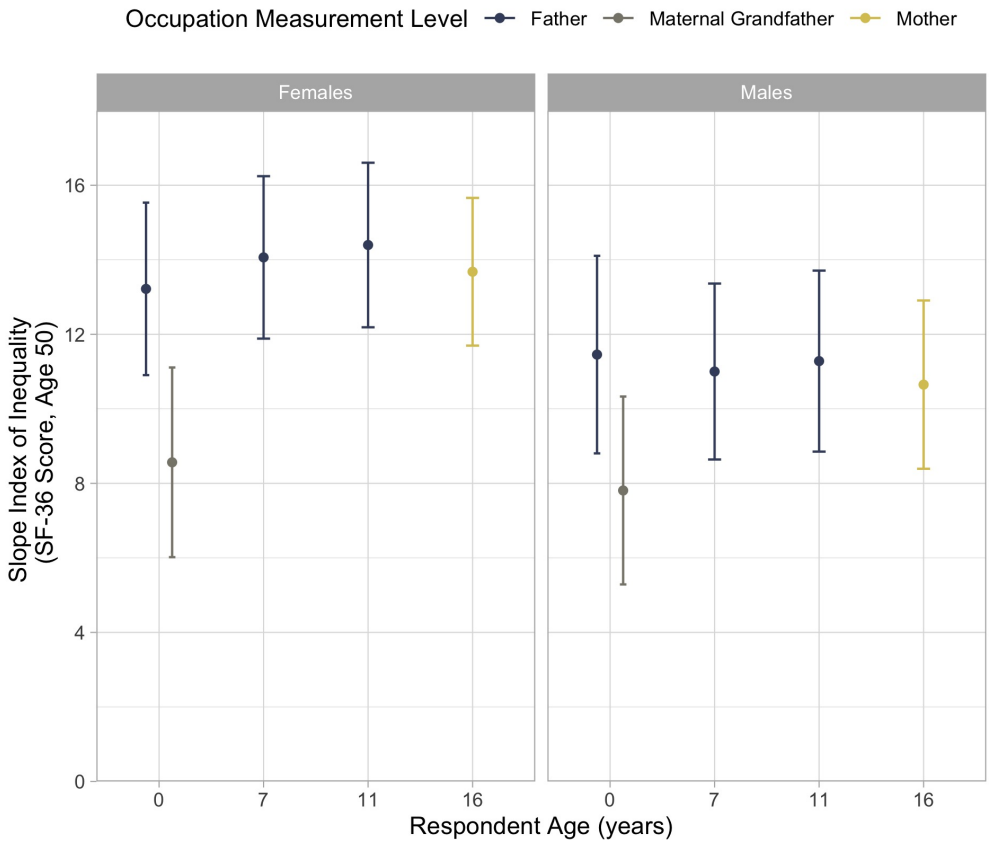
Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

Figure B.13: Impact of measurement timing on the relative index of inequality in SF-36 score at age 50 by parental education in childhood.



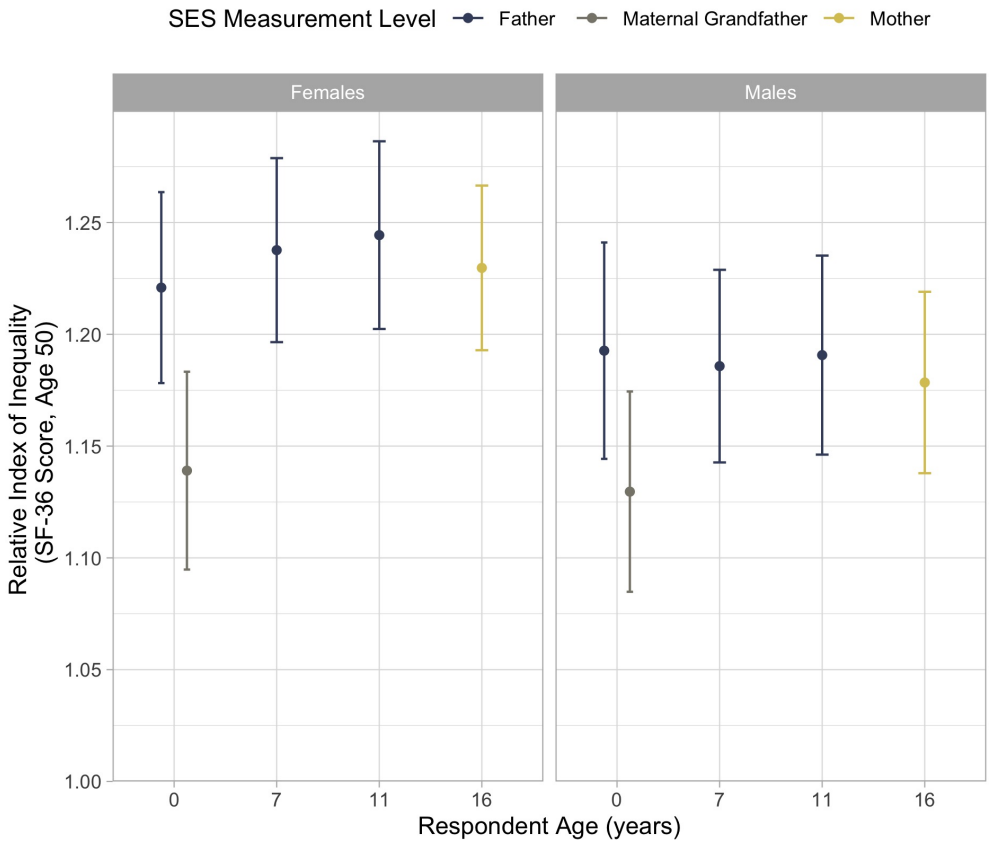
Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

Figure B.14: Impact of measurement timing on the slope index of inequality in SF-36 score at age 50 by parental occupation in childhood.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

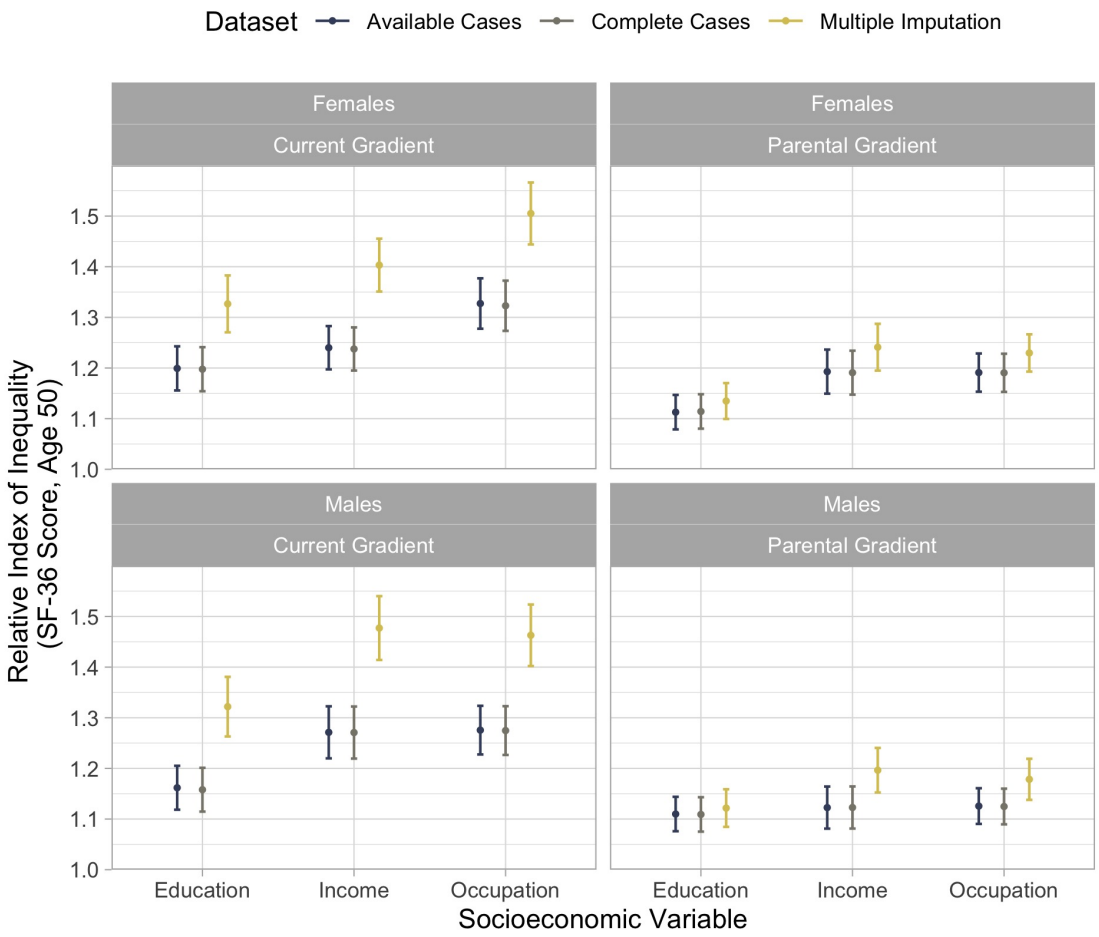
Figure B.15: Impact of measurement timing on the relative index of inequality in SF-36 score at age 50 by parental income in childhood.



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

Available and Complete Case Analyses

Figure B.16: Parental gradient estimates under multiple imputation, available case, and complete case analyses using the relative index of inequality



Data: 1958 National Child Development Study. Abbreviations: SF-36, 36-Item Short Form Survey.

B.3 Checking the Multiple Imputation Model

Building the Multiple Imputation Model

Selecting Auxiliary Variables
The MI strategy relies on the selection of auxiliary variables to inform the imputation model; the necessary auxiliary variables fall into three categories. First, all dependent and independent variables from all substantive models should be included in the imputation model, to limit bias in the analysis model (White, Royston, and Wood 2011). Second, variables predicting both non-response and the missing values of the outcome variable(s) should be included. The missing data guide for the 1958 NCDS describes a set of variables predicting non-response at the eighth sweep, age 50 Silverwood et al. 2020, pp. 62-63. From this set of variables, only those also significantly (e.g. $p<0.001$) associated with the outcome variable are included in the MI model. Other variables predicting non-response at age 50 but not the outcome variable are excluded because evidence suggests they do not provide additional information to the model. Third, variables predicting only the missing values of the outcome variable(s) should be included. These variables are selected based on data exploration and subject knowledge specific to the substantive models. The more complete a variable is, the better it will be able to inform the MI process. For this reason, the best auxiliary variables are likely to appear earlier sweeps of the study, where attrition is minimised. Considering these three categories, a list of all variables included in the imputation model can be found in Table B.1.

Checking the Imputation Model
To increase the reliability of estimates derived from multiple imputation datasets, several checks are recommended (Silverwood et al. 2020). The imputation model and checks, listed and explained below, are performed using the relevant Stata MI commands: mi and mi chained.

1. What is the extent of missing data for the relevant variables?
2. Have the variable mean and variance stabilised across iterations?
3. Has the variable composition changed across imputations?
4. Critically assess MI regression estimates:
    (a) Between, within, and total imputation variance.
    (b) Relative increases in variance due to nonresponse (RIV).

(c) Fraction of missing information (FMI).

(d) Relative efficiency (RE).

(e) Parameter-specific degrees of freedom (DF).

5. Comparison of results with non-imputed datasets.

   (a) Variable composition, variance and mean changes.

   (b) Complete and Available Case Analysis (CCA, ACA).

6. Missing not at Random (MNAR) sensitivity analysis.

To justify the use of MI in the first place, it is useful to (1) provide summary statistics on the substantive model variables to examine the extent of missingness (see Table B.1). Having selected variables for the imputation model, an initial limited model should be computed; in this case, a sample of 5 imputations and 10 iterations. This is the theoretical minimum and provides enough insight into the model to detect any issues before running the full model, which tends to be computationally intensive (White, Royston, and Wood 2011).

The initial model is used the check (2) whether the variable mean and variance have stabilised across iterations in each imputed dataset and (3) whether variable compositions are substantially different from non-imputed data. If either the variance or mean hasn't stabilised, it indicates more iterations are required for each imputation. If variable distributions deviate between non-imputed and imputed data it is not necessarily worrisome, however it should be noted explained where possible (Silverwood et al. 2020). For example, Mostafa et al. (2020) have shown the multiple imputation model exhibits lower educational attainment than the baseline dataset, however the imputed data more closely aligns with census data (Mostafa et al. 2020). It may therefore be hypothesised that less educated individuals may have been less likely to respond to survey questionnaires, which MI corrects. Variable distributions are assessed using histograms.

Having checked the initial model, the full model is used to calculate the imputation statistics. The full model consists of 50 iterations and 99 imputations, as dictated by the rules of thumb provided by Silverwood et al. (2020) and erring on the side of higher numbers of imputations and iterations since there is no disadvantage except time (White, Royston, and Wood 2011).

After running the full MI model, the relevant regressions are performed on the imputed dataset, providing the results accounting for missingness in the data.

Following these estimations, numerous (4) post-regression statistics are available to check the reliability of the analysis models (Buuren 2007; White, Royston, and Wood 2011).

a. Between, within, and total imputation variance

   Between-variance (V_b) is the variance of parameter estimates for each imputed dataset, in effect estimating the additional uncertainty from missing data across each imputed dataset. Within-variance (V_w) is the mean of the sampling variances: estimates the sampling variability we would have expected with no missing data, variance of parameters within each imputed dataset.

   Total variance is the sum of between- and within-variance, as is the standard case, however also includes an additional between-variance term accounting for the number of imputations conducted, m:

$$V_t = V_b + V_w + \frac{V_b}{m} \tag{1}$$

   So, the more imputations, m, the lower Vb/m will be and hence the lower the total variance (V_t). This will affect the parameter estimates standard errors (SE), since:

$$SE = \sqrt{V_t} \tag{2}$$

   Total, within, and between variance will inform the following indicators.

b. Relative Increases in Variance (RVI) measures the proportional increase in total sampling variance attributable to missing data.

$$RVI = V_b + \frac{V_b}{m} \tag{3}$$

   If imputation model variables are weakly correlated with the outcome, or if there's a large number of missing data, this will lead to a higher RVI. The closer this value is to zero, the better, since it indicates a lesser effect of missing data on standard errors.

c. Fraction of missing information (FMI) measures the amount of missing data in parameter estimates is attributable to nonresponse. In other words, the fraction of total sampling variance attributable to missing data.

$$FMI = \frac{RVI}{V_t} \tag{4}$$

So, if FMI = 0.5, 50% of the total sampling variance is attributable to missing data. The higher the FMI is, the greater the number of imputations required for reliable results. Consequently the FMI has also been suggested as a rule of thumb for indicating the sufficient number of imputations to ensure a MI model does not exceed the acceptable loss of efficiency. The rule of thumb suggests taking the percentage of the highest FMI as the minimum number of imputations; so, if FMI = 0.5, this would suggest at least 50 imputations are required.

d. Relative efficiency (RE) compares the efficiency of using the chosen number of imputations, rather than an infinite number.

$$RE = \frac{m}{(1 + FMI)} \tag{5}$$

For the purposes of this chapter, an efficiency greater than or equal to 98.5% is regarded as sufficient. A 95% efficiency level may also be attained with far fewer imputations however this analysis leans toward including more imputations to increase the reproducibility of results, as suggested by White et al. (2010) (White, Royston, and Wood 2011). Although statistical power and efficiency may be sufficient at m= 5, White et al. (2010) show results may still differ substantially when running a new MI model with the same specifications (White, Royston, and Wood 2011). This is because the Monte Carlo standard error – the standard deviation across repeated runs of the same imputation procedure – may still be high if m is low. Since the Monte Carlo standard error tends to zero the higher m is, more imputations should be favoured such that results are reproducible. From the formula it may also be inferred that a lower FMI will contribute to higher relative efficiency, since less missing data is attributable to nonresponse.

e. Parameter-specific degrees of freedom (DF) indicate the number of observations having account for the amount of parameters included in the model. Specifically, this chapter will use Barnard and Rubin's (1999) adjusted degrees of freedom for small samples. Indicates how much information I used in a parameter estimate; the more the better for statistical power. DF, or v, is defined as follows:

$$\nu = \frac{\nu_m * \nu_{obs}}{\nu_m + \nu_{obs}} \tag{6}$$

Where ν is the adjusted degrees of freedom, ν_m is Rubin's original degrees

of freedom formula in the presence of missing data (1987), ν_{obs} is degrees of freedom based on the observed (incomplete) data. These are defined as follows:

$$\nu_m = \frac{(m-1)}{\lambda^2} \tag{7}$$

$$\nu_{obs} = \frac{\nu_{com}+1}{\nu_{com}+3}\nu_{com}(1-\lambda) \tag{8}$$

$$\lambda = FMI = \frac{V_b + \frac{V_b}{m}}{V_t} \tag{9}$$

ν_{com} is degrees of freedom based on hypothetical complete data and FMI is as previously defined.

Following the full imputation procedure, the variable compositions, variance, and mean (5a) should also be checked to ensure the imputed data align with other empirical observations. Since the MI procedure is concerned with accurately representing standard errors for missing values, not uncovering the 'true' underlying values themselves, it is not necessarily concerning if variable bounds are exceeded and categorical variables take on continuous in the imputed datasets. Furthermore, as is standard practice (5b), Complete and Available Case Analyses (CCA, ACA) should also conducted and reasons for differences between results should be justified.

It has been shown that the MCAR hypothesis is not plausible for the 1958 NCDS (Mostafa et al., 2020), and MNAR and MAR hypotheses cannot be directly tested (Mostafa et al. 2020). Since MI relies on the MAR hypothesis for unbiased dataset estimation, it is pertinent to test the reliability of this assumption. In line with Silverwood et al. (2020), I apply a pattern-mixture model to conduct MNAR sensitivity analyses (Silverwood et al. 2020). This approach tests the MAR assumption by systematically changing data from the imputed datasets to replicate the influence of an unobserved missing data mechanism. For this study, MNAR sensitivity analyses will be conducted for each outcome variable (SF-36 and SAH) and primary SES variable (income, social class, and education at age 16 and 50), split by sub-samples into males and females.

Missing Data Prevalence for All Variables

Table B.1: Summary statistics, missing data prevalence, and 1958 NCDS data codes for variables included in the multiple imputation model.

Variable	**Mean**	**St.Dev.**	**Min.**	**Max.**	**Obs.**	**% Missing (Eligible)**	**% Missing (Total)**	**NCDS Variable Code(s)**
female	0.483	0.500	0	1	18,554	-6.54%	0.02%	n622
mothereduc_0	0.247	0.432	0	1	17,273	0.82%	6.92%	n537
fathereduc_7	15.069	2.056	13	39	13,994	9.28%	24.59%	n194, n195, n494
mothereduc_16	15.941	1.480	13	24	11,431	21.93%	38.40%	n2397
fathereduc_16	15.965	1.854	13	24	11,091	24.25%	40.24%	n2396
educ_50	2.585	1.417	0	5	9,783	0.06%	47.28%	ND8HNVQ
fatherclass_0	2.903	0.903	1	5	16,471	5.42%	11.25%	n492
fatherclass_7	3.243	1.246	1	6	14,114	8.50%	23.95%	n190
fatherclass_11	3.325	1.282	1	6	13,312	12.94%	28.27%	n1171
class_16	3.628	1.262	1	6	11,100	24.19%	40.19%	n2384, n2393
class_23	3.703	1.155	1	6	10,756	14.20%	42.04%	n6136
class_33	4.123	1.218	1	6	10,911	4.35%	41.21%	n540033, n540333
class_42	4.290	1.128	1	6	10,453	8.20%	43.67%	sc, sc12
class_46	4.057	1.219	1	6	8,264	13.32%	55.47%	n7sc
class_50	4.199	1.309	1	6	9,332	4.67%	49.71%	N8SC, N8PSC
nofinancialhardship_7	0.916	0.278	0	1	13,063	15.31%	29.61%	n315
nofinancialhardship_11	0.886	0.317	0	1	13,363	12.61%	27.99%	n1230
nofinancialhardship_16	0.895	0.307	0	1	11,366	22.37%	38.75%	n2441
income_equiv_16	12705	5794	254	34891	10,401	28.96%	43.95%	n2462-7, n2360
log_income_equiv_23	7.956	0.830	3	10	11,191	10.73%	39.70%	famnet, n5034
log_income_equiv_33	8.863	0.980	3	15	9,694	15.02%	47.76%	n500542, 36, n501038, 60, 66, n502613

log_income_equiv_42	9.223	0.987	2	15	9,618	15.54%	48.17%	cnetprd, -pay, pnetprd, -pay, ojnetpw, hhsize
log_income_equiv_46	9.051	0.745	2	14	6,690	29.83%	63.95%	n7cnetpd, n7cnetpy, nd7numhh
log_income_equiv_50	9.567	0.797	-1	14	7,675	21.60%	58.64%	N8CNETWK, N8PNETWK, ND8NUMHH
genhealth_23	0.904	0.295	0	1	12,524	0.10%	32.51%	n5739
genhealth_33	0.864	0.343	0	1	11,275	1.16%	39.24%	n503913
genhealth_42	0.818	0.386	0	1	11,374	0.11%	38.71%	hlthgen
genhealth_46	0.766	0.423	0	1	9,527	0.07%	48.66%	n7khlstt
genhealth_50	0.815	0.388	0	1	9,733	0.57%	47.55%	N8HLTHGN
genhealth_55	0.477	0.500	0	1	9,039	1.07%	51.29%	N9HLTHGN
mentalhealth_50	49.259	8.115	14	70	8,642	11.72%	53.43%	ND8WEMWB
sf36_50_gen	68.681	21.474	5	100	8,741	10.71%	52.90%	ND8GENH
vote_23	1.673	0.699	1	3	7,747	38.20%	58.26%	n5960
vote_33	1.784	0.697	1	3	8,161	28.46%	56.02%	n504636
econactivity_23	1.394	0.732	1	3	12,203	2.66%	34.24%	econstrg
healthcon_16	0.547	0.498	0	1	11,194	23.55%	39.68%	n123-131, n342
smoker_16	0.358	0.479	0	1	11,969	18.26%	35.50%	n2887
currentsmoker_50	0.227	0.419	0	1	9,735	0.55%	47.54%	ND8SMOKE
bmi_16_dummy	0.155	0.362	0	1	11,060	24.46%	40.40%	n1949, n1953
bmi_50_dummy	0.641	0.480	0	1	7,905	19.25%	57.40%	DVWT50, dvht23
bame_16	0.041	0.198	0	1	11,418	22.02%	38.47%	n2017
response_11	0.824	0.381	0	1	18,558	-21.37%	0.00%	n1845
response_16	0.789	0.408	0	1	18,558	-26.74%	0.00%	n1849
response_23	0.676	0.468	0	1	18,558	-48.04%	0.00%	n4118
response_33	0.615	0.487	0	1	18,558	-62.69%	0.00%	n500124
response_42	0.614	0.487	0	1	18,558	-62.98%	0.00%	unout
response_46	0.514	0.500	0	1	18,558	-94.65%	0.00%	n7dlasnt

mathsimp_42	0.274	0.446	0	1	11,354	0.29%	38.82%	mthimp
conduct_16	0.095	0.293	0	1	11,368	22.36%	38.74%	n2293
pregsmoke_0	0.411	0.492	0	1	17,349	0.38%	6.51%	n502
parentoutings_7	0.870	0.336	0	1	14,555	5.64%	21.57%	n181-2
disability_7	0.039	0.193	0	1	14,552	5.66%	21.59%	n281
disability_11	0.091	0.287	0	1	12,918	15.52%	30.39%	n1507
sepdivwid_23	0.037	0.188	0	1	12,532	0.03%	32.47%	n5113
nperroom_0	1.026	0.498	1	3	16,920	2.84%	8.83%	n512
social_7	0.879	0.506	-3	2	13,489	12.55%	27.31%	n455
social_11	8.496	8.987	0	70	14,156	7.42%	23.72%	n1008
cogability_11	42.940	16.144	0	80	14,131	7.59%	23.85%	n920
schoolattendance_16	2.320	1.347	-5	5	10,591	27.67%	42.93%	n1721
maths_16	12.755	6.997	0	31	11,920	18.59%	35.77%	n2930
socialcapital_33	15.250	5.676	1	24	10,324	9.49%	44.37%	n509531-54
birthwt_0	3295.127	579.746	312	6010	16,781	3.64%	9.58%	n574
infectiousdisease_7	2.506	1.143	0	6	14,404	6.62%	22.38%	n215-21
cogability_7	110.676	30.653	2	200	14,996	2.78%	19.19%	n90, n92, n1840
lifesat_33	7.422	1.724	0	10	10,629	6.82%	42.73%	n509772

Figure B.17: Proportion of missing data for SES analysis variables(age 16).

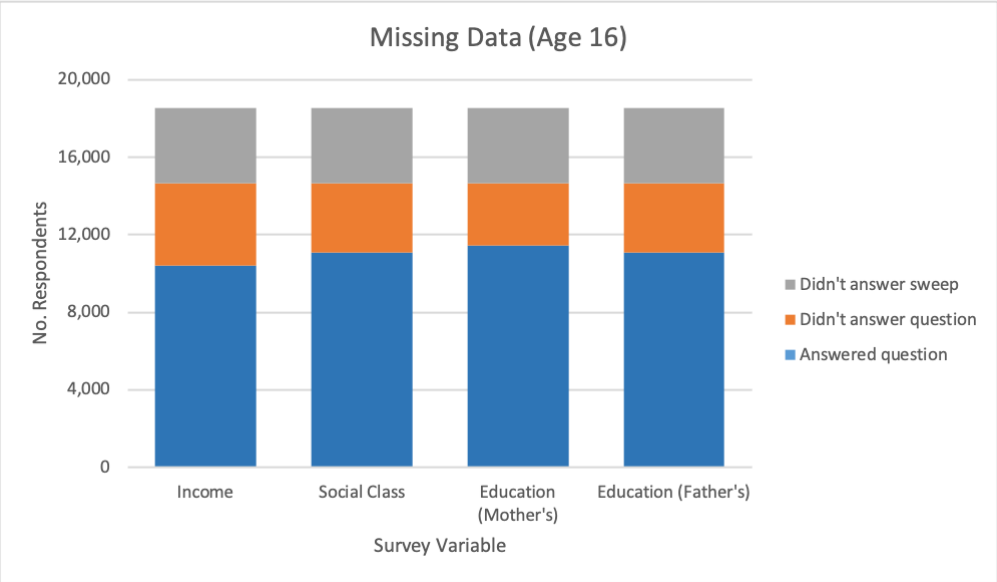
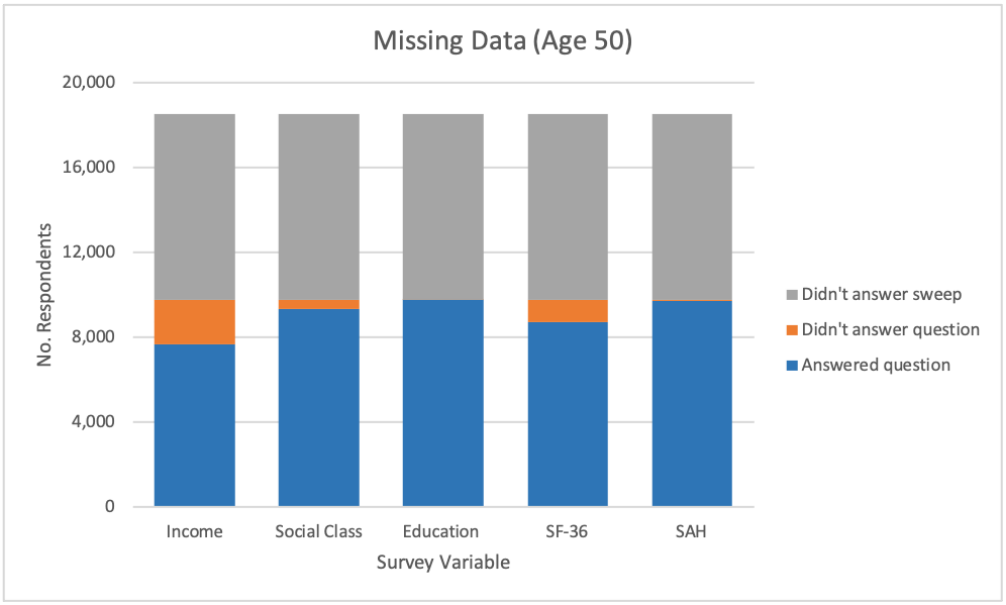


Figure B.18: Proportion of missing data for SES and health analysis variables (age 50).



Variable Stabilisation

Figure B.19: Variable stabilisation across MICE imputations: SF-36

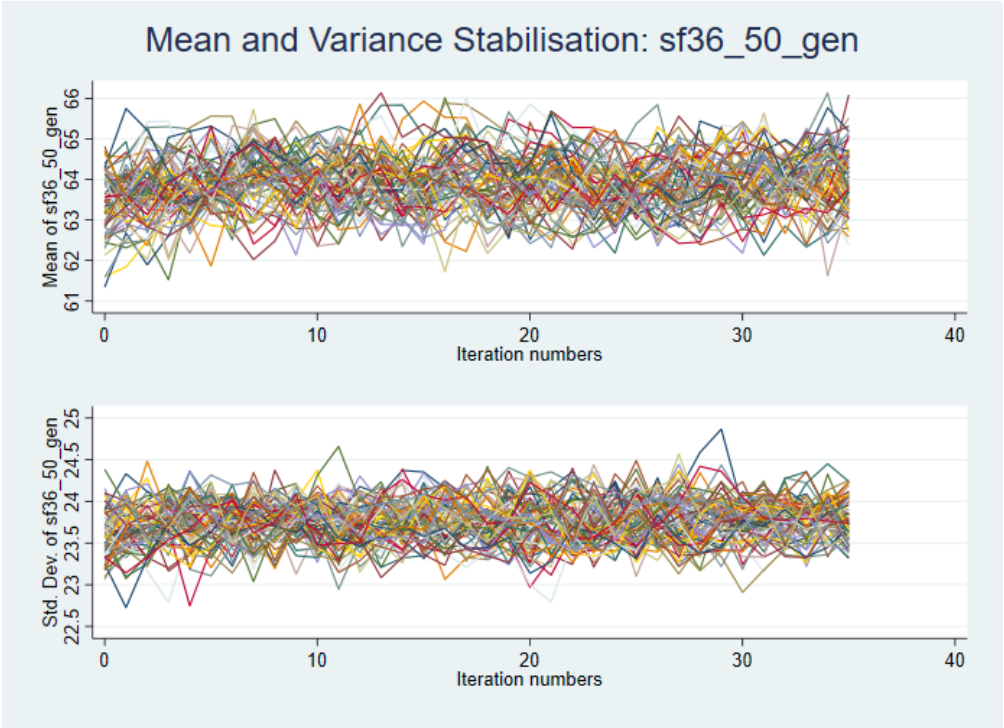


Figure B.20: Variable stabilisation across MICE imputations: self-assessed health

Figure B.21: Variable stabilisation across MICE imputations: income (age 50)

Figure B.22: Variable stabilisation across MICE imputations: income (age 16)

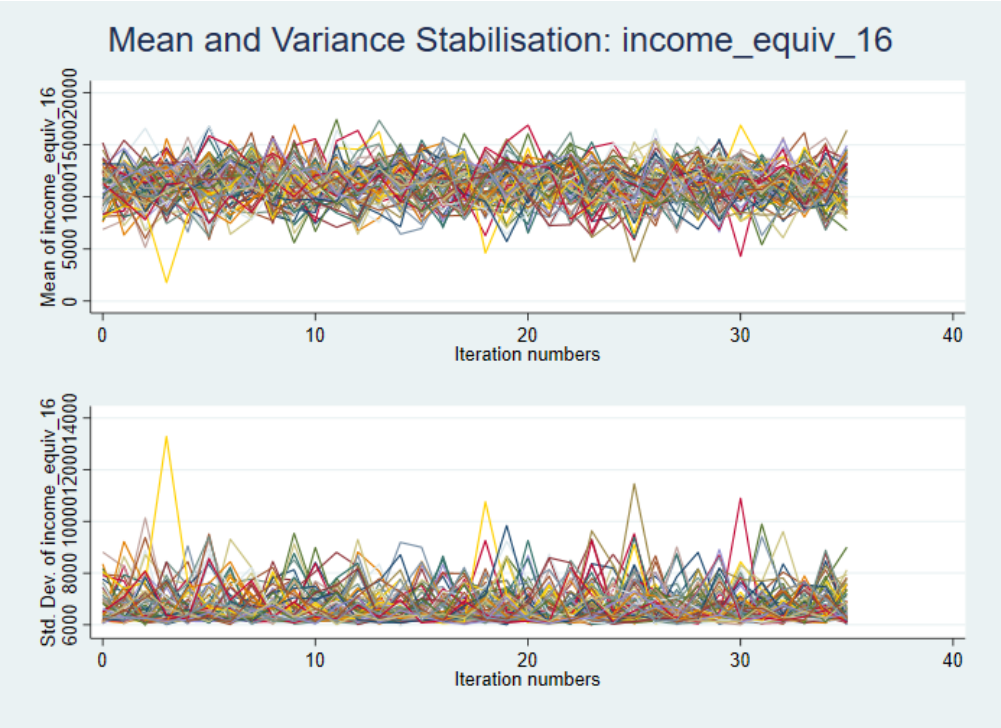


Figure B.23: Variable stabilisation across MICE imputations: financial hardship (age 16)

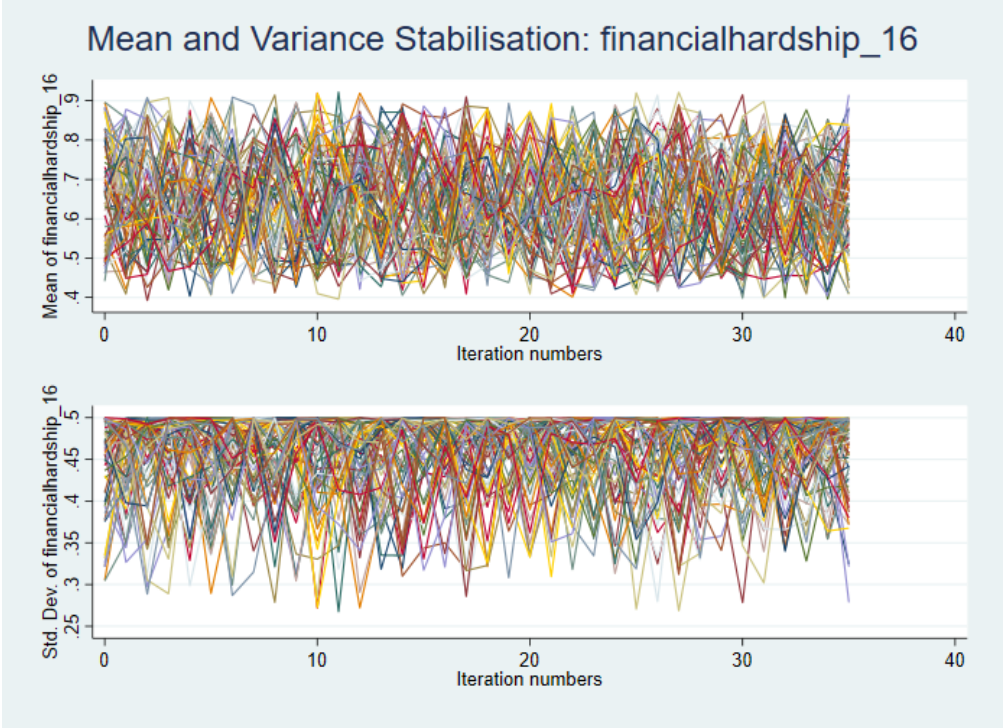


Figure B.24: Variable stabilisation across MICE imputations: financial hardship (age 11)

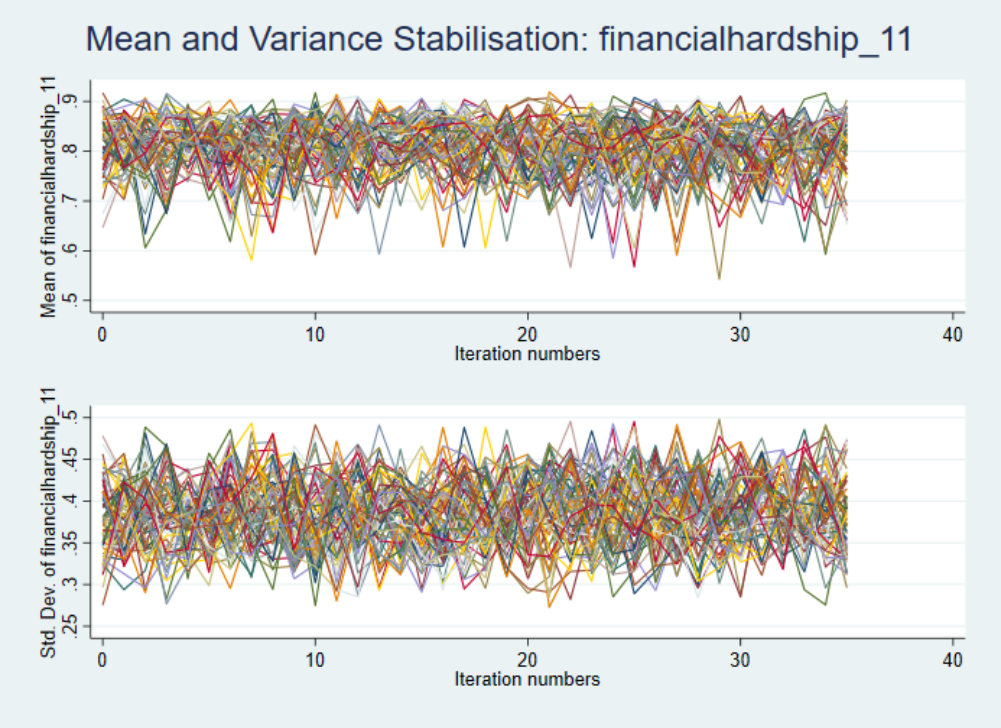


Figure B.25: Variable stabilisation across MICE imputations: financial hardship (age 7)

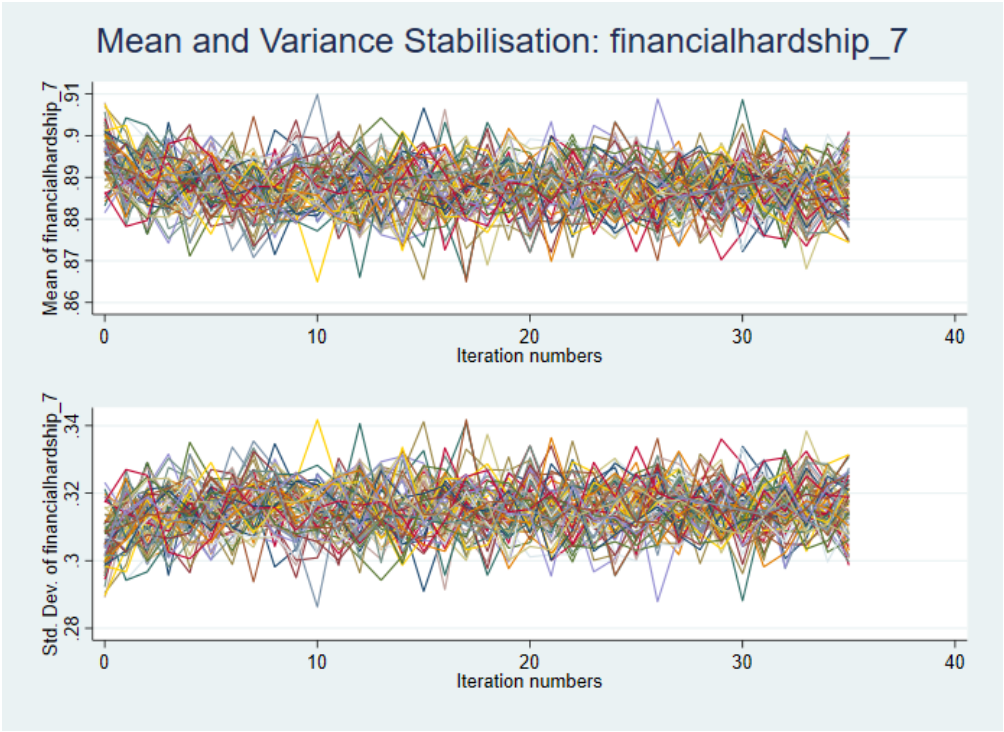


Figure B.26: Variable stabilisation across MICE imputations: social class (age 50)

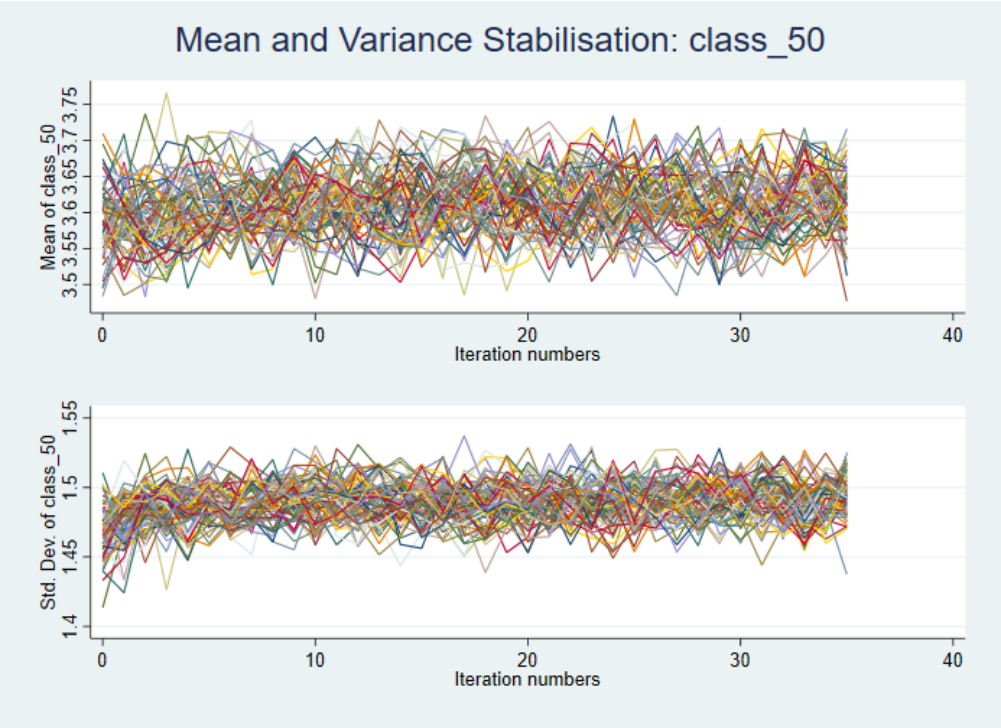


Figure B.27: Variable stabilisation across MICE imputations: social class (age 16)

Figure B.28: Variable stabilisation across MICE imputations: father's social class (age 11)

Figure B.29: Variable stabilisation across MICE imputations: father's social class (age 7)

Figure B.30: Variable stabilisation across MICE imputations: father's social class (age 0)

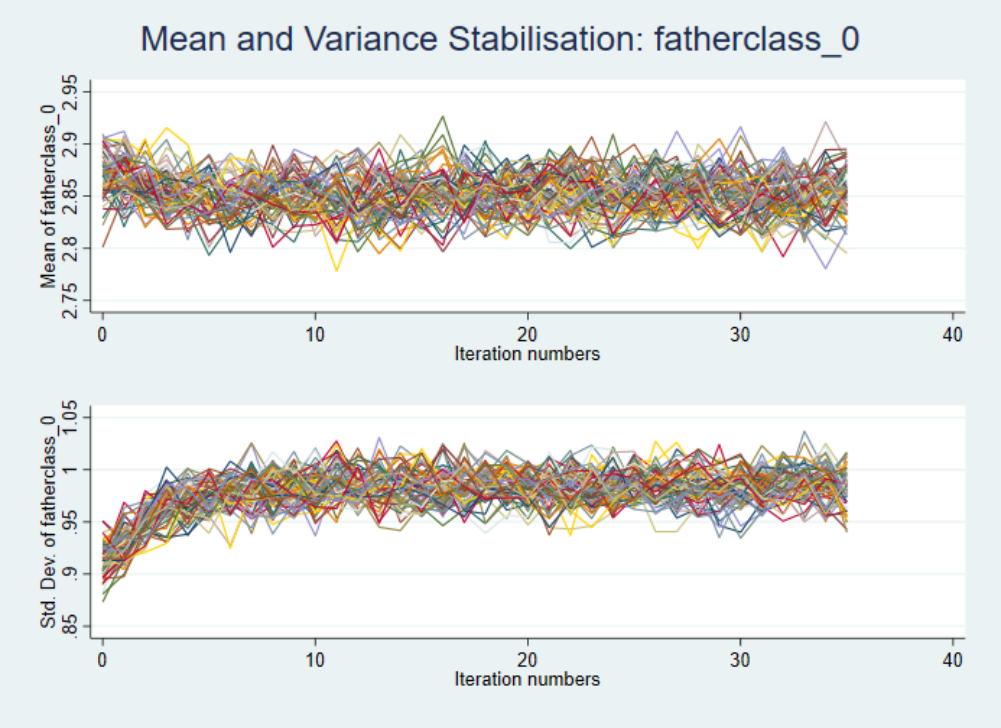


Figure B.31: Variable stabilisation across MICE imputations: education (age 50)

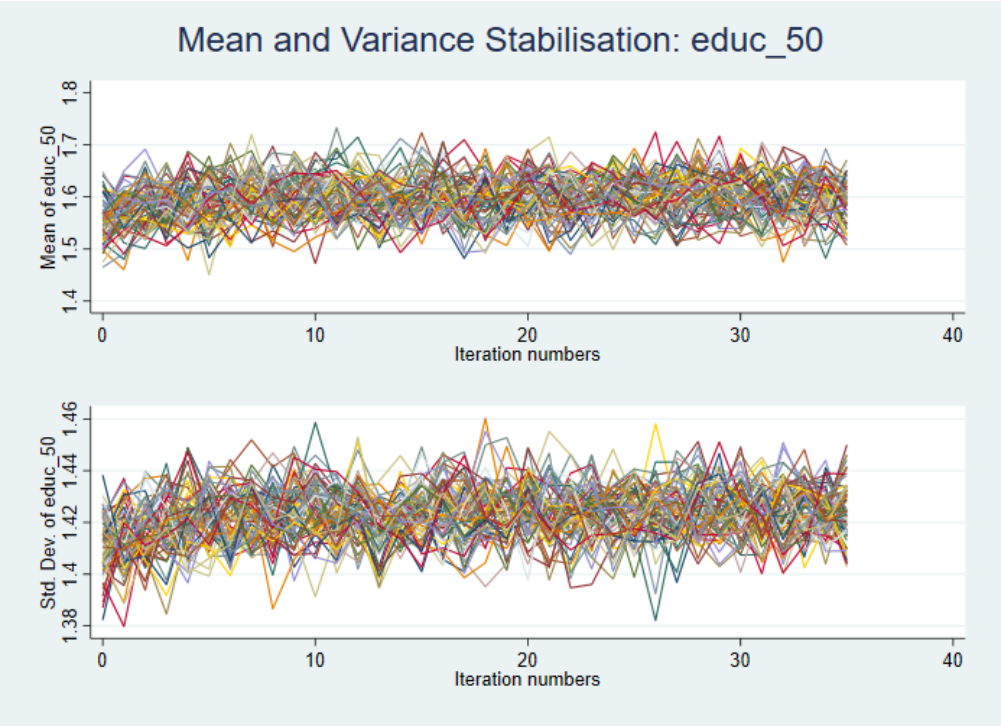


Figure B.32: Variable stabilisation across MICE imputations: mother's education (age 16)

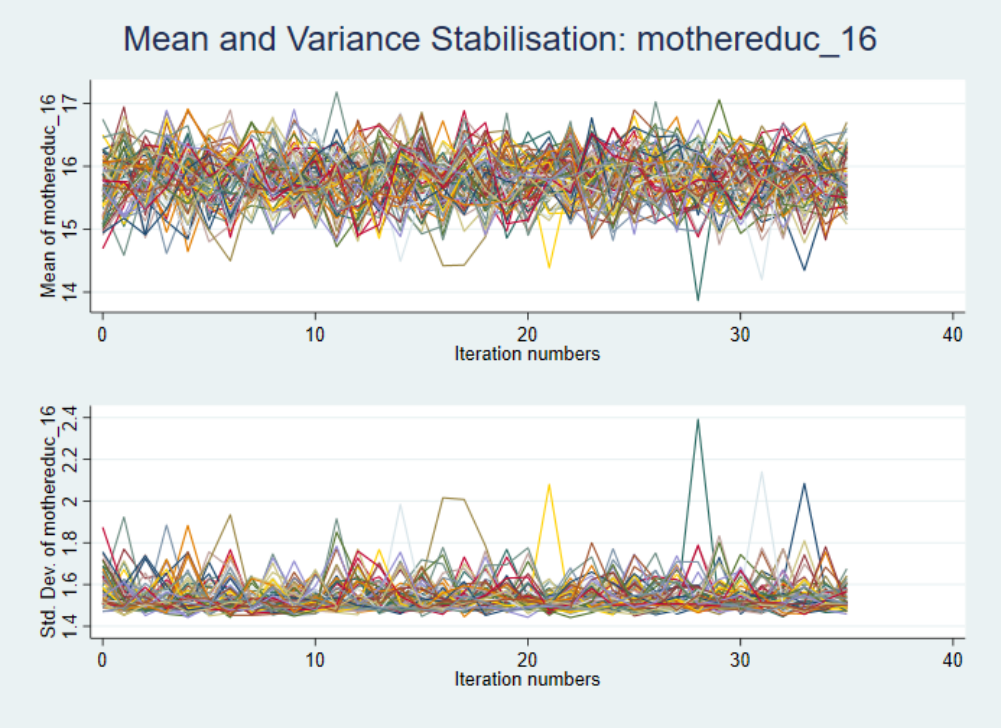


Figure B.33: Variable stabilisation across MICE imputations: father's education (age 16)

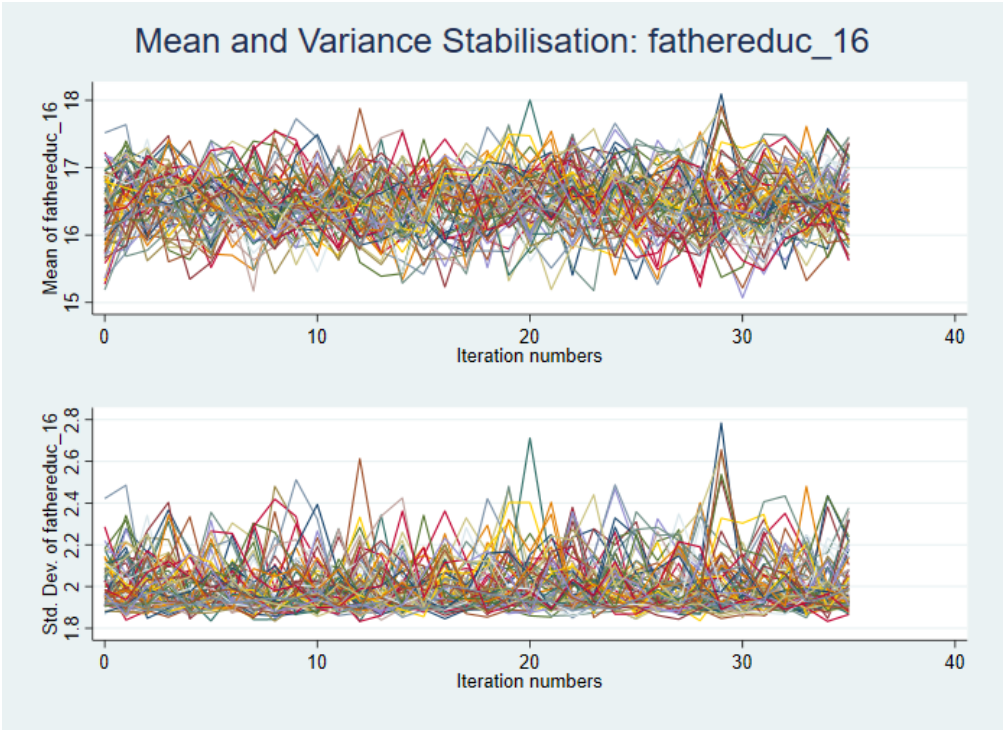


Figure B.34: Variable stabilisation across MICE imputations: father's education (age 7)

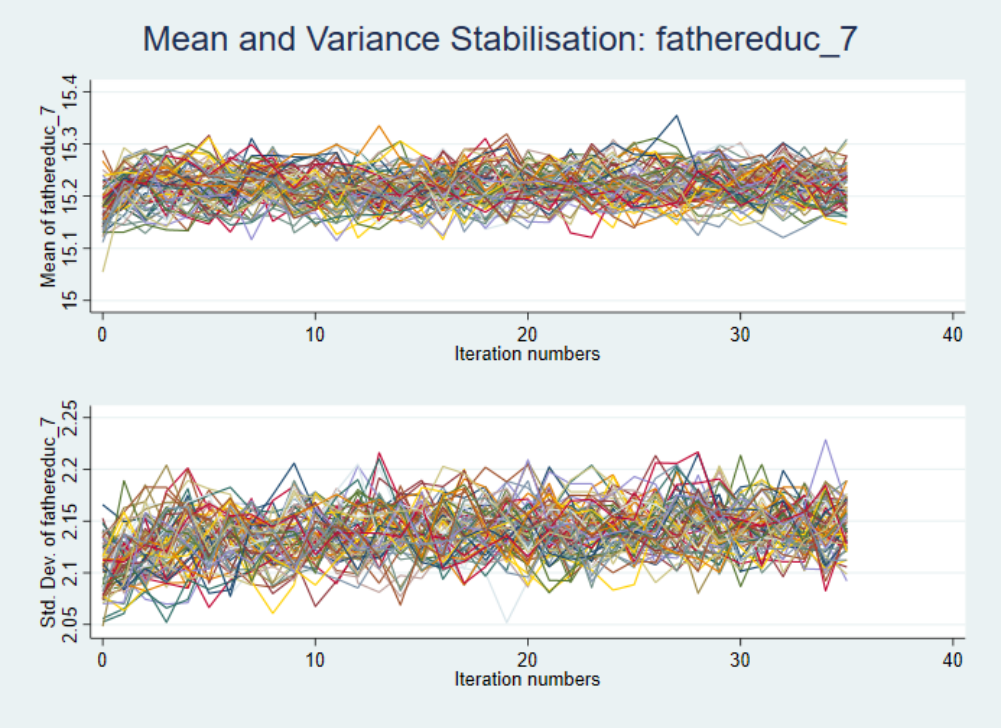
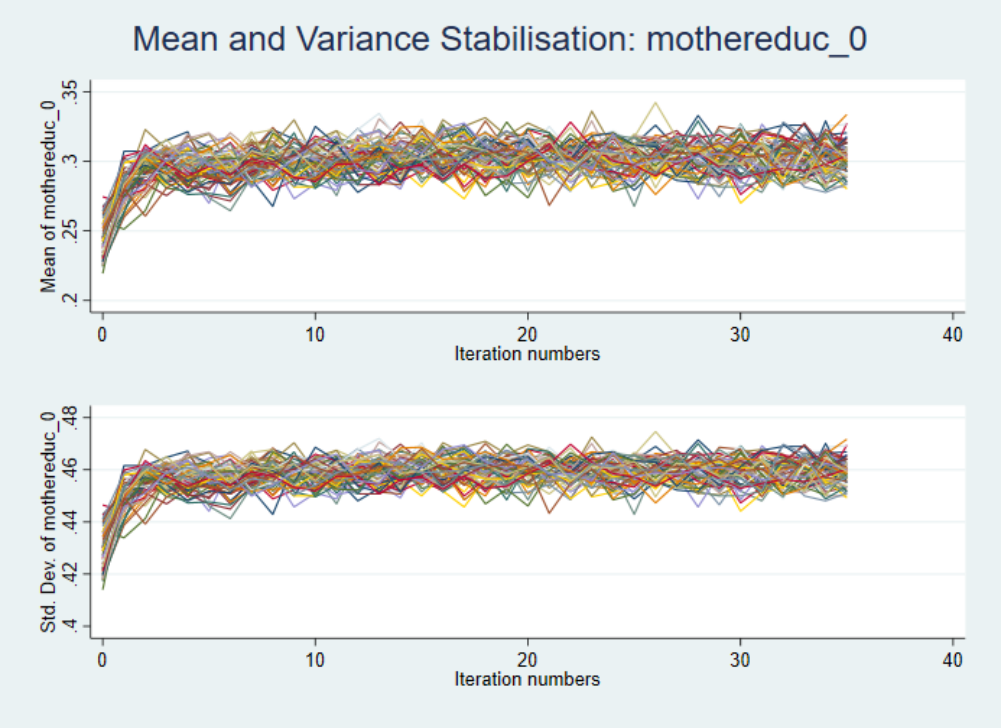


Figure B.35: Variable stabilisation across MICE imputations: mother's education (age 0)



Variable Composition

Figure B.36: Variable composition across MICE imputations: SF-36

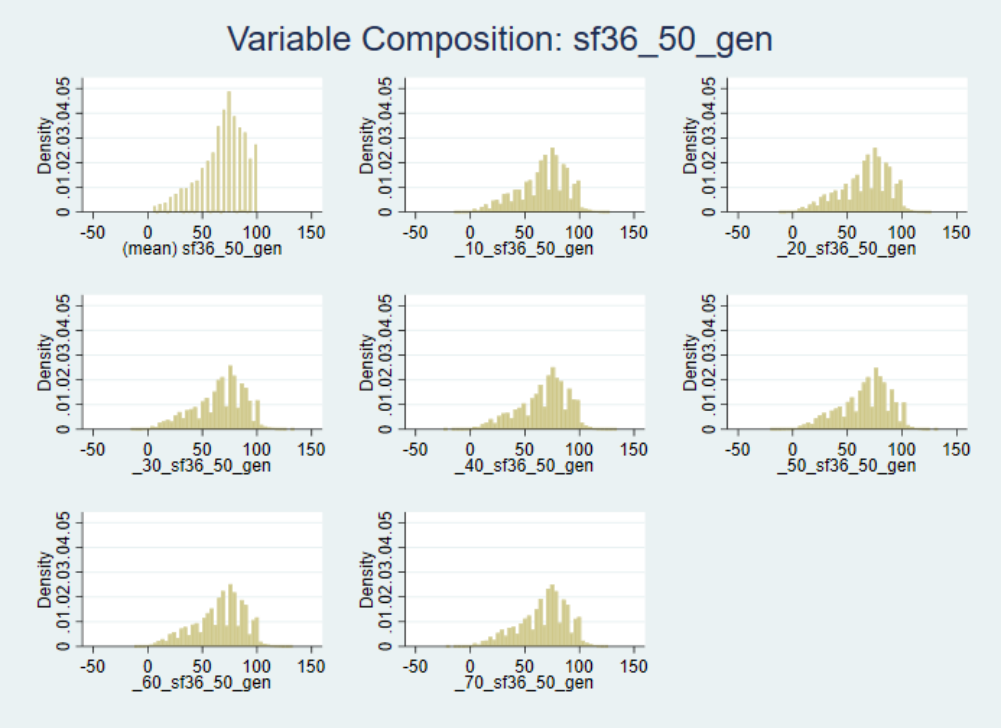


Figure B.37: Variable composition across MICE imputations: self-assessed health

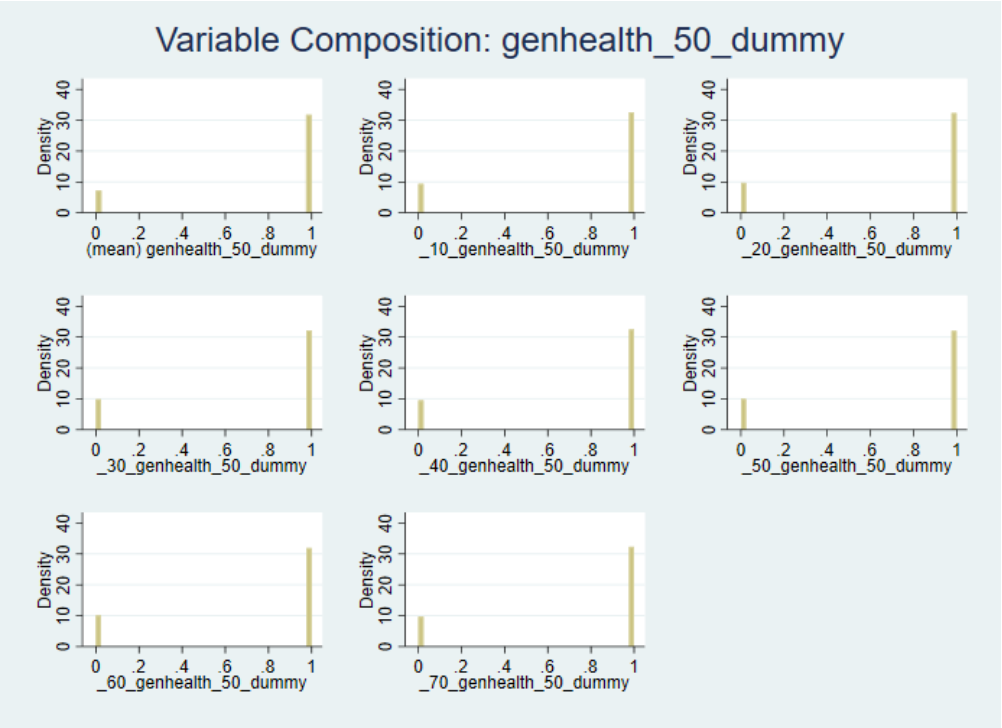


Figure B.38: Variable composition across MICE imputations: income (age 50)

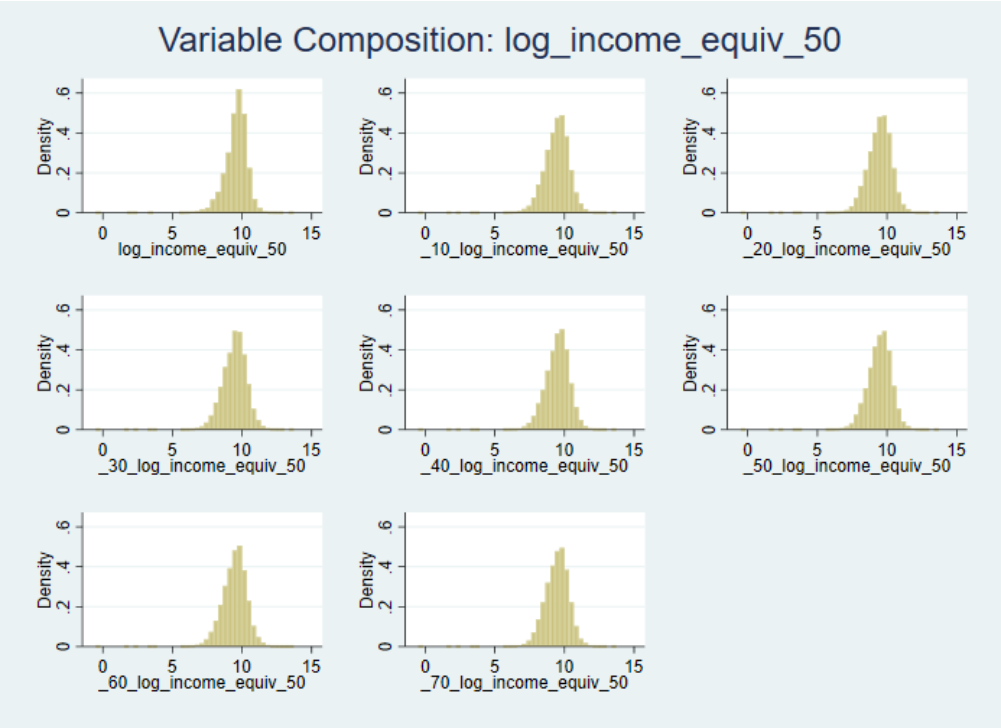


Figure B.39: Variable composition across MICE imputations: income (age 16)

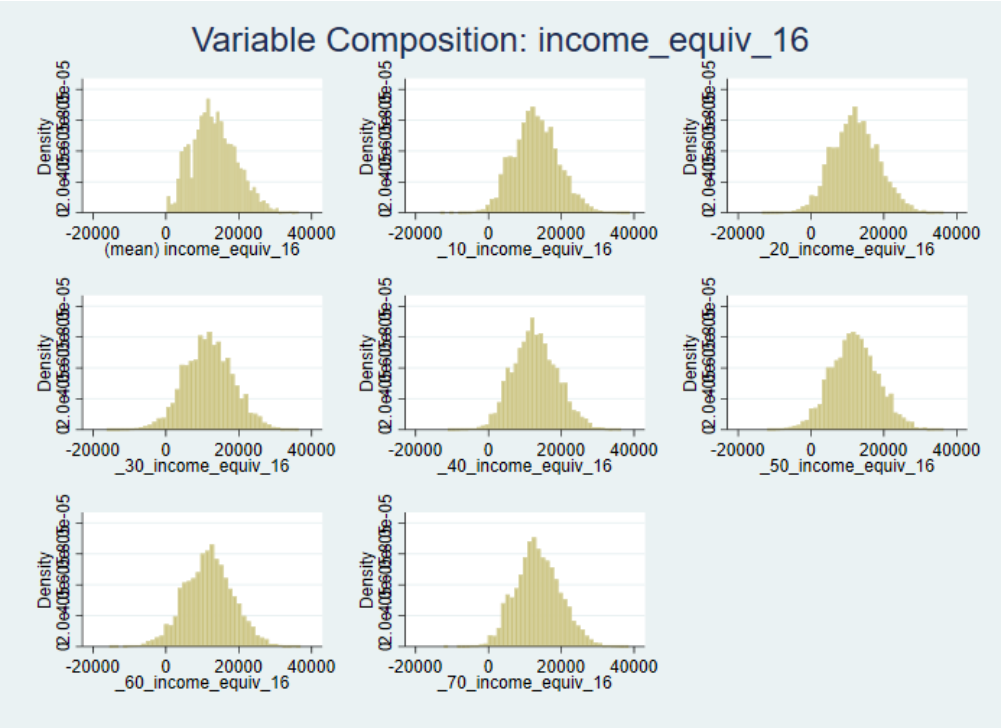


Figure B.40: Variable composition across MICE imputations: financial hardship (age 16)

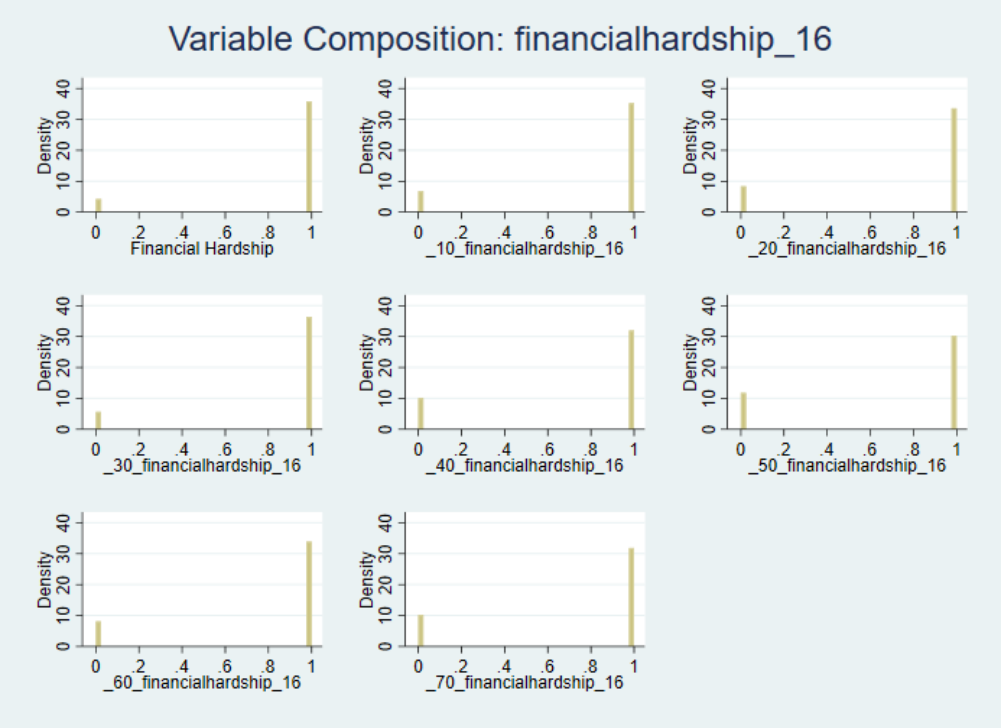


Figure B.41: Variable composition across MICE imputations: financial hardship (age 11)

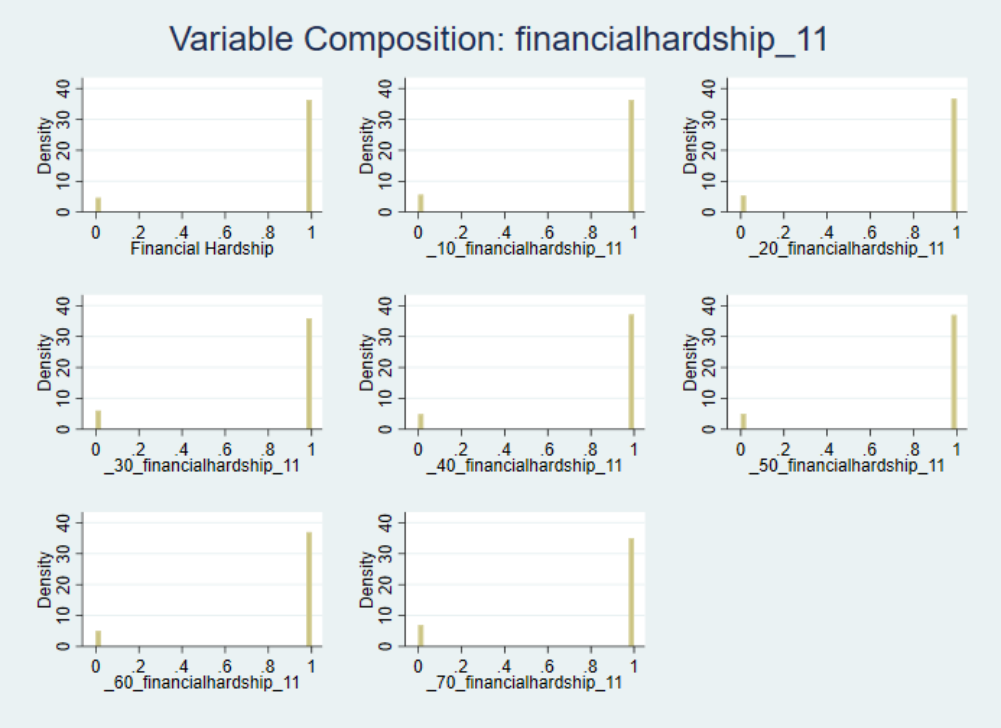


Figure B.42: Variable composition across MICE imputations: financial hardship (age 7)

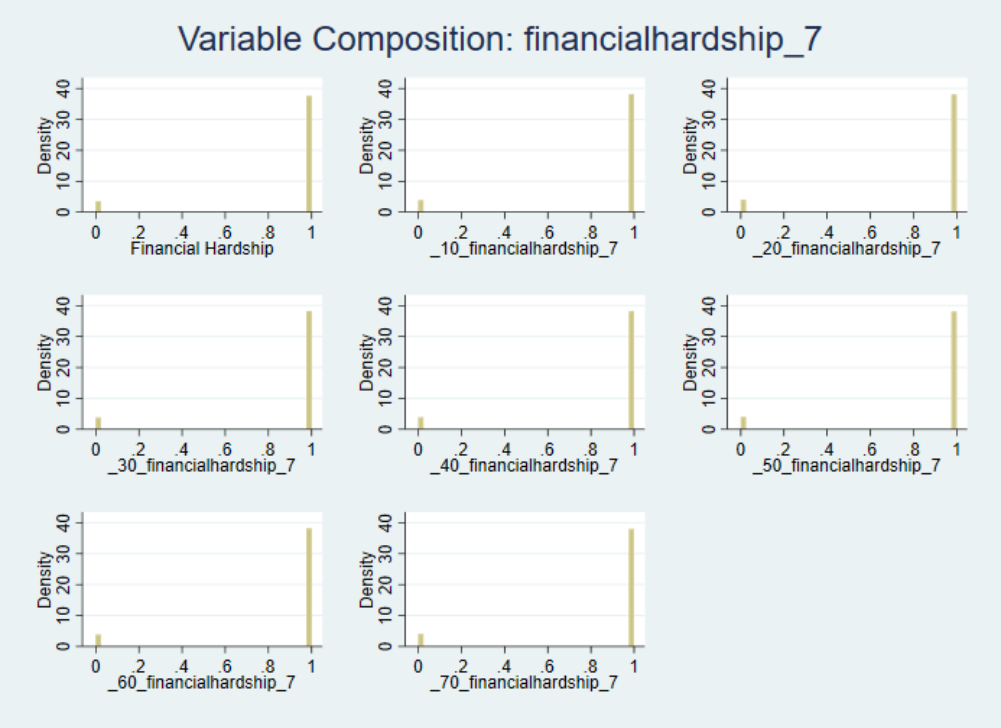


Figure B.43: Variable composition across MICE imputations: social class (age 50)

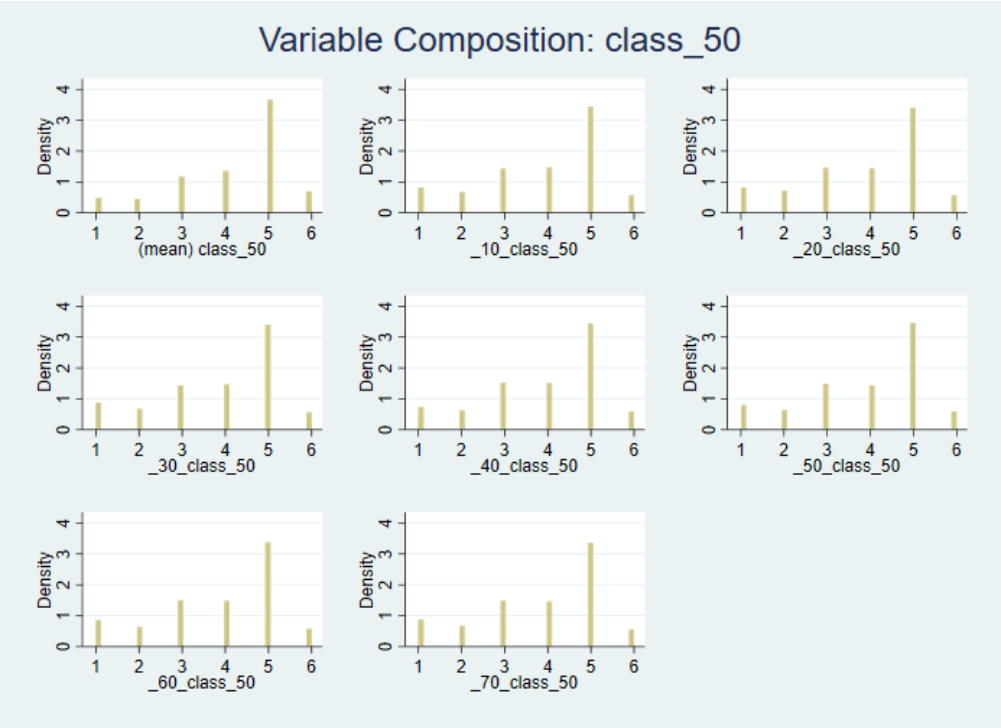


Figure B.44: Variable composition across MICE imputations: social class (age 16)

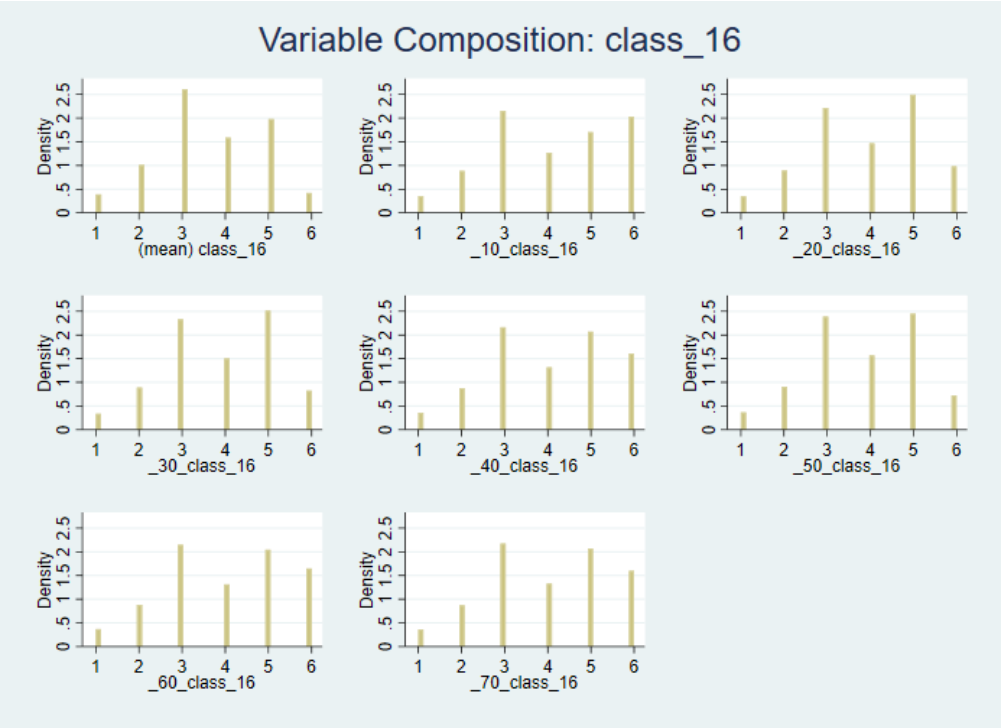


Figure B.45: Variable composition across MICE imputations: father's social class (age 11)

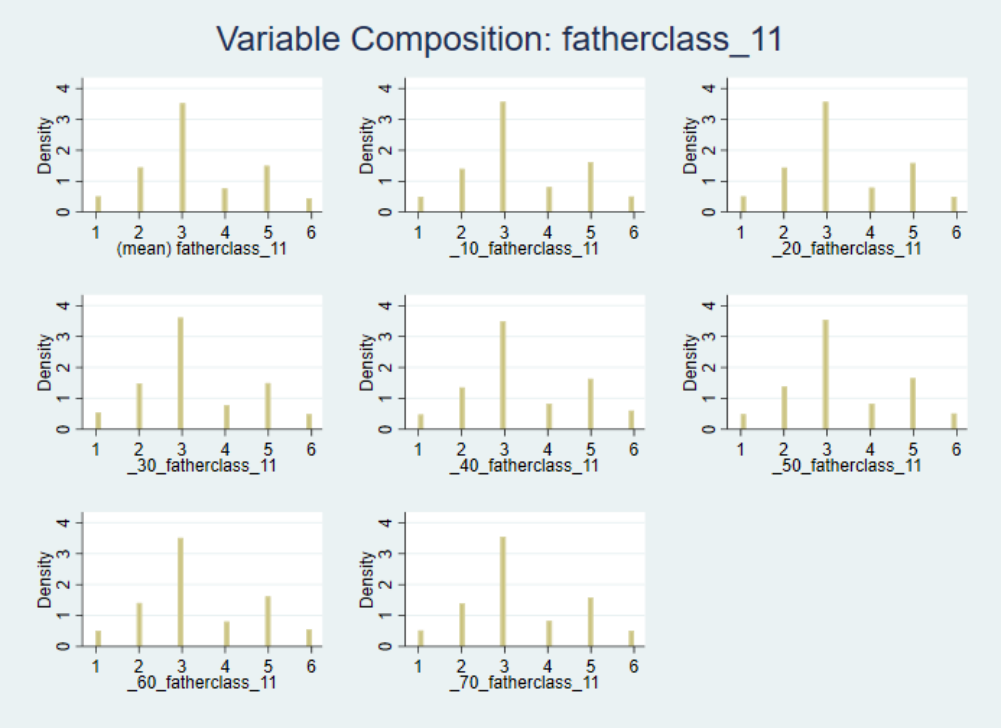


Figure B.46: Variable composition across MICE imputations: father's social class (age 7)

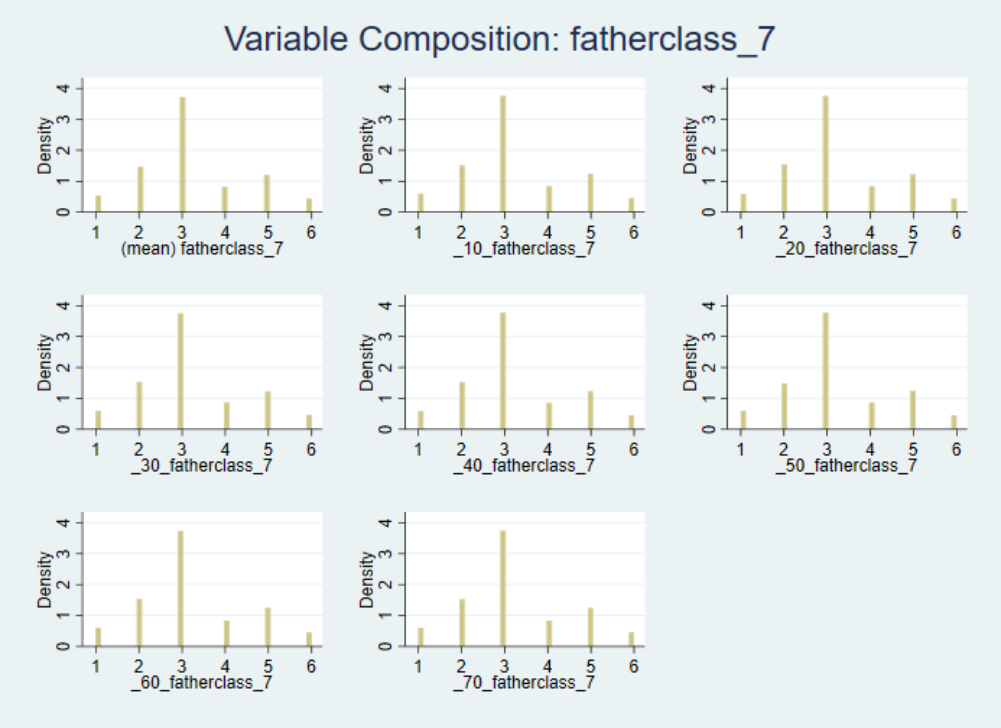


Figure B.47: Variable composition across MICE imputations: father's social class (age 0)

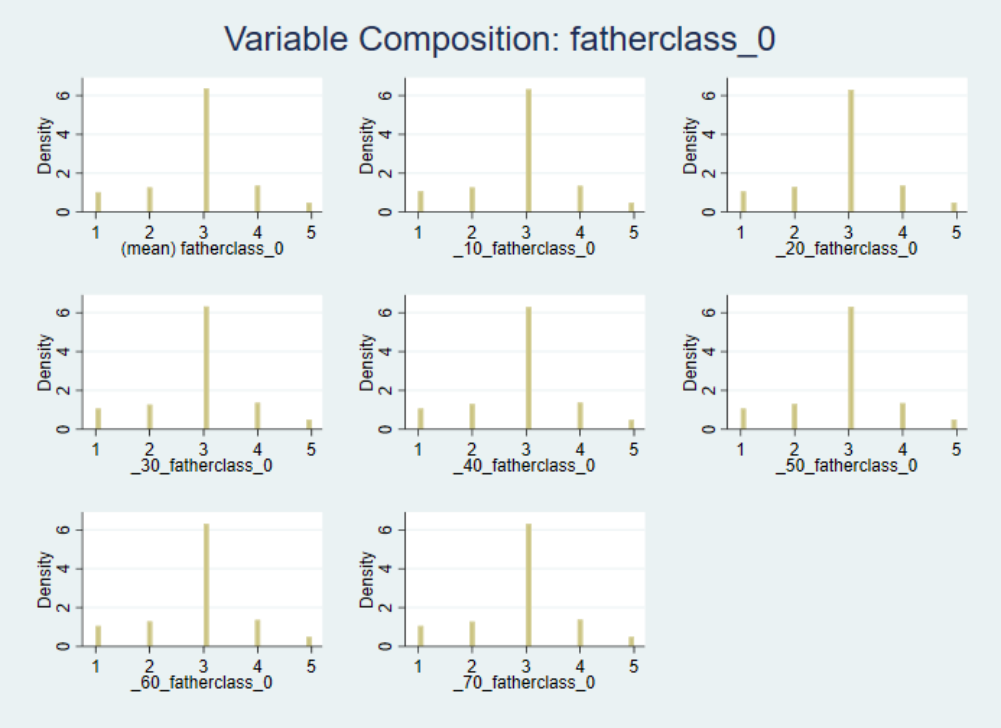


Figure B.48: Variable composition across MICE imputations: education (age 50)

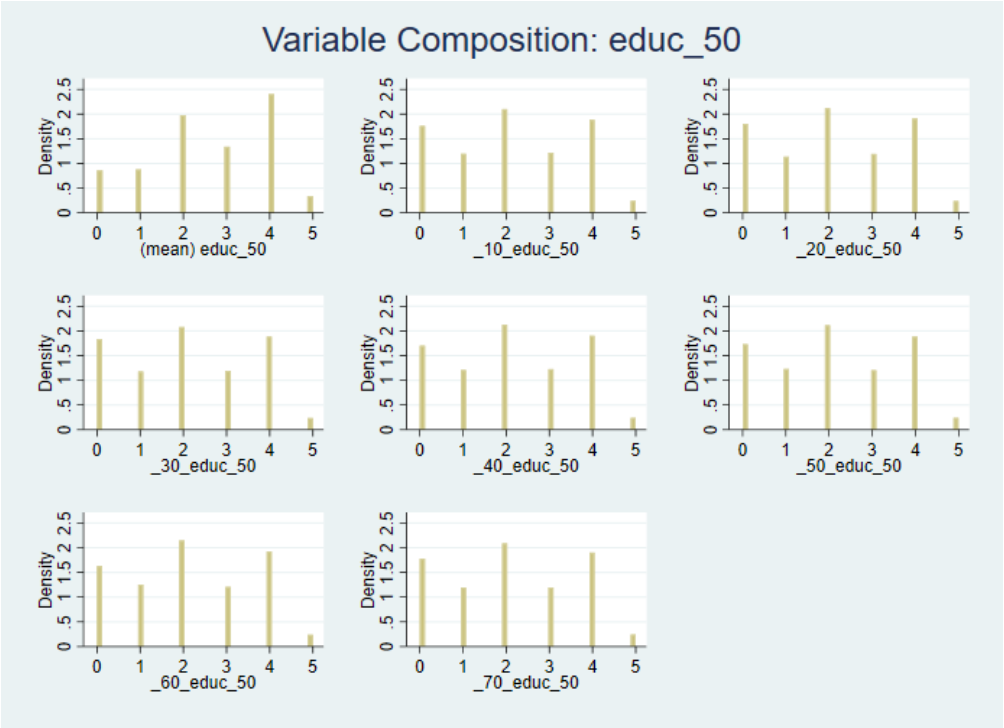


Figure B.49: Variable composition across MICE imputations: mother's education (age 16)

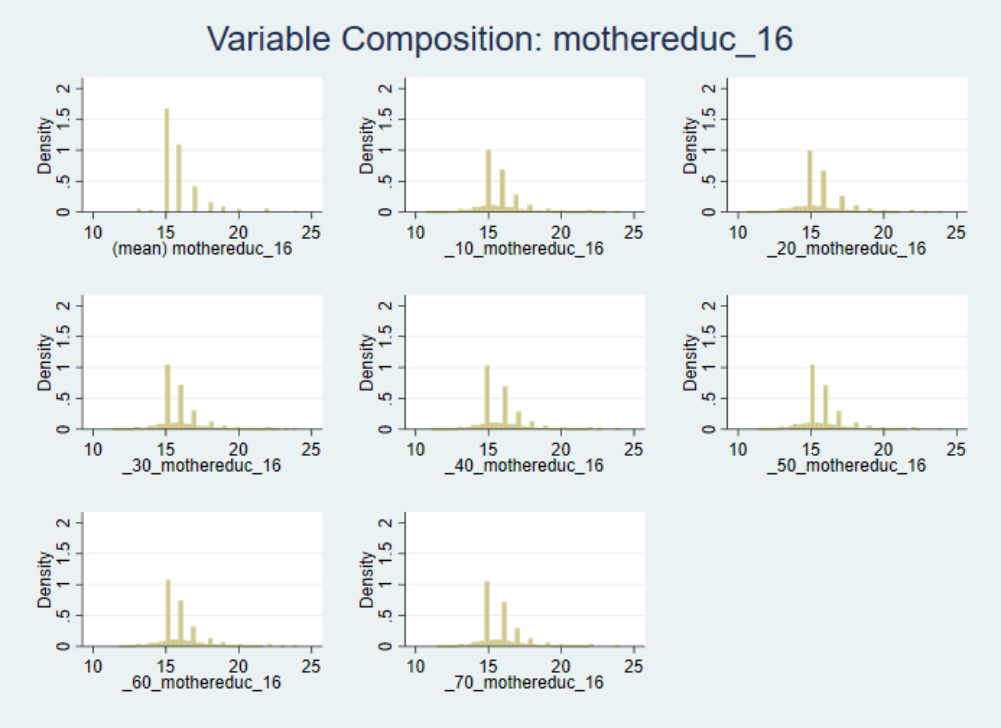


Figure B.50: Variable composition across MICE imputations: father's education (age 16)

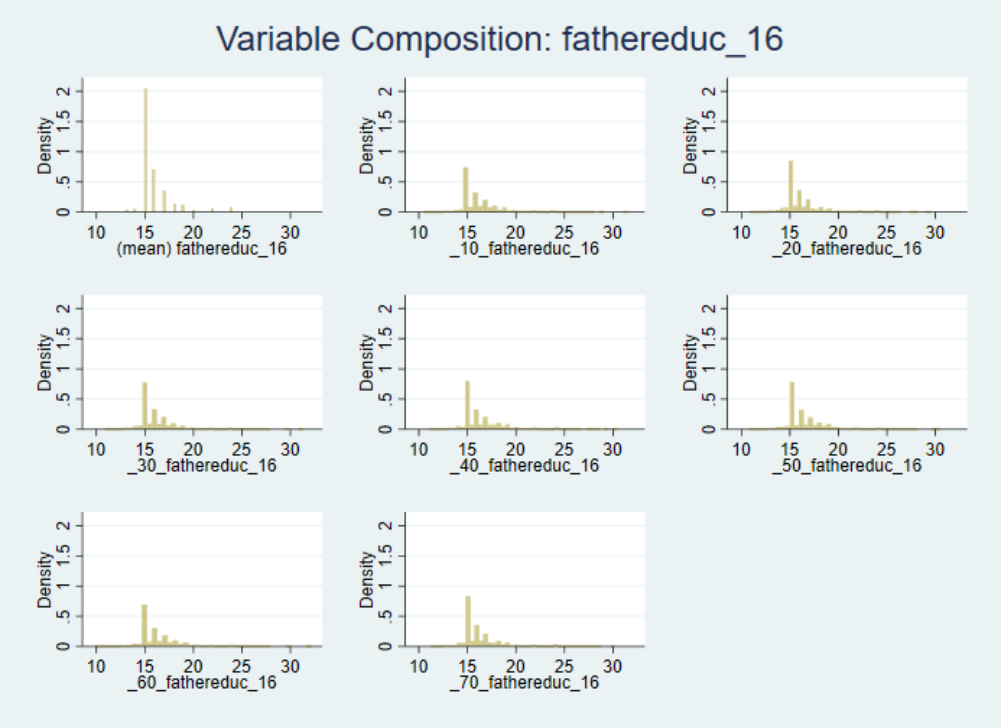


Figure B.51: Variable composition across MICE imputations: father's education (age 7)

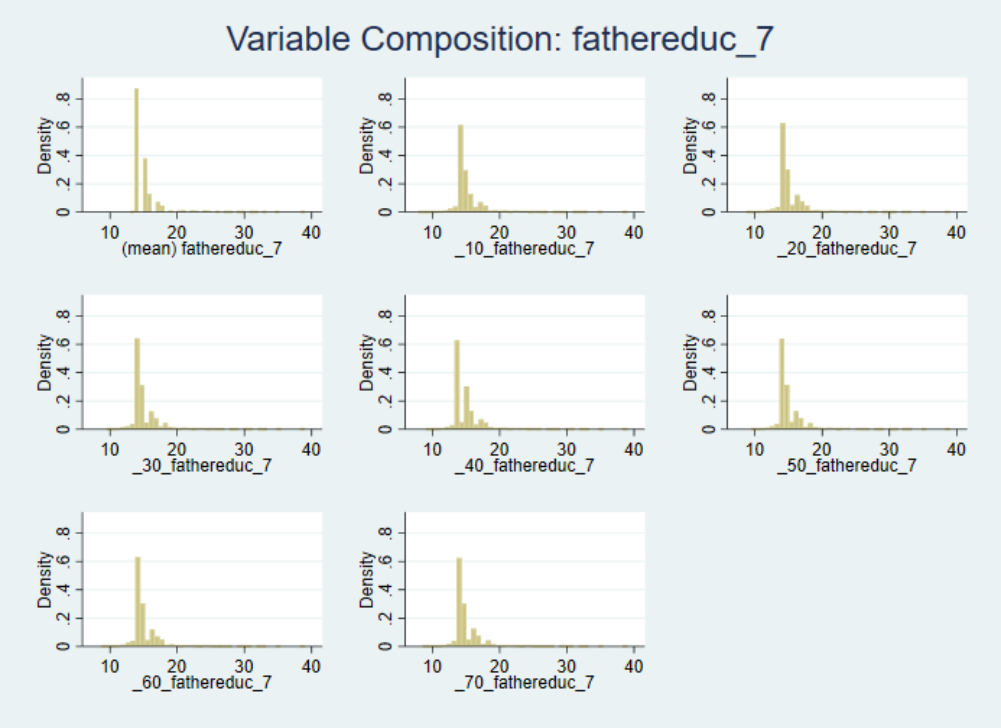
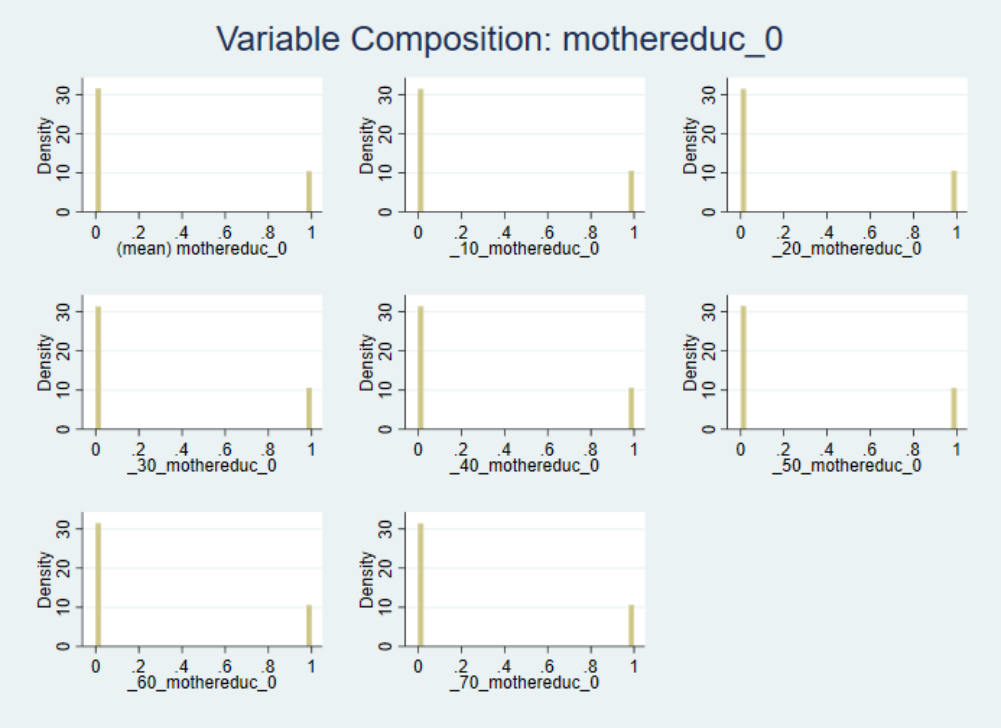


Figure B.52: Variable composition across MICE imputations: mother's education (age 0)



Post-regression Results

Available upon request. (i) Between, within, and total imputation variance, (ii) relative increase in variance, (iii) fraction of missing information, (iv) relative efficiency, and (parameter-specific degrees of freedom all indicate the reliability of the multiple imputation model, given 70 imputations (m) and 35 iterations, (i).

B.4 Cohort Study Reporting

Table B.2: Cohort study reporting checklist according to STROBE (Vandenbroucke et al. 2007).

Item	Recommendation		Page	Relevant text from manuscript
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	78	"Comparing Parental and Current Socioeconomic Gradients in Health: A Simple Inequality of Opportunity Metric Applied to a UK Birth Cohort"
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	78	""Aim: This study compares conventional current social gradients in health with parental social gradients: bivariate associations between fraction-ranked parental socioeconomic status in childhood and adult health. Methods: I calculate gradients using slope and relative indices of inequality in 36-item Short Form Health Survey (SF-36) scores at age 50 from the 1958 National Child Development Study, applying multiple imputation methods to account for attrition. Socioeconomic variables include income, education, and occupational class, measured across childhood, ages 0-16, and in adulthood at age 50.""
Introduction				

Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	79

"“Sophisticated measures of lifetime inequality of opportunity for health require undertaking the empirically challenging and ethically controversial task of disentangling the causal influence of circumstances and choices throughout an individual’s life, including unfortunate circumstances beyond individual control arising after childhood (Rosa Dias 2009).1 This raises the question about whether bivariate associations between childhood socioeconomic circumstances and adult health might be useful as a simple indicator of lifetime inequality of opportunity (IOp) for health.”"

Objectives	3	State specific objectives, including any prespecified hypotheses	79	"“In this chapter, I propose the parental gradient as a simple tool to measure and monitor inequality of opportunity for health. Parental gradients refer to bivariate associations between parental socioeconomic fraction-rank in childhood and adult health, while standard “current” gradients refer to bivariate associations between socioeconomic fraction-rank and health contemporaneously in adulthood. I use the 1958 National Child Development Study (NCDS) as an illustrative example to estimate parental gradients in respondent the 36-item Short Form Health Survey (SF-36) scores at age 50 by parental education, income, and occupation during respondents' childhood. I conduct robustness checks to investigate the potential impacts of: (i) intergenerational socioeconomic rank mobility; (ii) controlling for current socioeconomic fraction-rank; (iii) different timings of socioeconomic variable measurement during childhood; (iv) missing data methodology.”"
Methods				

Study design	4	Present key elements of study design early in the paper	91	"I calculate current and parental socioeconomic gradients in SF-36 scores at age 50, using both the slope and relative indices of inequality (SII; RII), since their results may differ in some instances (Pamuk 1988). Socioeconomic variables are fraction-ranked, with a 0-1 scale and mean of 0.5, aiding comparisons across socioeconomic variables and time. Fraction-ranking also aids interpretation of results, indicating the linear approximation of the difference or ratio between the highest and lowest socioeconomic advantage. I control for age through the cohort study design and, control for sex by conducting analyses on separate samples (Mackenbach et al. 1997; Moreno-Betancur et al. 2015)'"

	Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	81	"“The 1958 NCDS initially consisted of 17 415 children born 3rd-9th March 1958 in the UK, capturing key information through their life course (Power and Elliott 2006). Since then, nine follow-up sweeps have been conducted, most recently at age 55. Childhood follow-ups at age 7, 11, and 16 included cohort members, their parents, and teachers. These three childhood sweeps also supplemented the sample with immigrants born in the same week of March 1958, in response to emigration from the UK in the original sample, increasing the total cohort to 18 558 individuals. Although the most recent available sweep is at age 55, I use primary outcomes from the sweep at age 50, due to the availability of the 36-Item Short Form Health Survey (SF-36) questionnaire.”"
	Participants	6	Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	81-82	Table 4.1
	Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	83-90	Tables 4.2-4.3

Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	83-90	"Example: I use the ONS SOC-90 to categorise occupations into six occupational classes at age 50, ranging from 'I', professional occupations, to 'V', unskilled labour (ONS 2020b)."
Bias	9	Describe any efforts to address potential sources of bias	93-94	Robustness checks section.
Study size	10	Explain how the study size was arrived at	81-82	Missing data and multiple imputation approach.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	83-90	All variables described in data section and histograms provided in Appendix B.1.
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	92-94	4.3.2 Inequality measurements. 4.3.3 Robustness checks
		(b) Describe any methods used to examine subgroups and interactions	N/A	N/A
		(c) Explain how missing data were addressed	81-82	Missing data and multiple imputation approach.
		(d) If applicable, explain how loss to follow-up was addressed	N/A	Citation: (Power et al., 2006)
		(e) Describe any sensitivity analyses	93-94	4.3.3 Robustness checks
Results				

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	82	Table 4.1
		(b) Give reasons for non-participation at each stage	81	Citation: Power et al (2006)
		(c) Consider use of a flow diagram	N/A	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	83-90	Tables 4.2-4.3
		(b) Indicate number of participants with missing data for each variable of interest	83-90	Tables 4.2-4.3
		(c) Cohort study—Summarise follow-up time (eg, average and total amount)	82	Table 4.1
Outcome data	15*	Cohort study—Report numbers of outcome events or summary measures over time	174-176	Table B.1
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	94-97; 97-98	Unadjusted – 4.4.1 Parental and Current Socioeconomic Gradients in Health
		(b) Report category boundaries when continuous variables were categorized	83-90	Adjusted – 4.4.2 Robustness checks "Controlling for current socioeconomic status"

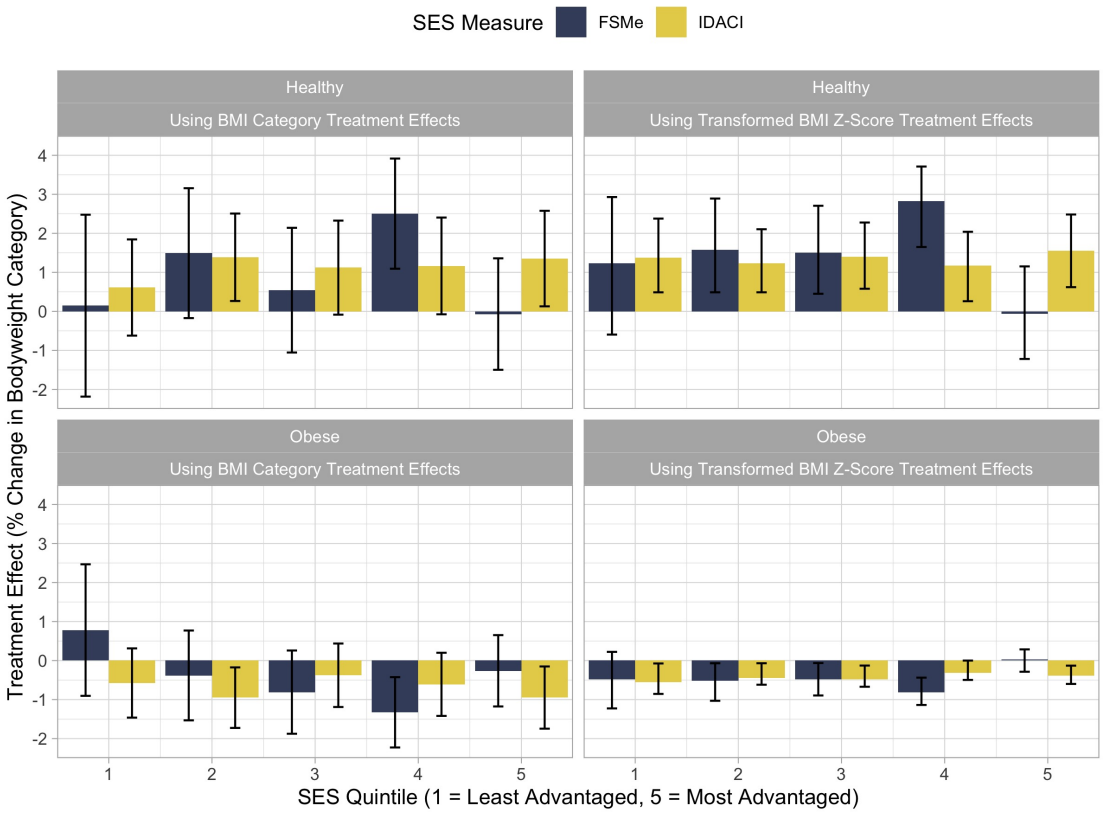
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	94-97	Tables 4.2-4.3
			160-168	Absolute and relative results provided in Section 4.4 and Appendix B.2
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	93-94	4.3.3 Robustness checks
Discussion				
Key results	18	Summarise key results with reference to study objectives	103-104	4.5.1 Principal findings
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	104-105	4.5.2 Strengths and Limitations
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	105	4.5.3 Policy Implications
External validity	21	Discuss the generalisability (external validity) of the study results	104	"Comparison with Previous Findings"
Other information				

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17	"“This work was supported by Wellcome Trust (Grant No. 205427/Z/16/Z) within the Senior Investigator Awards in Humanities and Social Science, titled Re-Engineering Health Policy Research for Fairer Decisions and Better Health of Professors Tim Doran and Richard Cookson.”"

C Appendix to Chapter 5

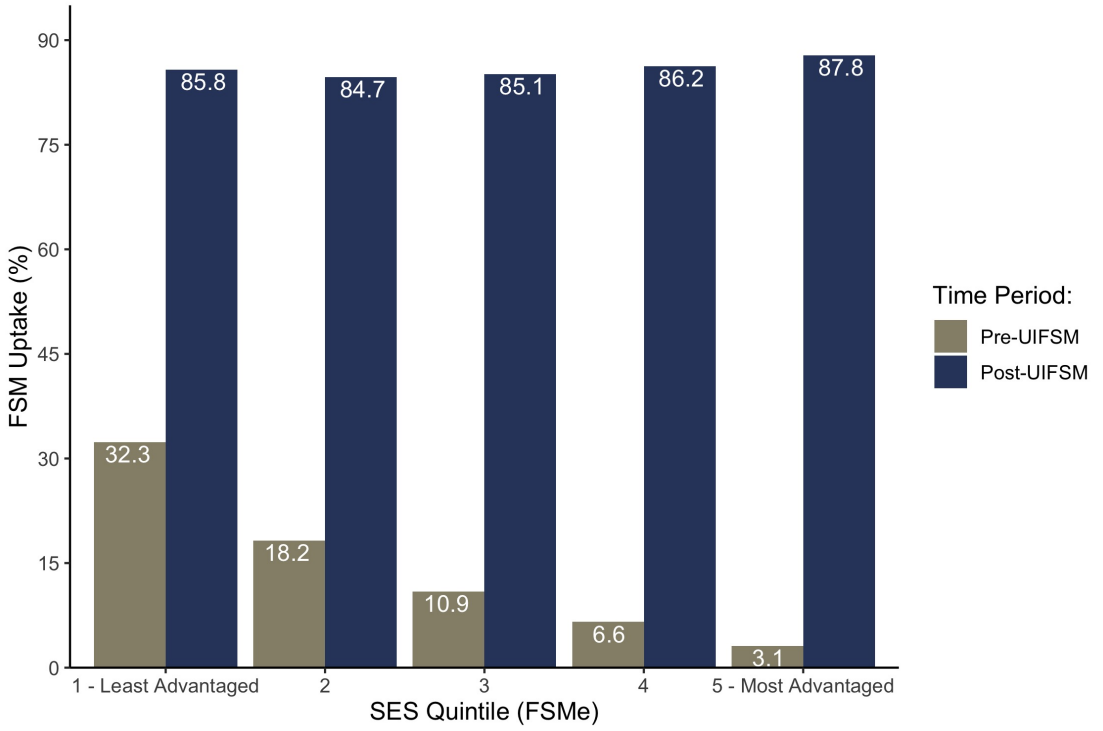
C.1 Universal Infant Free School Meal Statistics

Figure C.1: UIFSM treatment effects by BMI measure, socioeconomic measure, and socioeconomic quintile for healthy weight and obese groups.



Source: Holford and Rabe (2022). Abbreviations: UIFSM, Universal Infant Free School Meals; SES, socioeconomic status; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; BMI, body mass index.

Figure C.2: Proportion of FSM uptake by school-level FSM-eligibility quintiles pre- and post-UIFSM.



Data: DfE (2022). Pre-UIFSM: 2012/3-13/14; Post-UIFSM: 2014/15-15/16. Abbreviations: UIFSM, Universal Infant Free School Meals; SES, socioeconomic status; FSMe, Free School Meal eligibility.

C.2 Sensitivity Analyses: Maximum Cost-Effective Co-Fund

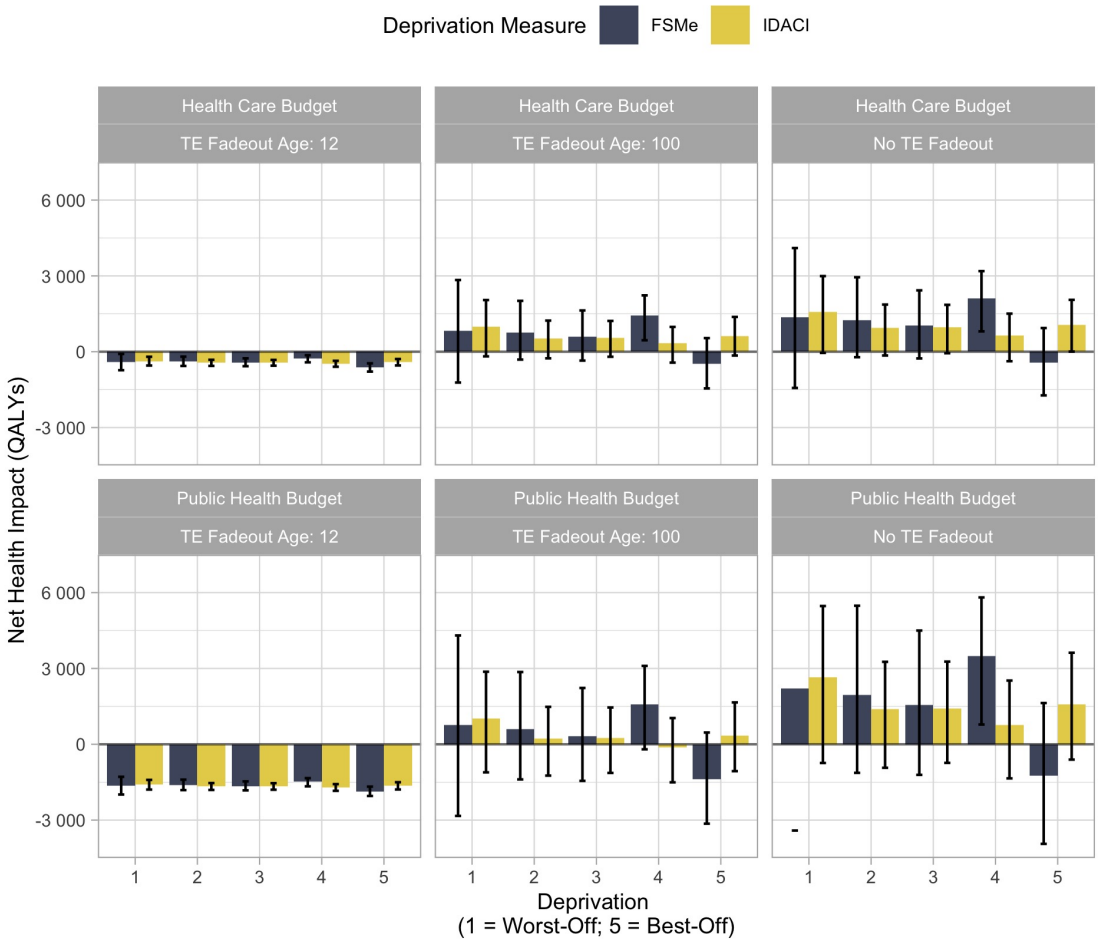
Table C.1: Maximum cost-effective UIFSM cofund

SES Measure	BMI Measure	Fadeout Age (Years)	Budget	Maximum Cost-Effective Co-Fund											
				£, millions						% Cohort UIFSM Costs					
				Mean TE	(	Min	-	Max	)	Mean TE	(	Min	-	Max	)
FSMe	BMI Z-Score	12	Public Health	5.2	(	0.5	-	9.8	)	1.1	(	0.1	-	2.1	)
			Health Care	13.7	(	1.4	-	26.1	)	2.9	(	0.3	-	5.6	)
		30	Public Health	17.4	(	1.7	-	33.4	)	3.7	(	0.4	-	7.1	)
			Health Care	38.4	(	3.8	-	73.6	)	8.2	(	0.8	-	15.6	)
		60	Public Health	36.0	(	3.1	-	69.6	)	7.7	(	0.7	-	14.8	)
			Health Care	66.9	(	6.0	-	128.9	)	14.2	(	1.3	-	27.4	)
		120	Public Health	58.7	(	4.5	-	114.3	)	12.5	(	0.9	-	24.3	)
			Health Care	92.9	(	7.6	-	180.0	)	19.7	(	1.6	-	38.3	)
		Never	Public Health	83.4	(	6.0	-	162.7	)	17.7	(	1.3	-	34.6	)
			Health Care	118.8	(	9.2	-	231.0	)	25.3	(	2.0	-	49.1	)
FSMe	BMI Category	12	Public Health	3.4	(	-3.7	-	10.4	)	0.7	(	-0.8	-	2.2	)
			Health Care	8.9	(	-10.0	-	27.8	)	1.9	(	-2.1	-	5.9	)
		30	Public Health	11.1	(	-13.7	-	35.8	)	2.4	(	-2.9	-	7.6	)
			Health Care	24.5	(	-30.0	-	79.1	)	5.2	(	-6.4	-	16.8	)
		60	Public Health	22.3	(	-33.4	-	78.1	)	4.7	(	-7.1	-	16.6	)
			Health Care	41.7	(	-59.9	-	143.3	)	8.9	(	-12.7	-	30.5	)
		120	Public Health	36.5	(	-60.8	-	133.8	)	7.8	(	-12.9	-	28.4	)
			Health Care	57.8	(	-90.9	-	206.5	)	12.3	(	-19.3	-	43.9	)
		Never	Public Health	52.4	(	-89.8	-	194.6	)	11.1	(	-19.1	-	41.4	)
			Health Care	74.1	(	-121.5	-	269.8	)	15.8	(	-25.8	-	57.4	)
IDACI	BMI Z-Score	12	Public Health	4.8	(	1.6	-	8.0	)	1.0	(	0.3	-	1.7	)
			Health Care	12.8	(	4.3	-	21.1	)	2.7	(	0.9	-	4.5	)
		30	Public Health	16.4	(	5.4	-	26.9	)	3.5	(	1.1	-	5.7	)
			Health Care	36.1	(	11.8	-	59.2	)	7.7	(	2.5	-	12.6	)
		60	Public Health	34.0	(	10.6	-	55.2	)	7.2	(	2.2	-	11.7	)
			Health Care	63.0	(	19.8	-	102.4	)	13.4	(	4.2	-	21.8	)
		120	Public Health	55.7	(	16.4	-	89.3	)	11.8	(	3.5	-	19.0	)
			Health Care	87.8	(	26.6	-	141.6	)	18.7	(	5.6	-	30.1	)
		Never	Public Health	79.4	(	22.9	-	126.7	)	16.9	(	4.9	-	26.9	)
			Health Care	112.6	(	33.4	-	180.7	)	23.9	(	7.1	-	38.4	)
IDACI	BMI Category	12	Public Health	4.5	(	-0.5	-	9.5	)	1.0	(	-0.1	-	2.0	)
			Health Care	12.0	(	-1.3	-	25.2	)	2.5	(	-0.3	-	5.4	)
		30	Public Health	15.6	(	-1.7	-	32.9	)	3.3	(	-0.4	-	7.0	)
			Health Care	34.4	(	-3.8	-	72.6	)	7.3	(	-0.8	-	15.4	)
		60	Public Health	34.3	(	-4.4	-	73.0	)	7.3	(	-0.9	-	15.5	)
			Health Care	62.9	(	-7.8	-	133.5	)	13.4	(	-1.7	-	28.4	)
		120	Public Health	59.2	(	-8.3	-	126.6	)	12.6	(	-1.8	-	26.9	)
			Health Care	91.0	(	-12.2	-	194.2	)	19.3	(	-2.6	-	41.3	)
		Never	Public Health	86.3	(	-12.5	-	185.0	)	18.3	(	-2.7	-	39.3	)
			Health Care	119.2	(	-16.6	-	254.9	)	25.3	(	-3.5	-	54.2	)

Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: UIFSM, Universal Infant Free School Meals; SES, socioeconomic status; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; BMI, body mass index; TE: treatment effect.

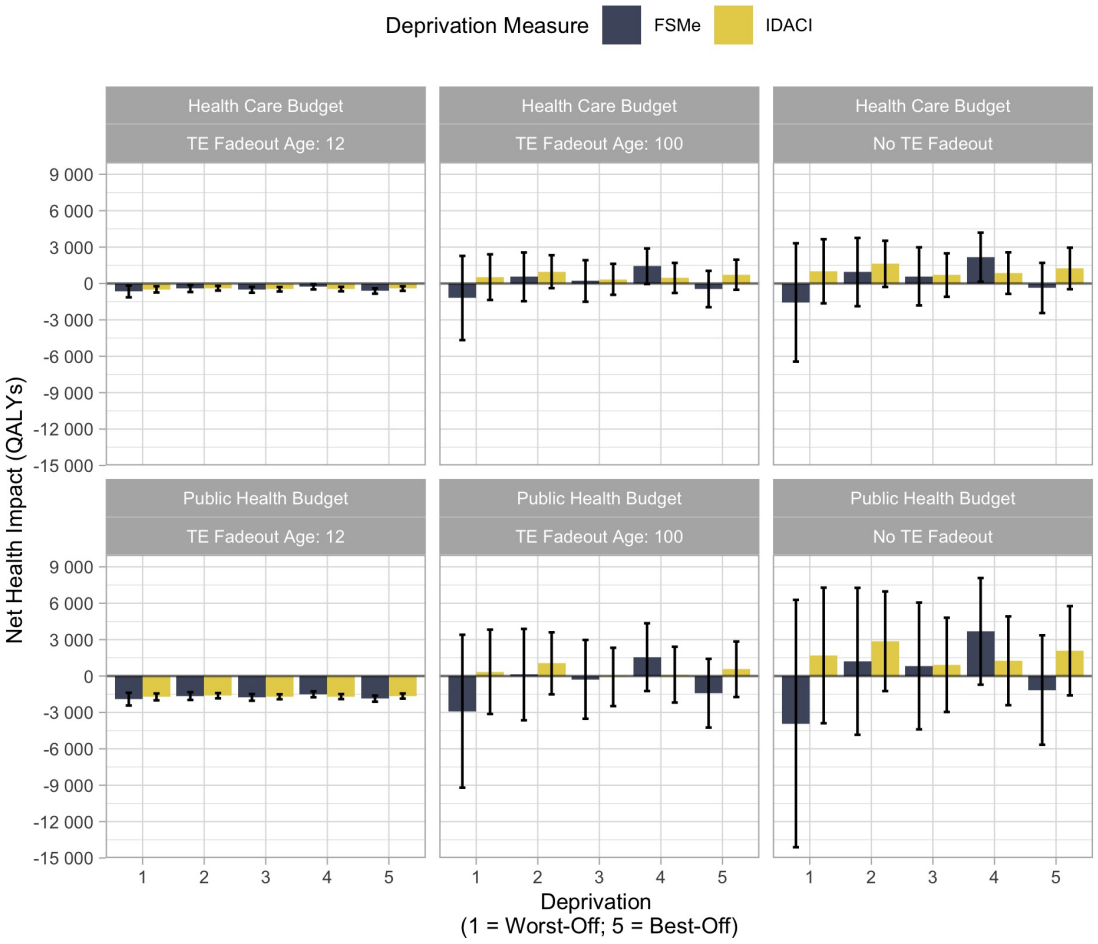
C.3 Sensitivity Analyses: Equity Impacts

Figure C.3: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI Z-Score treatment effects and assuming a 10% UIFSM co-fund and an equal distribution of health opportunity costs.



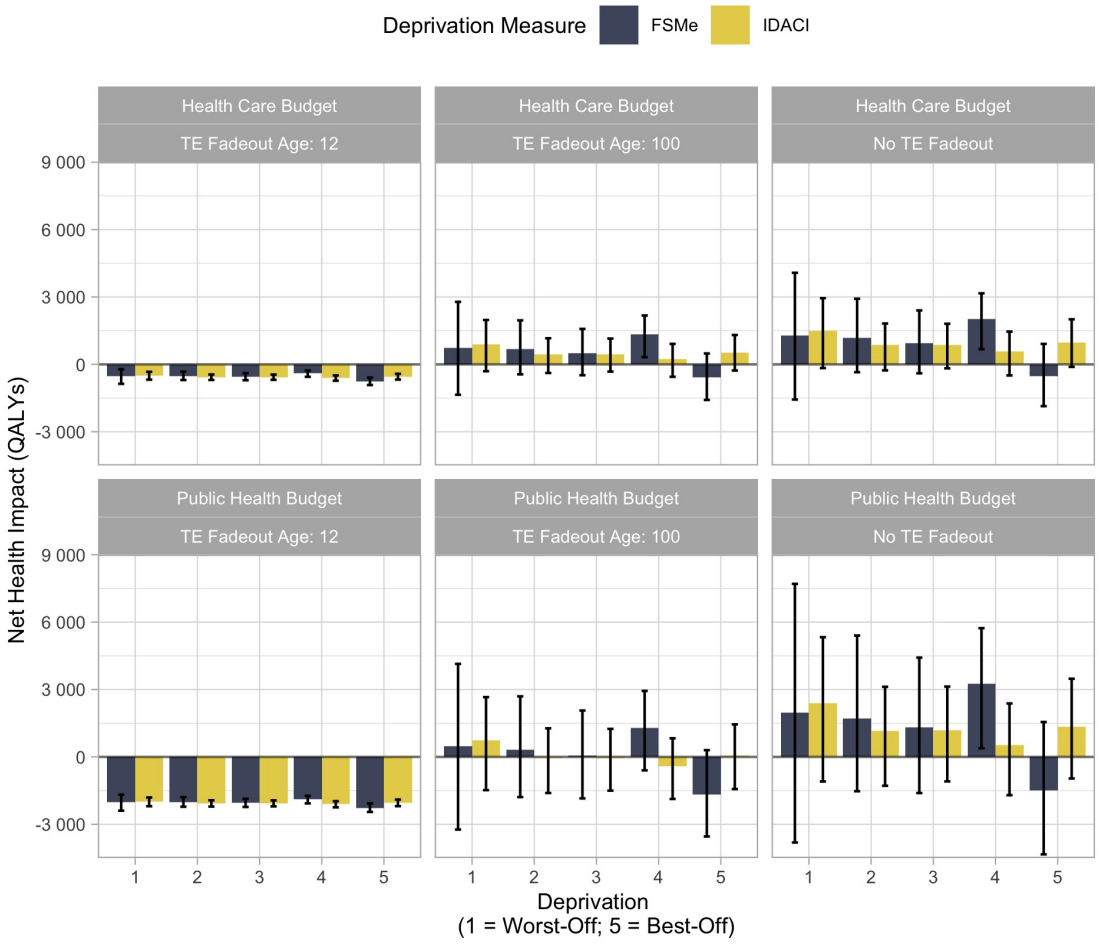
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.4: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI category treatment effects and assuming a 10% UIFSM co-fund and an equal distribution of health opportunity costs.



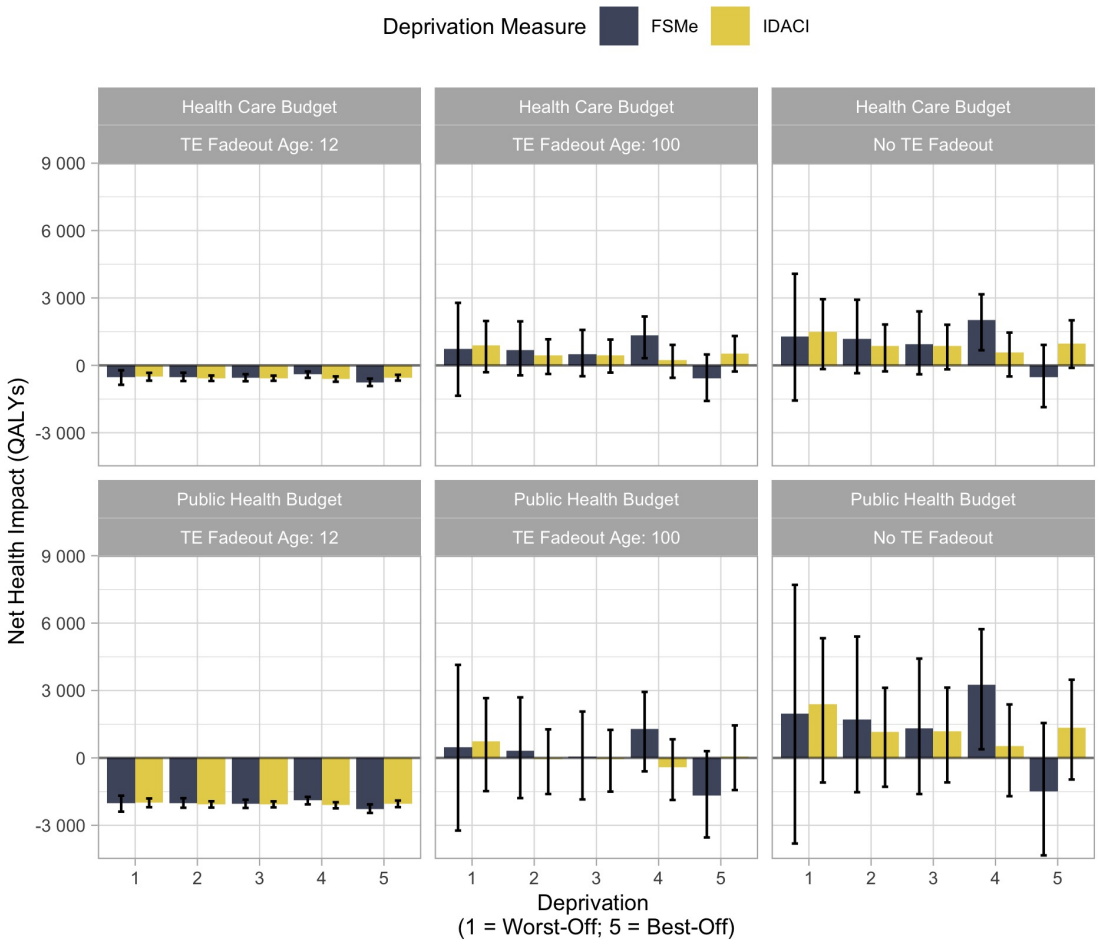
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.5: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI Z-Score treatment effects and assuming a 10% UIFSM co-fund and a pro-deprived distribution of health opportunity costs.



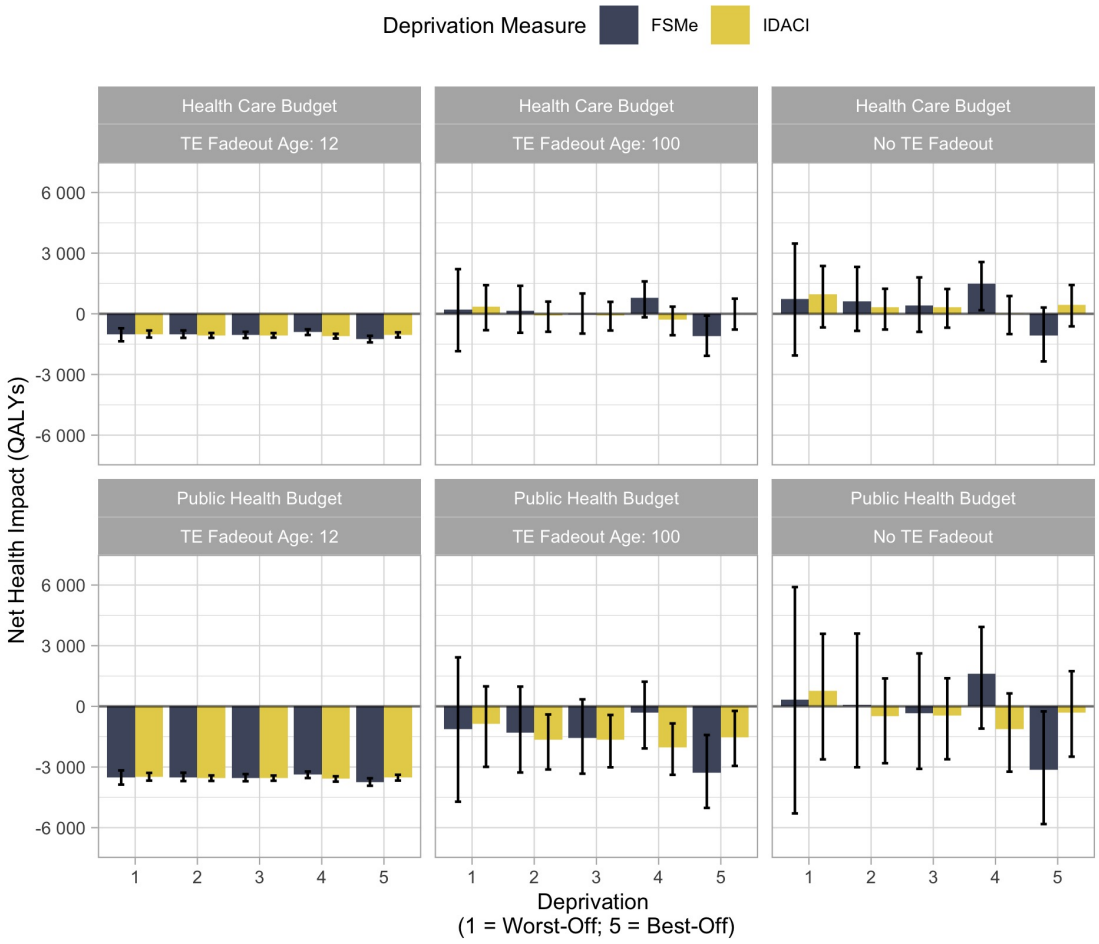
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.6: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI category treatment effects and assuming a 10% UIFSM co-fund and and a pro-deprived distribution of health opportunity costs.



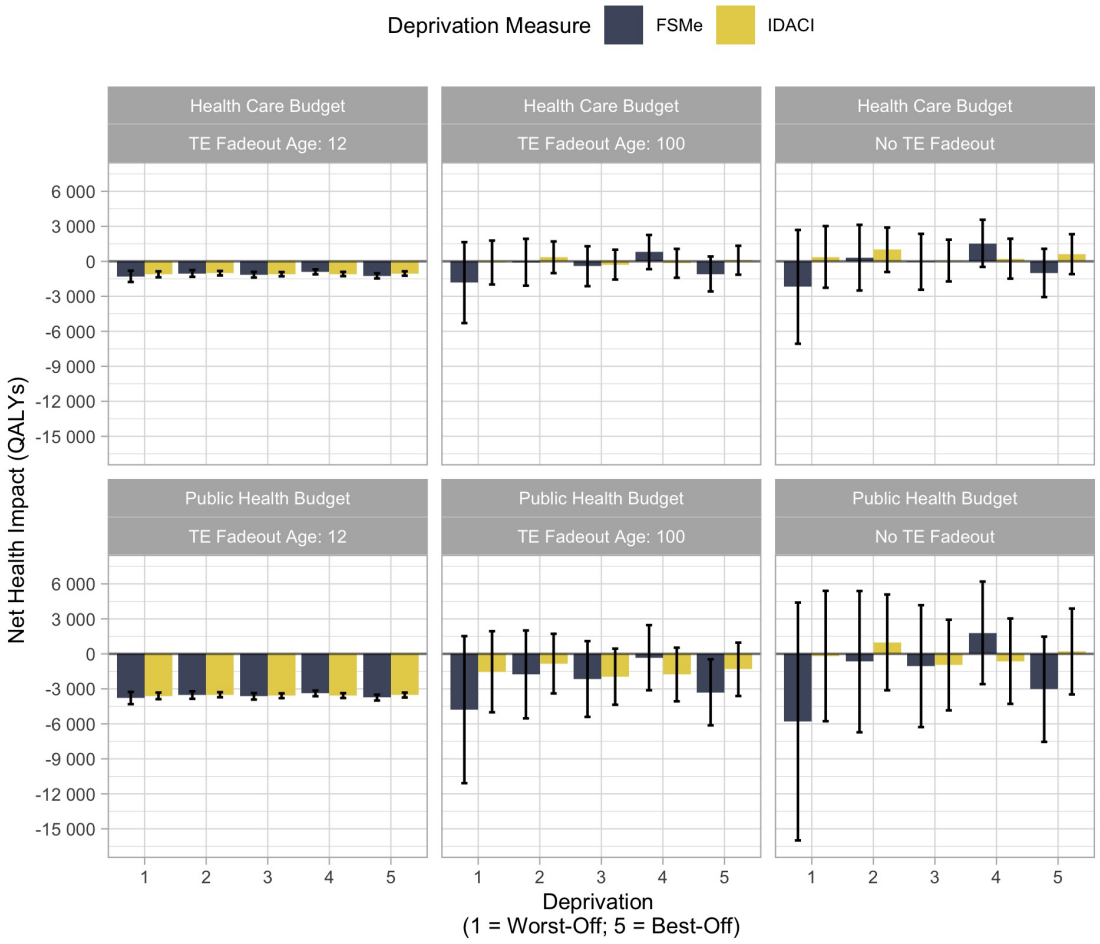
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.7: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI Z-Score treatment effects and assuming a 20% UIFSM co-fund and an equal distribution of health opportunity costs.



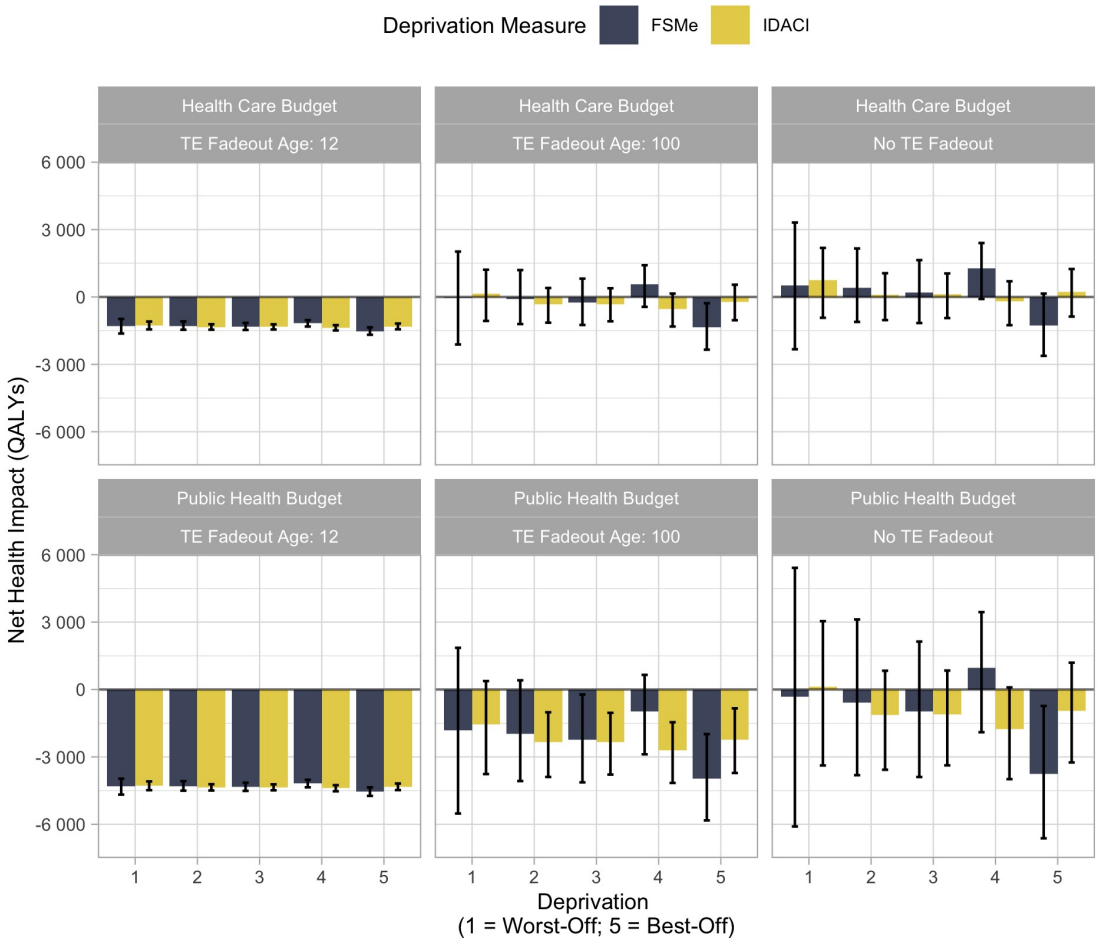
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.8: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI category treatment effects and assuming a 20% UIFSM co-fund and an equal distribution of health opportunity costs.



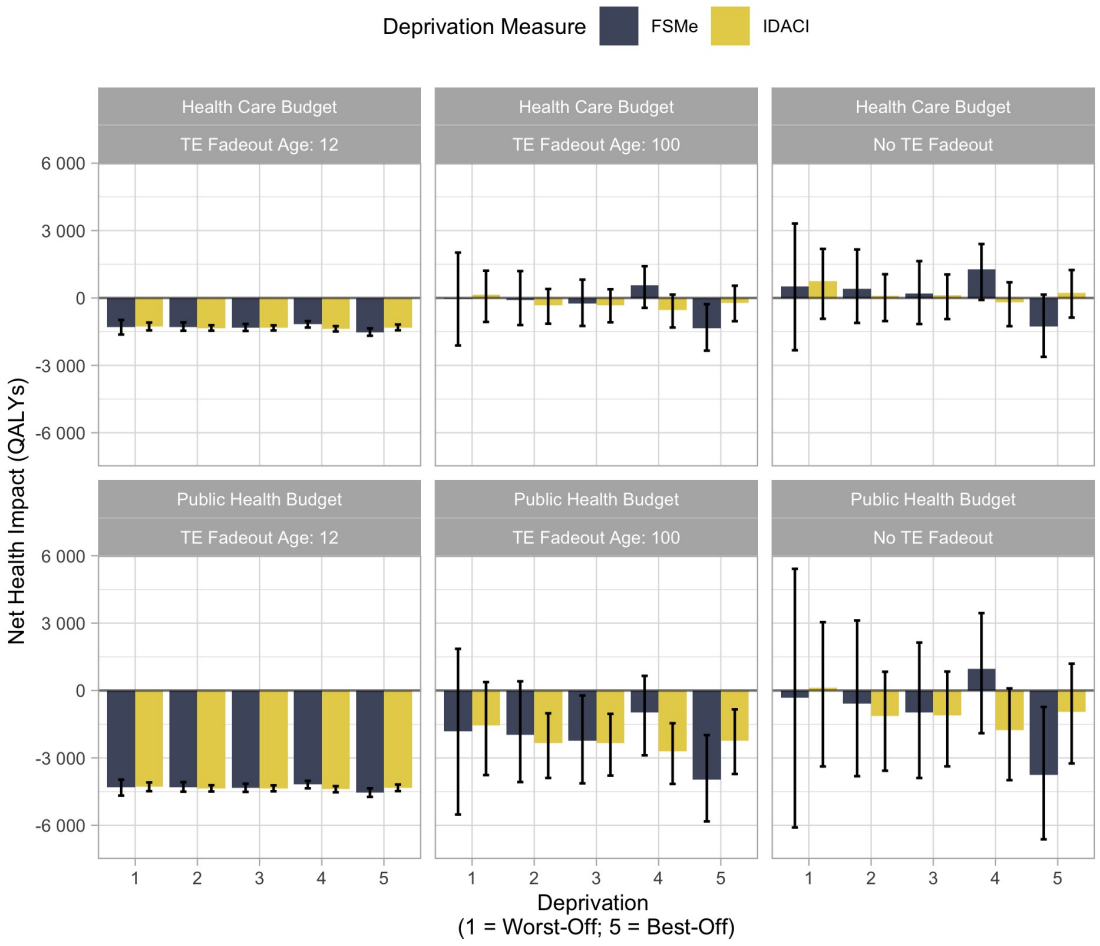
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.9: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI Z-Score treatment effects and assuming a 20% UIFSM co-fund and a pro-deprived distribution of health opportunity costs.



Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

Figure C.10: Net health benefits by deprivation measure, budget assumption, and treatment effect fadeout age using BMI category treatment effects and assuming a 20% UIFSM co-fund and and a pro-deprived distribution of health opportunity costs.

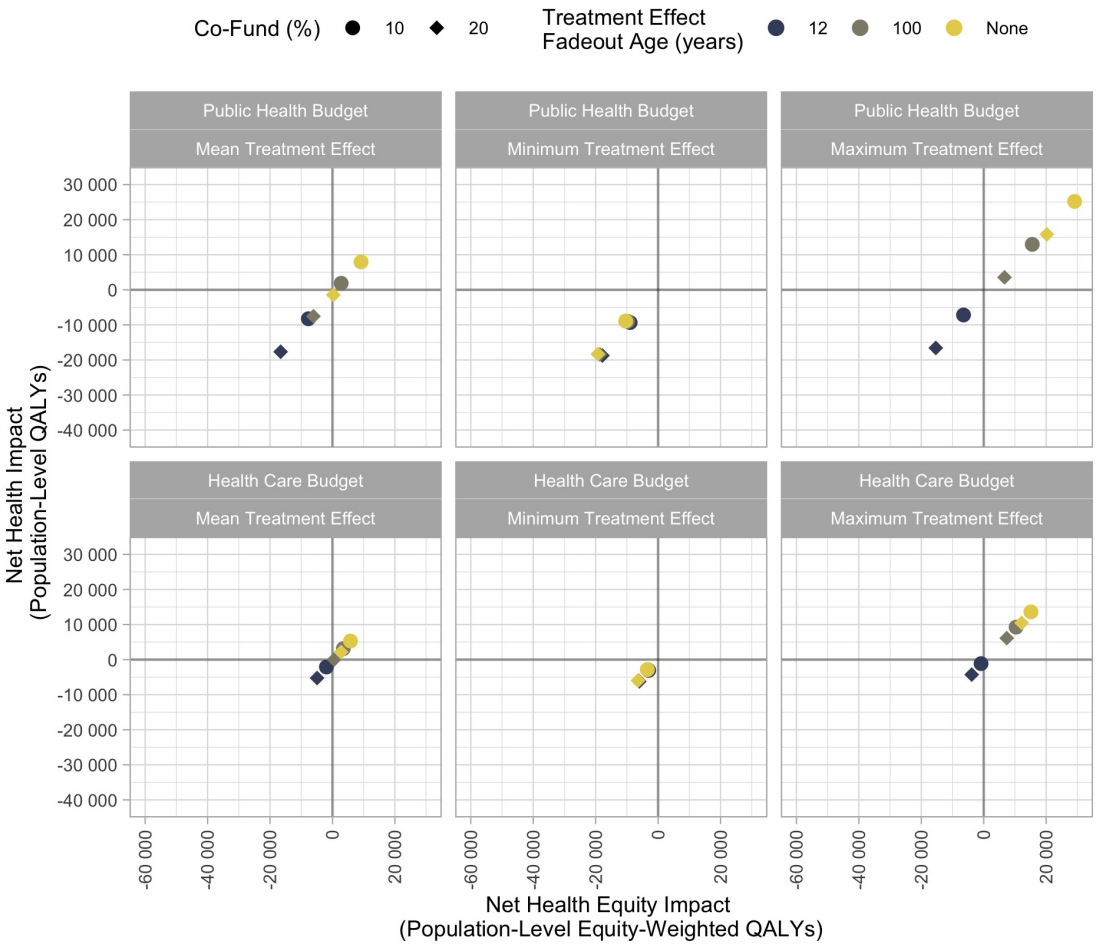


Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life year; UIFSM, Universal Infant Free School Meals; FSMe, Free School Meal eligibility; IDACI, Index of Multiple Deprivation Affecting Children Index; TE, treatment effect; HOC, health opportunity cost.

C.4 Sensitivity Analyses: Equity Trade-Offs

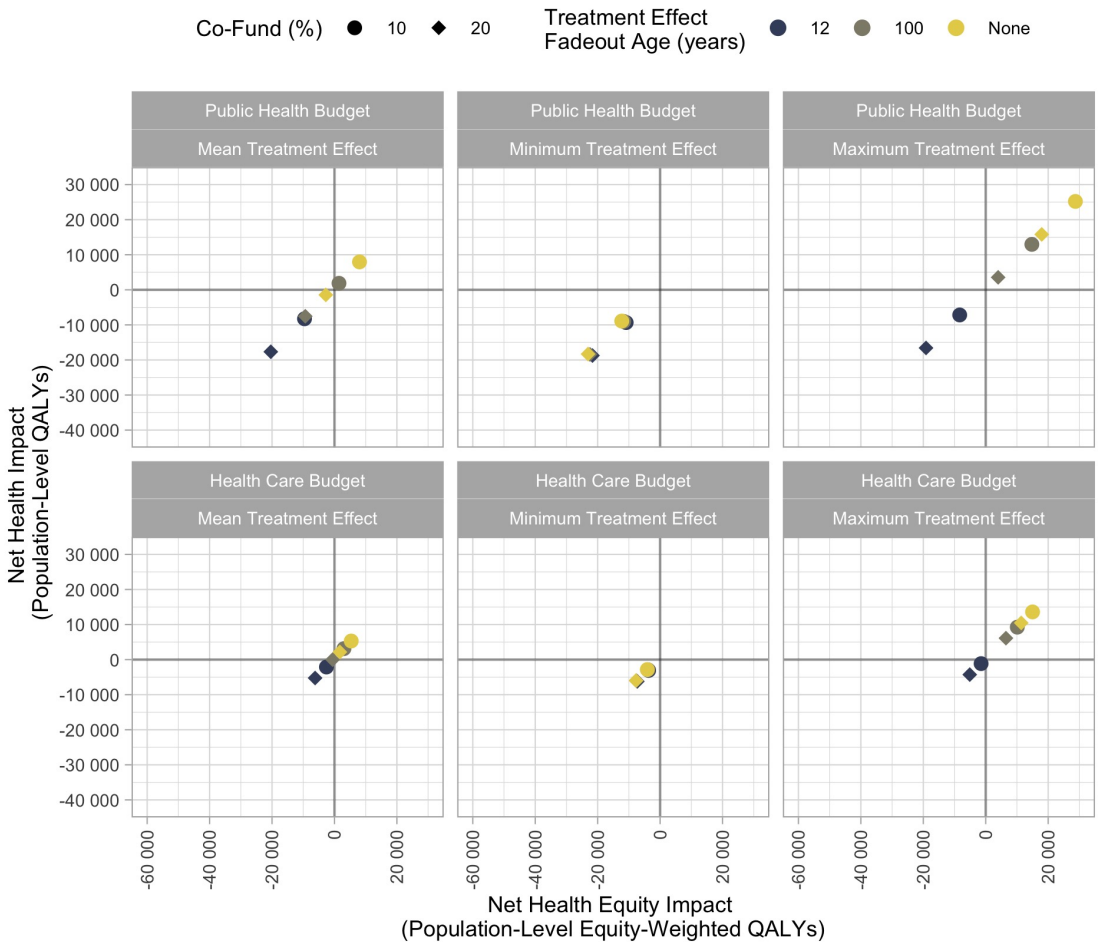
Equity Impact Plane

Figure C.11: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using FSMe quintiles, BMI Z-Score treatment effects, and a flat health opportunity cost gradient.



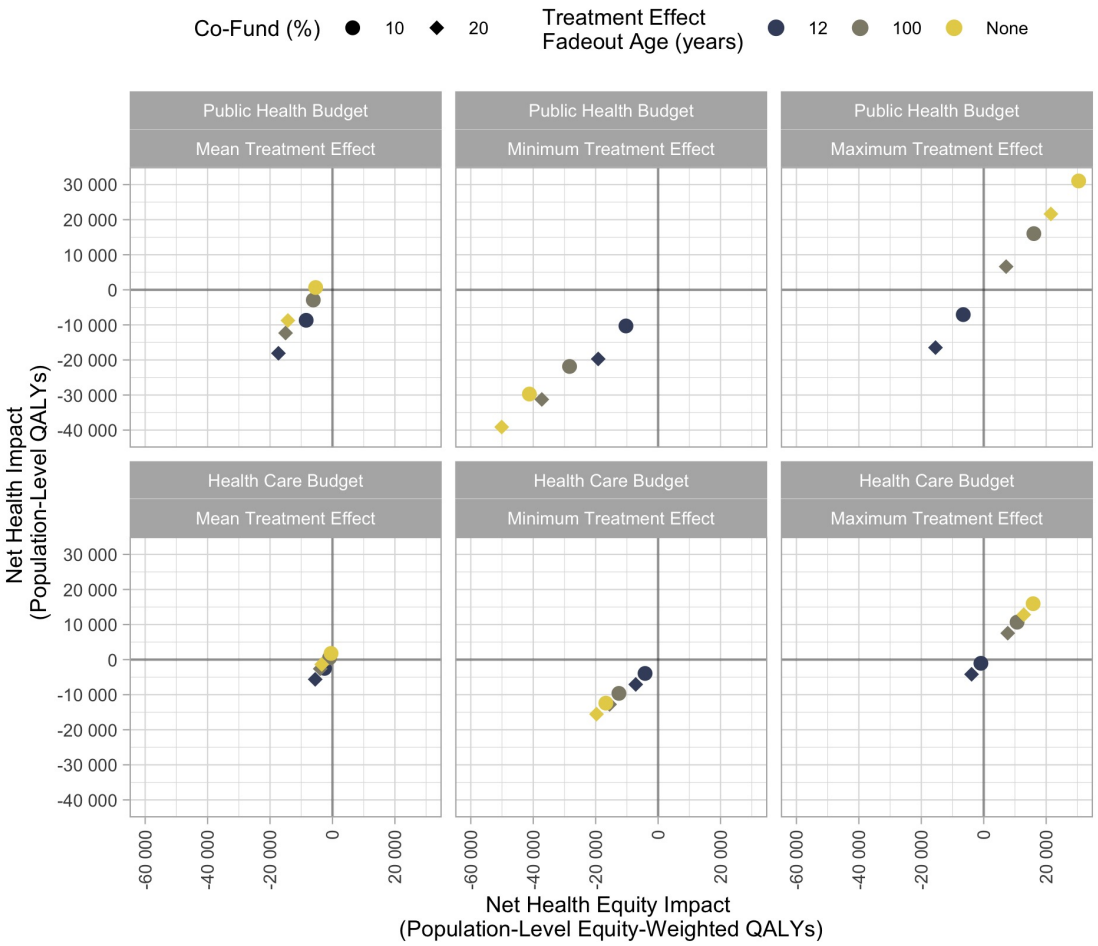
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.12: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using FSMe quintiles, BMI Z-Score treatment effects, and a pro-poor health opportunity cost gradient.



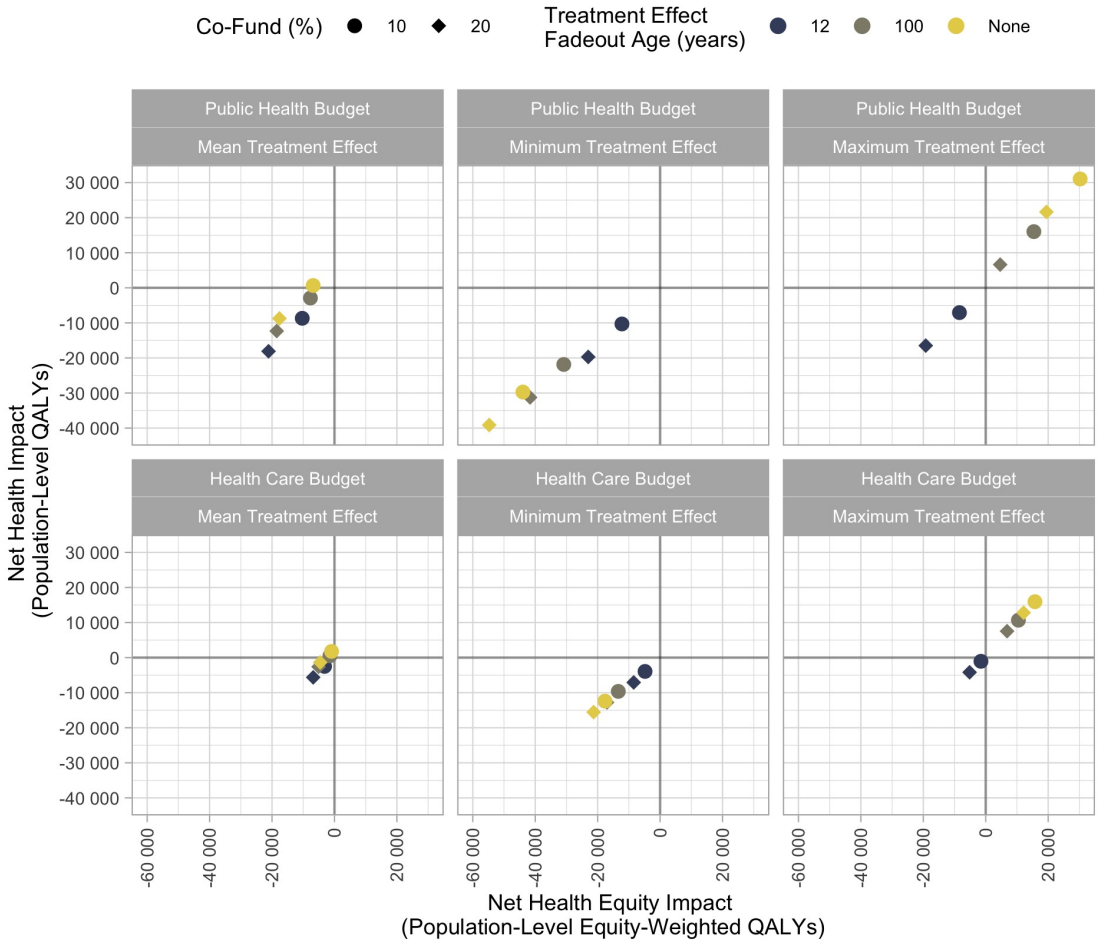
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.13: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using FSMe quintiles, BMI category treatment effects, and a flat health opportunity cost gradient.



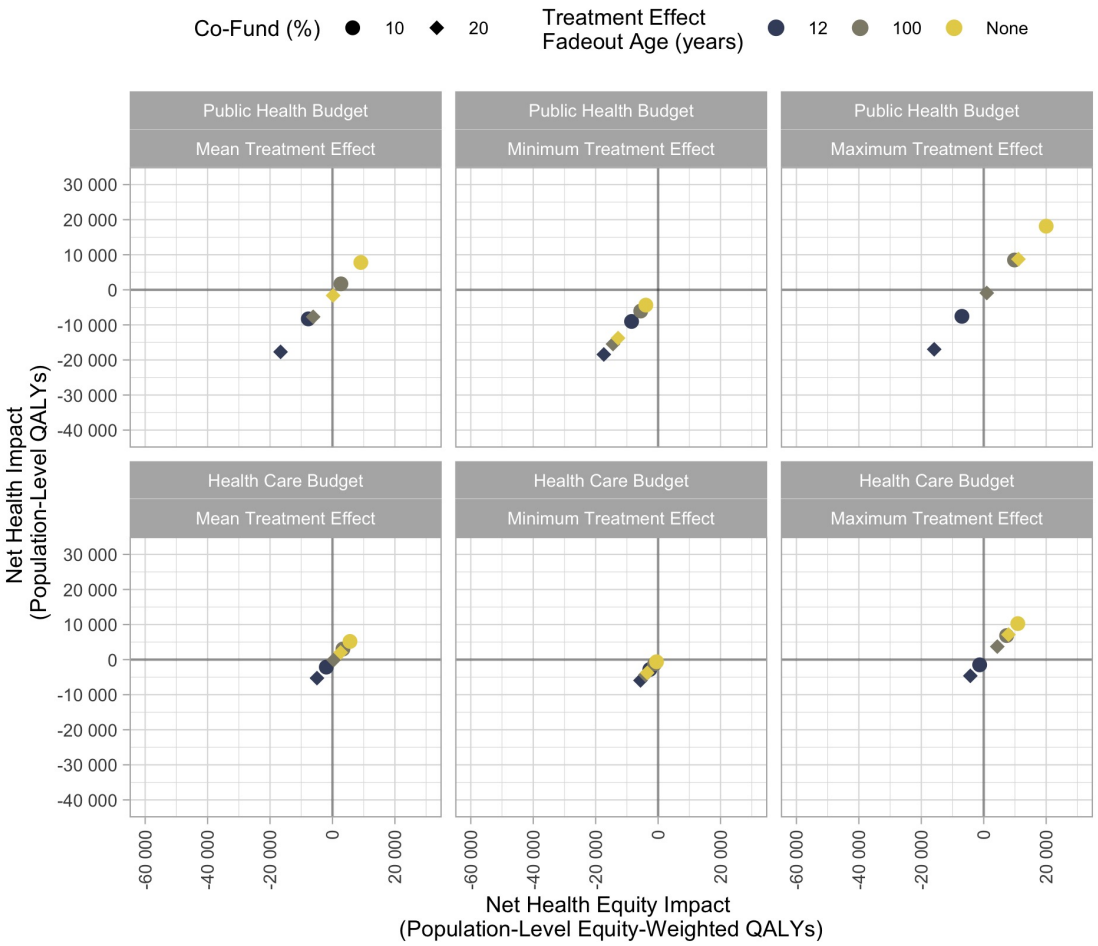
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.14: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using FSMe quintiles, BMI category treatment effects, and a pro-poor health opportunity cost gradient.



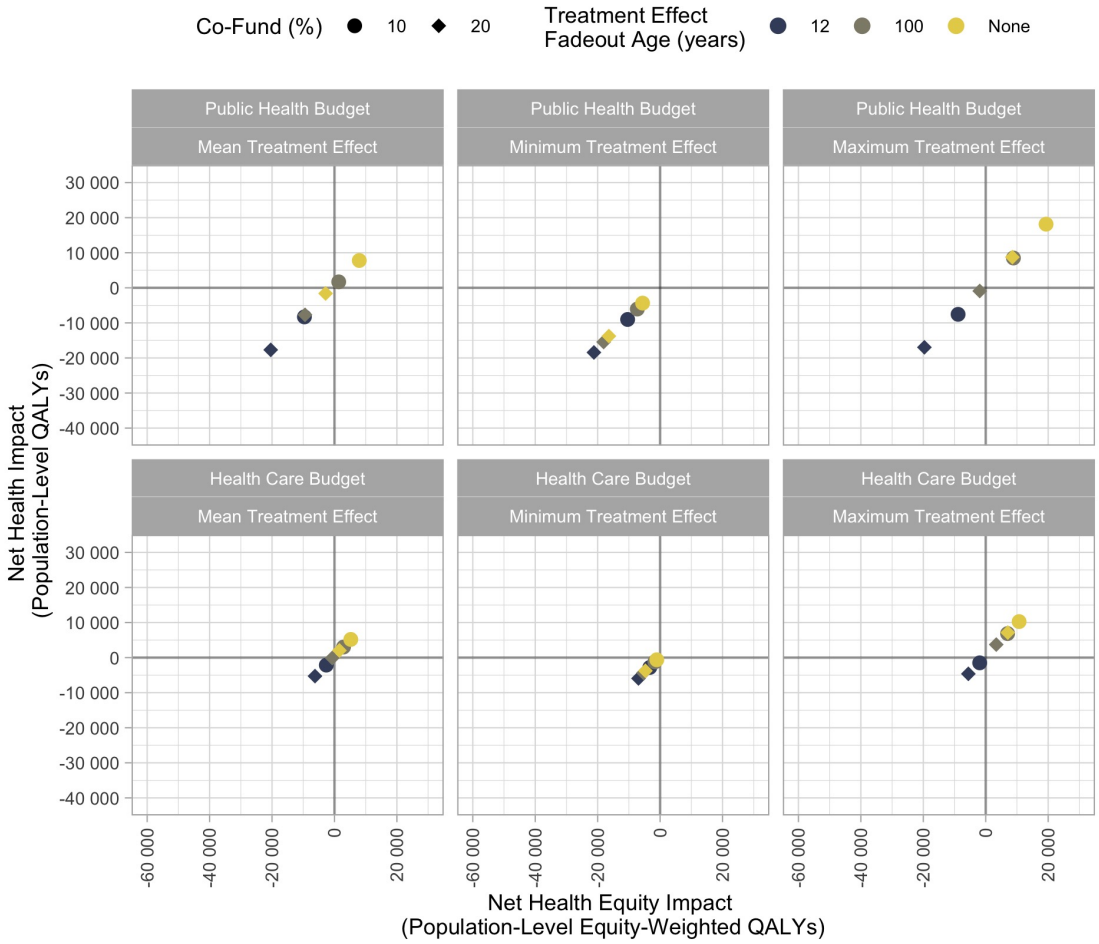
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.15: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using IDACI quintiles, BMI Z-Score treatment effects, and a flat health opportunity cost gradient.



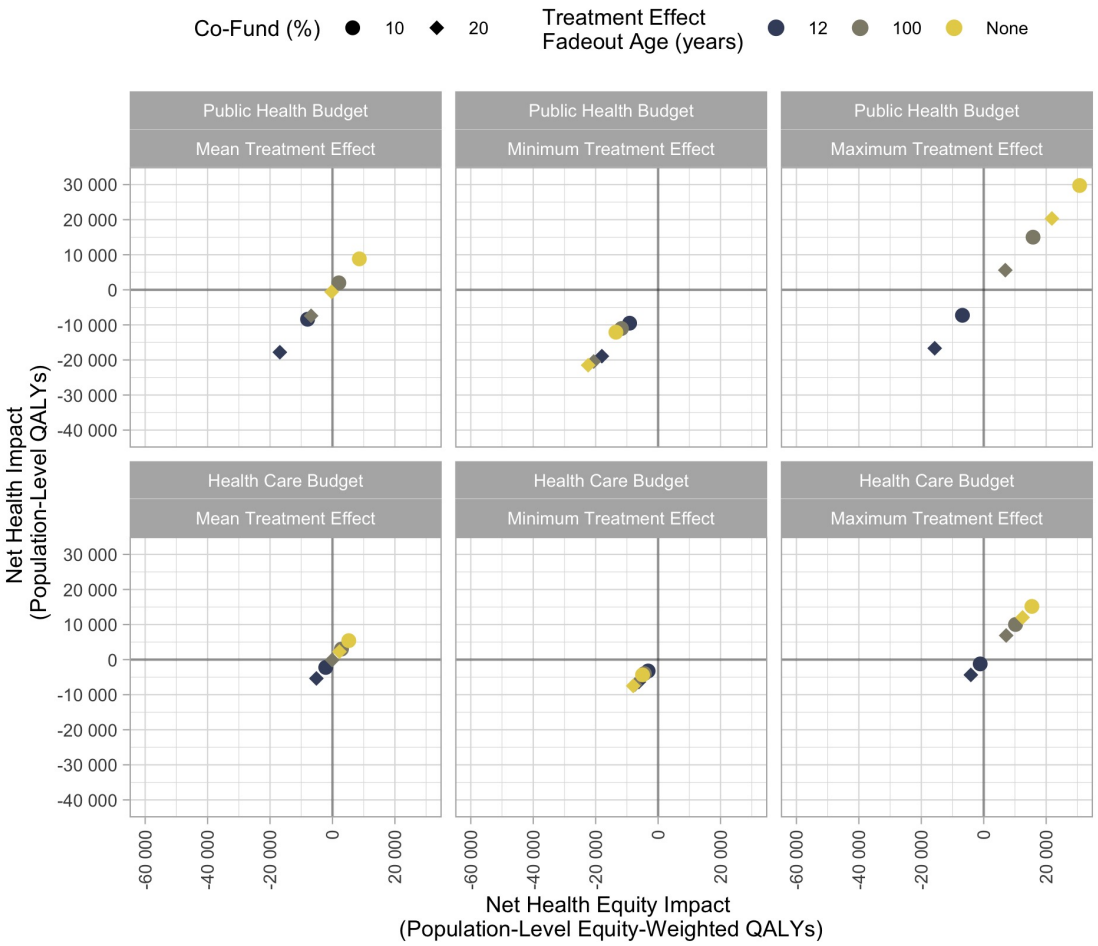
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.16: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using IDACI quintiles, BMI Z-Score treatment effects, and a pro-poor health opportunity cost gradient.



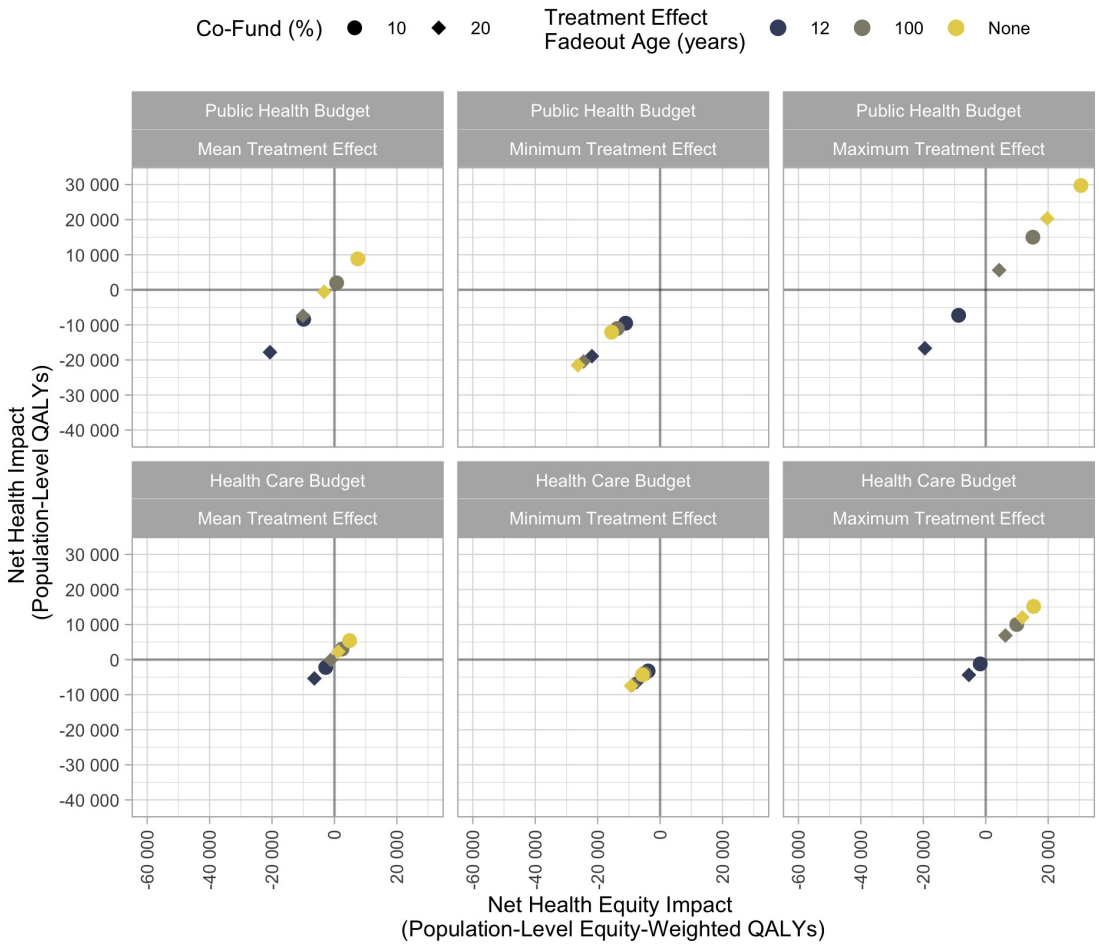
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.17: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using IDACI quintiles, BMI category treatment effects, and a flat health opportunity cost gradient.



Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

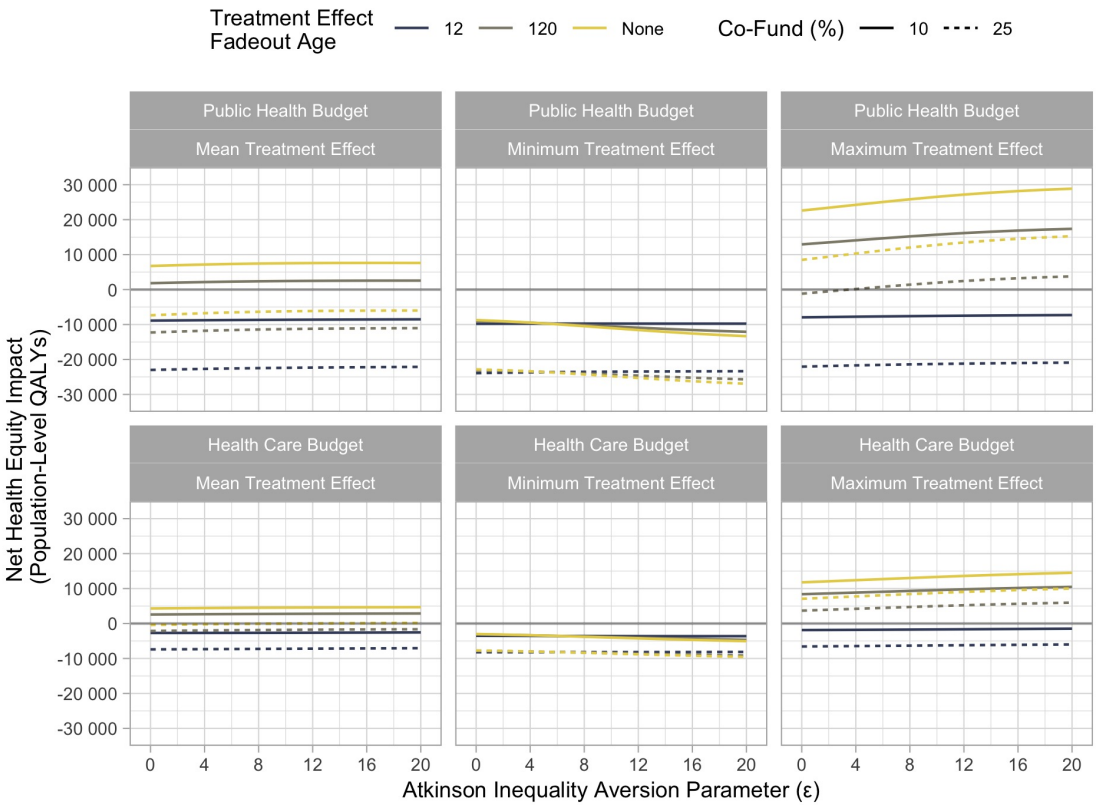
Figure C.18: Equity-efficiency impact plane by treatment effect fadeout age, co-funding level and treatment effect confidence interval, using IDACI quintiles, BMI category treatment effects, and a pro-poor health opportunity cost gradient.



Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

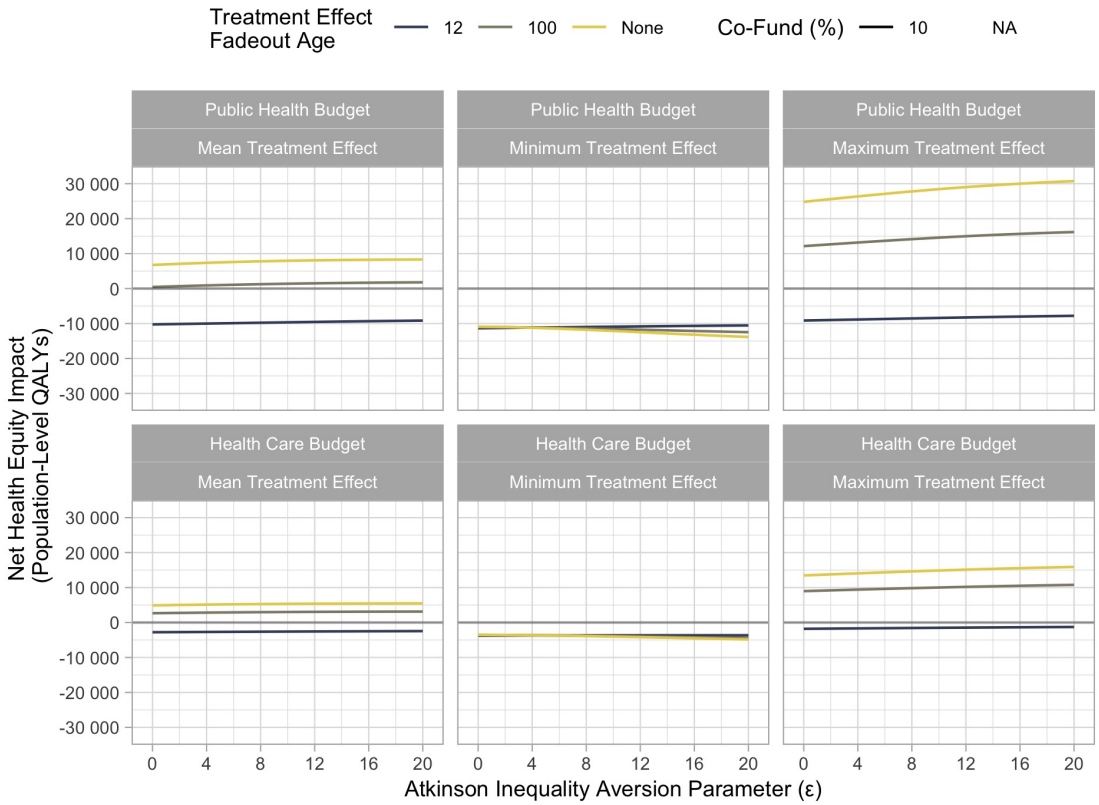
Atkinson Inequality Aversion Parameter

Figure C.19: Net equity-weighted QALYs by inequality aversion levels using BMI Z-Score treatment effects, FSMe quintiles and an equal distribution of health opportunity costs.



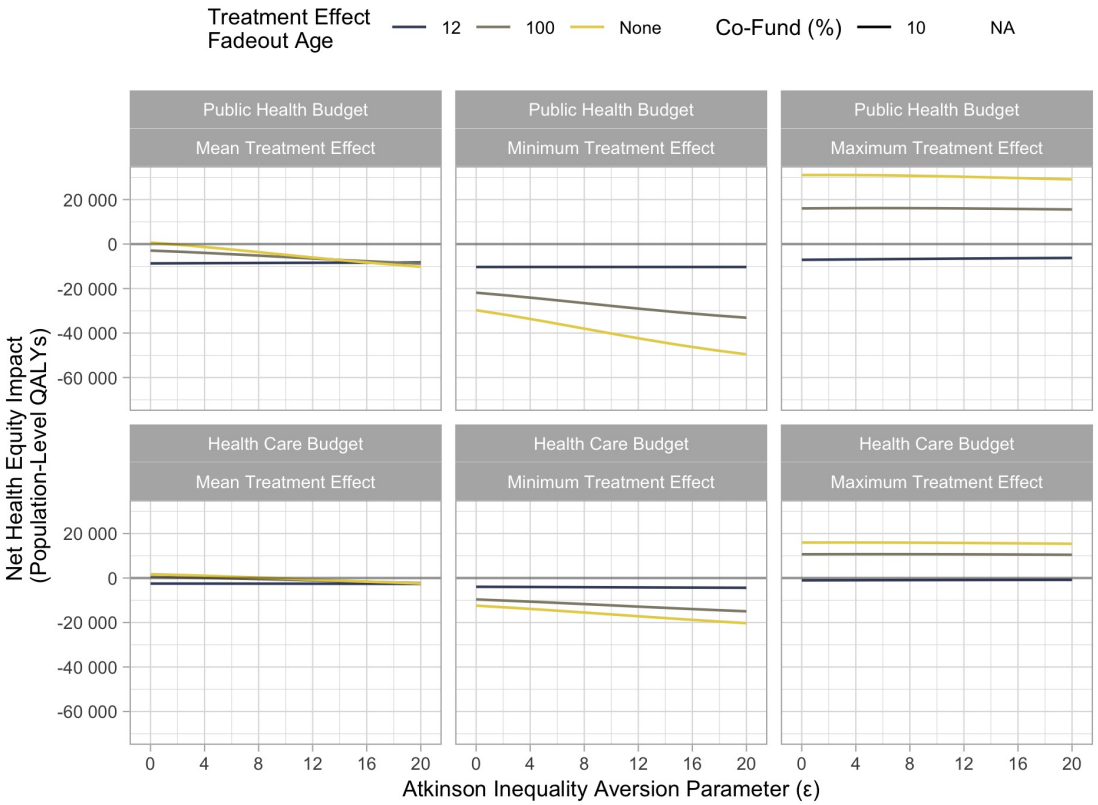
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.20: Net equity-weighted QALYs by inequality aversion levels using BMI Z-Score treatment effects, FSMe quintiles and a pro-poor distribution of health opportunity costs.



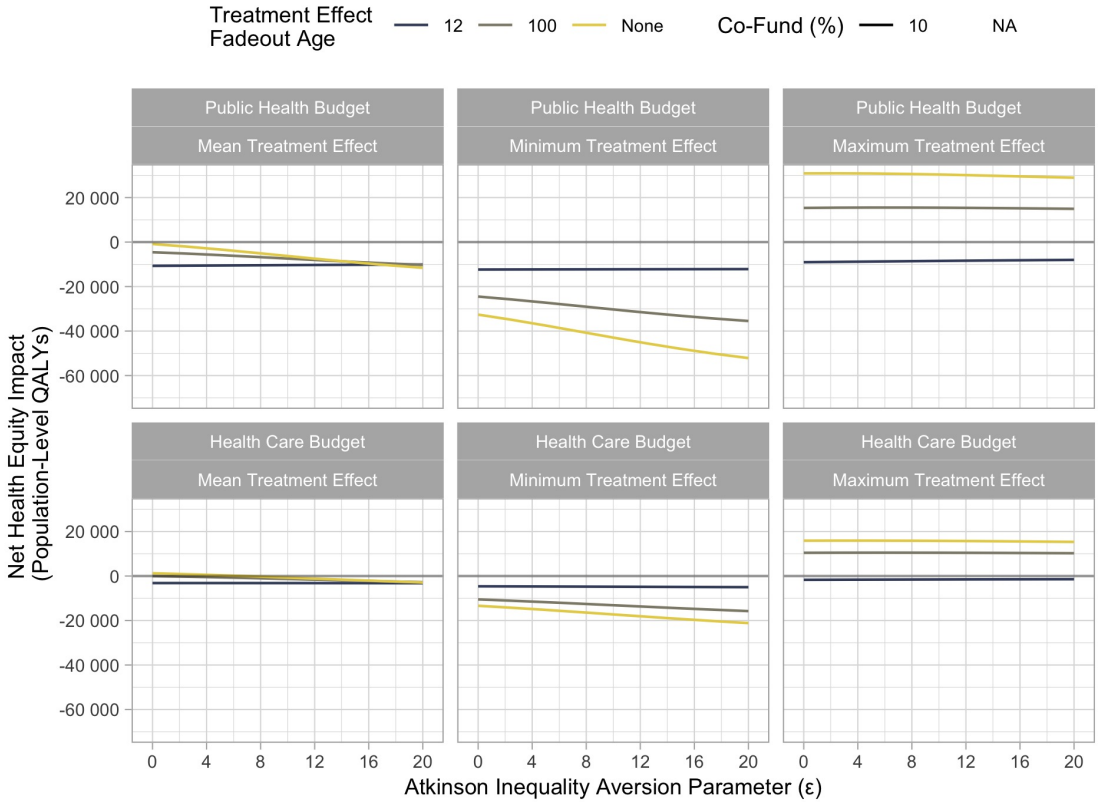
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.21: Net equity-weighted QALYs by inequality aversion levels using BMI category treatment effects, FSMe quintiles and an equal distribution of health opportunity costs.



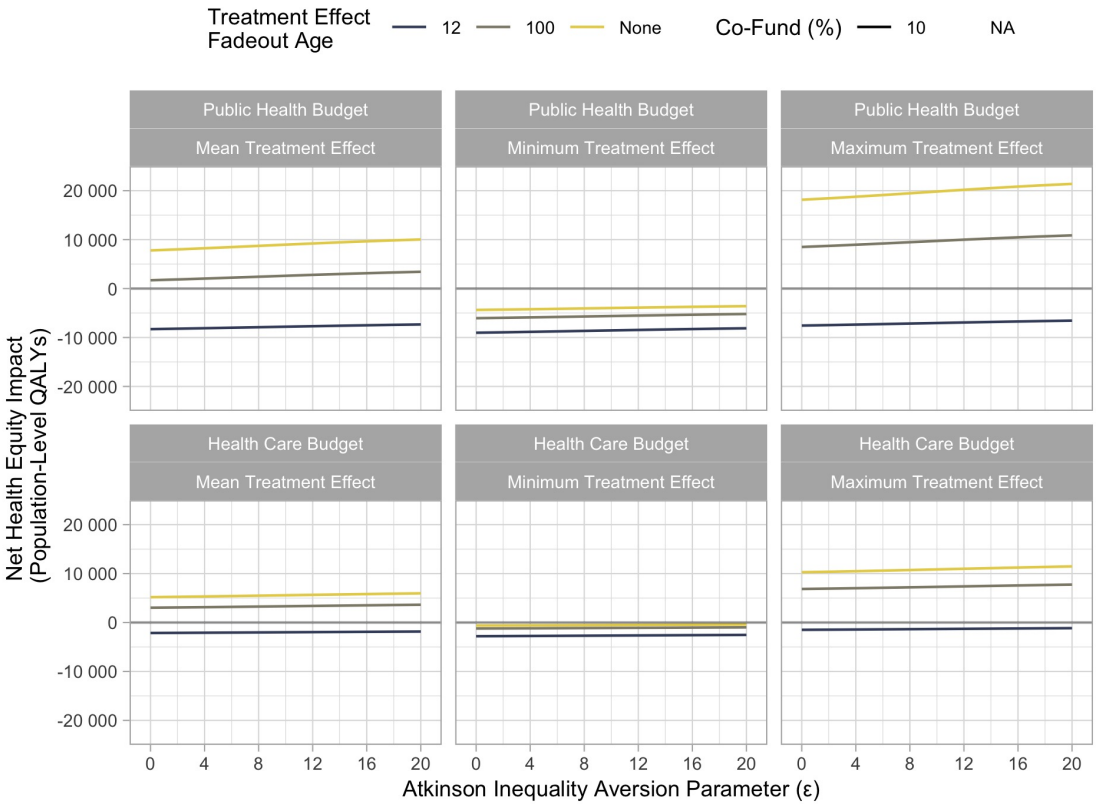
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.22: Net equity-weighted QALYs by inequality aversion levels using BMI category treatment effects, FSMe quintiles and a pro-poor distribution of health opportunity costs.



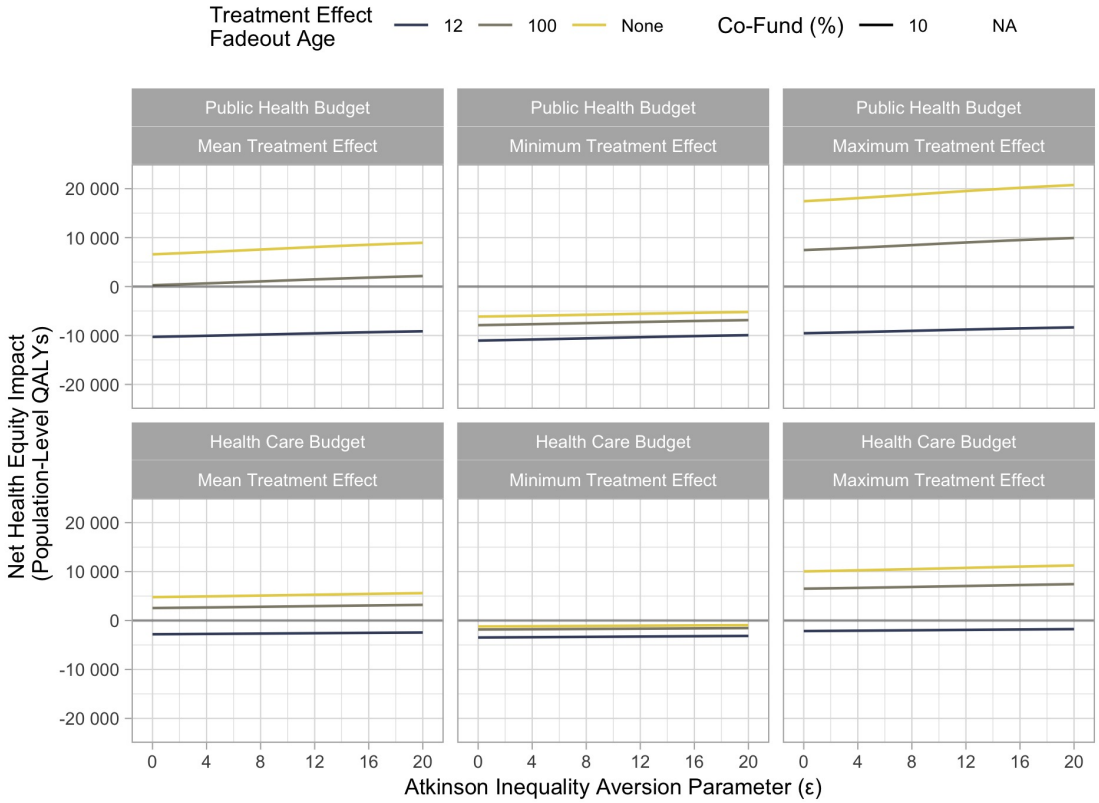
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.23: Net equity-weighted QALYs by inequality aversion levels using BMI Z-Score treatment effects, IDACI quintiles and an equal distribution of health opportunity costs.



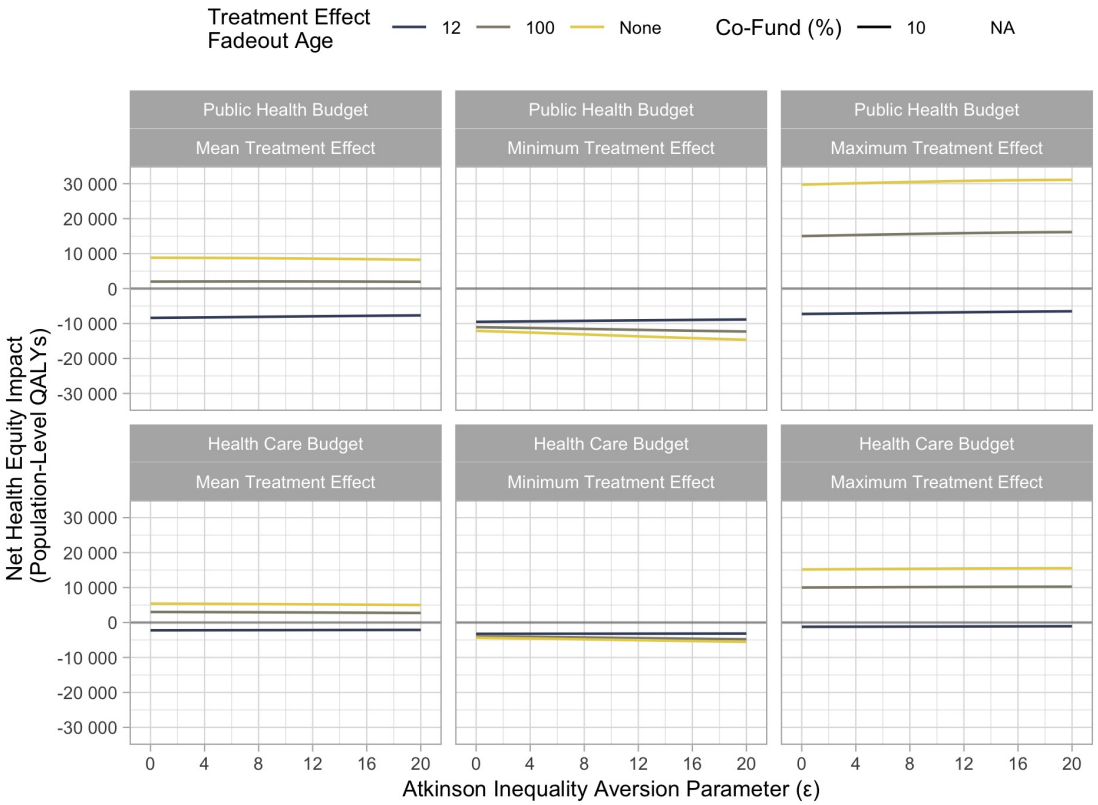
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.24: Net equity-weighted QALYs by inequality aversion levels using BMI Z-Score treatment effects, IDACI quintiles and a pro-poor distribution of health opportunity costs.



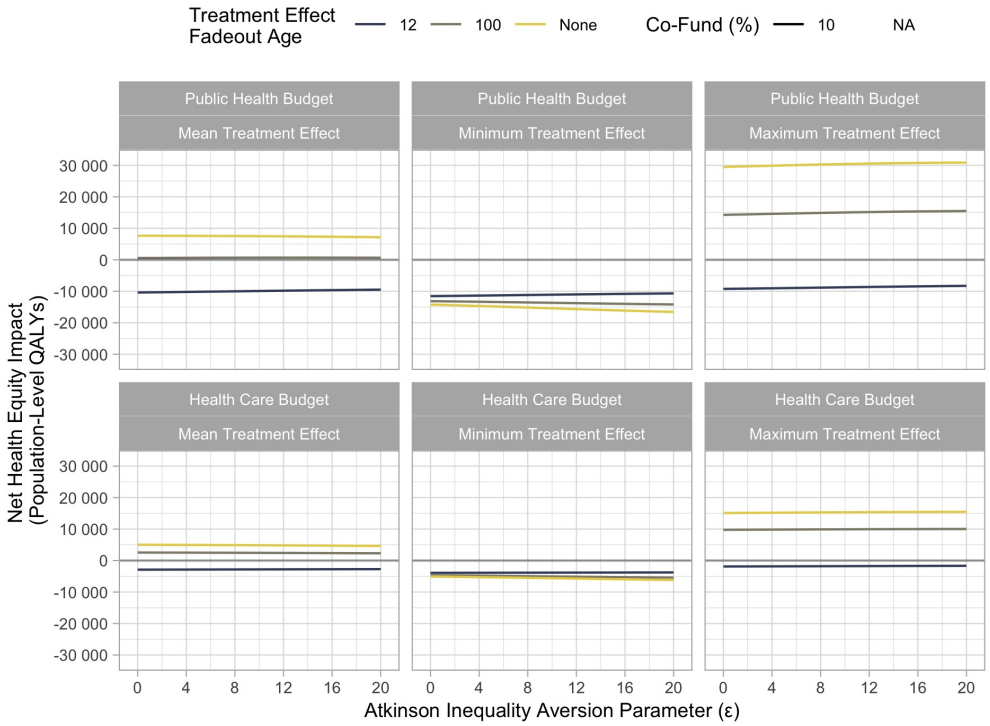
Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.25: Net equity-weighted QALYs by inequality aversion levels using BMI category treatment effects, IDACI quintiles and an equal distribution of health opportunity costs.



Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

Figure C.26: Net equity-weighted QALYs by inequality aversion levels using BMI category treatment effects, IDACI quintiles and a pro-poor distribution of health opportunity costs.



Notes: Public health (health care) budget assumes a QALY discount rate of 1.5% (3.5%) per year and a marginal productivity rate of £5 000 (£15 000) per QALY. Abbreviations: QALY, quality-adjusted life-year.

References

Albouy, Valerie and Laurent Lequien (2009). "Does compulsory education lower mortality?" In: *Journal of Health Economics* 28.1, pp. 155–168.

Allanson, Paul, Ulf-G. Gerdtham, and Dennis Petrie (2010). "Longitudinal analysis of income-related health inequality". In: *Journal of Health Economics* 29.1, pp. 78–86.

Almond, Douglas, Kenneth Y Chay, and David S Lee (2005). "The Costs of Low Birth Weight". In: *The Quarterly Journal of Economics* 120.3, pp. 1031–1083.

Almond, Douglas and Janet Currie (2011). "Human capital development before age five". In: *Handbook of Labor Economics*. Vol. 4 (Part B). Elsevier B.V. Chap. 15, pp. 1315–1486.

Almond, Douglas, Janet Currie, and Valentina Duque (2018). "Childhood circumstances and adult outcomes: Act II". In: *Journal of Economic Literature* 56.4, pp. 1360–1446.

Arceo, Eva, Rema Hanna, and Paulina Oliva (2016). "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City". In: *Economic Journal* 126.591, pp. 257–280.

Arneson, Richard J. (2002). "Luck Egalitarianism and Prioritarianism". In: *Ethics* 110.2, pp. 339–349.

Asada, Yukiko (2010). "On the Choice of Absolute or Relative Inequality Measures". In: *Milbank Quarterly* 88.4, pp. 616–622.

– (2013). "A Summary Measure of Health Inequalities: Incorporating Group and Individual Inequalities". In: *Inequalities in Health: Concepts, Measures, and*

Ethics. Ed. by Nir Eyal et al. New York: Oxford University Press. Chap. 3, pp. 37–51.

Asaria, Miqdad, Susan Griffin, and Richard Cookson (2016). "Distributional cost-effectiveness analysis: A tutorial". In: *Medical Decision Making* 36.1, pp. 8–19.

Atkinson, Anthony B (1980). "Horizontal equity and the distribution of the tax burden". In: *The Economics of Taxation*. Ed. by H J Aron and M J Boskins. Brookings, Washington D.C., pp. 3–18.

Attema, Arthur E., Werner B.F. Brouwer, and Karl Claxton (2018). "Discounting in Economic Evaluations". In: *PharmacoEconomics* 36.7, pp. 745–758.

Bailey, Drew et al. (2016). "Persistence and Fadeout in the Impacts of Child and Adolescent Interventions". In: *Journal of Research on Educational Effectiveness* 10.1, pp. 7–39.

Baird, Sarah et al. (2016). "Worms at Work: Long-Run Impacts of a Child Health Investment". In: *The Quarterly Journal of Economics* 131.4, pp. 1637–1680.

Bank of England (2020). *Inflation calculator*. URL: `https://www.bankofengland.co.uk/monetary-policy/inflation/inflation-calculator`.

Bann, David et al. (2018). "Socioeconomic inequalities in childhood and adolescent body-mass index, weight, and height from 1953 to 2015: an analysis of four longitudinal, observational, British birth cohort studies". In: *The Lancet Public Health* 3.4, e194–e203.

Bardin, Sarah, Liana Washburn, and Elizabeth Gearan (2020). "Disparities in the Healthfulness of School Food Environments and the Nutritional Quality of School Lunches". In: *Nutrients 2020, Vol. 12, Page 2375* 12.8, p. 2375.

Bates, Sarah et al. (2020). "A Systematic Review of Methods to Predict Weight Trajectories in Health Economic Models of Behavioral Weight-Management

Programs: The Potential Role of Psychosocial Factors". In: *Medical Decision Making* 2020.1, pp. 90–105.

Becker, Gary S (2009). *Human Capital*. 3rd ed. University of Chicago Press.

Beckfield, Jason and Nancy Krieger (2009). *Epi + demos + cracy: Linking political systems and priorities to the magnitude of health inequities-evidence, gaps, and a research agenda*.

Behrman, Jere R. and Mark R. Rosenzweig (2004). "Returns to birthweight". In: *Review of Economics and Statistics* 86.2, pp. 586–601.

Bidadanure, Juliana (2016). "Making sense of age-group justice: A time for relational equality?" In: *Politics, Philosophy and Economics* 15.3, pp. 234–260.

Black, Sandra E. and Paul J. Devereux (2011). "Recent developments in intergenerational mobility". In: *Handbook of Labor Economics*. Vol. 4 (part B). Elsevier B.V. Chap. 16, pp. 1487–1541.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes (2007). "From the Cradle to the Labour Market? The Effect of Birth Weight on Adult Outcomes". In: *The Quarteryly Journal of Economics* 122.1, pp. 409–439.

Bleakley, Hoyt (2007). "Disease and development: Evidence from hookworm eradication in the American South". In: *Quarterly Journal of Economics* 122.1, pp. 73–117.

Börsch-Supan, Axel et al. (2013). "Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)". In: *International Journal of Epidemiology* 42.4, pp. 992–1001.

Bound, John, Todd Stinebrickner, and Timothy Waidmann (2010). "Health, economic resources and the work decisions of older men". In: *Journal of Econometrics* 156.1, pp. 106–129.

Braveman, P. and S. Gruskin (2003). "Defining Equity in Health (Theory and Methods)". In: *Journal of Epidemiology & Community Health* 57.4, pp. 1–10.

Braveman, Paula, Nancy Krieger, and John Lynch (2000). "Health inequalities and social inequalities in health". In: *Bulletin of the World Health Organization* 78.2, pp. 232–233.

Brazier, John et al. (2022). "The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results". In: *Value in Health* 25.4.

Brekke, Kjell Arne and Snorre Kverndokk (2012). "Inadequate Bivariate Measures of Health Inequality: The Impact of Income Distribution". In: *Scandinavian Journal of Economics* 114.2, pp. 323–333.

Briggs, Andrew et al. (2003). "Missing....presumed at random: Cost-analysis of incomplete data". In: *Health Economics* 12.5, pp. 377–392.

Briggs, Andrew H., Karl. Claxton, and Mark J. Sculpher (2006). *Decision modelling for health economic evaluation*. Oxford, UK: Oxford University Press.

Brooks, Richard and Frank De Charro (1996). "EuroQol: the current state of play". In: *Health Policy* 37.1, pp. 53–72.

Broome, John (2004). *Weighing Lives*. New York: Oxford University Press.

Brown, Martin et al. (2013). "Managing overweight and obesity among children Report on Economic Modelling and Cost Consequence Analysis". In: *NICE Technical Report*.

Brunello, Giorgio et al. (2016). "The Causal Effect of Education on Health: What is the Role of Health Behaviours?" In: *Health Economics* 25.3, pp. 314–336.

Burgess, Simon, Carol Propper, and John Rigg (2004). "The Impact of Low-Income on Child Health: Evidence from a Birth Cohort Study". London.

Buuren, Stef van (2007). "Multiple imputation of discrete and continuous data by fully conditional specification". In: *Statistical Methods in Medical Research* 16.3, pp. 219–242.

Carlson, Kyle (2015). "Fear itself: The effects of distressing economic news on birth outcomes". In: *Journal of Health Economics* 41, pp. 117–132.

Carpenter, James and Michael Kenward (2013). *Multiple Imputation and Its Application*. 2nd ed. West Sussex, UK: John Wiley & Sons, Ltd.

Case, Anne and Angus Deaton (2005). "Broken down by work and sex: How our health declines". In: *Analyses in the Economics of Aging*. Ed. by David A Wise. University of Chicago Press. Chap. 6, pp. 185–212.

– (2021). *Deaths of Despair and the Future of Capitalism*. Princeton University Press.

Case, Anne, Angela Fertig, and Christina Paxson (2005). "The lasting impact of childhood health and circumstance". In: *Journal of Health Economics* 24.2, pp. 365–389.

Case, Anne, Christina Paxson, and Kimberly Bryan (2010). "Causes and Consequences of Early-Life Health". In: *Demography* 47.Supplement, pp. 65–85.

Cavelaars, Aejm et al. (1998). "Morbidity differences by occupational class among men in seven European countries: an application of the Erikson-Goldthorpe social class scheme". In: *International Journal of Epidemiology* 27, pp. 222–230.

Cawley, John and Christopher J. Ruhm (2011). *The Economics of Risky Health Behaviors*. Vol. 2. Elsevier B.V., pp. 95–199.

Chay, Kenneth Y. and Michael Greenstone (2003). "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession". In: *Quarterly Journal of Economics* 118.3, pp. 1121–1167.

Cheng, Adam et al. (2016). *Reporting guidelines for health care simulation research*.

Chetty, Raj et al. (2014). "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States". In: *The Quarterly Journal of Economics* 129.4, pp. 1553–1623.

Chiteji, Ngina (2010). "Time preference, noncognitive skills and well being across the life course: Do noncognitive skills encourage healthy behavior?" In: *American Economic Review: Papers & Proceedings* 100.2, pp. 200–204.

Clark, Damon and Heather Royer (2016). "The Effect of Education on Adult Mortality and Health : Evidence from Britain". In: *American Economic Review* 103.6, pp. 2087–2120.

Claxton, K et al. (2010). "Appropriate Perspectives for Health Care Decisions".

Claxton, Karl Philip, M J Sculpher, and Anthony J Culyer (2007). "Mark versus Luke? Appropriate methods for the evaluation of public health interventions". In: *CHE Research Paper* 31.

Cobb-Clark, Deborah A. and Stefanie Schurer (2012). "The stability of big-five personality traits". In: *Economics Letters* 115.1, pp. 11–15.

Cohen, Juliana F.W. et al. (2021). "Universal School Meals and Associations with Student Participation, Attendance, Academic Performance, Diet Quality, Food Security, and Body Mass Index: A Systematic Review". In: *Nutrients 2021, Vol. 13, Page 911* 13.3, p. 911.

Cole, T. J., J. V. Freeman, and M. A. Preece (1995). "Body mass index reference curves for the UK, 1990." In: *Archives of Disease in Childhood* 73.1, pp. 25–29.

Connelly, Roxanne and Lucinda Platt (2014). "Cohort Profile: UK Millennium Cohort Study (MCS)". In: *International Journal of Epidemiology* 43.6, pp. 1719–1725.

Conti, Gabriella et al. (2010). "The Education-Health Gradient". In: *American Economic Review: Papers & Proceedings* 100.May, pp. 234–238.

Contoyannis, Paul and Andrew M. Jones (2004). *Socio-economic status, health and lifestyle*. Vol. 23. 5, pp. 965–995.

Cookson, Richard (2013). “Can the NICE "end-of-life premium" be given a coherent ethical justification?” In: *Journal of Health Politics, Policy and Law* 38.6, pp. 1129–1148.

Cookson, Richard et al. (2017). “Using Cost-Effectiveness Analysis to Address Health Equity Concerns”. In: *Value in Health* 20.2, pp. 206–212.

Cookson, Richard et al. (2020). *Distributional Cost-Effectiveness Analysis: Quantifying Health Equity Impacts and Trade-Offs*. Oxford, UK: Oxford University Press.

Cookson, Richard et al. (2021a). “Equity-informative methods of health services research”. In: *Journal of Health Organization and Management* 35.6, pp. 665–681.

Cookson, Richard et al. (2021b). “Quality adjusted life years based on health and consumption: A summary wellbeing measure for cross-sectoral economic evaluation”. In: *Health Economics (United Kingdom)* 30.1.

Cookson, Richard et al. (2021c). “The inverse care law re-examined: a global perspective”. In: *The Lancet* 397.10276, pp. 828–838.

Coveney, Max et al. (2020). “Thank goodness for stickiness: Unravelling the evolution of income-related health inequalities before and after the Great Recession in Europe”. In: *Journal of Health Economics* 70, p. 102259.

Cowell, Alexander J. (2006). “The relationship between education and health behavior: Some empirical evidence”. In: *Health Economics* 15.2, pp. 125–146.

Cowell, F.A. (2000). “Measurement of inequality”. In: *Handbook of Income Distribution*. Elsevier, pp. 87–166.

Crawford, Claire and Ellen Greaves (2013). *A comparison of commonly used socio-economic indicators: their relationship to educational disadvantage and relevance to Teach First*. Tech. rep. London, UK: Institue for Fiscal Studies.

Culyer, Anthony J and Adam Wagstaff (1993). "Equity and Equality in Health and Health Care". In: *Journal of Health Economics* 12, pp. 431–457.

Cunha, Flavio and James J. Heckman (2007). "Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling". In: *Labour Economics* 14.6, pp. 870–893.

– (2008). "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation". In: *Journal of Human Resources* 43.4, pp. 738–782.

Cunha, Flavio, James J. Heckman, and Susanne Schennach (2010). "Estimating the Technology of Cognitive and Noncognitive Skill Formation". In: *Econometrica* 78.3, pp. 883–931.

Cunha, Flavio et al. (2006). "Interpreting the Evidence on Life Cycle Skill Formation". In: *Handbook of the Economics of Education*. Chap. 12, pp. 697–812.

Currie, Janet (2009). "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development". In: *Journal of Economic Literature* 47.1, pp. 87–122.

– (2011). "Inequality at Birth: Some Causes and Consequences". In: *American Economic Review: Papers & Proceedings* 101.3, pp. 1–22.

Currie, Janet and Rosemary Hyson (1999). "Is the impact of health shocks cushioned by socioeconomic status? The case of low birthweight". In: *American Economic Review* 89.2, pp. 245–250.

Currie, Janet and Enrico Moretti (2003). "Mother's education and the intergenerational transmission of human capital: Evidence from college openings". In: *Quarterly Journal of Economics* 118.4, pp. 1495–1532.

– (2007). "Biology as destiny? Short- and long-run determinants of intergenerational transmission of birth weight". In: *Journal of Labor Economics* 25.2, pp. 231–263.

Currie, Janet and Matthew Neidell (2005). "Air pollution and infant health: What can we learn from California's recent experience?" In: *Quarterly Journal of Economics* 120.3, pp. 1003–1030.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder (2009). "Air pollution and infant health: Lessons from New Jersey". In: *Journal of Health Economics* 28.3, pp. 688–703.

Currie, Janet and Mark Stabile (2003). "Socioeconomic status and child health: Why is the relationship stronger for older children?" In: *American Economic Review* 93.5, pp. 1813–1823.

– (2006). "Child mental health and human capital accumulation: The case of ADHD". In: *Journal of Health Economics* 25.6, pp. 1094–1118.

– (2009). "Mental Health in Childhood and Human Capital". In: *The Problems of Disadvantaged Youth: An Economic Perspective*. Ed. by Jonathan Gruber. Chicago: University of Chicago Press. Chap. 4, pp. 115–148.

Currie, Janet and Reed Walker (2011). "Traffic congestion and infant health: Evidence from E-Zpass". In: *American Economic Journal: Applied Economics* 3.1, pp. 65–90.

Currie, Janet et al. (2010). "Child health and young adult outcomes". In: *Journal of Human Resources* 45.3, pp. 517–548.

Cutler, David M. and Adriana Lleras-Muney (2010). "Understanding differences in health behaviors by education". In: *Journal of Health Economics* 29.1, pp. 1–28.

Cutler, David M, Adriana Lleras-Muney, and Tom Vogl (2008). "Socioeconomic Status and Health: Dimensions and Mechanisms". Cambridge, MA.

Dahlgren, Göran and Margaret Whitehead (1991). *Policies and strategies to promote social equity in health.*

Daniels, Norman (1988). *Am I My Parents' Keeper? An Essay on Justice Between the Young and the Old.* New York, NY: Oxford University Press.

Daniels, Norman (2008). *Just Health*. New York, NY: Cambridge University Press.

Deaton, Angus (1999). "Inequalities in Income and Inequalities in Health".

Deaton, Angus and Christina Paxson (1994). "Intertemporal Choice and Inequality". In: *Journal of Political Economy* 102.3, pp. 437–467.

DfE (2022). *Schools, Pupils, and their Characteristics*. URL: `https://www.gov.uk/government/collections/statistics-school-and-pupil-numbers`.

Diderichsen, Finn, Timothy Evans, and Margaret Whitehead (2009). "The Social Basis of Disparities in Health". In: *Challenging Inequities in Health: From Ethics to Action*.

Ding, Weili et al. (2009). "The impact of poor health on academic performance: New evidence using genetic markers". In: *Journal of Health Economics* 28.3, pp. 578–597.

DLUHC (2022). *DCLG Indices of Deprivation 2015*. URL: `https://opendatacommunities.org/def/concept/folders/themes/societal-wellbeing`.

Dodge, Kenneth A. (2020). "Annual Research Review: Universal and targeted strategies for assigning interventions to achieve population impact". In: *Journal of Child Psychology and Psychiatry* 61.3, pp. 255–267.

Doorslaer, Eddy van and Xander Koolman (2004). "Explaining the differences in income-related health inequalities across European countries". In: *Health Economics* 13.7, pp. 609–628.

Drummond, Michael F et al. (2015). *Methods for the Economic Evaluation of Health Care Programmes*. 4th ed. Oxford. Chap. 6.

Dunn, Caroline G. et al. (2020). "Feeding Low-Income Children during the Covid-19 Pandemic". In: *New England Journal of Medicine* 382.18.

Eikemo, Terje Andreas et al. (2008). "Welfare state regimes and differences in self-perceived health in Europe: A multilevel analysis". In: *Social Science & Medicine* 66.11, pp. 2281–2295.

Elliott, Jane and Peter Shepherd (2006). "Cohort Profile: 1970 British Birth Cohort (BCS70)". In: *International Journal of Epidemiology* 35.4, pp. 836–843.

Erreygers, Guido (2009). "Can a single indicator measure both attainment and shortfall inequality?" In: *Journal of Health Economics* 28.4, pp. 885–893.

Erreygers, Guido and Roselinde Kessels (2017). "Socioeconomic status and health: A new approach to the measurement of bivariate inequality". In: *International Journal of Environmental Research and Public Health* 14.7.

Erreygers, Guido and Tom Van Ourti (2011). "Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice". In: *Journal of Health Economics* 30.4, pp. 685–694.

European Union (2021). *Provision of school meals across the EU: An overview of rationales, evidence, facilitators and barriers*. Tech. rep. Luxembourg: European Commission.

Evans, Charlotte E.L. et al. (2010). "A comparison of British school meals and packed lunches from 1990 to 2007: meta-analysis by lunch type". In: *British Journal of Nutrition* 104.4, pp. 474–487.

Eyal, Nir (2018). "Inequality in Political Philosophy and in Epidemiology: A Remarriage". In: *Journal of Applied Philosophy* 35.1, pp. 149–167.

Fluharty, Meg E. et al. (2021). "Socioeconomic inequalities across life and premature mortality from 1971 to 2016: Findings from three British birth cohorts born in 1946, 1958 and 1970". In: *Journal of Epidemiology and Community Health* 75.2.

French, Eric (2005). "The effects of health, wealth, and wages on labour supply and retirement behaviour". In: *Review of Economic Studies* 72.2, pp. 395–427.

Friedman, Milton (1957). "The Permanent Income Hypothesis". In: *Theory of the Consumption Function*. Calcutta: Princeton University Press. Chap. 3, pp. 20–37.

Frijters, Paul, John P. Haisken-DeNew, and Michael A. Shields (2005). "The causal effect of income on health: Evidence from German reunification". In: *Journal of Health Economics* 24.5, pp. 997–1017.

Fuchs, Victor (1986). *The Health Economy*. 1st ed. Cambridge, MA: Harvard University Press.

Gakidou, E E, C J L Murray, and J Frenk (2001). *Measuring Preferences on Health System Performance Assessment*. Tech. rep. World Health Organization.

Galama, Titus and Arie Kapteyn (2011). "Grossman's missing health threshold". In: *Journal of Health Economics* 30.5, pp. 1044–1056.

Galama, Titus J. and Hans van Kippersluis (2018). "A theory of socio-economic disparities in health over the life cycle". In: *Economic Journal* 129.617, pp. 338–374.

Galama, Titus J, Adriana Lleras-Muney, and Hans van Kippersluis (2018). "The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence".

Galobardes, Bruna et al. (2006a). *Indicators of socioeconomic position (part 1)*.

– (2006b). *Indicators of socioeconomic position (part 2)*.

García-Gómez, Pilar (2011). "Institutions, health shocks and labour market outcomes across Europe". In: *Journal of Health Economics* 30.1, pp. 200–213.

García-Gómez, Pilar, Andrew M. Jones, and Nigel Rice (2010). "Health effects on labour market exits and entries". In: *Labour Economics* 17.1, pp. 62–76.

Garcia-Gomez, Pilar and Angel Lopez-Nicolas (2006). "Health shocks, employment and income in the Spanish labour market". In: *Health Economics* 15.9, pp. 997–1009.

García-Gómez, Pilar et al. (2013). "Long- term and spillover effects of health shocks on employment and income". In: *Journal of Human Resources* 48.4, pp. 873–909.

GBD 2015 Obesity Collaborators (2017). "Health Effects of Overweight and Obesity in 195 Countries over 25 Years". In: *New England Journal of Medicine* 377.1, pp. 13–27.

Goettler, Andrea, Anna Grosse, and Diana Sonntag (2017). "Productivity loss due to overweight and obesity: A systematic review of indirect costs". In: *BMJ Open* 7.10.

Golberstein, Ezra, Gilbert Gonzales, and Ellen Meara (2019). "How do economic downturns affect the mental health of children? Evidence from the National Health Interview Survey". In: *Health Economics* 28.8, pp. 955–970.

Griffin, Susan et al. (2012). *Univariate Assessment of Health Inequalities – Technical Report*. Tech. rep. February. York: Centre for Health Economics, University of York, pp. 1–52.

Grossman, Michael (1972). "On the Concept of Health Capital and the Demand for Health". In: *The Journal of Political Economy* 80.2, pp. 223–255.

Guh, Daphne P. et al. (2009). "The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis". In: *BMC Public Health* 9, pp. 1–20.

Hamilton, D., A. Dee, and I. J. Perry (2018). "The lifetime costs of overweight and obesity in childhood and adolescence: a systematic review". In: *Obesity Reviews* 19.4, pp. 452–463.

Hardt, Jochen and Michael Rutter (2004). *Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence.*

Harper, Sam et al. (2010). "Implicit Value Judgments in the Measurement of Health Inequalities". In: *Milbank Quarterly* 88.1, pp. 4–29.

Heathcote, Jonathan et al. (2018). “Two Views of Inequality over the Life Cycle”. In: *Journal of the European Economic Association* 3.2, pp. 765–775.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006). “The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior”. In: *Journal of Labor Economics* 24.3.

Heymsfield, Steven B. and Thomas A. Wadden (2017). “Mechanisms, Pathophysiology, and Management of Obesity”. In: *New England Journal of Medicine* 376.3.

Hirose, Iwao (2005). “Intertemporal Distributive Judgement”. In: *Ethical Theory and Moral Practice* 8.4, pp. 371–386.

Hoffmann, Rasmus, Hannes Kröger, and Eduwin Pakpahan (2018). “Pathways between socioeconomic status and health: Does health selection or social causation dominate in Europe?” In: *Advances in Life Course Research* 36.November 2017, pp. 23–36.

Holford, Angus (2015). “Take-up of Free School Meals: Price Effects and Peer Effects”. In: *Economica* 82.328, pp. 976–993.

Holford, Angus and Birgitta Rabe (2020). *Impact of the Universal Infant Free School Meal Policy*. Tech. rep. Wivenhoe Park, UK: Institute for Social and Economic Research.

– (2022). “Going universal. The impact of free school lunches on child body weight outcomes”. In: *Journal of Public Economics Plus* 3, p. 100016.

Hollingworth, W. et al. (2012). “Economic evaluation of lifestyle interventions to treat overweight or obesity in children”. In: *International Journal of Obesity* 36.4, pp. 559–566.

Holtug, Nils and Kasper Lippert-Rasmussen (2007). “An Introduction to Contemporary Egalitarianism”. In: *Egalitarianism: New Essays on the Nature*

and Value of Equality. Ed. by Nils Holtug and Kasper Lippert-Rasmussen. 1st ed. Oxford: Oxford University Press. Chap. 1, pp. 1–37.

Horton, Nicholas J. and Stuart R. Lipsitz (2001). "Multiple imputation in practice: Comparison of software packages for regression models with missing variables". In: *American Statistician* 55.3, pp. 244–254.

Hoynes, Hilary, Doug Miller, and David Simon (2015). "The Earned Income Tax Credit, Poverty, and Health". In: *American Economic Journal: Economic Policy* 7.1, pp. 172–211.

Johnson, Rucker C. and Robert F. Schoeni (2011). "The influence of early-life events on human capital, health status, and labor market outcomes over the life course". In: *B.E. Journal of Economic Analysis and Policy* 11.3.

Jones, Andrew M. and Angel López Nicolás (2004). "Measurement and explanation of socioeconomic inequality in health with longitudinal data". In: *Health Economics* 13.10, pp. 1015–1030.

Jones, Andrew M., John E. Roemer, and Pedro Rosa Dias (2014). "Equalising opportunities in health through educational policy". In: *Social Choice and Welfare* 43.3, pp. 521–545.

Jusot, Florence and Sandy Tubeuf (2019). "Equality of Opportunity in Health and Healthcare". In: *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Kakwani, N C (1984). "On the measurement of tax progressivity and redistributive effect of taxes with applications to horizontal and vertical equity". In: *Advances in Econometrics* 3, pp. 149–168.

Kakwani, Nanak (1977). "Measurement of Tax Progressivity: An International Comparison". In: *The Economic Journal* 87.345, pp. 71–80.

Kaplan, Robert M. and Ron D. Hays (2022). *Health-Related Quality of Life Measurement in Public Health*.

Kenkel, Donald (1991). "Health Behavior, Health Knowledge, and Schooling". In: *Journal of Political Economy* 99.2, pp. 287–305.

Kenney, Erica L. et al. (2022). "Costs, Reach, and Benefits of COVID-19 Pandemic Electronic Benefit Transfer and Grab-and-Go School Meals for Ensuring Youths' Access to Food During School Closures". In: *JAMA Network Open* 5.8, e2229514.

Kessler, Ronald C. et al. (1995). "Social consequences of psychiatric disorders, I: Educational attainment". In: *American Journal of Psychiatry* 152.7, pp. 1026–1032.

Kinge, Jonas Minet et al. (2019). "Association of Household Income With Life Expectancy and Cause-Specific Mortality in Norway, 2005-2015". In: *JAMA* 321.19, pp. 1916–1925.

Kippersluis, Hans van, Owen O'Donnell, and Eddy van Doorslaer (2011). "Long-Run Returns to Education". In: *Journal of Human Resources* 46.4, pp. 695–721.

Kippersluis, Hans van et al. (2009). "Health and income across the life cycle and generations in Europe". In: *Journal of Health Economics* 28.4, pp. 818–830.

Kjellsson, Gustav, Ulf G. Gerdtham, and Dennis Petrie (2015). "Lies, Damned Lies, and Health Inequality Measurements: Understanding the Value Judgments". In: *Epidemiology* 26.5, pp. 673–680.

Knittel, Christopher R, Douglas L Miller, and Nicholas J Sanders (2016). "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health". In: *The Review of Economics and Statistics* 98.2, pp. 350–366.

Kolotkin, R. L. and J. R. Andersen (2017). "A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life". In: *Clinical obesity* 7.5.

Kotler, Lisa A. et al. (2001). "Longitudinal relationships between childhood, adolescent, and adult eating disorders". In: *Journal of the American Academy of Child and Adolescent Psychiatry* 40.12.

Kreimeier, Simone and Wolfgang Greiner (2019). "EQ-5D-Y as a Health-Related Quality of Life Instrument for Children and Adolescents: The Instrument's Characteristics, Development, Current Use, and Challenges of Developing Its Value Set". In: *Value in Health* 22.1, pp. 31–37.

Kreiner, Claus Thustrup, Torben Heien Nielsen, and Benjamin Ly Serena (2018). "Role of income mobility for the measurement of inequality in life expectancy". In: *Proceedings of the National Academy of Sciences* 115.46, pp. 11754–11759.

Kuehnle, Daniel (2014). "The causal effect of family income on child health in the UK". In: *Journal of Health Economics* 36.1, pp. 137–150.

Kuh, D et al. (2003). "Life Course Epidemiology". In: *Journal of Epidemiology & Community Health* 57.10, pp. 778–783. URL: http://link.springer.com/10.1007/978-0-387-09834-0.

Lambert, Peter and Buhong Zheng (2011). "On the consistent measurement of attainment and shortfall inequality". In: *Journal of Health Economics* 30.1, pp. 214–219.

Lauderdale, Diane S. (2001). "Education and survival: Birth cohort, period, and age effects". In: *Demography* 38.4, pp. 551–561.

Leonardi-Bee, Jo, Mirriam Lisa Jere, and John Britton (2011). "Exposure to parental and sibling smoking and the risk of smoking uptake in childhood and adolescence: a systematic review and meta-analysis". In: *Thorax* 66.10, pp. 847–855.

Levy, D. T. et al. (2011). "Simulation models of obesity: a review of the literature and implications for research and policy". In: *Obesity Reviews* 12.5, pp. 378–394.

Lhachimi, Stefan K. et al. (2012). "DYNAMO-HIA–A Dynamic Modeling Tool for Generic Health Impact Assessments". In: *PLOS ONE* 7.5, e33317.

Lien, Diana S. and William N. Evans (2005). "Estimating the impact of large cigarette tax hikes: The case of maternal smoking and infant birth weight". In: *Journal of Human Resources* 40.2, pp. 373–392.

Lin, Ming Jen and Jin Tan Liu (2009). "Do lower birth weight babies have lower grades? Twin fixed effect and instrumental variable method evidence from Taiwan". In: *Social Science and Medicine* 68.10, pp. 1780–1787.

Lindeboom, Maarten, Ana Llena-Nozal, and Bas van der Klaauw (2009). "Parental education and child health: Evidence from a schooling reform". In: *Journal of Health Economics* 28.1, pp. 109–131.

Lippert-Rasmussen, Kasper (2008). "Measuring the disvalue of inequality over time". In: *Theoria* 69.1-2, pp. 32–45.

Lleras-Muney, Adriana (2005). "The relationship between education and adult mortality in the United States". In: *Review of Economic Studies* 72.1, pp. 189–221.

Lobstein, Tim et al. (2015). "Child and adolescent obesity: part of a bigger picture". In: *The Lancet* 385.9986, pp. 2510–2520.

Lomas, James, Stephen Martin, and Karl Claxton (2019). "Estimating the Marginal Productivity of the English National Health Service From 2003 to 2012". In: *Value in Health* 22.9, pp. 995–1002.

Lorenz, M. O. (1905). "Methods of measuring the concentration of wealth". In: *Publications of the American Statistical Association* 9.70.

Love-Koh, James et al. (2015). "The Social Distribution of Health: Estimating Quality-Adjusted Life Expectancy in England". In: *Value in Health* 18.5, pp. 655–662.

Love-Koh, James et al. (2020). "Estimating Social Variation in the Health Effects of Changes in Health Care Expenditure". In: *Medical Decision Making* 40.2, pp. 170–182.

Lübker, Christopher and Fabrice Murtin (2022). *Educational Inequalities in Longevity Among OECD Countries Around 2016*. Tech. rep. Paris, France: Organisation for Economic Co-operation and Development.

Lundborg, Petter, Dan-Olof Rooth, and Jesper Alex-Petersen (2022). "Long-Term Effects of Childhood Nutrition: Evidence from a School Lunch Reform". In: *The Review of Economic Studies* 89.2, pp. 876–908.

Mackenbach, Johan et al. (2015). *Measuring Educational Inequalities in Mortality Statistics*. Tech. rep. Paris: Organisation for Economic Co-operation and Development.

Mackenbach, Johan P. (2011). "Can we reduce health inequalities? An analysis of the English strategy (1997-2010)". In: *Journal of Epidemiology and Community Health* 65.7.

– (2012). "The persistence of health inequalities in modern welfare states: The explanation of a paradox". In: *Social Science & Medicine* 75.4, pp. 761–769.

Mackenbach, Johan P and Anton E Kunst (1997). "Measuring the magnitude of socio-economic inequalities in health: An overview of available measures illustrated with two examples from Europe". In: *Social Science & Medicine* 44.6, pp. 757–771.

Mackenbach, Johan P et al. (1997). "Socioeconomic inequalities in morbidity and mortality in western Europe". In: *The Lancet* 349.9066, pp. 1655–1659.

Mackenbach, Johan P. et al. (2008). "Socioeconomic inequalities in health in 22 European countries". In: *New England Journal of Medicine* 358.23, pp. 2468–2481.

Mackenbach, Johan P. et al. (2018). "Trends in health inequalities in 27 European countries". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.25, pp. 6440–6445.

MacMahon, S. et al. (2009). "Body-mass index and cause-specific mortality in 900 000 adults: Collaborative analyses of 57 prospective studies". In: *The Lancet* 373.9669, pp. 1083–1096.

Maheswaran, Hendramoorthy et al. (2013). "Estimating EQ-5D utility values for major health behavioural risk factors in England". In: *J Epidemiol Community Health* 67.2, pp. 172–180.

Marchi, Margaret and Patricia Cohen (1990). "Early Childhood Eating Behaviors and Adolescent Eating Disorders". In: *Journal of the American Academy of Child and Adolescent Psychiatry* 29.1.

Marmot, Michael and Richard Wilkinson, eds. (2005). *Social Determinants of Health*. 2nd ed. Oxford: Oxford University Press.

Marmot, Michael et al. (2008). "Closing the gap in a generation: health equity through action on the social determinants of health". In: *The Lancet* 372.9650, pp. 1661–1669.

Marmot, Michael et al. (2020). *Health Equity in England: The Marmot Review 10 Years On*. Tech. rep. Institute of Health Equity.

Marsh, Kevin et al. (2016). "Multiple Criteria Decision Analysis for Health Care Decision Making - Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force". In: *Value in Health* 19.2.

Martin, Stephen, James Lomas, and Karl Claxton (2020). "Is an ounce of prevention worth a pound of cure? A cross-sectional study of the impact of English public health grant on mortality and morbidity". In: *BMJ Open* 10.10, e036411.

McCrary, Justin and Heather Royer (2011). "The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth". In: *American Economic Review* 101.1, pp. 158–195.

McDowell, I (2006). *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd ed. Oxford: Oxford University Press, pp. 649–666.

McKerlie, Dennis (1989). "Equality and Time". In: *Ethics* 99.3, pp. 475–491.

McLeod, Jane D and Karen Kaiser (2004). "Childhood Emotional and Behavioral Problems and Educational Attainment". In: *American Sociological Review* 69.5, pp. 636–658.

McMahan, Jeff (2002). "Identity". In: *The Ethics of Killing: Problems at the Margins of Life*. New York, NY: Oxford University Press. Chap. 1, pp. 3–95.

McPherson, Klim and Martin Brown (2007). *Tackling Obesities: Future Choices - Modelling Future Trends in Obesity and the Impact on Health 2nd Edition*. Tech. rep. Foresight Programme Government Office for Science.

Mead, E. et al. (2016). "Predicting future weight status from measurements made in early childhood: A novel longitudinal approach applied to Millennium Cohort Study data". In: *Nutrition and Diabetes* 6.3, e200–e200.

Mead, Emma et al. (2017). "Diet, physical activity and behavioural interventions for the treatment of overweight or obese children from the age of 6 to 11 years". In: *Cochrane Database of Systematic Reviews* 2017.6.

Miech, Richard A. et al. (1999). "Low socioeconomic status and mental disorders: A longitudinal study of selection and causation during young adulthood". In: *American Journal of Sociology* 104.4, pp. 1096–1131.

Mindell, Jennifer et al. (2012). "Cohort Profile: The Health Survey for England". In: *International Journal of Epidemiology* 41.6, pp. 1585–1593.

Mocan, Naci and Duha T. Altindag (2014). "Education, cognition, health knowledge, and health behavior". In: *European Journal of Health Economics* 15.3, pp. 265–279.

Modigliani, Franco (1966). "The life cycle hypothesis of saving, the demand for wealth and the supply of capital". In: *Social Research* 33.2, pp. 160–217.

Moreno-Betancur, Margarita et al. (2015). "Relative Index of Inequality and Slope Index of Inequality". In: *Epidemiology* 26.4, pp. 518–527.

Mostafa, T et al. (2020). "Improving the plausibility of the missing at random assumption in the 1958 British birth cohort: A pragmatic data driven approach". London.

Mostafa, Tarek and Richard D. Wiggins (2015). "The impact of attrition and non-response in birth cohort studies: A need to incorporate missingness strategies". In: *Longitudinal and Life Course Studies* 6.2, pp. 131–146.

Mostafa, Tarek et al. (2021). "Missing at random assumption made more plausible: evidence from the 1958 British birth cohort". In: *Journal of Clinical Epidemiology* 136, pp. 44–54.

Muntaner, Carles et al. (2011). "Politics, welfare regimes, and population health: controversies and evidence". In: *Sociology of Health & Illness* 33.6, pp. 946–964.

Murray, C J L, E E Gakidou, and J Frenk (2000). "Reply to Braveman et al". In: *Bulletin of the World Health Organization* 78.2, pp. 234–235.

Murray, C. J.L., E. E. Gakidou, and J. Frenk (1999). "Health inequalities and social group differences: What should we measure?" In: *Bulletin of the World Health Organization* 77.7, pp. 537–543.

Murtin, Fabrice et al. (2022). "Educational inequalities in longevity in 18 OECD countries". In: *Journal of Demographic Economics* 88.1.

NCMP (2020). *National Child Measurement Programme, England 2019/20 School Year*. Tech. rep. NHS Digital.

Neeland, Ian J., Paul Poirier, and Jean Pierre Després (2018). "Cardiovascular and Metabolic Heterogeneity of Obesity: Clinical Challenges and Implications for Management". In: *Circulation* 137.13.

Neidell, Matthew J. (2004). "Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma". In: *Journal of Health Economics* 23.6, pp. 1209–1236.

NHS (2019). *The NHS Long Term Plan*. Tech. rep. National Health Service.

NICE (2013). *Guide to the methods of technology appraisal 2013: Process and methods*. Tech. rep. National Institute for Health and Care Excellence.

– (2020). *CHTE Methods Review: Discounting*. Tech. rep. National Institute for Health and Care Excellence.

O'Brien, Katie M. et al. (2017). "Predictors and long-term health outcomes of eating disorders". In: *PLoS ONE* 12.7.

O'Donnell, Owen, Eddy Van Doorslaer, and Tom Van Ourti (2015). *Health and inequality*. 1st ed. Vol. 2. Elsevier B.V., pp. 1419–1533.

OECD (2013). *What are Equivalence Scales? OECD Project on Income Distribution and Poverty*. Tech. rep. Paris: Organisation for Economic Co-operation and Development.

– (2015). *In It Together: Why Less Inequality Benefits All*. Tech. rep. Paris: Organisation for Economic Co-operation and Development.

– (2019). *The Heavy Burden of Obesity: The Economics of Prevention*. Tech. rep. Paris, France: Organisation for Economic Co-operation and Development.

OHID (2022). *Office for Health Improvement and Disparities*. URL: `https://www.gov.uk/government/organisations/office-for-health-improvement-and-disparities`.

Okunogbe, A et al. (2022). "Economic impacts of overweight and obesity: current and future estimates for 161 countries". In: *BMJ Global Health* in press.

Olsen, Jan Abel (2017). "Beyond cost-effectiveness: priority setting". In: *Principles in Health Economics and Policy*. 2nd ed. New York: Oxford University Press, pp. 249–260.

O'Neill, P. (1999). "Clinical evidence: Acute otitis media". In: *BMJ* 319.7213, pp. 833–835.

ONS (2019). *Average household income, UK: Financial year ending 2018*.

– (2020a). *Population by Index of Multiple Deprivation (IMD), England, 2001 to 2019 – Office for National Statistics*. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/12386populationbyindexofmultipledeprivationimdengland2001to2019.

– (2020b). *SOC 2000 - Office for National Statistics*. URL: https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/socarchive.

– (2021). *Health state life expectancies, UK: 2017 to 2019*. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/healthstatelifeexpectanciesuk/2017to2019.

Owen, Lesley et al. (2018). "The cost-effectiveness of public health interventions examined by NICE from 2011 to 2016". In: *Journal of Public Health* 40.3, pp. 557–566.

Pamuk, Elsie R. (1988). "Social-class inequality in infant mortality in England and Wales from 1921 to 1980". In: *European Journal of Population* 4.1, pp. 1–21.

Parfit, Derek (1986). *Reasons and Persons*. New York: Oxford University Press.

Parnham, J. C. et al. (2020). "Half of children entitled to free school meals did not have access to the scheme during COVID-19 lockdown in the UK". In: *Public Health* 187.

Pearce, Anna et al. (2019). "Pathways to inequalities in child health". In: *Archives of Disease in Childhood* 104.10, pp. 998–1003.

Pearl, Judea (2000). *Causality: Models, Reasoning and Inference*. Vol. 29. Cambridge, MA: MIT Press.

– (2001). "Direct and Indirect Effects". In: *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., pp. 411–420.

– (2009). "Causal inference in statistics: An overview". In: *Statistics Surveys* 3, pp. 96–146.

Plotnick, Robert (1981). "A Measure of Horizontal Inequity". In: *The Review of Economics and Statistics* 63.2.

Pollack, Craig Evan et al. (2007). "Should Health Studies Measure Wealth?: A Systematic Review". In: *American Journal of Preventive Medicine* 33.3, pp. 250–264.

Popkin, Barry M. et al. (2020). "Individuals with obesity and COVID-19: A global perspective on the epidemiology and biological relationships". In: *Obesity Reviews* 21.11.

Power, Chris and Jane Elliott (2006). "Cohort profile: 1958 British birth cohort (National Child Development Study)". In: *International Journal of Epidemiology* 35.1, pp. 34–41.

Rawls, John (1972). *A Theory of Justice*. Oxford: Oxford University Press.

Regidor, Enrique (2004). "Measures of health inequalities: part 2". In: *Journal of Epidemiology and Community Health* 58.11, pp. 900–3.

Reilly, J. J. and J. Kelly (2010). "Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review". In: *International Journal of Obesity 2011 35:7* 35.7, pp. 891–898.

Reiss, Franziska (2013). "Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review". In: *Social Science and Medicine* 90, pp. 24–31.

Renard, Françoise et al. (2019). "Monitoring health inequalities when the socio-economic composition changes: are the slope and relative indices of inequality appropriate? Results of a simulation study". In: *BMC Public Health*.

Robson, Matthew, Tim Doran, and Richard Cookson (2019). *Estimating and decomposing conditional average treatment effects: the smoking ban in England*. Tech. rep. York: Health, Econometrics and Data Group (HEDG) Working Papers.

Robson, Matthew et al. (2017). "Eliciting the Level of Health Inequality Aversion in England". In: *Health Economics* 26.10, pp. 1328–1334.

Roemer (1998). *Theories of Distributive Justice*. Cambridge, MA: Harvard University Press.

Roemer, John E. and Alain Trannoy (2015). *Equality of Opportunity*. 1st ed. Vol. 2. Elsevier B.V., pp. 217–300.

Rosa Dias, Pedro (2009). "Inequality of Opportunity in Health: Evidence from a UK Cohort Study". In: *Health Economics* 18, pp. 1057–1074.

– (2010). "Modelling Opportunity in Health Under Partial Observability of Circumstances". In: *Health Economics* 19, pp. 252–264.

Roux, C. W. le et al. (2018). "Obesity and healthcare resource utilization: results from Clinical Practice Research Database (CPRD)". In: *Obesity Science & Practice* 4.5, pp. 409–416.

Royer, Heather (2009). "Separated at girth: US twin estimates of the effects of birth weight". In: *American Economic Journal: Applied Economics* 1.1, pp. 49–85.

Salm, Martin and Daniel Schunk (2012). "The Relationship between Child Health, Developmental Gaps, and Parental Education: Evidence from Administrative Data". In: *Journal of the European Economic Association* 10.6, pp. 1425–1449.

Sanchis-Gomar, Fabian et al. (2020). "Obesity and Outcomes in COVID-19: When an Epidemic and Pandemic Collide". In: *Mayo Clinic Proceedings* 95.7.

Scanlon, T. M. (2018). *Why Does Inequality Matter?* Oxford: Oxford Univeristy Press.

Sculpher, Mark J et al. (2014). "Research to inform resource allocation in health and social care. Results of a scoping study".

Sellen, Peter et al. (2018). *Evaluation of Universal Infant Free School Meals*. Tech. rep. Education Policy Institute.

Sen, Amartya (1979). "Equality of what?" In: *The Tanner Lecture on Human Values*. Stanford University, pp. 197–220.

– (1995). "Welfare Economics and Inequality". In: *Inequality Reexamined*. Oxford: Oxford University Press. Chap. 6, pp. 89–101.

Silverwood, R et al. (2020). *Handling missing data in the National Child Development Study*. Tech. rep. London: University College London.

Simmonds, M. et al. (2016). "Predicting adult obesity from childhood obesity: A systematic review and meta-analysis". In: *Obesity Reviews* 17.2, pp. 95–107.

Simmonds, Mark et al. (2015). "The use of measures of obesity in childhood for predicting obesity and the development of obesity-related diseases in adulthood: a systematic review and meta-analysis". In: *Health Technology Assessment* 19.43, pp. 1–336.

Skarda, Ieva, Miqdad Asaria, and Richard Cookson (2022). "Evaluating childhood policy impacts on lifetime health, wellbeing and inequality: Lifecourse distributional economic evaluation". In: *Social Science and Medicine* 302.

Smith, James P. (2009). "The impact of Childhood Health on Adult Labour market Outcomes". In: *The Review of Economics and Statistics* 91.3, pp. 478–489.

Spence, Suzanne et al. (2013). "The impact of food and nutrient-based standards on primary school children's lunch and total dietary intake: A natural experimental evaluation of government policy in England". In: *PLoS ONE* 8.10.

Spence, Suzanne et al. (2014). "A repeat cross-sectional study examining the equitable impact of nutritional standards for school lunches in England in 2008 on the diets of 4-7y olds across the socio-economic spectrum". In: *International Journal of Behavioral Nutrition and Physical Activity* 11.1.

Spence, Suzanne et al. (2019). "A natural experimental evaluation of the effect of universal infant free school meals on key stage 1 pupil's dietary intake in northeast England: a pilot study". In: *The Lancet* 394, S87.

Spencer, Elizabeth A et al. (2002). "Validity of self-reported height and weight in 4808 EPIC–Oxford participants". In: *Public Health Nutrition* 5.4, pp. 561–565.

Steptoe, Andrew et al. (2013). "Cohort Profile: The English Longitudinal Study of Ageing". In: *International Journal of Epidemiology* 42.6, pp. 1640–1648.

Sterne, Jonathan A.C. et al. (2009). *Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls.*

Taylor, Chris (2018). "The Reliability of Free School Meal Eligibility as a Measure of Socio-Economic Disadvantage: Evidence from the Millennium Cohort Study in Wales". In: *British Journal of Educational Studies* 66.1.

Temkin, Larry (1993). *Inequality.* New York: Oxford University Press.

Thokala, Praveen et al. (2016). "Multiple criteria decision analysis for health care decision making - An introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force". In: *Value in Health* 19.1.

Trasande, Leonardo (2017). "How Much Should We Invest In Preventing Childhood Obesity?" In: *Health Affairs* 29.3, pp. 372–378.

Tremmel, Maximilian et al. (2017). "Economic burden of obesity: A systematic literature review". In: *International Journal of Environmental Research and Public Health* 14.4.

UN (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. Tech. rep. United Nations.

– (2021). *State of School Feeding Worldwide 2020*. Tech. rep. Rome, Italy: World Food Programme.

Van Ourti, Tom, Eddy van Doorslaer, and Xander Koolman (2009). "The effect of income growth and inequality on health inequality: Theory and empirical evidence from the European Panel". In: *Journal of Health Economics* 28.3, pp. 525–539.

Vandenbroucke, Jan P. et al. (2007). *Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration.*

Verguet, Stéphane, Jane J. Kim, and Dean T. Jamison (2016). "Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial". In: *PharmacoEconomics* 34.9, pp. 913–923.

Wachtler, Benjamin et al. (2020). "Socioeconomic inequalities and COVID-19 – A review of the current international literature". In: *Journal of health monitoring* 5.Suppl 7, pp. 3–17.

Wadsworth, Michael et al. (2006). "Cohort profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development)". In: *International Journal of Epidemiology* 35.1.

Wagstaff, Adam, Pierella Paci, and Eddy van Doorslaer (1991). "On the Measurement of Inequalities in Health". In: *Soc. Sci. Med.* 33.5, pp. 545–557.

Walker, Simon et al. (2019). "Striving for a Societal Perspective: A Framework for Economic Evaluations When Costs and Effects Fall on Multiple Sectors and Decision Makers". In: *Applied Health Economics and Health Policy* 17.5.

Weatherly, Helen et al. (2009). *Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations.*

Webbink, Dinand et al. (2012). "The Effect of Childhood Conduct Disorder on Human Capital". In: *Health Economics* 21.8, pp. 928–945.

White, Ian R., Patrick Royston, and Angela M. Wood (2011). "Multiple imputation using chained equations: Issues and guidance for practice". In: *Statistics in Medicine* 30.4, pp. 377–399.

Whitehead, Margaret (1991). "The Concepts And Principles Of Equity And Health". In: *Health Promotion International* 6.3.

Whitehead, Margaret, Alex Scott-Samuel, and Göran Dahlgren (1998). "Setting targets to address inequalities in health". In: *Lancet* 351.9111.

WHO (1981). *Global Strategy for Health for All by the Year 2000.* Tech. rep. Geneva: World Health Organisation.

– (2022). *Health Inequality Monitor.* URL: `https://www.who.int/data/inequality-monitor`.

Wilkinson, R G (1992). "Income Distribution and Life Expectancy". In: *BMJ* 304.6820, pp. 165–168.

Withrow, D. and D. A. Alter (2011). "The economic burden of obesity worldwide: A systematic review of the direct costs of obesity". In: *Obesity Reviews* 12.2.

Woods, Beth et al. (2016). "Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research". In: *Value in Health* 19.8.

Yang, Jun, Jiahui Hu, and Chunyan Zhu (2021). "Obesity aggravates COVID-19: A systematic review and meta-analysis". In: *Journal of Medical Virology* 93.1.

Yang, Yang and Kenneth C Land (2013). *Age-Period-Cohort Analysis: New Models, Methods and Empirical Applications.* Boca Raton, FL: Taylor and Francis Group.

Zaninotto, Paola and Camille Lassale (2019). "Socioeconomic trajectories of body mass index and waist circumference: results from the English Longitudinal Study of Ageing". In: *BMJ Open* 9.4, e025309.