Thesis title

From Margins to Mainstream: Revising Health Economics for Mental Health Equity and Societal Value

Candidate

James Grant Lathe

Degree

PhD in Health Economics

Institution

MRC Unit for Lifelong Health and Ageing at UCL, Department of Population Science & Experimental Medicine, Faculty of Population Health Sciences, UCL, London, WC1E 7HB, UK

Declaration

I, James Grant Lathe confirm that the work presented in my thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis

Acknowledgements

This doctoral thesis is dedicated to Ann McClenaghan, whose moral integrity will forever serve as an inspiration. I want to thank my partner for their patience, my family for their wavering support, and my supervisors for teaching me that economics might not have the answers to all questions. I also thank all my colleagues, peers, and friends whose thoughts have contributed to the development of this thesis.

Abstract

This thesis investigates critical issues in applied mental health economics across five chapters.

Chapter 1 provides a background in mental health economics, highlighting the global rise in mental health disorders, the disparity between the burden of these disorders and the funding allocated to their research and treatment, and the necessity for evaluative methods that are sensitive to mental health's unique characteristics.

Chapter 2 critiques current economic evaluations, arguing they might inadequately capture the true societal value of mental health. It calls for comprehensive evaluation methods that consider spillover effects and patient values.

Chapter 3 utilises dynamic panel methods to estimate the intersectoral costs of mental and physical health improvements. It showcases the utility of advanced econometric methods in observational data to address relevant topics in health economics research. It also demonstrates the necessity for long-term data to enhance the robustness of these findings and suggests future research directions.

Chapter 4 presents a detailed visual framework to standardise terminology and elucidate the mechanisms of interindividual health spillovers.

Chapter 5 summarises the key takeaways of the previous chapters and discusses the two distinct priorities identified by this thesis for advancing research and policy. First, improving the capture of multi-sectoral spillovers, and second, considering the relevance of interindividual spillovers to the decision rules of economic evaluations in health.

Overall, this thesis calls for a holistic, interdisciplinary approach to ensure that methods for economic evaluations adequately characterise the societal value of improving mental health.

Impact statement

This thesis considers how to characterise the value of improving mental health within evaluative frameworks. Given the scarcity of evidence and applied resources in mental health economics, this research lays the foundational groundwork necessary to eventually effect instrumental impact—specifically, to transform practice and enhance societal decision-making. Therefore, this thesis primarily represents conceptual impact. It enhances awareness and understanding of key issues, advances the application of methodologies, and fosters discourse to promote knowledge exchange and capacity building within health economics.

Following an overview of mental health economics, my thesis begins by providing a comprehensive review of social value judgments, creating the first thorough resource for mental health stakeholders and health economists. This work, published in The Lancet Psychiatry, has been well-received by academics and healthcare professionals, informing their applied methods and decision-making processes. Following publication, I was invited as a guest on the Researching Happy podcast to reach a broader, non-specialist audience.

Second, this thesis then demonstrates the utility of modern econometric approaches to analyse health impacts across various policy-relevant sectors. As the first of its kind, this study underlines the difficulties of causal estimation of value in observational settings and the need for broader data collection to inform resource allocation and policy decisions. Third, it elucidates the distinct mechanisms for interindividual spillovers, highlighting the broader implications of mental health improvements beyond simple interpretations of caregiver burden. This work provides a framework for understanding how mental health interventions, or mental health as a domain across all interventions, can mechanistically impact interindividual health.

Fourth, the thesis contributes to the ongoing discourse on multi-sectoral spillovers by proposing policy and practice modifications. Last, through interdisciplinary exposure, this thesis has enabled me to develop contextual knowledge and a diverse skillset that extends beyond the scope of a typical economics doctoral program, positioning me to explore these topics further in my post-doctoral career.

In summary, this thesis advances the understanding of mental health economics and lays the groundwork for more comprehensive and inclusive evaluative frameworks, setting the stage for future practical applications that can influence policy and improve societal health.

UCL Research Paper Declaration Form

1. For a research manuscript that has already been published

a) **What is the title of the manuscript?**
Examining How Well Economic Evaluations Capture the Value of Mental Health

b) **Please include a link to or doi for the work**
10.1016/S2215-0366(23)00436-4

c) **Where was the work published?**
The Lancet Psychiatry

d) **Who published the work?** (e.g. OUP)
Elsevier

e) **When was the work published?**
January 2024

f) **List the manuscript's authors in the order they appear on the publication**
James Lathe, Richard Silverwood, Alun Hughes, and Praveetha Patalay

g) **Was the work peer reviewed?**
Yes

h) **Have you retained the copyright?**
Not for the published version, but yes for the accepted version and earlier versions.

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4527290

2. For multi-authored work, please give a statement of contribution covering all authors

JL conceived and oversaw the project, conducted the literature searches, drafted the original manuscript, selected the references, designed the original figures, compiled the supplementary material, and revised the final manuscript. PP, RJS, and ADH contributed to the manuscript's drafting, figure design, and revision.

3. In which Chapter(s) of your thesis can this material be found?

Chapter 2

4. e-Signatures confirming that the information above is accurate

Candidate
James Lathe

Date:
29/04/24

Supervisor/ Senior Author (where appropriate)
Praveetha Patalay

Date
17/06/2024

Table of contents

Chapter 1. A background in applied mental health economics

"*No health economist can stay on top of two entire literatures. Cultivate the ability to be selective—selective in the seminars and conferences you attend, selective in the review articles that you read, selective in the experts you consult. The goal is to capture most of what is valuable and relevant in new mainstream economics and in medicine*."
- Victor R. Fuchs[1]

The global burden of mental and addictive disorders, and its relative share of total disability-adjusted life years (DALYs), has continued to increase in the last decades, at least in part due to stigma and treatment gap,[2] and recently exacerbated by societal impacts of COVID-19.[3] Half of individuals will develop one or more mental disorders by age 75, which predominantly emerge in childhood, adolescence, or young adulthood.[4] There are calls for improved practice and increased funding for psychological therapies in response to this public health concern.[5]

Such calls have also been met by several proposed research agendas covering the breadth of the mental health sciences.[6–11] From a funding perspective, the share of all causes DALYs attributable to mental disorders is not only greater than their share of associated healthcare expenditures but also much greater than the share of public funding for health research dedicated to mental health (Finland 9.7%, France 4.1%, Spain 5.7%, UK 4.0%), inclusive of not-for-profit funding in countries with a large charity sector for health research such as the UK.[12] The first comprehensive analysis of global grant funding for mental health research demonstrates that this share plummets even further from a global perspective,[13] and Patel's insightful commentary is correct in its title of *Mental health research funding: too little, too inequitable, too skewed*.[14]

Some conversely argue that now is the time for translating evidence into practice;[15] for example, we know some treatments work to a degree - a network meta-analysis of psychotherapies for adult depression found that most treatments were efficacious with little difference between them.[16] However, some argue that therapies neither benefit nor harm on average, and their clinical benefit requires further study.[17] From an economic perspective, treatments such as CBT are likely cost-effective whether in-person,[18] or internet-delivered.[19]

However, if we *were* to translate evidence into practice, it would require increased spending; current spending on mental health only accounts for less than 2% of governmental median health expenditure globally.[20] This is not surprising when, of the low relative spend on mental health research, 50% is devoted to biological and aetiological research, and less than 8% is allocated to health services research and prevention each, respectively, and around 5% to screening and detection.[13] Work is required to ensure that the value of research in mental health and associated healthcare is well established so that funding bodies and grant investigators may invest adequately.

Since the late Carl Taube's seminal *The Future of Mental Health Services Research*,[21] the field of health economics has grown exponentially, as have the varied fields examining mental health across medicine, psychology, epidemiology, and public health. However, in part due to the now broader domains of the field, the intersection between economics and mental health has not driven forward at quite the same pace, a significant constraint being perhaps a lack of incentive or demonstrable worth to grant issuers or receivers. For example, many clinical trials now complete a piggyback economic evaluation; however, there may be competition between clinical and economic study objectives[22], and for clinicians, the resources required to make further use of these economic data may be infeasible. There are calls for economic evaluation as a standard across RCTs globally,[23] however, even in the UK, where the National Institute of Health Research (NIHR) now mandates some form of economic evaluation in RCT grant applications, these analyses are frequently limited to within-trial analysis. There is little scope for exploratory work to inform parameter synthesis in models or to make full use of the rich economic data collected. This is particularly relevant for mental health economics, where developing the case for comprehensive data is crucial.[24] These constraints may have led to stagnant areas of research in health economics, such as mental health promotion and mental disorder prevention.[25]

For applied (and theoretical) research, the most highly-cited reference sources have given little consideration to mental health and focused on the broader issues of insurance markets, moral hazard, adverse selection, and other traditional economic approaches of less applicability to public healthcare systems.[26–30] Furthermore, while there has been recognition by the psychological community of the importance of economic evaluation,[31] and there have been an increasing number of attempts to introduce economic evaluation to mental healthcare professionals or non-economists,[32,33] the predominant research resources for economic evaluation, clinical trials, and decision modelling, do not touch on mental health.[34–38] Unsurprisingly, generalist health economics texts do not focus on mental health, as health economics aims to provide comparable frameworks across disease areas. Mental health resources are similarly constrained in their consideration of health economics and decision-making.[39]

For physicians and policymakers, the book *Mental Health Economics*,[40] presents a much-needed introduction to general methodological topics and reviews cost-effectiveness evidence across various mental disorders. *The Economic Case for the Prevention of Mental Illness*,[41] and *Economics and mental health: the current scenario,*[23] describe how economic evaluation may aid the development of mental health policies and elaborate upon the difficulties currently facing the field, such as lack of cost-effectiveness evidence, the capture of relevant costs, and silo budgeting across sectors. Yet there exist further gaps in the knowledge base which fall outside of the scopes of these works, and beyond a handful of key texts, there are no resources designed explicitly for economists who wish to undertake research at this intersection (Figure A).[42–44]

Figure A. The intersection between economics and the health sciences.

While the knowledge base in health economics is expanding, this does not necessarily lead to practical changes. Bridging the gap between research and real-world application remains a challenge, underscoring the importance of fostering robust collaborative efforts across disciplines to ensure that advancements in knowledge effectively translate into improved health outcomes and economic policies. Economics has a role to play beyond the evaluation and economic consequences of mental disorders. It should investigate the suitability of current methods for assessing the value of improving mental health, working across fields to provide causal insight into the mechanistic processes and later outcomes of mental health across its entire spectrum and myriad of presentations.

Although this thesis focuses on the value of mental health in evaluative frameworks, i.e., health economics as a healthcare science, the mental health intersection with health economics has been underexamined across all fundamental domains – from how economics typically defines mental health, allocative efficiency and marginal productivity of mental health-related expenditure, to health economics as a behavioural science and models for intertemporal choice.[1,45–47]

The importance of economic evaluations to mental health, what makes mental health different, and the literature gap

If cost-effectiveness dictates health technology assessment (HTA) decisions,[48] then, healthcare systems will necessarily adopt treatments that favour health domains more effectively valued by the current methods for economic evaluation (HTA criteria). In this sense, effective valuation means how well outcomes are captured, their attributed weights, and their correlation with relevant indirect resource use. HTA criteria also shape the scope and valuation of benefits and costs, making funding decisions endogenous to these criteria, particularly in countries with robust health regulators, such as the UK, Canada, Australia, and parts of Europe. Theoretically, the value of research and development should reflect the societal benefits of innovations.[49] However, HTA regulations can distort this value both absolutely (by limiting what matters for approval or reimbursement) and relatively (by incentivising adherence to standard evaluation methods that may not capture true societal value).

In *The future of health economics*, Fuchs discusses health economics as both an input into health policy and services research and as a behavioural science. He argues that knowledge in economics is *rarely sufficient to be an effective health economist.*[1] I argue that this is exemplified within mental health economics. No prominent works discuss the suitability of current methods for capturing the diverse value of mental health. Indeed, one of the few generalist texts which mention mental health, the Handbook of Health Economics, states:

"…*mental health can claim no special methodology*" [26]*p.895*

I quote this not as a criticism, as the generation of this thesis is entirely built on the shoulders of others,[50] and indeed, the methods for economic evaluation cannot be unique to any one disease area. However, current discourse on prominent methodological issues overlooks mental health. We should recognise that in a similar way to how healthcare diverges from classical economic markets,[51] mental health and subsequent care further diverge in numerous characteristics from physical health, from aetiology of disorders, epidemiology of mental wellbeing, its effect on risk preferences and help-seeking, social contagion,[52–57] societal cost burden,[23,58] judgement by both the individual and society,[59] the necessity of treatment through social determinants,[60] and , in the same vein as Arrow, to increased uncertainty.[26] Mental health problems can also be caused by, be a cause of, or share common causes with physical health conditions.[61] As highlighted in the background, mental health accounts for only a fraction of global health expenditures.[12,20] Consequently, this thesis extends its focus beyond clinical mental health interventions, addressing mental health as a critical health domain within all economic evaluations.

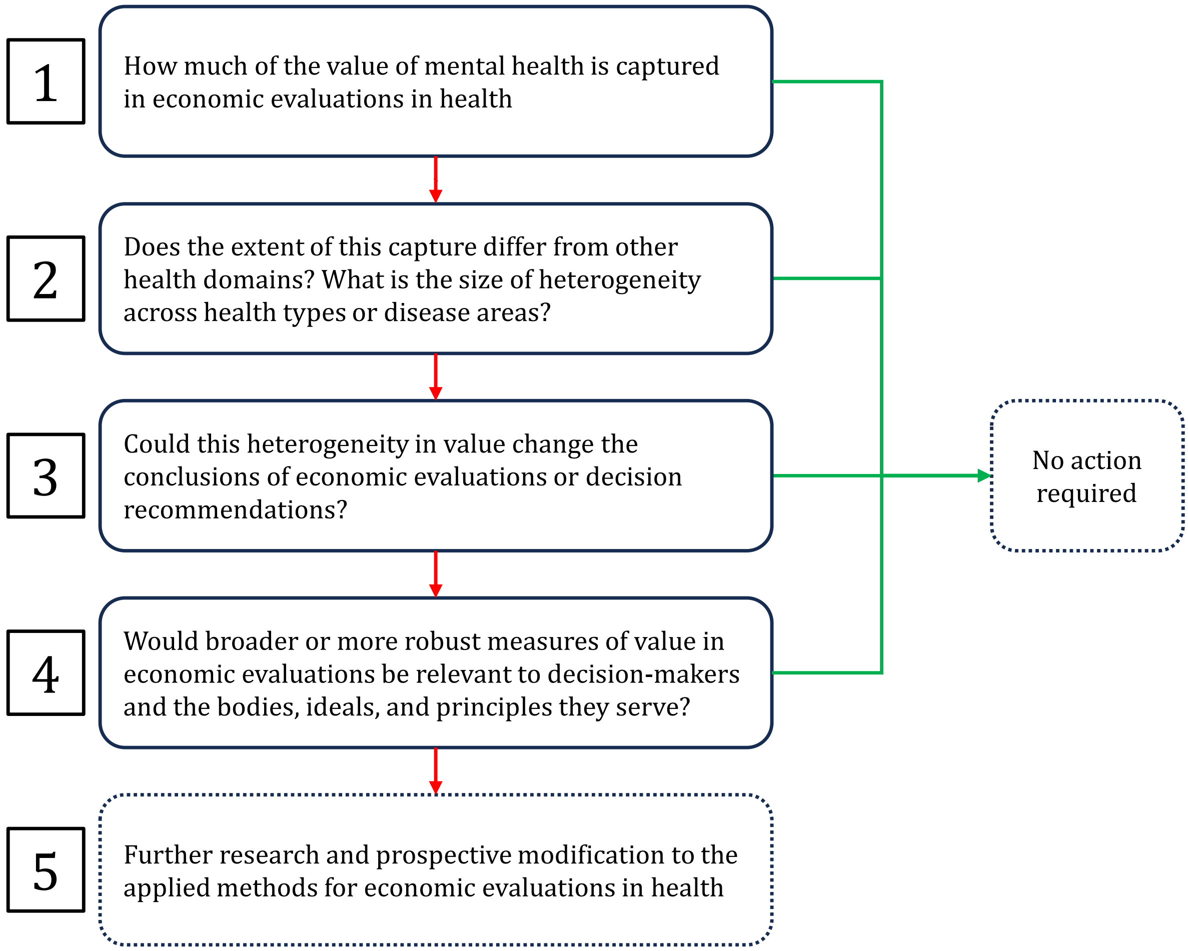
Some researchers have sought to address the unique challenges of mental health in economic evaluations, such as the suitability of the EQ-5D as an outcome measure.[62] However, this represents only one dimension of measuring and attributing value in economic evaluations. Only recently have areas particularly applicable to mental health begun to receive attention in the literature, such as externalities or spillovers. For example, there has been a push within the health technology assessment community to consider better interindividual effects to tackle antimicrobial resistance.[63] While this focus is encouraging, much of the discussion broadly applies to other disease areas where interindividual spillovers in health are sizeable, such as mental health, degenerative diseases, and cancer.[24,57,64,65] Another example is resource use measurement, whose diverse methodologies lack a universally endorsed approach.[66] Efforts to refine self-reported healthcare resource use have resulted in standardised modular measures,[67,68] with similar progress in sectors like education and criminal justice.[69] However, multi-sectoral costs particularly relevant to mental health, such as welfare benefits or tax revenue, remain underexplored.

From the perspective of mental health, the limited interrogation of the methods for economic evaluation may stem from several factors. Drugs for mental health conditions, while representing significant societal value, historically accrued a smaller proportion of this value for manufacturers compared to treatments in other disease areas.[70] Additionally, there has been little growth in the development of mental health technologies or pharmacological interventions in the last 40 years.[71] This lack of innovation, combined with health economics often focusing on technology adoption over efficiency,[72] has feasibly contributed to the insufficient examination of evaluative methods through less exposure. However, as novel treatments and technologies emerge,[73–80] addressing these gaps is critical. Our methods and judgments should be ready for pragmatic and contextual evaluation of new treatments,[81] i.e., reforming methods and processes to keep pace with innovation.[82]

Thesis structure

If capturing the actual benefits and costs of mental health across the lifecourse is desirable, then the current practices for evidence generation must be interrogated. The Chapters of this thesis contribute to the literature by investigating the links between the methods for economic evaluation and the value of improving mental health. It follows a structured approach of necessary conditions that would be required to recommend further research or a change in applied practices (Figure B).

Figure B. Conditional requirements for expanding the methods for economic evaluations to capture the value of improving mental health.



Chapter 2 considers the extent to which current guidelines for and applied applications of economic evaluations in health capture the value of mental health (addressing point 1 in Figure B). Chapter 3 empirically examines the multi-sectoral value of improving mental and physical health (point 2). Chapter 4 conceptualises interindividual spillovers through a multidisciplinary lens. Chapter 5 summarises the main findings of the thesis and offers a general discussion (points 3, 4, and 5). A glossary of terms is available in the appendix.

Chapter 2. Examining how well economic evaluations capture the value of mental health.

The content of this Chapter is now available as a Health Policy article in the Lancet Psychiatry.[24] Following copyright guidance, I present the accepted article (except for Box 1) before any formatting, editing, or amendments by the Lancet Psychiatry.

Abstract

Health economics informs healthcare decision-making but has historically paid insufficient attention to mental health. Economic evaluations in health must define an appropriate scope for benefits and costs and decide how to value them. This Health Policy article provides an overview of these processes and considers to what extent they capture the value of mental health. We suggest that although current practices are both transparent and justifiable, there are distinct limitations for mental health. Most social value judgements, such as the exclusion of interindividual outcomes and multi-sectoral costs, diminish the value of improving mental health, and this may be disproportionate compared to other types of health. Economic analyses may have disadvantaged interventions which comparatively improve mental health, but research is required to test the size of such differential effects and any subsequent impact on decision-making, such as health technology assessment. Collaboration between health economics and the mental health sciences is crucial for achieving mental–physical health parity in evaluative frameworks and ultimately improving population mental health.

Mental health, health economics, and economic evaluation

The global burden of mental disorders is increasing, as is their share of total disability-adjusted life years (DALYs).[2] However, expenditure on mental health accounts for less than two percent of governmental health expenditure globally.[20] The World Health Organization has argued that we must 'deepen the value given to mental health',[83] and others have called for mental health to be considered in all policies beyond health.[60,84] The development of such evidence-based policymaking worldwide requires ways to measure and assign value. Government bodies and economists use established frameworks to value health and inform policymaking, but across the mental health sciences there is little awareness of health economics and its role in healthcare decision-making.[40] Equally, economics has historically neglected mental health, and while that is slowly being rectified, considerable knowledge gaps remain.[23,40,41]

Panel. Scope of this health policy paper.

We discuss

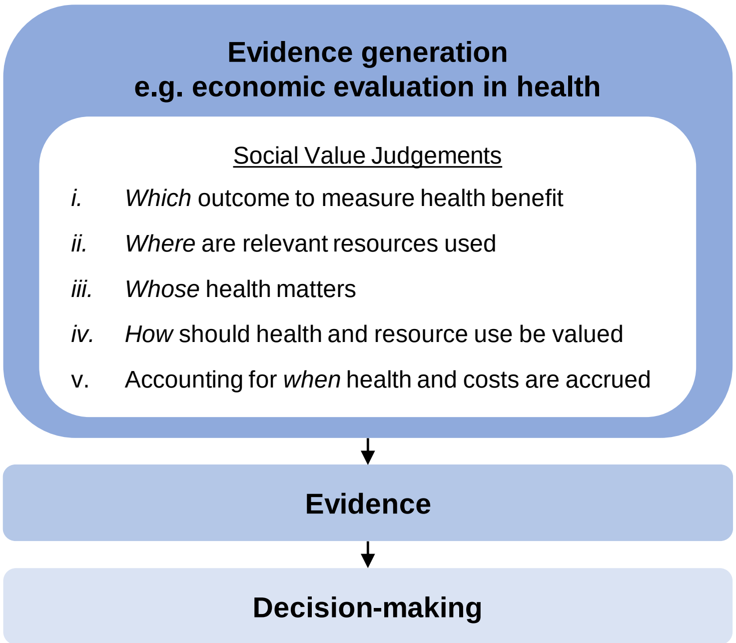
- ✓ How economic evaluations in health measure and attribute value.
- ✓ Guidelines for (and practical applications of) economic evaluations in health and health technology assessments (HTAs).
- ✓ Cross-disciplinary evidence of the benefits and costs of mental health and comparisons to other forms of health where available.
- ✓ How the impacts of mental health may conflict with current guidelines and practices.

We do not discuss

- × Cost-benefit approaches, conceptual frameworks for examining cross-sectoral value for money, or developments in multicriteria decision analysis.
- × Behavioural considerations that influence the quality of economic evaluations e.g., the effects of mental health on engagement with healthcare services or item non-response and attrition in observational studies.

This Health Policy article addresses one such gap: to what extent do economic evaluations in health capture the value of improving mental health? (Panel). We do not aim to address all the problems mental health faces in economics, nor the numerous difficulties faced when translating evidence into policy, i.e., health economics informs, but does not make, health policy (Figure C).[23,41,85] Instead, we aim to raise awareness of the profound impact of health economics on how mental health is valued and provoke discussion around the underlying principles and judgements. Examples cited are not an indictment of any actor, or the role of health economics in decision-making; the issues highlighted often result from iterative and well-intentioned developments in evaluative practices.

Figure C. Social value judgements in economic evidence generation and decision-making.



National Health Technology Assessment (HTA) systems (e.g. NICE in England and Wales),[86–88] inform local to national-level decisions on treatment provision, spanning healthcare technologies, clinical guidelines, and public health guidance. Their recommendations are based on evidence of clinical and cost-effectiveness, which in turn rely upon deciding what benefits and costs are important and how to value them, collectively termed social value judgements (Figure C). Such appraisals have global relevance. Decisions made and methods used by NICE can influence healthcare decision-making worldwide.[89] Social value judgements are also relevant beyond economic evaluations in health: from the non-economic use of generic health status instruments,[90] the productivity losses estimated by cost of illness studies,[91] to the discount rates used widely across governments. Although the practices of HTAs and economic evaluations in health do not directly dictate such use, they undoubtedly influence them.

HTA bodies usually apply social value judgements equally to all interventions, regardless of the disease area. However, mental health differs from physical health in several important ways, such as the challenges of measurement and diagnosis, societal stigma, the contribution of sectors beyond healthcare to outcomes, and the interconnected nature of mental disorders and other health conditions.[92] Crucially, mental health has multiple downstream impacts beyond the individual. These include effects on interpersonal relationships and family cohesion, employment and finance, and wider impacts on social services and the criminal justice system. This is not to say that physical health conditions cannot have such impacts, but we argue that these consequences are often more pronounced for mental health conditions. If parity is sought between mental and physical health, social value judgements must account for these impacts. For example, the introduction of NHS Talking Therapies, formerly known as Improving Access to Psychological Therapies (IAPT), is widely considered to be pioneering. However, economic evaluation of such schemes required a more comprehensive collection of outcomes and resource-use measures to demonstrate the value of reducing functional impairment, the incidence of harmful behaviours, or multi-sectoral costs.[93,94]

We categorise five key themes subject to social value judgements: (i) Which outcome to measure health benefit, (ii) Where are relevant resources used, (iii) Whose health matters, (iv) How should health and resource use be valued, and (v) Accounting for when health and costs are accrued. For each, we provide an overview of current practices and weigh evidence on the extent to which they capture the value of mental health. We then discuss the possible effects on decision-making and propose the next steps for research and policy development.

Which outcome to measure health benefit

The choice of outcome depends on numerous judgements, such as: the purpose of treatment, what we mean by health, and how to quantify mental health adequately. Medical decision-making incurs an opportunity cost: treating one individual means the same resources cannot be used to treat another;[35] therefore, a single summary figure of health is helpful to compare the benefits of different treatments. Economic evaluations employ such a standardised health measure as a generic outcome that is sensitive to health change in different disease areas as a complement to disease-specific outcomes.[95] The applicability of such generic measures as outcomes in mental health has been scrutinised,[43] and Brazier et *al.*[42] provide a rigorous guide to their use in economic evaluations. Of these, the EQ-5D, a generic five-dimensional health status instrument, is the most used in health-economic appraisals worldwide.[96] The EQ-5D has one item related to mental health (the self-identified presence of depression or anxiety). This adequately captures these common mental disorders,[97,98] however, the EQ-5D lacks sensitivity to other mental health conditions such as psychosis, schizophrenia, and bipolar disorders,[99–101] and the composite nature of the question leads to an under-reporting of problems.[102] Recognising that such generic instruments may favour physical over mental health, there are calls for developing a better instrument for use in mental health populations.[97]

Other generic or condition-specific preference-based measures can be used to compare mental health interventions when the EQ-5D is unsuitable,[87,103] but none fully cover the dimensions that are important to individuals with mental health problems.[104] Recent attempts to improve content validity include the Recovering Quality of Life (ReQoL) measure,[105] the CORE-6D,[106] the Mental Health Quality of Life Questionnaire (MHQoL).[107] Moreover, changing the health status measure for mental health settings is potentially problematic as it could impinge upon comparability within healthcare systems and across research. Such a change would also imply that disorder-based criteria can define mental health and neglect the central role of mental health in the lives of individuals.[92] Capturing mental health is essential in all healthcare settings because although mental healthcare requires some threshold for intervention (e.g. disorder-based criteria), mental health (i.e., psychopathology and psychological differences between individuals) exists on a continuous spectrum.[108] As such, interventions do not have to treat mental disorders directly to improve mental health, particularly when mental health problems can also be caused by, be a cause of, or share common causes with physical health conditions.[61]

Therefore, generic measures are likely to remain integral to equitable evaluation, especially if mental health is to be accounted for in all policymaking.[60,84] To this end, there are pushes in economics to expand the evaluative space towards subjective wellbeing.[109,110] The EQ-5D currently captures little of the variation caught by instruments for mental wellbeing,[111] and the lack of social domains in the EQ-5D has been noted as a particular barrier to demonstrating the value of treating behaviour problems in childhood and adolescence.[112] Recently developed generic instruments may offer improvements, such as the EQ Health and Wellbeing (EQ-HWB) instrument, which combines health and wellbeing domains to facilitate cross-sectoral comparisons,[113,114] or the Investigating Choice Experiments Capability Measure for Adults (ICECAP-A).[115] However, wellbeing evaluation has some challenges, such as adaptation, i.e., a permanent improvement in an outcome may only be temporarily associated with improved wellbeing.[110]

Where are relevant resources used

The relevance of resource use is inherently associated with a decision, is highly context-dependent,[116] and can take different perspectives ranging from the individual or payer to society as a whole. Evaluations tend to follow the guidance of executive public bodies (e.g. NICE),[86,87] or a consensus of experts such as the Second Panel on Cost-Effectiveness in Health and Medicine,[117] or the taskforces of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).[118] Globally, most HTA guidelines recommend a healthcare or payer perspective as their reference case,[119] but some countries, such as the Netherlands, consider all costs to be relevant regardless of where they are accrued.[120] In practice, most evaluations have followed a narrow healthcare perspective.[119] Such a perspective is recommended because these bodies do not set healthcare budgets; instead, they offer guidance on what represents an efficient use of healthcare resources without necessarily reflecting on where all the consequences of treatment lie.[23]

Many economic evaluations report additional analysis under a broader societal perspective. Although the term societal implies the capture of multi-sectoral spillovers, most costs beyond the healthcare sector are infrequently captured.[119,121] In practice, societal perspectives are usually limited to one form of productivity loss: the time off work due to ill-health (absenteeism).[122] Poor mental health is associated with sizeable absenteeism costs but also influences the workplace productivity of an individual when present at work (presenteeism).[123,124] These costs are rarely included in economic evaluations,[125] despite evidence that improving mental health leads to greater reductions in presenteeism than absenteeism costs.[58]

Mental health has bidirectional relationships with multi-sectoral resource use, and the societal burden of mental illness and psychosocial problems exceeds, and extends beyond, healthcare costs and absenteeism.[126] Primary care and mental health services bear only a fraction of the costs of mental ill-health. In adolescents, most are borne by frontline or special education.[58,127,128] In adults, they also lie across the criminal justice and welfare sectors.[40,58,124,129,130] This is a problem because while one treatment may appear less expensive than another, the costs may have moved to another sector where they are not measured.[131] For instance, in the United States, investing in community-based mental health programs, while costly, is significantly outweighed by the potential cost savings from averting individuals' involvement in the criminal justice system.[132] In the UK, Layard and Clark offer a tangible example that scaled-up evidence-based psychotherapies can pay for themselves if costs, such as welfare benefits or increased tax revenue, are considered.[94] Encouragingly, there is recognition that more comprehensive cost collection is essential,[121] and the PECUNIA consortium has recently developed a questionnaire to aid the collection of health-related multi-sectoral resource use.[69]

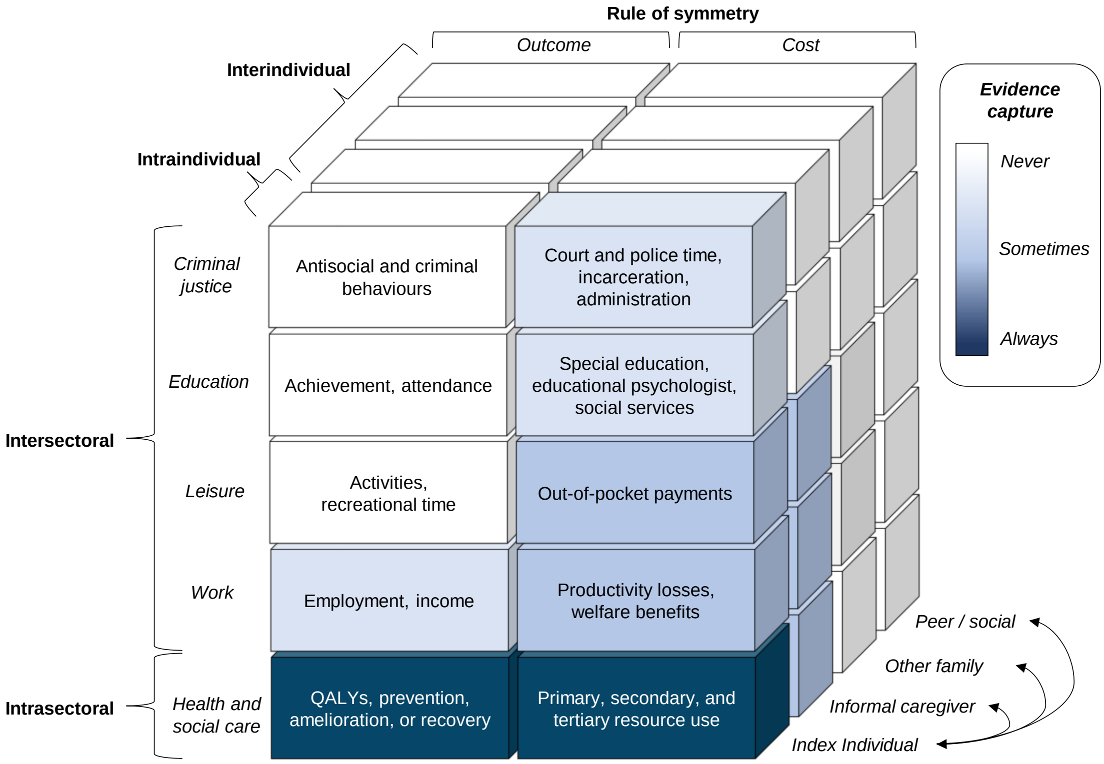
Whose health matters

If societal perspectives intend to support optimal societal decisions, impacts on the health of others may be as important as multi-sectoral costs. For example, mental health problems are often implicated in criminal behaviours, and the health impact of physical and emotional harm to victims exceeds the costs to the criminal justice system and productivity losses;[130] however, the value of averting adverse events is unlikely to be captured in primary data collection.[133] More generally, families, friends, and broader networks of people interact dynamically as a complex system, and a lack of social network weighting in generic health instruments has been noted as a barrier to demonstrating the full value of improvements in mental health.[112] This is overlooked because such instruments measure individual rather than collective health.[134] For instance, NICE guidance over the past decade indicated that: "the perspective on outcomes should be all direct health effects, whether for patients or other people".[86] What constitutes direct effects is ill-defined and has been generally interpreted as effects on informal caregivers alone. Similarly, HTA bodies worldwide vary in their recommendations, ranging from the guidelines of Canada and Australia, which specify that health beyond the individual should not be included in base case analysis, to those of the Netherlands, which considers the health of all impacted individuals to be relevant.[120]

Informal carers for those with mental health problems are invaluable to society and are integral to the health and social care system. Poor mental health causes and is a consequence of caregiver burden, wherein caregiving affects the psychological health of the carer to a greater degree than their physical health,[135] and caregivers of people with mental illness experience a higher subjective burden than those caring for people with a somatic illness.[136] In practice, informal care (when relevant) is usually included as a cost, not an outcome,[137] and carer health is rarely included, even in disease areas where informal caregiving is common.[120,138]

A focus on caregiver burden, and not the wider social network, overlooks the fact that informal caregiving is not dichotomous and may underestimate the benefits of improving mental health; for example, family illness leads to significant decrements in mental health among family members, independent of carer status.[139] This may be because the current interpretation of caregiver burden neglects other forms of transmission, and despite recognition of interindividual effects in tackling antimicrobial resistance,[63] there has been no similar call for mental health. This is surprising when the mental health sciences have long acknowledged the communicable nature of mood and mental health.[52] For example, caring about a family member may have just as much impact as caring for them,[140] and poor peer health increases mood problems.[141] Spillovers may be greatest within families, and parent-child relationships contribute to significant intergenerational effects.[142] Emotional contagion may spread up to three degrees of separation,[143] and although such effects are likely context-specific, there is moderate evidence for the contagion of anxiety and depression.[53] The inclusion of the health of non-caregiving family members in applied economic evaluation is extremely rare (primarily found in vaccination studies) and has mainly been investigated using the EQ-5D.[144]

Figure D. A summary of the evidence captured by economic evaluations in health. Dark blue boxes represent direct benefits and resource use. Boxes in lighter shades or white represent indirect effects or spillovers.



Although this section has predominantly tackled health, spillovers may culminate in resource use by others.[134] From a societal perspective, which individual used resources is irrelevant to governmental budgets, barring distributional concerns.[121] Figure D shows a dimensional breakdown of current evidence capture in economic evaluations, inspired by the Impact Inventory Template produced by the Second Panel on Cost-Effectiveness in Health and Medicine.[117] If a benefit is included within an evaluation, the cost should be included, and vice versa. This practice is termed the rule of symmetry, or internal consistency;[145] however, future unrelated (indirect) costs and benefits are often treated asymmetrically.[146] Over time, a broader outcome measure may capture the aggregate effects of all multi-sectoral outcomes. Figure D also does not include socially desirable outcomes, such as pro-sociality or environmental behaviours, which do not easily fit within the remit of the specified sectors.

How should health and resource use be valued

Once health status is measured, responses are weighted by a societal tariff to represent the relative value the public places on different health states. This produces a health-related quality of life (HRQoL) index, which is typically anchored between zero (death) and one (full health). There are numerous ways to elicit preferences,[147] and the development of HRQoL value sets are methodologically complex and subject to considerable scrutiny.[95,148] Whose preferences matter is a pivotal question that significantly influences the value attributed to health states.[149,150] Such values differ across communities and populations, so country-specific tariffs are used where available.[87,117] But there are other well-documented problems, from participant's immediate preoccupation when values are elicited to health states impacting individuals differently from how they imagine them.[151]

From the perspective of mental health, patients give a higher weight to mental health dimensions compared to physical health dimensions than do members of the general population,[43] and mental health states may be more difficult to understand than their physical counterparts,[151] but alternative approaches to preference elicitation, such as using subjective wellbeing data,[152] may overcome such drawbacks. Any comparisons of approaches are inextricably tied to the items of HRQoL instruments and what individuals are asked to value. For example, the moderate or extreme "anxious or depressed" states in the EQ-5D-3L may impact on subjective wellbeing more than their stated preferences indicate.[152] This links to the broader debate about whether health states should be valued in terms of the activities they permit or their subjective wellbeing,[153,154] and whether these values should reflect those of patients or the general population.[150]

Economic evaluation typically values resource use, such as healthcare, through attachment to established unit costs; however, the valuation of other forms of resource use is not always straightforward. Because we can only give value to what is captured, here we discuss the productivity losses to which societal perspectives in economic evaluation are usually constrained,[119,122] and which are frequently employed by cost-of-illness studies.[125] All of these approaches assume that one monetary unit of productivity loss is equivalent to one in healthcare benefit or cost. The most frequently used method is the human capital approach, which assumes that gross wages represent the productivity of an individual, i.e., their time off work due to ill-health is multiplied by their pro-rata wage.[129,155] Evaluations that derive absenteeism costs frequently use the national average or median wage instead to avoid disadvantaging individuals with severe mental illness.[38] An alternative method is the friction cost approach, which limits absenteeism costs to the time it takes to hire and train a replacement worker, i.e., previous levels of productivity return after a friction period.[155]

Newer approaches use compensation and multiplier effects to better represent real-world production losses attributable to absenteeism.[156] These reflect that an individual's absence often has a larger (multiplied) or smaller (compensated) impact than their wage indicates. Notably, for mental health, presenteeism multipliers may be equal to or higher than absenteeism multipliers.[157] Given the substantial costs, there is no doubt that the capacity to work is essential,[124,129] but focusing on technical dimensions may overlook the normative dimensions of social relevance.[158] For example, mental health has a causal impact on employment status,[159] and employment itself may be considered a critical mental health intervention,[124,160] but current methods give no value to the gain or loss of employment, i.e., we value averting productivity losses but not productivity gains. Overall, productivity losses are not an opportunity cost in governmental spending unless through the channel of tax revenue, which would also attribute value via gains in employment.

Accounting for when health and costs are accrued

Economic evaluations adjust lifespan for life quality, such that (the index of) weighted HRQoL responses are transformed to quality-adjusted life years (QALYs), generally through the linear interpolation of these HRQoL indices at each observed time point. QALYs and costs accrued beyond the first year are discounted (compounded annually) to account for the opportunity cost of investment and time preferences for health, i.e., people preferring current health over future health.[35] The rationales and methodologies of discounting within economic evaluation vary globally,[161] with a discount rate of 5%

being the most common.[162] NICE and the UK government specify a constant discount rate of 3.5% for benefits and costs,[86,87] which means a QALY (and a cost) in ten years is worth around 70% of one now, falling to 25% after 40 years, and 13% after 60 years. Discounting will always be required to some degree; however, there are problems such as double discounting,[162] or decision-makers also preferring current health,[23] which further reduce the value of long-term benefits. The practice also relies on several assumptions;[162] for example, that returns to spending will be as, or more, efficient in the future.

Discounting distorts the perceived effectiveness of interventions with long-term or cumulative consequences, so the value of prevention may be affected more than treatment. Most mental health problems emerge in adolescence and;[4] therefore, the value of the most effective avenues for intervention (earlier) may be the most affected by discounting. In recognition, NICE now supports a lower rate of 1.5% in some scenarios, such as for treatments whose benefits are sustained over a long period.[87] Longitudinal evidence, where available, highlights the long-term impacts of mental health on later health, social, and economic outcomes,[163] and mental health spillovers onto the health of others do not appear to decrease over time.[164] There is also evidence of greater long-term adverse consequences for earlier life poor mental health compared to poor physical health,[165] and that improvements in mental health are preventative for all-cause mortality or suicide.[166] However, it is uncertain whether improving mental health in adulthood has longer-term benefits than other health improvements.

Can social value judgements affect decision recommendations?

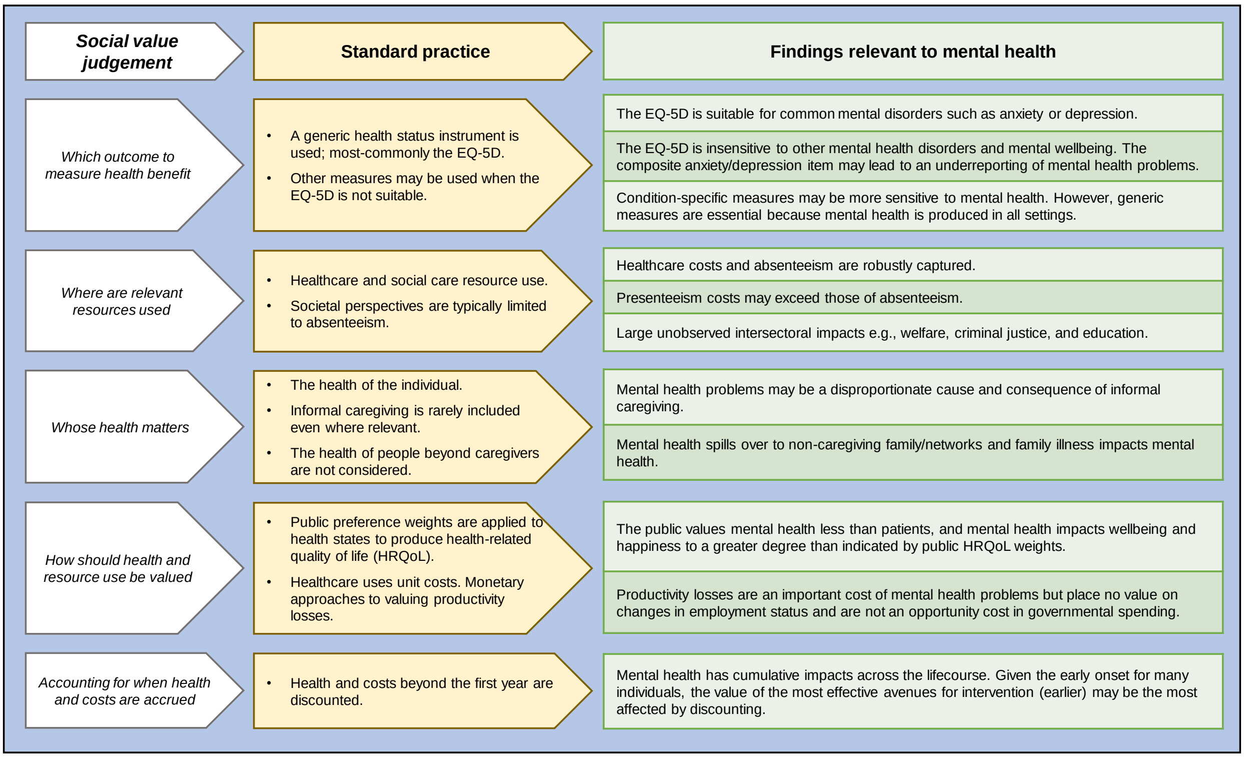
There is a paucity of data on whether methodological change can impact the conclusions of economic evaluations, but some limited evidence exists on the inclusion of multi-sectoral costs and interindividual outcomes. Including spillovers, such as family health or societal costs, generally makes interventions more cost-effective (more health is produced per monetary unit).[120,144,167,168] When economic evaluations compare treatments, a minor change in the cost-effectiveness calculation, such as the inclusion of social costs, can affect the value of each treatment differently and alter decision recommendations.[138,169–171] This means that the optimal treatment judged by healthcare perspectives can be suboptimal from a societal perspective. Broadening perspectives is likely to increase the relative effectiveness of interventions which improve health domains or symptomologies with greater spillover than those without.[172–174] We argue that mental health is one such domain and that it is particularly subject to bidirectional spillovers, e.g. conditions such as depression affect the health of others to a greater degree than some physical health conditions,[164,175] and family illness leads to severe decrements in mental health; comparable evidence is not found for other health dimensions.[139,164]

Discussion: towards capturing the value of mental health

Current practices for economic evaluation are justifiable but have limitations. Current social value judgements underestimate the value of improving mental health to some degree and possibly do so disproportionately compared to other forms of health (Figure E). Differential undervaluation may not always be a problem, as mental health is correlated with other health, and many healthcare decisions would likely remain the same under other perspectives. However, much of the value of mental health is not captured in economic evaluations, and minor changes in perspective can impact on their conclusions. This suggests that, at times, decision-making based on economic evidence may have disadvantaged interventions which comparatively improve mental health domains (Box 1). HTA criteria and methodologies also influence the value of treatment, which may shape the research priorities of (and innovation by) the private health research sector.

Whether mental health is disadvantaged in economic evaluations should be empirically tested, and we need to ask questions such as: “Do recommendations align with the comparator that maximises mental health symptomology?”, “How frequently are healthcare perspectives suitable surrogates for wider perspectives?”, or “Could mental health serve as such a surrogate?”. However, such research depends on data availability and is limited by whether studies disaggregate their reported health and social data.

Figure E. A summary of the extent to which social value judgements capture the value of mental health.

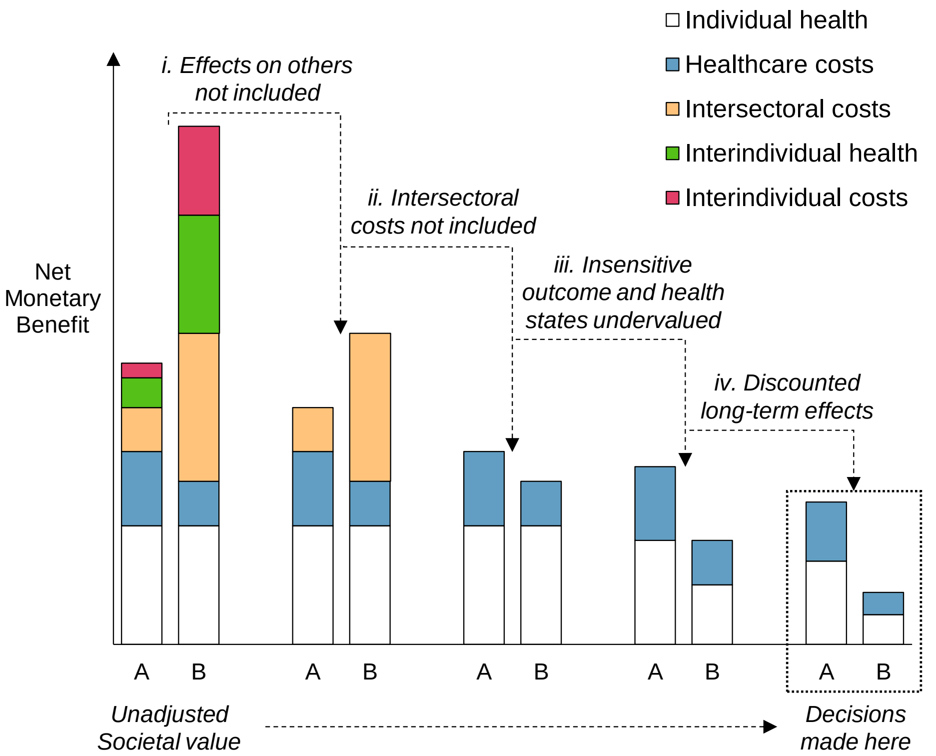


These considerations also apply to conditions and symptomologies beyond mental health. Health economics does not undertake nosology, but it should identify conditions with downstream consequences where social value judgements may similarly impact on equitable evaluation. For example, including presenteeism costs increases the value of improving mental health but would also help to capture the value of improving chronic somatic diseases.[123]

We believe that societal perspectives should move from individual- to population-centric and that an all-government approach to health should be adopted.[176] Interindividual health and multi-sectoral costs are compatible with current value frameworks in healthcare, are integral to understanding the ramifications of spending (or not spending) on population health,[145,146,173] and may help to address inequalities, e.g. mental health spillovers are greater in lower-income households.[139] Others have also argued for the inclusion of certain spillovers, such as caregiver and family effects,[120,137,140,171,174,177,178] and the public also views these effects as important,[178] but there are opposing arguments which span a range of normative and technical domains.[179] There are some scenarios in which a population-centric approach introduces moral quandaries; for example, if interindividual effects are included, does this imply that the health of individuals with larger social networks is worth more?[178] We believe engaging with public opinion on such dilemmas and further empirical examination of the size of spillovers across disease areas is essential.[173,178] However, at present, the data required to examine such questions are rarely collected.[171]

Box 1. A hypothetical example of the differential impact of social value judgements in economic evaluation on recommendations and decision-making.

Net monetary benefit (NMB) calculations facilitate comparisons of benefits and costs by assigning monetary value to all factors. Consider two interventions where both produce equivalent total individual health. Relatively, A improves non-mental health domains more than B, and B improves mental health more than A. Social value judgements at each step decreases the value of B more than A (see *i.-iv.* below). This may lead to the adoption of intervention A which appears more cost effective than B. Such a decision is not societally optimal and reduces public mental health under the counterfactual.



Amending HTA criteria to require more comprehensive benefits and costs would raise minimum data requirements, but the consequences of any changes would need to be considered carefully. Accordingly, this paper aims to motivate mental health practitioners, researchers, and funders to consider broader data collection in randomised studies and to encourage the reporting of disaggregated health economics data to permit secondary analysis.[180] Many observational studies, such as national surveys and panel datasets, do not collect broader health economic data, but linkage to administrative data offers an opportunity to enrich these datasets and the opportunity to improve the integration of economics with lifecourse epidemiology.

We recognise that extra data collection comes at a cost, whether in terms of money, participant burden, item completion, or clinical objectives.[22] A broader trade-off between precision in healthcare versus societal costs should be considered. Although we argue that all spillovers are important, that does not mean they are equally observable. For example, presenteeism costs are notoriously difficult to measure,[125] whereas the receipt of welfare benefits is less subjective. The acceptability of collecting such self-reported data from participants compared with other less burdensome approaches (e.g., administrative records) should be investigated. The development of multipliers to account for interindividual effects would be a pragmatic way to inform economic evaluation;[173] such methods are already used to adjust individual self-reports to reflect production losses better.[156]

Top-down evidence generation, which accounts for the total effects of governmental expenditure, may offer a more accurate assessment of the wider value of improvements in mental health. For example, area-level data largely account for inter-individual spillovers. Such evidence is exceptionally scarce; therefore, population-level research should be a priority to generate practice-based evidence in mental health. This includes analysis of the productivity of mental health expenditure versus other forms of government spending,[176] which may require investment into data linkage across the health, education, welfare, and criminal justice sectors.

Economic evaluations could use additional instruments that demonstrate better psychometric properties to address outcome sensitivity. However, because generic outcome measures are used to set priorities across healthcare systems, other approaches could be explored for mental health, such as a bolt-on to the EQ-5D.[113] This should be paired with an examination of how we attribute value. Encouragingly, Euroqol wishes to support research that examines the basis/rationale of value sets by patient groups,[96] and evidence suggests that the public supports asking those who have experienced heath-states to inform policy.[151] HTA bodies have a long history of involving the public and patients to inform guidance, but research is required across settings, countries, and cultures to reflect population and societal diversity. There are other approaches to the outcome and valuation problems, such as using a broader generic outcome;[113–115] and mental health experts and health economists should work together to explore the validity of these instruments with transdiagnostic classification systems and modern conceptualisations of mental health.[181,182] Mental health practitioners and researchers are well placed to increase patients' voices in all of the issues raised by this paper.[95]

In conclusion, mental health has far-reaching consequences not captured by economic evaluations in health. Progress requires interdisciplinary collaboration between economics and the mental health sciences. While funding for mental health promotion, treatment, and research is essential, a broader focus on evaluative processes may strengthen the economic case for mental health and so benefit population mental health. Everyone is a stakeholder in health economics and economic evaluation.

Search strategy and selection criteria

This review did not address any explicit question about which results might be appraised and synthesised, leaving any systematic review or a proceeding umbrella review inherently uninformative.[183] A field-wide review also stretches beyond the remit of broader scoping reviews and invalidates many existing systematic approaches; first, relevant information is spread across the numerous subfields within economics, psychiatry, psychology, and epidemiology, among others, each with their terminology and heterogeneous interpretation. Second, if differences could be harmonised and terminologies agreed upon for systematic purposes, the quantity of extractable data would be so large that any rigorous attempt to do so would have been infeasible within our funding and time constraints.

Therefore, we took five distinct steps: first, we selected the key common-knowledge references for any paper discussing the value of mental health in economic evaluations, e.g., institutional guidance for economic evaluations in health or international guidance for best practice. Second, we included further articles based on our firsthand knowledge of the literature on economic evaluations and mental health. Third, we searched Google Scholar using the following set of terms individually:

<u>Mental Health and Economic Keywords</u>
("mental health" OR psychiat*) AND "economic evaluation"
("mental health" OR psychiat*) AND "cost-effectiveness"
("mental health" OR psychiat*) AND ("quality of life" OR "health-related quality of life")
("mental health" OR psychiat*) AND (cost OR "resource use")
("mental health" OR psychiat*) AND productivity
("mental health" OR psychiat*) AND discounting

<u>Economic Evaluation Keywords</u>
"economic evaluation" AND spillover
"health spillover"
"economic evaluation" AND productivity
"economic evaluation" AND discounting

We judged whether the articles of the first ≈15 pages of results were relevant to specific social value judgement areas or the discussion by title, abstract, and then full text read. The relevance of the number of pages varied by search term. Fourth, we recursively examined the articles that cite or are cited by the included papers. This resulted in around 400 relevant papers. Because of the limitations of our methods, the fifth step was receiving detailed feedback from two experts in the methods for economic evaluation and one in mental health to ensure completeness in the narrative overview, an accurate representation of sources, and intelligibility across disciplines. Expert opinion led to significant changes to this paper's content and form. We underline that the goal was to capture most of what was valuable, which does not mean that all relevant papers are discussed or referenced. Further references that could not be included because of reference limits are available in the supplementary reading list.

Contributors

JL conceived and oversaw the project, conducted the literature searches, drafted the original manuscript, selected the references, designed the original figures, compiled the supplementary material, and revised the final manuscript. PP, RS, and AH contributed to the manuscript's drafting, figure design, and revision. All authors approved this final version.

Declaration of interests

We declare no conflicts of interest.

Acknowledgements

We thank the independent experts Prof David St Clair (Psychiatry), Prof Hareth Al-Janabi (Health Economics), and Prof Werner Brouwer (Health Economics) for their time, feedback, and insights. We also thank the anonymous reviewers whose comments greatly contributed to the final manuscript. The views expressed in this paper represent those of the authors and not those of the external experts or funders. This work was supported by a Medical Research Council doctoral studentship [MC_ST_LHA_2019]. The Medical Research Council funds the MRC Unit for Lifelong Health and Ageing at UCL [MC_UU_12019/1], and The Centre for Longitudinal Studies is supported by the Economic & Social Research Council [ES/W013142/1]. The funders played no role in the study design, in the collection, analysis or interpretation of data, in the report's writing, or in the decision to submit the article for publication.

Chapter 3. The societal value of mental and physical health: developing multi-sectoral cost profiles using a dynamic approach in a British general population panel study

Abstract

Economic evaluation in health depends upon defining an appropriate scope for benefits and costs for inclusion. These perspectives focus on individual health, healthcare resource use, and, occasionally, productivity losses. However, mental health has far-reaching multi-sectoral impacts, which are generally not considered. If health economics aims to inform optimal societal decisions, estimating the relative size of multi-sectoral spillovers attributable to different health domains is essential.

Although there is growing attention to mental health spillovers and qualitative approaches to the relevance of multi-sectoral costs, quantitative longitudinal evidence is scarce and faces problems of endogeneity biases outside of randomised settings. This study attempts to overcome such problems by applying a dynamic panel approach to the latest seven waves of the UK Household and Longitudinal Survey (UKHLS) to examine the relative value of improving mental or physical health across sectors beyond healthcare. Shocks are applied to the estimated system to derive the cumulative value in healthcare costs, welfare benefits, and labour tax revenue.

Results indicate significant autoregression of physical health deficits, labour tax revenue, welfare benefits, and healthcare costs. Cumulative results were uncertain; mean estimates suggest that increases in either physical or mental health deficits reduce labour tax revenue, and increases in mental health deficits increase welfare costs. Secondary analysis using a just-identified model, including (valued) health-related quality of life, is showcased to demonstrate the utility of dynamic profiles of net monetary benefit.

Dynamic panel approaches offer many advantages over traditional longitudinal modelling. Reanalysis, once successive waves are released, would significantly improve

the power of this study by addressing the limited sample size dictated by model order and instruments required for valid estimation. Further research across settings, populations, and methodologies is also recommended. Unlike disparate multi-sectoral outcomes, multi-sectoral costs do not require valuation and are, therefore, immediately compatible with current value frameworks in healthcare.

Background

Economic evaluation in healthcare depends upon defining an appropriate scope for benefits and costs for inclusion, typically focusing on individual health, healthcare costs, and productivity losses.[119,122,184] Various arguments have been made for the more general inclusion of spillovers in economic evaluations.[177] Indeed, multi-sectoral and cross-sectoral action may be requisite to health equity (brief definitions are presented in Box 2).[185] A recent review of economic evaluations in mental healthcare indicates that including spillovers significantly impacts conclusions around cost-effectiveness,[186] but in Chapter 2 I posit that disregarding these wider effects could disproportionately and adversely affect the value of improving mental health compared to other forms of health in all economic evaluations,[24] e.g., because "*the substantive costs of mental health disorders do not come from treating them, but rather from not treating them*".[187] Evidence is required of the relative size of multi-sectoral costs across health domains,[188] and importantly in economic evaluations, time.[189] This paper highlights that empirical approaches are possible in observational cohort or panel studies.

Some have supported cost-benefit approaches and conceptual frameworks to examine multi-sectoral value for money,[176] or multicriteria decision analysis as a complement to cost-effectiveness analysis.[190] Conversely, despite acknowledgements that wider costing perspectives are essential,[121,191] multi-sectoral resource use has received little attention within the remit of *health economics and outcome research*. Multi-sectoral costs represent low-hanging fruit because, unlike multi-sectoral outcomes, they are not contingent on valuing disparate outcomes; rather, disparate costs are comparable, all representing an opportunity cost in governmental spending, and are therefore compatible with current value frameworks.

Box 2. Defining sectoral considerations or interactions.

These terms are used to describe any consideration or interaction that spans multiple sectors. However, these are frequently used synonymously in the literature.

- **Cross-sectoral**

Efforts that break down boundaries between sectors, involving shared initiatives and resource exchange.

- **Intersectoral**

Consideration of, or collaboration at, the juncture between different sectors to address complex issues.

- **Multi-sectoral**

Involvement of multiple sectors working towards a common goal, often in parallel rather than in an integrated way, e.g., measuring indirect treatment costs in multiple sectors.

For health economists, the two driving questions are which are most impactful if included and how feasible they are to observe.[178] However, because most costing perspectives are limited to healthcare,[119] and societal perspectives to absenteeism,[122] there is a lack of data from randomised studies to examine such questions, and other approaches are required where possible. Recent work has addressed these questions through expert surveys,[192,193] leading to the development of a cross-sectoral resource use instrument.[69] Beyond healthcare costs and employment status, they agree with estimates that education and criminal justice sector resource use are important cost drivers in health.[192] This is broadly in line with cost drivers in mental health, which are borne across the welfare, criminal justice, and employment sectors.[24] Layard and Clark employed an alternative approach and used simple empirical modelling using estimates from causal studies to suggest that welfare benefits and tax revenue are essential cost drivers for mental health.[194]

To compare multi-sectoral value, measures of physical health, mental health, and healthcare resource use are essential to capture. However, typically considered educational costs, such as primary or secondary special education,[192] are irrelevant in adult populations, and some multi-sectoral costs, such as those to the criminal justice system, are not readily available in observational studies. Furthermore, although extremely important,[192] the effects of mental health on the propensity to commit crime are unlikely to be observed, nor are the impacts on those affected by crime.[133] Therefore, this paper considers two costs available in observational studies: welfare benefits and labour tax revenue. Welfare benefits are particularly relevant in the current discourse, e.g., the share of those on disability benefits because of a mental disorder has been increasing in OECD countries,[124] and the debate around the impacts of welfare conditionality.[195]

Employment capacity is important to individuals and policymakers alike; however, typical methods such as productivity losses give no value to gain or loss of employment, nor do they represent an opportunity cost in government spending. Labour-related tax revenue is advantageous because it can be used as a proxy for employment function, representing a straightforward way to value gains/losses of employment and otherwise represents money that governments could spend, e.g., income tax and national insurance contributions represent most of the UK government's direct tax receipts.[196]

This paper empirically estimates the value of improving mental and physical health across healthcare costs, welfare benefits, and labour tax revenue, alongside comparing their multi-sectoral cost profiles. The secondary analysis extends analyses to an actual healthcare perspective by including valued health-related quality of life. Section 2. briefly conceptualises the essential determinants and characteristics of the relationships between multi-sectoral spillovers – using this to discuss the advantages of specific modelling approaches. Section 3. introduces the data set and variables. Section 4. presents the maintained statistical model, steps for model fitting, and implementation of impulse response functions. Section 5. reports the results of the primary and secondary analyses. Section 6. discusses the findings, limitations, and conclusions.

Methods

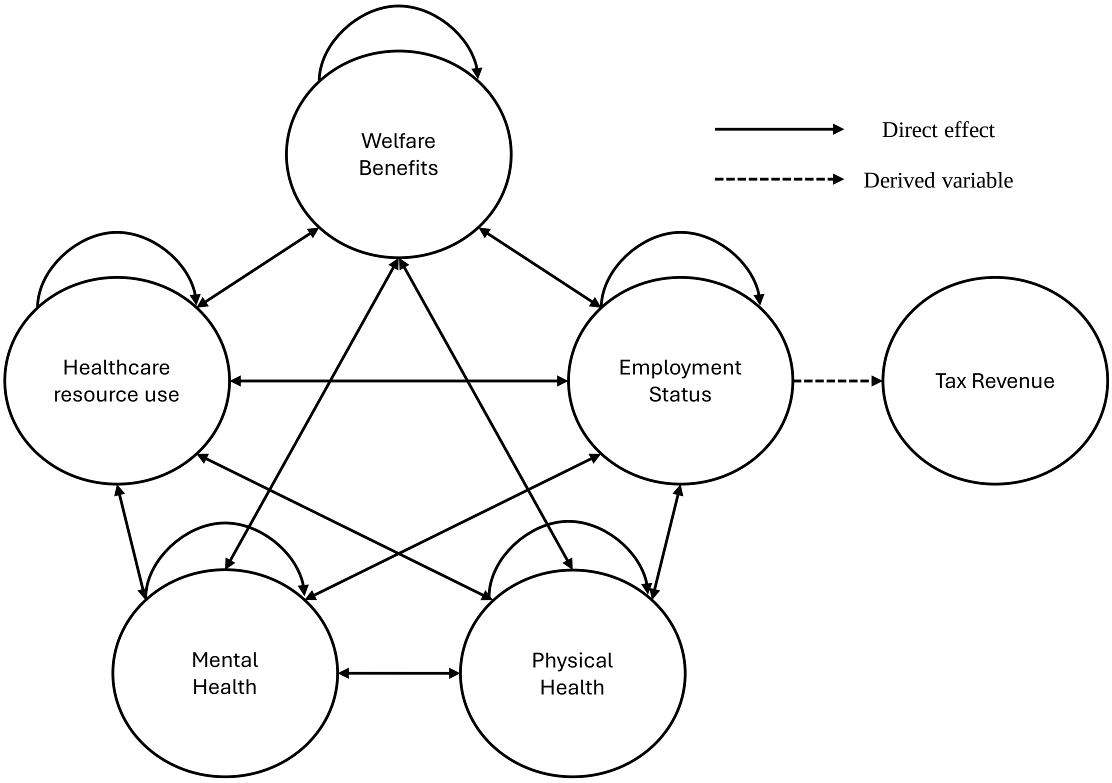
Conceptualising the relation between multi-sectoral spillovers

An individual's health is intricately linked to various aspects of their environment and society, resulting in a complex interplay between mental health, overall health, and resource utilisation across sectors. Although empirical evidence of these relationships exists, it remains relatively limited.[24,41,127,197] This section aims not to hypothesise the existence of these relationships,[198] which are well-established, or delve into specific associations like mental health problems and mental health-related disability allowances. Instead, the focus is on briefly elucidating the driving characteristics of these relationships to inform modelling strategy.

1. Bidirectional relation

The interplay between health, wellbeing, socioeconomic factors, and government services underscores the bidirectional nature of their relationships (Figure F). Mental health influences an individual's employment status, income, health outcomes, healthcare service utilisation, receipt of welfare benefits, and educational attainment.[40,58,94,124,129,130,159,199–201] An individual's overall health status profoundly impacts their mental wellbeing, economic prosperity,[156,202] and access to healthcare services.[203] Mental and physical health are intrinsically linked in their contribution to overall health status.[61,71,92,181] Healthcare service utilisation directly intervenes on an individual's health and mental wellbeing, with potential indirect effects across other sectors. Welfare benefits are closely tied to healthcare resource utilisation,[204,205] mental health,[206] and employment status.[207] Similarly, employment status influences an individual's health,[208] mental wellbeing, eligibility for welfare benefits,[124,209–212] and impacts public healthcare use.[213,214]

Figure F. Causal diagram of health and economic factors.



2. Autoregressive

The health production function posits that current health status is contingent upon past health stock and investments,[215] reflecting autoregressive properties. Mental health and wellbeing also exhibit autoregressive tendencies, such as the long-standing discussion around the hedonic treadmill and inertial/habitual channels.[216–218] Healthcare in the UK operates as an integrated system with referrals, and behaviourally, past utilisation serves as a crucial predictor of current use.[214] This autoregressive nature extends to economic factors, such as welfare benefits, which may either alleviate temporary hardship or foster dependence, or tax revenue (as an indicator of employment status economic prosperity), e.g., prolonged unemployment reduces control-regaining efforts,[219] or marginal income above sufficiency contributes to increased wealth generation and development opportunities.[220]

3. Dynamic interaction as a system

There are lagged direct effects between outcomes (past outcomes influencing current values in other outcomes),[199] in line with broader epidemiological life course theories.[221–223] Moreover, because of characteristics 1 & 2, effects propagate through other outcomes (the indirect effects not commonly captured by current costing methods),[224] as changes in health reverberate through an individual's economic prosperity and use of government services, thereby creating a complex system of dynamic reciprocal influences.

Prospective estimators

Because of these characteristics, and because broader cost data are not collected in RCTs,[24] the examination of multi-sectoral spillovers which reflect complex real-world interactions is challenging,[225] and poses the distinct challenge of endogeneity. Endogeneity bias takes many forms, such as common-method variance, measurement errors, omitted variables/selections, and of importance in mental health research, dynamic endogeneity and simultaneity, or reverse causality. Such bias leads typical econometric methods for evaluation to struggle in mental health research using observational cohorts.[226] Established methods such as regression discontinuity design or difference-in-differences cannot be applied due to a lack of plausible exogenous shocks or weak instruments for mental health, which may have led to its under-examination.

For this purpose, panel data can be exploited to examine within-group change where fixed effects remove unobserved time-invariant heterogeneity, relevant to mental health research where *"Some evidence suggests that selection into mental health is almost entirely based on time-invariant characteristics".*[159] Knapp et al.[227] note that analysis using classical fixed effects in the mental health setting may suffer from poor identification, poor instruments, and other issues, and subsequently used other approaches, such as propensity-score matching, to examine the effects of unpaid caregiving on health. While fixed effects may control for selection into mental health, it does not address dynamic endogeneity bias/simultaneity, wherein general estimators, and lags of the independent variable (making the static panel dynamic) can be employed using lagged values of the instrumented dependent variables (internal instruments). These methods do not assume the availability of exogenous instruments, so lags are the best instrument for outcomes (e.g., assuming they are correlated with the present value but uncorrelated with the error term). GMM is such a class of semi-parametric estimators (like OLS, GMM is a minimum distance estimator; it minimises the weighted sum of the squared sample moments) and is consistent in the presence of unobserved heterogeneity, simultaneity, and dynamic endogeneity.[228,229] In layperson's terms, moments refer to statistical characteristics of the data, such as variable mean, variance, and covariance. GMM estimation derives model parameters such that predictions from said model match the specific statistical characteristics of the data as closely as possible. GMM requires fewer assumptions to be consistent at the cost of some efficiency.

Unless moments are uncorrelated and of equal variance, GMM becomes inefficient, i.e., including two instruments with the same mean but one with a much larger variance, effectively wasting the information in the other.[229] This is particularly relevant in this study's complex model, which has variables whose moment conditions are likely highly correlated. Two-step GMM can overcome such concerns, weighting moments in inverse proportion to their variances and covariances. The model is initially estimated using the identity matrix, and then the weighting matrix for the moment conditions is iteratively updated to minimise the asymptotic variance of the estimators.

GMM estimation also does not require data to be normally distributed (relevant to all this paper's outcomes of interest); rather, it depends on moment conditions, which are equations involving the parameters of the model and the data that should hold regardless of the distribution of the data. GMM produces consistent and efficient parameter estimates if the moment conditions are correctly specified (valid) and underlying assumptions are met. However, it is worth noting that the data distribution could influence the efficiency of GMM estimates. The GMM estimator becomes more efficient as more valid moment conditions are included.

Notably, GMM, unlike typical fixed-effects estimation, does not require strict exogeneity assumptions, and overcomes the identification challenges of two-way fixed effects models,[230] because it removes fixed effects by internally transforming the data through either first differences (FD) or forward orthogonal deviations (FOD). Arellano and Bover initially proposed FOD as an alternative to first differencing;[231] where FD subtracts the previous value of a variable from its current value, FOD instead subtracts the average of all available future observations from its current value.[229] This allows the untransformed variable in levels to be valid instruments for current transformed values (as opposed to the second lag for FD) because only forward realisations are used in the transformation. I note that system-GMM which instruments levels with change (as opposed to instrumenting changes with levels), might be a superior estimator in certain circumstances but further discussion can be found elsewhere.[232] Technical evidence suggests that levels perform well as instruments;[233] however, it is not clear whether levels in outcomes, such as mental health, are indicative of change more than a change in mental health is indicative of a state.

FOD offers four distinct advantages. First, while FOD and FD should be equivalent in large cross-sections, gaps due to non-response over time are magnified when using FD (e.g., if y_{it} is missing, then y_{it} and $y_{i,t+1}$ are missing in the transformed data but using FOD only y_{it} would be missing).[234] Furthermore, if they differ, FOD may have better finite sample properties.[233,235] Second, FD require a longer time dimension than FOD, which may be an issue when fitting panel VAR models using short panels.[234] Third, as individuals may revert to a long-run state, e.g., hedonic adaptation,[216,236] or flow vs lifetime utility,[237] this transformation reflects an individual's deviation, as opposed to relative contemporary flux, i.e., FD. Fourth, unlike FD, FOD does not automatically introduce first-order autocorrelation in the error term. Last, averaging across future observations somewhat overcomes random measurement error and regression dilution bias,[238] reducing the downward regression slope bias towards the null, particularly prevalent in FD models (which may also increase the noise-to-signal ratio).

GMM approaches are increasingly accessible,[228,229] and their methods are increasingly advanced. For example, the introduction of non-linear moment conditions can overcome the weak instruments problem or deviations from mean stationarity, which plagues GMM estimation of linear dynamic panel data models while yielding substantial efficiency gains.[239] Equally, there have been significant advances in maximum likelihood structural equation models (ML-SEM),[240,241] which might be less biased and more efficient than GMM in many circumstances,[242] addressing a principal limitation of traditional fixed-effects estimators and being able to control for unmeasured time-invariant variables when their effects change over time. It also avoids the typical problem of initial conditions and reduces the downward bias of the Arellano & Bond

(AB/FD) estimator for the specified autoregressive parameter.[241] It may also produce unbiased estimates of dynamic effects if reverse causality is present and does not suffer from misspecification of lags.[243] Unlike the AB estimator, ML-SEM assumes multivariate normality of exogenous and endogenous variables; however, it produces consistent estimators even when this assumption is violated.[244] However, if ML-SEM were to model simultaneous, reciprocal causation, it would require further instrumental variable techniques, something of a challenge in mental health. Additionally, sharing a drawback with typical GMM, cross-lagged ML-SEM only permits univariate analysis. A different approach is required to untangle the coevolution of mental health, HRQoL, and resource use within the same population.

Panel vector autoregression

Equation-by-equation GMM yields consistent estimates of the coevolution between health and resource use, but efficiency can be gained by fitting the model as a system of equations. Panel vector autoregression (PVAR) models extend typical Vector Auto Regression (VAR), popular for multivariate time series because it explicitly models the interdependencies between different time series variables, to employ the GMM estimator (instead of OLS) as a system and treats all variables as endogenous. Such methods represent a bridge between time series methods and applied microeconometrics.[245] Usefully, PVAR enables impulse response functions (IRFs) to estimate the dynamic causal effects of improving health across all outcomes. IRFs are more common in applied macroeconomics and are distinct from local projections that would otherwise rely on exogenous shocks to mental health.

For clarity, I present the advantages of my chosen methodological approach in Table 1.

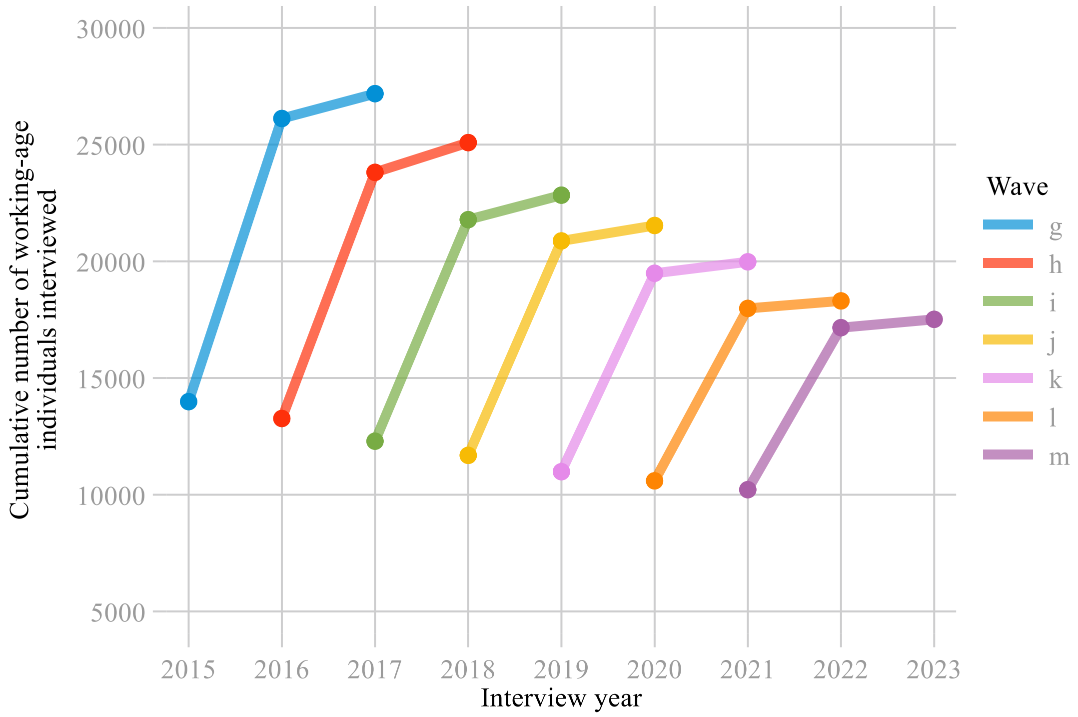
Table 1. Summary of the challenges in longitudinal analyses in health and the strengths of panel vector autoregression using a generalised method of moments estimator.

Challenges	Strengths
Non-normal distribution of variables	Generalised Method of Moments (GMM) provides robust parameter estimation by relying on the moment conditions of transformed data.
Time invariant confounding	Forward Orthogonal Deviations (FOD) removes time-invariant individual components (fixed effects).
Reverse causality	The PVAR GMM model structure ensures that changes in independent variables precede those in dependent variables and controls for lagged dependent variables.
Omitted variable bias, dynamic endogeneity and simultaneity	The identification strategy uses instrumental variables (IVs) that are uncorrelated with the error term. When valid, IVs address omitted variable bias, dynamic endogeneity introduced by lagged dependent variables, and simultaneity bias.
Statistical efficiency	PVAR GMM improves efficiency compared to an equation-by-equation approach. However, using instrumental variables in GMM estimation primarily improves the consistency of estimates when facing endogeneity issues. Although IVs often increase the estimator's variance, they are crucial for yielding unbiased and consistent results where simpler methods, such as OLS, fail due to endogenous relationships within the data.
Serial correlation and heteroscedasticity	GMM estimators can be more robust against heteroscedasticity and certain types of serial correlation than those used in typical longitudinal data analysis. This robustness is achieved through appropriate weighting matrices in the estimation process and appropriate unit-specific error clustering.
Model validity	The framework supports rigorous model and moment testing, such as overidentifying restriction tests, which ensure the validity of the instrumental variables and the correctness of the overall model specification.
Indirect dynamic effects	Impulse Response Functions (IRFs) estimate causal dynamic impacts, quantifying both cumulative direct and indirect effects of an impulse across a system of equations over time.

Data and population sampling

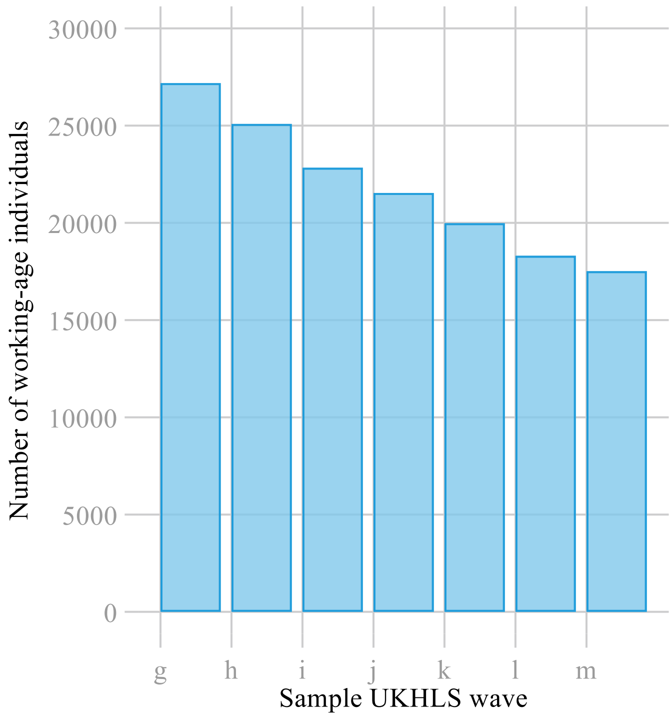
This study uses the UK Household Longitudinal Study (UKHLS), also known as Understanding Society. UKHLS is a representative study of the UK, featuring a robust sample design covering a range of social, economic, behavioural, and health domains.[246,247] UKHLS started in 2009, building upon the previous British Household Panel Survey (BHPS), with around 40,000 households, of which 8000 are a continuation of BHPS and include an ethnic, immigrant, and ethnic boost sample. Because this study targets tax revenue (generated through labour income), the analytic sample was restricted to working-age adults aged 25-64 years, although observations for individuals who aged in or out of this age range are retained. UKHLS includes entire households, wherein all generations, siblings, and spouses are followed up at approximately 12-month increments, relative to the sample month of each household, and each wave is sampled over 24 months (Figure G). UKHLS provides helpful user guides and documentation of the rotating modules of the household and individual questionnaire.[248,249]

Figure G. The cumulative number of working-age individuals interviewed in each UKHLS wave by calendar year.



This study utilises seven years (Waves g [7] through m [13]) of the Understanding Society Mainstage dataset,[250] freely available to researchers from the UK Data Service. Increased periods of observation T would be possible by including the entire length of the harmonised BHPS; however, the collection of robust healthcare service use data was only implemented in Wave g (7) of UKHLS. One strategy would be to use individuals who participated throughout all seven waves; however, if mental health influences the likelihood of study attrition, analyses might underestimate observed relationships. Therefore, this study examines available case data representing 152,448 observations for 34,179 individuals over this period (Figure H).

Figure H. Count of individuals participating in each UKHLS wave.



Analysis used all estimable panels within this population; however, this precise number fluctuates based upon lag lengths, i.e., at least three time periods are required to estimate a panel with two lags, and estimands with fewer observations cannot be estimated. The estimated panels for each model are reported in the results.

As discussed, the topic of spillovers is generally sparse, so, unsurprisingly, there have been few attempts using UKHLS to examine health spillovers explicitly. However, there are some relevant quantitative outputs, such as trajectories in mental health and socio-spatial conditions,[251] within-individual behavioural spillover attributable to the single-use carrier bag charge,[252] local government spending on adult social care and carers' subjective wellbeing,[253] ADHD on the health and wellbeing of ADHD children and their siblings,[254] the health and economic impacts of unpaid care by young people,[227] mental health on employment,[159] and further qualitative work investigating social-norm induced effects of neighbours work on wellbeing.[255]

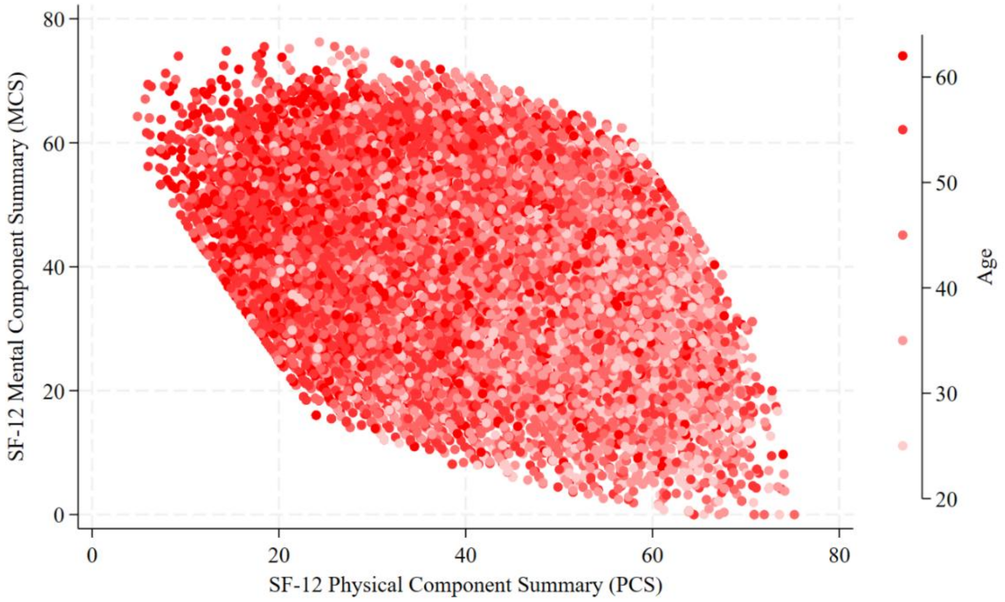
Outcomes

Mental and physical health

The Short-Form 12 Health Survey (SF-12) is a measure that asks individuals to self-report their physical and mental health 'in the last four weeks'.[256] The measure has been validated across ethnic samples in the UK general population,[257] and cross-validated with its parent, the SF-36, in a multinational setting, where physical and mental component scores between each instrument are nearly identical in longitudinal studies.[258] The SF-12 instrument also permits the derivation of mental and physical component summary scores (PCS / MCS), standardised using US population data.[259] Therefore, both are on a continuous scale where 0 represents low functioning, and 100 represents high functioning (Figure I).

Standardisation (and inbuilt normalisation) permit analyses to examine change relative to the population distribution, which offers many advantages.[260] Importantly, there is no underlying causal equivalence in health terms of a one-point improvement in MCS or PCS scores; however, we can pay attention to their respective shares of total monetary benefit.

Figure I. Scattered MCS and PCS scores by age bins.



Resource use

Healthcare

UKHLS includes participant self-reports of general practitioner (GP) visits, outpatient services, and inpatient bed-days. Although such data are not granular, the use of average unit costs is common and will be representative at the population level in large samples. This resource use is assumed to capture the major healthcare cost drivers, including all-cause resource use, because of the extensive links between mental health and comorbidities (such as medical offset),[61,261–264] health behaviours, decision-making, and resource use preferences.[265] Modern empirical evidence has shown that the cost savings of preventive care in individuals with severe mental illness are much more extensive in secondary than primary care.[266] Psychotherapies reduce hospital utilisation by individuals with long-term chronic conditions,[267] and otherwise reduce physical healthcare costs.[94] Individual-level costs across services were aggregated to form the variable for healthcare costs. Unit cost sources and methods for valuing healthcare resource use are presented in Table 2.

Table 2. Healthcare costing

Type	Timeframe	Technical notes
General practitioner (GP)	GP and outpatient visits in the last 12 months were available in the bandings ‘none’, ‘one or two’, ‘three to five’, ‘six to ten’ or ‘more than ten’. This study followed the procedures of Knapp et al.,[227] taking the mid-point of each range, aside from ‘more than ten’ for which I took the low point of 11.	Because COVID-19 distinctly impacted the UK healthcare system and might have contemporaneously affected unit costs, GP contacts were multiplied by their relevant Personal Social Services Research Unit (PSSRU) 2019/20 UK unit cost of £39,[268] including direct care staff and qualifications costs (born by the state). UKHLS does not collect medication data, so a prescription cost per consultation (actual cost) of £30.90 was added to each GP contact. These costs were inflated to 2021/22 values using Oct 2019-Oct 2021 CPI index rates provided by the Office for National Statistics (ONS) for Medical Services (6.91%).[269] The inflated total cost per appointment equals £74.73. GP appointments will have transitioned to remote appointments during (and frequently after) the pandemic. However, such appointments were attached to the same unit cost as those accrued before this period to avoid introducing non-stationarity and ensure that the same impacts on health and behaviours will be treated equally.
Outpatient visits		Similarly, the PSSRU 2019/20 weighted average cost of all outpatient attendances of £135 was inflated from 2019/20 to 2021/22 values using the ONS rates for Hospital Services (8.67%),[269] resulting in an inflated cost of £146.70.
Inpatient bed-days	The last 12 months were self-reported as a continuous outcome.	NHS elective and non-elective combined excess bed-day costs were unavailable from the 2018/19 cost collection onwards due to NHS England and NHS Improvement moving towards patient-level costing, and as such, no longer hold such data. Accordingly, this study uses the excess bed day cost from the archived 2017/18 reference costs of £345.76 (average elective and non-elective excess bed day costs relative to their total activity, respectively).[270] This was again inflated to 2021/22 values using ONS rates for Hospital Services,[269] deriving a rate of 16.9% from the midpoint of the NHS financial year of October

		2017 to October 2021 and an inflated cost of £404.19 There are recent methods designed to capture the actual opportunity cost of bed-days because most costs associated with hospital stays are accrued during the initial period following admission.[271] Excess bed costs are presumed to reflect the cost per bed-day at the national margin.

Work and welfare

UKHLS provides extensive documentation for their income data and procedures.[272] Cost components and technical notes are available in Table 3.

Table 3. Welfare benefits and tax revenue valuation

Variable	Components
Welfare benefits	The measure of welfare benefits was compiled from net monthly social benefit income; all state benefits and allowances, rebates, and credits: severe disablement allowance, disability living allowance, incapacity benefit, income support, job seeker's allowance, working tax credit (includes disabled person's tax credit), council tax benefit, employment and support allowance, return to work credit, in-work credit for lone parents, other disability-related benefit or payment, income from any other state benefit, universal credit, and personal independence payments. Because the purpose of this study is to estimate the total value, proceeding analysis assumes that the sample restriction to the working-age population, the removal of time-invariant fixed effects, and the use of control variables remove the noise of the following benefits if they are not related to the impact of changes in health: state retirement (old age) pension, a widow's or war widow's pension, a widowed mother's allowance / widowed parent's allowance, pension credit (includes guarantee credit & saving credit), industrial injury disablement allowance, attendance allowance, carer's allowance (formerly invalid care allowance), war disablement pension, child benefit (including lone-parent child benefit payments), child tax credit, maternity allowance, housing benefit, foster allowance/guardian allowance, rent rebate (NI only), rate rebate (NI only – offset against rates).

Labour tax revenue	Calculated by subtracting the derived variable net monthly labour income from *gross monthly labour* income, including usual pay, self-employed income, and pay in a second job. Analysis did not consider other income components for which changes (and crucially, effects on tax revenue) are less probable to be related to health within each panel sampling period; net miscellaneous income (educational grants, payments from non-domiciled other family members, or any other payment), net private benefit income (trade union payments, maintenance or alimony, sickness or accident insurance), net investment income (savings, investments, lodger rent), and net pension income (pension from previous employer or spouse previous employer).
Technical notes	Incomes and welfare benefits accrued in different years were inflated to the last year of data collection (from October of their reporting year to October 2021, using the all-item CPI* index from the Office for National Statistics (ONS).[269] G2015/16- 13.26% , H2016/17- 12.25%, I2017/18- 9.02%, J2018/19- 6.47%, K2019/20- 4.89%, L2020/21- 4.12%, and M2021/22 – no adjustment. Incomes, benefits, and derived-tax revenues were only available for the survey month, so their values were multiplied by twelve to generate annualised estimates.

**CPI reflects changes in the consumption value of income and may not align precisely with wage growth over this period. I judged this distinction relevant to analysing relationships with other economic and health-related factors.*

Controls

There are a variety of mediators and moderators within the UKHLS dataset which could be considered, such as caregiving, adverse events, and social support. Variables such as these were not included as controls because they are within the causal chain between health and multi-sectoral costs, e.g., the propensity to give care linked to one's health and free time, adverse events (such as being a victim of criminal behaviour) impact on all outcomes, poor mental health may be both a cause and a consequence of social isolation. Furthermore, adding more endogenous variables to PVAR increases the number of instruments exponentially, resulting in unidentifiable models.

Age, as a time-variant control, was specified in simple continuous form. Despite possible non-linear relationships with all outcomes (given the non-linearity of human ageing itself)[273], age impacts are not the target of this study; analysis examines the working-age population, and the addition of polynomials would increase model complexity. Conversely, time-invariant fixed effects are removed by FOD transformation, which applies to typical controls such as sex and educational attainment, for which, like other studies, the restriction of the sample to working-age adults allows educational attainment, also serving in this capacity as a measure of adult socioeconomic position, to be considered time-invariant.[274]

Secondary analysis

For an accurate comparison against the healthcare perspective, costs should also be compared to valued HRQoL itself, but it presents the distinct drawback that measures of HRQoL and those of health will capture similar constructs and impinge on causal inference in analyses that lack distinct exogenous shocks. If primary analyses are acceptably robust in their estimates, secondary analyses will examine measures of HRQoL and mental health.

HRQoL

To derive HRQoL, the SF-12 was used to construct the SF-6D: a generic preference-based single measure of health which generates an index by applying preference weights to a reduced form of the SF-12 (or SF-36). These weights had been elicited through the standard gamble approach on a sample of the UK general population.[147] The source of the tariff, model 4 (parsimonious consistent model), and a discussion of limitations such as floor effects can be found in Brazier and Roberts.[275] The reduced form of the SF-12 is produced by excluding the general health items and combining the two role limitation dimensions. Furthermore, the number of items per dimension is condensed to one (wherein multiple items per dimension of health had previously been tapping into the same underlying construct). These changes were manually administered to the SF-12 long-form data from UKHLS participants.

Current evidence supports the psychometric validity and responsiveness of the SF-6D in common mental health and personality disorders,[98] making it particularly suitable for this study. Arguments have also been made for its specific use in estimating spillover effects, such as on the caregivers or parents of children with ASD.[276] Unlike the ceiling effects of the EQ-5D, the SF-6D suffers from floor effects (wherein the lowest possible health state is 0.345), making it more suitable for examining milder conditions, such as mental health continuum, as opposed to severe conditions, which may favour an instrument with ceiling rather than floor effects. However, this is less of an issue in cohorts such as UKHLS, where individuals with severe and chronic illness might be underrepresented,[277] and reflects an issue with selection into enrolment.

Net monetary benefit

The net monetary benefit (NMB) framework scales HRQoL and resource use to a common unit of costs.[278] It estimates population health impact, including opportunity costs, but is predicated upon establishing a willingness-to-pay (WTP) per outcome. Generally, net monetary benefit combines a threshold per quality-adjusted life year (QALY) with costs to form an aggregate summary statistic which represents the value of an intervention or policy. NMB is a typical method used within economic evaluation, whether RCT or modelling-based, but is rarely applied to observational data. This is likely due to the inapplicability of net benefit regressions outside of randomised settings,[279] which would struggle to disentangle the complex non-linear relationships that inform the net benefit construction. There is considerable heterogeneity in the value of a QALY within governmental departments and between countries,[280] but here I focus on the United Kingdom. There is debate around what a QALY is worth and whether it should be that recommended by NICE (20-30k),[86,87] or the marginal productivity/opportunity cost of the NHS (5-15k).[281–284] Because this study seeks to estimate the relative size of the multi-sectoral value, the monetary benefit was derived using the lower benchmark value of £20k. Because each panel period represents one year (within household sampling), change in HRQoL is assumed to represent an average across one year, which permits valuation using willingness-to-pay per QALY.

Mental health

Because the SF-6D is derived from the SF-12, the respective mental and physical component scores could not be included in modelling due to their linear relation to SF-6D unity. The 12-item General Health Questionnaire (GHQ-12) is used instead to measure mental health. It is one of the most widely used measures to detect psychological distress, with little to no evidence of retest effects in general population samples,[285] or age-related bias,[286] which would otherwise confound the analysis of lifecourse data. Although the GHQ-12 is a well-validated and commonly employed measure of mental health and screening tool for (non-psychotic) mental illness, evidence shows that the latent constructs assessed by the GHQ-12 and common measures to capture mental wellbeing, such as the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS), overlap to a large extent.[111]

Specifically, this study uses the derived GHQ Likert variable where GHQ-12 responses are encoded to a single scale of 0 to 3, summed to produce an index value from 0 (the least distressed) to 36 (the most distressed). The GHQ-12 asks individuals to report recent symptom frequency, not severity. However, there is increasing evidence of equivalence in psychometric and measurement properties of frequency vs severity item response options, and they do not substantially affect the measurement of constructs such as depression and generalised anxiety disorders.[287] The use of the GHQ-12 is intended to tap into the underlying construct outside of the linear relation of instruments (SF-12 to SF-6D); therefore, SF-6D autoregression might be attributable to physical health deviations, which would invalidate the inclusion of the PCS alone.

Variable specification

Conceptualising the production of health to inform modelling uses an applied economic framework. Although it is traditional to follow approaches such as the Grossman model to inform model specifications for generic measures of health or health capital over time,[215] the proposed health deficit model may better align with theory,[288–290] wherein health deficits are productive instead of health stock depreciating which is particularly suitable when examining mental health across the lifecourse.[47,291]

Therefore, PCS and MCS scores were subtracted from 100, and the valued SF-6D (SF-12) index was subtracted from 1. This specification allows health deficits to be self-productive instead of depreciating over time. For this same reason, I did not follow the standard practice of similar studies that inverted the GHQ-12 (a measure of psychological distress) to become a measure of wellbeing.[199] Mental health variables (MCS deficit and GHQ-12) were not transformed to a binary indicator because mental health is increasingly established as a continuum,[108,292] and arbitrarily defining some threshold value as a proxy for diagnosis leaves inference towards underlying constructs questionable.[293] From a statistical perspective such threshold classification also throws away valuable information.

In an ideal world, the model would analyse costs separately across GP visits, outpatient visits, and inpatient stays. However, this increases model complexity dramatically through the required interactions alongside the increased number of instruments. Additionally, healthcare services in the UK are part of an integrated system, with each step often a referral from the previous, posing further endogeneity concerns.

Although conversion to the natural log is a typical strategy for examining non-Gaussian distributed variables,[294–296] annualised cost data were not transformed because these outcomes must retain their cardinal properties for interpretation and comparison following cumulative dynamic estimation (i.e., these data cannot be returned to a natural scale by simply inverting the transformation), and such equal treatment helps avoid specification mining. GMM estimation performs well with non-normally distributed data as it fits around moments of the data, not the entire distribution.

Outliers

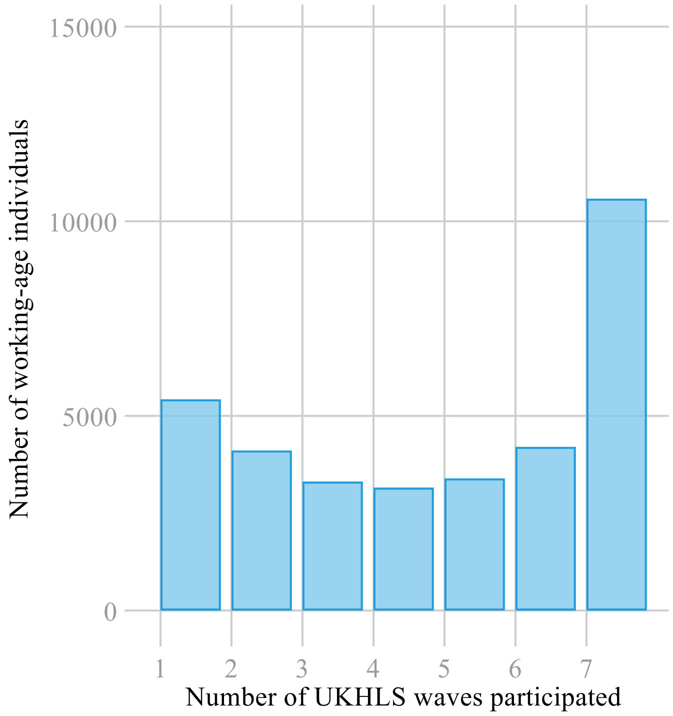
Although GMM estimation is relatively robust to non-normal distributions, and data are transformed to deviations (not levels), the levels of variables are still used for instrumentation. Outliers can skew sample moments, leading to biased or inconsistent parameter estimates if these outliers are not appropriately handled. This is partly addressed using two-step GMM, which inversely weights the covariance matrix of the moment conditions, but outliers can disproportionately affect the covariance matrix, leading to a suboptimal weighting and possibly distorting the GMM estimator.

The annualised cost data are zero-inflated, left-skewed data and feature long right tails, so outlier detection cannot use Mahalanobis distance, which assumes multivariate normal distributions. Therefore, I apply one-sided Winsorization and replace observations greater than the 99th percentile with the 99th percentile for each annualised cost variable, retaining the observation but reducing the impact of extreme values. This substitution affects 1,533 observations of annualised labour tax revenue and welfare benefits. A greater number of observations for annualised healthcare costs were impacted (6,427) because of derived unit costing and heavier tail, i.e., >1% of individuals have healthcare costs >=99th percentile.

Missing data and attrition

UKHLS is an unbalanced panel, and attrition is not the same as item missingness. Unlike birth cohorts with fixed membership, panels change over time and may involve individuals and households leaving, and so others joining, to retain representative properties (Figure J). The initial samples of these harmonised datasets (comparing UKHLS and BHPS) demonstrate representativeness and findings on attrition bias were reassuring (similar patterns across datasets, no association with health status). Some differential in attrition was detected with fewer young people aged 16 to 19 (compared to the 60-69 age group) still participating in UKHLS after six years compared to the BHPS.[297]

Figure J. Count of working-age individuals by number of UKHLS wave participation.



The handling of item-level missing data depends upon mechanisms for missingness, generally broken into the categories of missing completely at random (MCAR), where no relationship exists between the probability of missingness and observed or unobserved data, missing at random (MAR), where this probability depends upon observed data, and missing not at random (MNAR), where this probability depends upon unobserved data (such as the missing value itself). The assumption of MCAR is untenable in observational cohorts, and applied research under MAR is heterogeneous,[298] and typically either do not attempt to deal with missing data or undertake either Multiple Imputation of Chained Equations (MICE), which build into their models the inherent uncertainty associated with the missing data; specifying a separate conditional distribution for each imputed variable,[299–302] or Full Information Maximum Likelihood (FIML) which like MICE is superior to all simple imputation strategies.[303] FIML and MI parameter estimation are equivalent when data missingness is <25%, yet FIML may be more sensitive to model misspecification.[304] FIML is superior to MI in the presence of upper-level dependencies, i.e. households.[305] However, these approaches are limited or are challenging to carry out for dynamic systems.[306] UKHLS already employ methods to deal with item-level missing data, particularly for income-related data (e.g., attributing individual-level benefits as a proportion of household benefits) and is considered the gold standard by the UK government.[272]

Observations for a handful of missing age items were interpolated using preceding or proceeding age responses. Hospital days were set to zero for individuals who either specified they were both not an inpatient in the last twelve months or did not respond to the question but reported 'not applicable' for inpatient hospital bed days. Thirty-four observations of negative tax revenue had their values set to zero.

UKHLS allows people to participate in future waves even if they missed participation in in prior waves, and sample attrition is not always monotonic.[307] Therefore, one advantage of the employed FOD method is that it includes more observations than traditional FD because it does not require the presence of future observations at every time point and is more robust to gaps in the data; thus, using most available case data where any future observation is available. This transformation is also beneficial as estimates will not be biased if non-responders differ in time-invariant characteristics; rather, the assumption narrows to being that the relationship between changes in exposure and changes in other outcomes is the same. Additionally, comparing the share of multi-sectoral costs by MCS/PCS narrows this down to the differences in exposure to relative change between other outcomes within individuals. Discussion of the relevance of representativeness can be found elsewhere.[308,309]

Alternatively, the problem of missing observations when using longer lags as instruments can be circumvented by using GMM-style instruments,[310] which can permit all relevant lags to be used as instruments. GMM-style instruments substitute zeros for missing observations (for all missing data whether attributable to attrition or item non-response). However, it can be problematic in unbalanced panels (wherein many instruments would be created), particularly in small T settings, and can increase the likelihood of instrument proliferation because it may overfit endogenous variables. The GMM estimates would approach those from OLS, fail to expunge their endogenous components and impair the validity of Hansen's overidentification statistic.[229]

The survey is representative when using all data in the mainstage dataset, but longitudinal weights are recommended to address attrition and ensure representativeness; however, such weights are incompatible with GMM estimation and, therefore, panel vector autoregressions. Therefore, to demonstrate the potential of observational data to estimate multi-sectoral cost profiles, this study follows the practice of similar studies using this dataset, which analyses data as provided (inclusive of data imputed by the UKHLS team).[199]

Modelling strategy

Maintained statistical model (MSM)

This paper's data management and analyses used Stata/MP.[311] Panel vector autoregressions and subsequent IRFs were implemented using the st0455 package by Abrigo and Love.[234] Their guidance provides a thorough and detailed overview for undertaking PVAR,[234] alongside Roodman's pedagogical paper for GMM,[229] applied GMM guidance by Kripfganz,[239] and both Andrews & Lu's,[312] and Kiviet's,[313] guidance for model specification and selection.

I present the estimated system of equations below (equation 1), where MCS_{it} = MCS deficit, PCS_{it} = PCS deficit, HC_{it} = annualised healthcare costs, WB_{it} = annualised welfare benefits, and TR_{it} = annualised tax revenue for individual *i* at time *t* using the (later defined) optimal maximum lag length L. C_i represents a vector of time-invariant controls, i.e., sex and education, and X_{it} is a vector of time-variant exogenous controls, i.e., age and time dummies. The idiosyncratic error term is represented by ε_{it} and variable-specific panel fixed effects errors are represented by u_i.

(1)

$$\begin{cases} MCS_{it} = \alpha + \sum_{t-L}^{t-1} b_{it}\, MCS_{it} + \sum_{t-L}^{t-1} b_{it}\, PCS_{it} + \sum_{t-L}^{t-1} b_{it}\, HC_{it} + \sum_{t-L}^{t-1} b_{it}\, WB_{it} + \sum_{t-L}^{t-1} b_{it}\, TR_{it} + b_{it}.C_i + b.X_{it} + \varepsilon_{it} + u_i \\ PCS_{it} = \alpha + \sum_{t-L}^{t-1} b_{it}\, MCS_{it} + \sum_{t-L}^{t-1} b_{it}\, PCS_{it} + \sum_{t-L}^{t-1} b_{it}\, HC_{it} + \sum_{t-L}^{t-1} b_{it}\, WB_{it} + \sum_{t-L}^{t-1} b_{it}\, TR_{it} + b_{it}.C_i + b.X_{it} + \varepsilon_{it} + u_i \\ HC_{it} = \alpha + \sum_{t-L}^{t-1} b_{it}\, MCS_{it} + \sum_{t-L}^{t-1} b_{it}\, PCS_{it} + \sum_{t-L}^{t-1} b_{it}\, HC_{it} + \sum_{t-L}^{t-1} b_{it}\, WB_{it} + \sum_{t-L}^{t-1} b_{it}\, TR_{it} + b_{it}.C_i + b.X_{it} + \varepsilon_{it} + u_i \\ WB_{it} = \alpha + \sum_{t-L}^{t-1} b_{it}\, MCS_{it} + \sum_{t-L}^{t-1} b_{it}\, PCS_{it} + \sum_{t-L}^{t-1} b_{it}\, HC_{it} + \sum_{t-L}^{t-1} b_{it}\, WB_{it} + \sum_{t-L}^{t-1} b_{it}\, TR_{it} + b_{it}.C_i + b.X_{it} + \varepsilon_{it} + u_i \\ TR_{it} = \alpha + \sum_{t-L}^{t-1} b_{it}\, MCS_{it} + \sum_{t-L}^{t-1} b_{it}\, PCS_{it} + \sum_{t-L}^{t-1} b_{it}\, HC_{it} + \sum_{t-L}^{t-1} b_{it}\, WB_{it} + \sum_{t-L}^{t-1} b_{it}\, TR_{it} + b_{it}.C_i + b.X_{it} + \varepsilon_{it} + u_i \end{cases}$$

Equation 1 can be condensed to a single system representation, and FOD transformation removes the effects of time-invariant controls C, and the variable-specific panel fixed effects errors u_i.

$$O_{it}^{*} = \alpha + \sum_{t-L}^{t-1} b_{it}\, O_{it}^{*} + c.X_{it} + \varepsilon_{it} \qquad (2)$$

Equation 2 presents the reduced-form model of an L-order (L lags of each variable in the system) panel VAR, and the L lags of transformed outcomes are instrumented by to be specified untransformed levels, where O is a vector of endogenous outcome variables (MCS deficit, PCS deficit, annualised healthcare costs, annualised welfare benefits, and annualised tax revenue), with b representing their coefficients. X is a vector of assumed exogenous controls (age and time dummies), c are the control variable coefficients, ε is the idiosyncratic error term. Following standard notation, * is used to indicate the variable or vector transformed during estimation, i.e., the orthogonal deviation of the MCS deficit is represented in equation 3.

$$MCS_{it}^{*} = (MCS_{it} - \overline{MCS_{\iota t}})\sqrt{T_{it}/(T_{it} + 1)} \tag{3}$$

Where T_{it} is the number of available future observations for individual *i* at time *t,* and $\overline{MCS_{\iota t}}$ is the average of all available future observations of the MCS deficit.

Assumption testing

In GMM estimation of linear dynamic panel models, it is common to test each variable for non-stationarity using unit-root tests, meaning that its statistical properties (moments), such as mean and variance, change over time (e.g., one such form of unit root is where the autoregressive coefficients of the lagged value of the series are equal to 1, such that reversion to the mean is improbable, therefore leading to non-stationarity). Data must be strictly stationary when analysing macro panels (T>N). However, most unit root tests only suit large T and small N samples,[314] and in this micro-econometric panel (small T, large N), [non]stationarity is not as much of an issue. As discussed by Woolridge:

"*we do not need to restrict the dynamic behavior of our data in any way because we are doing fixed-T, large-N asymptotics…. Here, a large cross section and relatively short time series allow us to be agnostic about the amount of temporal persistence*.".[315]p.175

To err on the side of caution, if the panel GMM model includes individual fixed effects or time-specific effects that capture the unobserved heterogeneity, then the issue of non-stationarity due to unit roots can be addressed without explicitly testing for it. This is because fixed and time-specific effects can help remove the effects of any trends or time-specific shocks that may be present in the data, ensuring that inference of a lag on current outcomes is comparable no matter which specific year it is.

Because of computing power constraints, I did not include a time dummy for each year (increasing the number of dummies exponentially increases the time needed to estimate barely identified GMM models and can lead to models failing to converge).

Furthermore, in unbalanced panels, time dummies are only time-variant for individuals who appear in the dataset in the reference period, typically the first year in the sample, e.g., 2015. Individuals who appear after the reference period (>=2016) will have a time-invariant value (1) for all their time dummies across later periods. This lack of variation can lead to biased estimates and diminished capacity to control for period-specific heterogeneity in the model.

Instead, I specified time dummies for periods with a strong prior belief of structural breaks, i.e., known impacts on outcomes of interest. COVID-19 and subsequent lockdowns significantly impacted healthcare resource use, public health, and the broader economy (confirmed by simple OLS regressions). Therefore, I specified two dummies using the year of sampling unit interview, one representing the entire length of COVID lockdown periods across the UK between March 2020 and October 2021, and one representing the post-lockdown period (and transition therein) from November 2021 to the end of Wave g (7) in 2023 (only 573 interviews in 2023). These dummies can be treated as strictly exogenous.[239] The reference period is the five years prior, thus ensuring that the time dummies will be time-variant for all individuals. Similarly, because these data represent individuals, not firms, an individual's age might contribute to non-stationarity. This is because dynamic modelling requires a certain number of observations; therefore, the estimation sample age of this study must inevitably trend upwards over time, even if the sample in each wave is representative. As discussed, age is included as an exogenous control variable.

Other forms of unit root can be present, such as if the variables exhibit a random walk, are problematic in typical implementations of time series analysis, e.g., pricing assets or stocks, but can be disregarded on theoretical grounds in this application to individual health and resource use, e.g., the established characteristics of the variables examined in this study, such as autoregression or reversion to long term states, contradict the assumptions of a random walk.

Cross-sectional independence is a frequently overlooked assumption, as GMM estimators also assume that idiosyncratic disturbances are uncorrelated across individuals. This is relevant for variables such as welfare benefits, which are often allocated on a household basis (although adjusted for by the UKHLS team),[272] and to allow for intra-group correlation where spillovers may exist, which otherwise would result in model misspecification and unadjusted heteroscedasticity.[245] Although time dummies make this assumption more likely to hold,[229] the MSM also clusters by the time-invariant primary sampling unit (PSU) i.e., household, to which each participant belongs.[249] This represents the highest level of aggregation; these individuals will always relate over time. This analysis does not account for stratified sampling, which will not affect estimated coefficients but might slightly overestimate standard errors.[249] This clustering also leads to standard errors that are [more] robust to heteroscedasticity and autocorrelation.

Optimal model moment and selection criteria (MMSC)

The validity of PVAR estimates is dependent on choosing the optimal lag order in both model specification and moment condition.[234] Selection-order statistics, used to identify optimal moments and model lag order, were estimated using repetitions of the *Pvarsoc* subcommand with different lags specified as instruments.

A trade-off exists between lag, instrument length, and t examination periods. Efficiency is gained through more moments (instruments) but leads to reduced observations and an average number of t per panel. For example, second-order FOD estimation instruments the t-1 and t-2 lags with the same lags in levels (that is, untransformed); therefore, because this study uses 7 waves of data, specifying the 4^{th} or beyond lag as an instrument (at most, i.e., presuming the same level as the instrument) reduces the average number of t per panel to 2 (assuming no missing observations).

GMM requires equal or more moment conditions than parameters; otherwise, the model is not identified and has no unique solution. I tested instrument combinations 1&2, 2&3, 1-3, and 2-4 to maximise sample size in this small-T panel. I specified the maximum order tested for each subcommand as the number of instruments minus 1 (Table 4). This restriction avoids estimating just-identified models where the number of instruments equals the number of endogenous variables i.e., lags. Although such higher-order models can be estimated, their instrument validity cannot be tested (MMSC statistics). Therefore, I do not consider them for selection.

Table 4. Outline of model identification by order and instruments

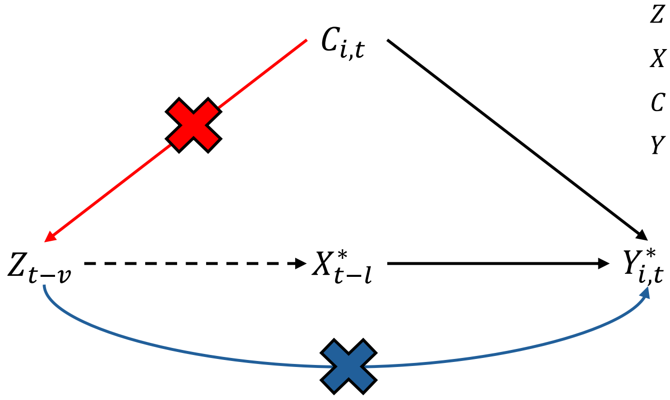
Order	Lags used as instruments					
	1	2	1 - 2	2 - 3	1 - 3	2 - 4
1	Just-identified	Just-identified	Identified	Identified	Identified	Identified
2	Unidentified	Unidentified	Just-identified	Just-identified	Identified	Identified
3	Unidentified	Unidentified	Unidentified	Unidentified	Just-identified	Just-identified

Of these identified models, initial optimal selection traditionally follows the two following principles to derive optimal values *v* and *l* in Figure K:

- Does not reject the hypothesis (using Hansens J statistic) that the chosen instruments are valid, and the model is not overidentified. Rejection of the test (J pvalue <0.05) indicates model misspecification, or crucially, that some instruments are correlated with the model residuals. For clarity, I note that the statistic might suggest, but cannot explicitly test, violations of exchangeability (assumption B). Assumption C also cannot be explicitly tested, and in this setting, it is feasible that longer lags are correlated with current values controlling for shorter lags. However, this is why this study employs the GMM method, which offers many other analytical strengths. More robust instruments are not available, and longer lags are more likely to relate to closer time points of independent variables.

- Minimises the MMSC, which are adjusted for degrees of freedom, as proposed by Andrews and Lu:[312] Modified Bayesian Information Criterion (MBIC), Modified Akaike Information Criterion (MAIC), and Modified Quasi-likelihood Information Criterion (MQIC).

Figure K. Assumptions in instrumental variable analyses



Z = Variable levels used as instruments
X = FOD-transformed independent variables
C = Confounders (Model residuals)
Y = FOD-transformed dependent variable

A. Z is not independent of X (relevance)
B. Z is independent of C (exchangeability)
C. Z is independent of Y conditional on X and C (exclusion restriction)

Table 5. Order-selection criteria of panel VAR models by lag length of the endogenous variables as instruments.

Instrument lengths	Order	CD	J	J pvalue	MBIC	MAIC	MQIC
Lags 1 to 2 (sample 3 to 6) n=53007 (n panels=18124 & $\bar{T}$=2.93)	1	0.999	197.55	9.08E-29	-74.40	147.55	78.21
Lags 2 to 3 (sample 4 to 6) n=34883 (n panels=14421 & $\bar{T}$=2.42)	1	0.964	20.38	0.727	-241.12	-29.62	-97.00
Lags 1 to 3 (sample 4 to 6) n=34451 (n panels=13989 & $\bar{T}$=2.46)	1	1.000	351.26	7.34E-47	-171.10	251.26	116.63
	2	0.999	82.84	4.07E-08	-178.34	32.84	-34.48
Lags 2 to 4 (sample 5 to 6) n=20462 (n panels=11285 & $\bar{T}$=1.81)	1	0.998	63.89	0.090	-432.43	-36.11	-184.62
	2	0.997	11.01	0.993	-237.15	-38.99	-105.35

Order = length of lags, CD = coefficient of determination (roughly equivalent to R-squared in OLS), J = Hansens overidentification statistic, MBIC/MAIC/MQIC = model fit statistics.

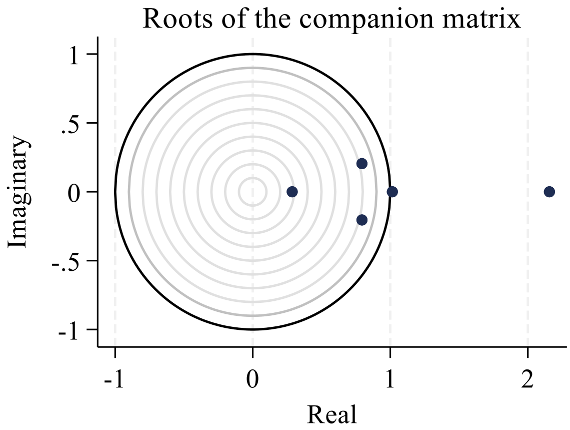
MMSCs are presented for each model in Table 5. The three models highlighted in red reject Hansen's overidentification restriction at the 5% alpha level, and the model highlighted in orange rejects Hansen's test at the 10% level. The two models in green, which report more acceptable J pvalues, also report acceptable values across all MMSCs.

Stability

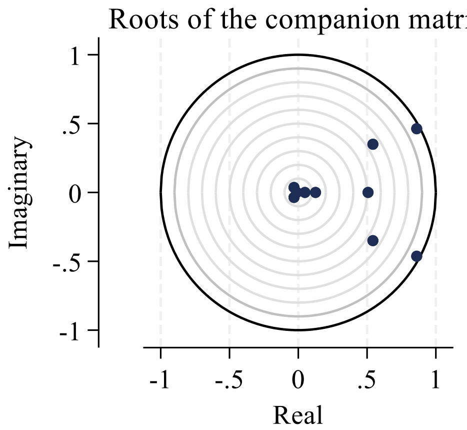
The chosen model must also be dynamically stable, and its solutions converge over time, i.e., it must be valid in understanding long-term relationships among variables. A model is considered stable if the modulus of each eigenvalue of the fitted model is less than 1. The postestimation command *Pvarstable* was used to test the stability condition of the two acceptable models (Figure L).

Figure L. Eigenvalue stability conditions by tested model.

First-order model with lags 2 to 3 as instruments



Second-order model with lags 2 to 4 as instruments



Only the second-order model using 2 through 4 lags as instruments had all eigenvalues in the unit circle (i.e., values <1), satisfying the stability condition. The second-order model also aligns with the literature that suggests that two lags and longer may be necessary if a regressor is endogenous,[229] and are sufficient to capture the persistence of the dependent variable.[199,228]

Causality tests

The *Pvargranger* command was then used to produce Granger causality statistics for the acceptable model, presented in Table 6, under the hypothesis that all coefficients on the lag of variable x are jointly zero in the equation for variable n, i.e., that variable x does not Granger-cause variable n). However, I note that these tests target predictability, not causation, in a traditional sense.

Table 6. Granger causality statistics

Equation	Excluded independent variable	chi2	df	p> chi2
Physical health deficit	Mental health deficit	1.92	2	0.38
	Annualised labour tax revenue	3.85	2	0.15
	Annualised welfare benefits	2.24	2	0.33
	Annualised healthcare costs	0.90	2	0.64
	ALL	10.86	8	0.21
Mental health deficit	Physical health deficit	0.65	2	0.72
	Annualised labour tax revenue	1.14	2	0.57
	Annualised welfare benefits	2.82	2	0.25
	Annualised healthcare costs	0.59	2	0.75
	ALL	5.78	8	0.67
Annualised labour tax revenue	Physical health deficit	2.32	2	0.31
	Mental health deficit	4.16	2	0.13
	Annualised welfare benefits	2.32	2	0.31
	Annualised healthcare costs	6.62	2	0.04
	ALL	10.05	8	0.26
Annualised welfare benefits	Physical health deficit	2.05	2	0.36
	Mental health deficit	0.77	2	0.68
	Annualised labour tax revenue	2.28	2	0.32
	Annualised healthcare costs	2.39	2	0.30
	ALL	4.69	8	0.79
Annualised healthcare costs	Physical health deficit	0.52	2	0.77
	Mental health deficit	0.22	2	0.89
	Annualised labour tax revenue	0.08	2	0.96
	Annualised welfare benefits	3.82	2	0.15
	ALL	4.11	8	0.85

I cannot reject the hypothesis that the considered variables do not Granger-cause each other. Although the model is validated in terms of instrument appropriateness and overall specification, and so a strong foundation for inference, the lack of rejection in the Granger causality tests implies that within my analysis, there is insufficient evidence to claim that one variable causes another in the Granger sense (prediction). This does not mean there is no causal relationship. Instead, this method does not detect Granger-causality in this sample. I could consider whether additional or different lags of the variables would impact Granger causality results, but this would impinge on the validity of the instruments considered. There is no guarantee that a model that passes all specification tests can be found, and judgement is needed. Lack of rejection is not a reason to exclude individual variables; they should be included if they are of interest and there are strong theoretical arguments. The second-order model using 2 to 4 lags as instruments was formalised as the final specification.

Impulse-response functions

Impulse-response functions (IRFs) are typically used in economics to estimate the result of shocks in vector autoregressive models. It is particularly useful to examine the value of improving health: it uses the fitted system to describe the evolution of variables over a specified time-horizon, where otherwise estimating the total effect of a shock to such a complex system is difficult, i.e., it introduces a shock in the variable of interest and allows it to propagate back and forth through the estimated system over several iterations.

Generally, causality cannot be inferred from simple IRFs because any variable shock is likely contemporaneously accompanied by a shock in other variables.[234] The Cholesky decomposition is a way to produce orthogonalised IRFs (OIRFs); however, it is not unique and depends on the ordering of variables. Cholesky ordering does not affect PVAR estimates but does affect the post-estimation of OIRFs (variables earlier in Cholesky order are allowed to affect the later ones contemporaneously, but not vice versa); therefore, such ordering should be based on solid theoretical ground, i.e., causal assumptions about the timing of relationships between variables.[234] The validity of OIRFs depends on strong identification strategies to decide the Cholesky Order. However, there is no clear Cholesky order for the considered system.

Conversely, simple IRFs do not require such assumptions and are not subject to misspecification in this way. Simple IRFs are also valid when shocks are assumed independent, in this study's case, that of targeting the relative value of exogenous intervention on mental or physical health.

Therefore, this study estimates (using the *Pvarirf* subcommand) simple IRFs and cumulative IRFs using 100-iteration Monte Carlo draws estimated 95% confidence intervals using a Gaussian approximation. It should be emphasised that the relative size and shape of IRF estimates represent this study's target of interest, and it is not to examine what a change in health is worth or test the null hypothesis (which is indicated by significance).[316] IRFs ran for three steps (three-year forecast horizon) because this small-T panel model may not accurately reflect more extended temporal dynamics. For visual comparison, tabulated cumulative IRF estimates were line fit on a single graph per specified impulse. Graphs and estimates are interpretable as the cost of decrements or the value of improvements in the MCS or PCS, respectively. All models' tax revenue was multiplied by -1 to equate it to a governmental opportunity cost, i.e., all other costs are sought to be minimised, whereas tax revenue would be maximised.

Sensitivity analyses

Further fixed-effects panel analyses using the maintained statistical model were estimated for comparison with GMM estimates. Model estimation used the xtreg command and followed the same specification for each equation within the finalised GMM model. This ensured the preservation of temporal ordering to address simultaneity, where changes in the independent variables precede changes in the dependent variable. Like the GMM model, fixed-effects estimation addresses time-invariant heterogeneity but does so through within-unit estimation, making it particularly suitable for unbalanced data.[315] However, it cannot include lags of dependent variables, as this would violate endogeneity assumptions. The differences in estimates between the GMM and fixed-effects models are attributable to the absence of lagged dependent variables, FOD transformation, and instrumentation. Cluster robust standard errors were generated at the primary sampling unit level.

The correlation between u_i (unobserved individual effects) and Xb (explanatory variables) was low across all models, suggesting that random-effects could be considered for increased efficiency. However, the Hausman test (using non-cluster robust estimates) rejected the null hypothesis that the unobserved individual effects are uncorrelated with the explanatory variables, meaning that the difference in coefficients between the fixed-effects (FE) and random-effects (RE) models are systematic. This implies that random-effects models, and population-averaged panel-data models by using generalised estimating equations, are inconsistent due to systematic differences in coefficients. Therefore, the fixed-effects models, which permit correlation between unobserved individual effects and the explanatory variables, are preferable for consistent results.

Results

Descriptive statistics by UKHLS wave are available in the Appendix (Tables A1-A7). The main results of the panel vector autoregression are reported in Table 7. The size of coefficients might appear extremely variable at first glance, so it is important to remember that extreme discordance exists between the likely values that would be multiplied against these coefficients. For example, a change in health will be bounded <1, and a change in any of the costs may range into the thousands.

Only a few coefficients are significant at the 5% level, and these all occur at the t-1 time point, where past increases in physical health deficits are significantly correlated with later increases in physical health deficits [0.59 (CI 0.05, 1.12)] and decrements in mental health reduce future tax revenue [-468.07 (CI -920.39, -15.74)]. Tax revenue [0.724 (CI 0.4, 1.04)], welfare benefits [0.746 (CI 0.3, 1.2)], and healthcare costs [0.82 (CI 0.01, 1.64)] all demonstrate significant autoregression. Age also significantly increases physical health deficits, whereas the lockdown periods correspond to improvements in physical health.

*Table 7. Results of a second-order panel vector autoregression (two-step GMM) using lags 2-4 as instruments.**

	ΔPhysical health (PCS deficit)	ΔMental health (MCS deficit)	ΔTax revenue (£)	ΔWelfare benefits (£)	ΔHealthcare costs (£)
ΔPhysical health (PCS deficit)					
t- 1	**0.59**	-0.27	-419.87	-77.54	9.45
	0.05, 1.12	-1.01, 0.47	-971.66, 131.94	-334.04, 178.96	-72.4, 91.39
t- 2	0.05	0.02	9.75	29.28	-4.42
	-0.03, 0.13	-0.09, 0.13	-53.99, 73.48	-16.79, 75.34	-17.22, 8.37
ΔMental health (MCS deficit)					
t- 1	0.12	0.53	**-468.07**	17.44	-14.35
	-0.40, 0.64	-0.17, 1.24	-920.39, -15.74	-249.47, 284.36	-89.39, 60.68
t-2	0.03	-0.01	24.4	10.22	-0.61
	-0.02, 0.08	-0.08, 0.06	-21.42, 70.23	-14.88, 35.31	-8.21, 7
ΔTax revenue (£)					
t-1	2.2E-04	4.2E-05	**0.724**	-0.089	-0.004
	-2.8E-05, 4.7E-04	-3E-04, 3.8E-04	0.4, 1.04	-0.21, 0.03	-0.04, 0.03
t-2	3.6E-07	2.7E-05	-0.017	0.003	-0.0003
	-4.9E-05, 5E-05	-3.9E-05, 9.6E-05	-0.1, 0.06	-0.02, 0.02	-0.01, 0.01
ΔWelfare benefits (£)					
t-1	5.1E-04	-8.0E-04	0.395	**0.746**	0.103
	-1.6E-04, 1.9E-03	-1.7E-03, 1.9E-04	-0.16, 0.95	0.3, 1.2	-0.01, 0.22
t-2	-2.8E-05	2.2E-05	0.016	0.064	-0.012
	-1.8E-04, 1.1E-04	-1.7E-04, 2.1E-04	-0.09, 0.12	-0.04, 0.17	-0.03, 0.01
ΔHealthcare costs (£)					
t-1	-1.2E-03	-2.1E-03	0.96	-1.37	**0.82**
	-6E-03, 3.6E-03	-8.8E-03, 4.5E-03	-2.86, 4.79	-4.28, 1.52	0.01, 1.64
t-2	7.1E-05	2.5E-04	0.17	0.33	-0.08
	-7.5E-04,8.9E-04	-8.7E-04, 1.4E-03	-0.48, 0.84	-0.17, 0.84	-0.22, 0.06
Age					
	1.06	1.05	-492.47	250.71	-34.06
	0.01, 2.1	-0.39, 2.49	-1,533, 548	-303, 805	-189, 121
Lockdown					
	-1.63	-0.81	475.78	-155.25	-119.77
	-2.95, -0.31	-2.63, 1.01	-847, 1,799	-836, 525	-310, 71
Post lockdown					
	-2.69	-4.37	2027.04	-1353.05	38.27
	-7.01, 1.62	-10.23, 1.48	-1,978, 6,032	-3,669, 963	-599, 675

** n = 20462 (n panels = 11285 & $\bar{T}$ = 1.81). The correlation coefficients between past changes in independent variables (rows) at t-1 or t-2 and changes in dependent variables (columns) at time t. 95% confidence intervals are presented immediately below the fitted coefficients for each time lag. P-values were derived from Wald tests (using z-statistics), and coefficients significant at the 5% level are reported in bold.*

Sensitivity analyses

The estimates of the fixed-effects panel sensitivity analyses (Table 8) confirm the significant lagged correlations between all variables, aside from tax revenue and welfare benefits on health, and healthcare costs on the other economic variables. All models fit the data adequately with significant overall explanatory power (Prob > F = <0.0001). However, the R-squared indicated that the fixed-effects models explained little variation in the dependent variable. Therefore, the unobserved individual heterogeneity substantially contributes to variation.

Table 8. Results of five panel fixed-effects linear regression models.

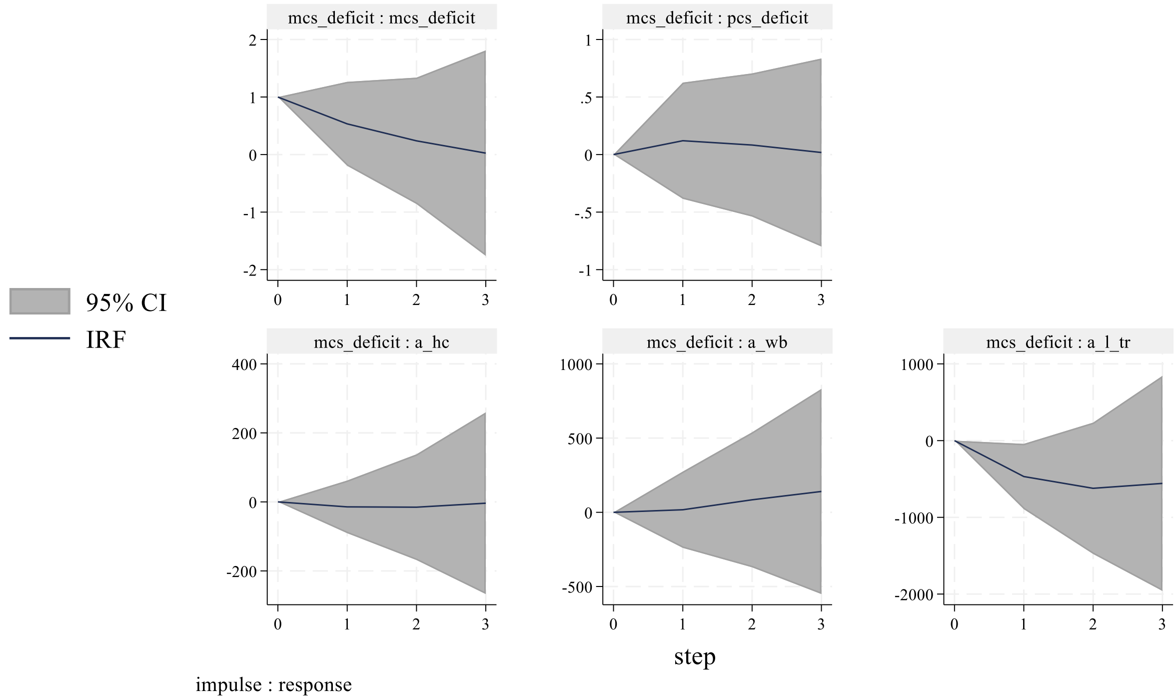
Variable	(1) Physical health (PCS deficit)	(2) Mental health (MCS deficit)	(3) Tax revenue (£)	(4) Welfare benefits (£)	(5) Healthcare costs (£)
Physical health (PCS deficit)					
t-1	-	**0.081**	**-15.9**	**13.9**	**8.81**
	-	(0.0070)	(3.75)	(2.78)	(0.85)
t-2	-	**0.066**	-5.56	2.74	**-4.40**
	-	(0.0067)	(3.50)	(2.70)	(0.78)
Mental health (MCS deficit)					
t-1	**0.062**	-	-0.24	**5.43**	**2.85**
	(0.0042)	-	(2.90)	(1.95)	(0.56)
t-2	**0.052**	-	0.79	**4.67**	**-2.09**
	(0.0039)	-	(2.73)	(1.87)	(0.54)
Tax revenue (£)					
t-1	-5.1e-06	**0.000017**	-	**-0.0050**	**0.0020**
	(5.1e-06)	(6.8e-06)	-	(0.0018)	(0.00062)
t-2	-1.5e-06	9.9e-06	-	**-0.0034**	0.00064
	(5.1e-06)	(6.6e-06)	-	(0.0016)	(0.00060)
Welfare benefits (£)					
t-1	3.8e-06	-3.1e-06	**-0.021**	-	**-0.0051**
	(0.000012)	(0.000014)	(0.0055)	-	(0.0016)
t-2	-7.8e-06	-9.9e-06	**-0.018**	-	**-0.0060**
	(0.000012)	(0.000014)	(0.0054)	-	(0.0016)
Healthcare costs (£)					
t-1	-0.000047	**-0.00016**	-0.051	0.014	-
	(0.000044)	(0.000050)	(0.027)	(0.022)	-
t-2	**-0.00019**	**-0.00016**	-0.033	0.029	-
	(0.000042)	(0.000049)	(0.026)	(0.022)	-
Age	**0.28**	**0.20**	**-50.7**	**-74.9**	-0.78
	(0.031)	(0.040)	(24.6)	(13.2)	(3.97)
Lockdown	**-0.73**	**0.26**	**-197**	**186**	**-101**
	(0.076)	(0.10)	(60.7)	(32.5)	(10.3)
Post lockdown	**-0.27**	-0.28	2.87	**169**	**-91.9**
	(0.13)	(0.17)	(107)	(54.4)	(16.7)
Observations	73,784	73,784	75,702	75,702	75,702
Number of panels	21,512	21,512	22,038	22,038	22,038
R-squared	0.073	0.010	0.045	0.037	0.013

Robust standard errors in parentheses. Coefficients significant at the 5% level are reported in bold.

Impulse response functions

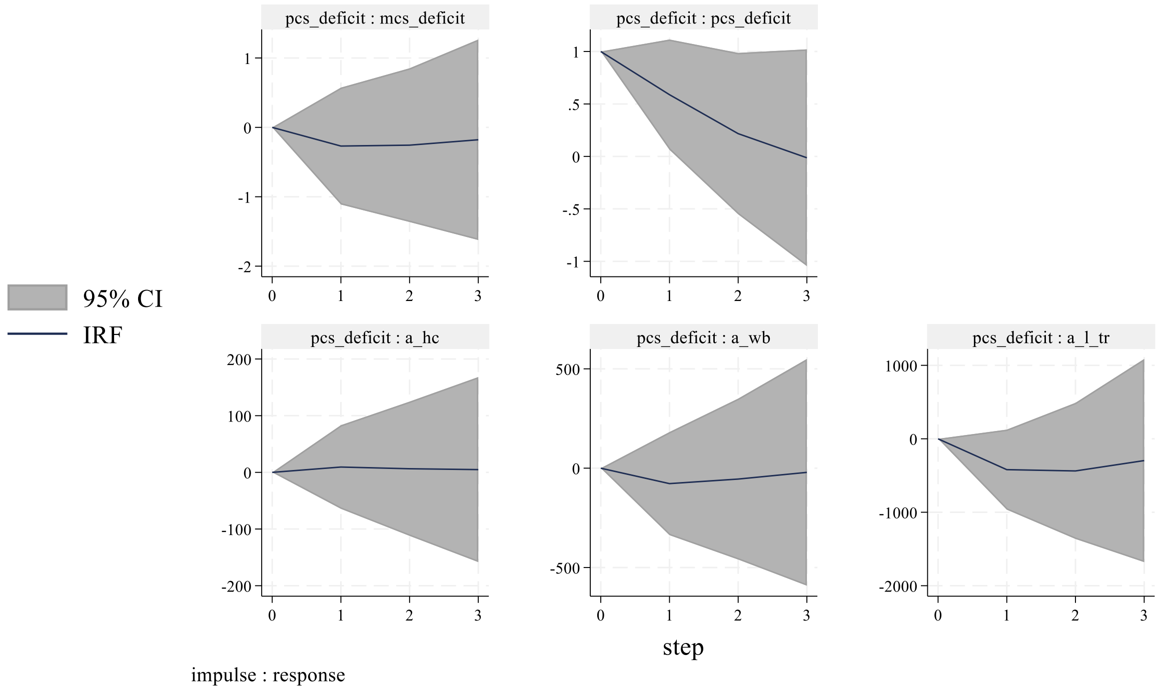
The results of impulse response functions are shown in Figures M, N, O, and P. Figures M and N show the change in the impacts of the impulse variable (health) on the response variables (all other outcomes) across time. For example, the MCS deficit was specified as the impulse in Figure M, so the coefficient of the MCS on itself at t=0 is equal to one; then, we observe the diminishing impact (on average) of the increase in MCS deficit across the three steps (years).

Figure M. Impulse response function estimates using an impulse of one unit increase in MCS deficit.



IRF confidence intervals are substantial; therefore, the discussed results focus on mean trends. An increase in MCS deficit (Figure M) leads to an increase in physical health deficit, exhibiting diminishing returns over time. Healthcare costs appear relatively insensitive to changes in mental health, whereas welfare benefits show modest increases, trending upwards over time. Tax revenue experiences a negative impact initially but stabilises subsequently.

Figure N. Impulse response function estimates using an impulse of one unit increase in PCS deficit.



An increase in the PCS deficit (Figure N) further increases the physical health deficit, which diminishes over time, and slightly decreases the mental health deficit. Healthcare costs appear insensitive to changes in physical health. Welfare benefits slightly decrease before trending towards zero, and tax revenue is initially reduced, but this impact diminishes over time.

This study aims to examine the relative size of spillovers, and while Figures M and N provide insight into these relationships over time, comparisons between perspectives are inherently concerned with the total value attributable to changes—Figures O and P present cumulative estimates.

Figure O. Cumulative monetary benefit estimated by cumulative impulse response functions using an impulse of one unit increase in MCS deficit.

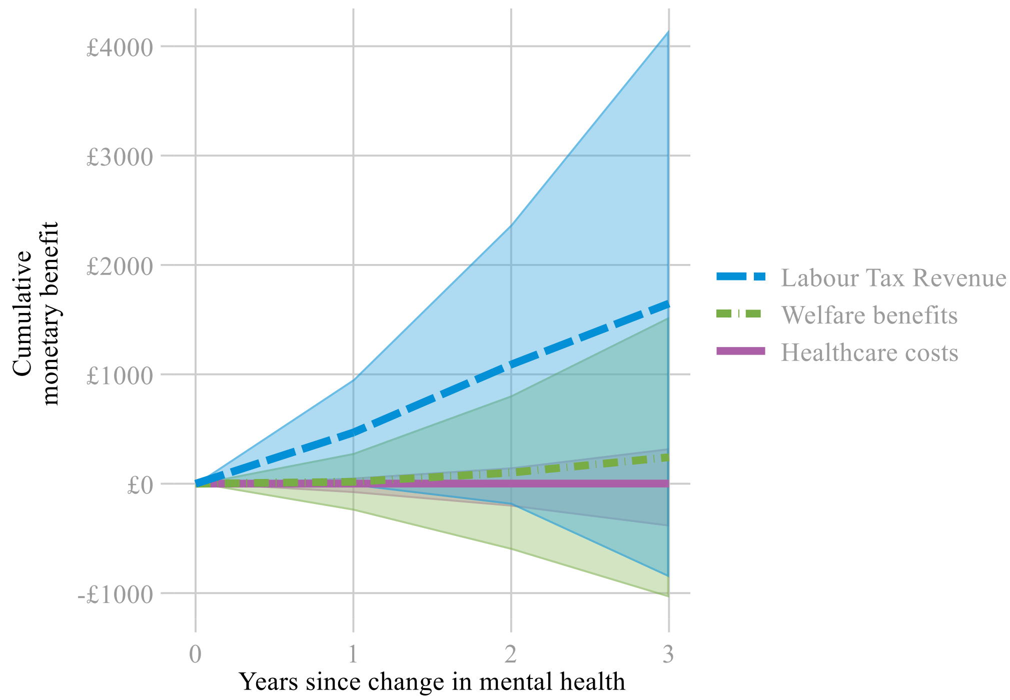
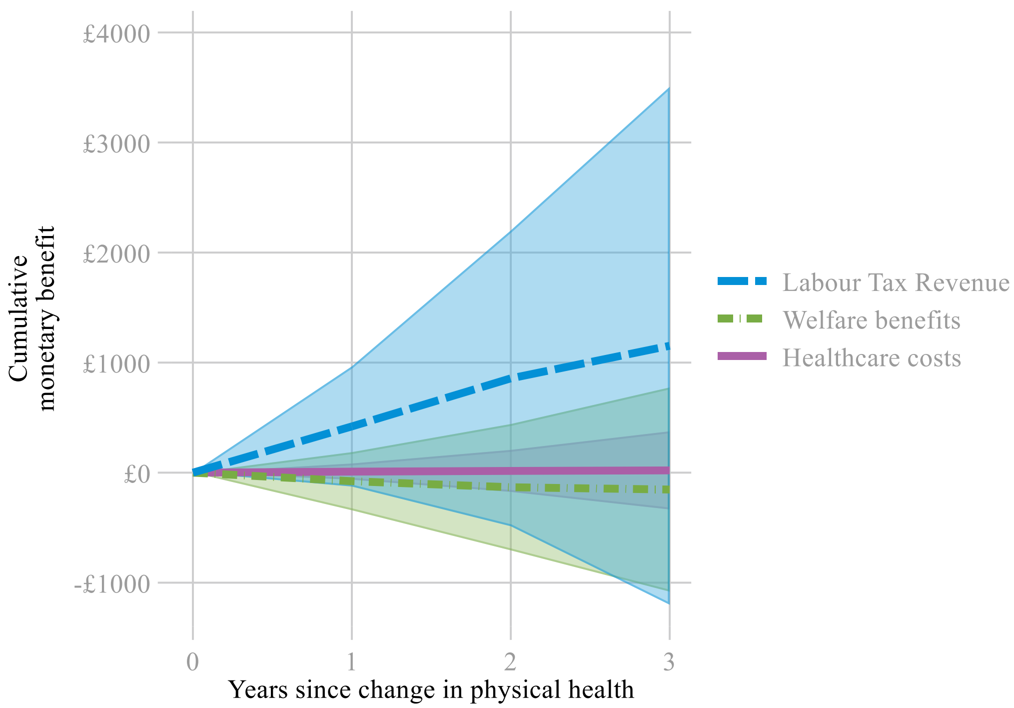


Figure P. Cumulative monetary benefit estimated by cumulative impulse response functions using an impulse of one unit increase in PCS deficit.



Cumulative IRFs for mental and physical health deficit impulses (Figures O & P) indicate a larger attributable value (on average) to changes in labour tax revenue compared to healthcare costs and welfare benefits. Cumulative healthcare cost estimates tend to be null for both health domains, while mental health deficits might increase welfare benefits over time.

Secondary analyses

Because of the lack of power of the maintained statistical model, further analyses introducing the measure of HRQoL are not informative. I use the planned secondary analysis variables (HRQoL and lnGHQ) to present the visual utility of such multi-sectoral cost profiles under less uncertainty, using a just-identified model to examine a healthcare perspective (Figure Q). The PVAR results used to estimate the OIRFs are available in the Appendix (Table A8), but inference should not be made from either of these graphs or results.

Figure Q. Back-transformed by $\hat{\beta}\ ln(1.01)$ cumulative monetary benefit and share of mean cumulative monetary benefit estimated by cumulative orthogonalised impulse response functions using an impulse of one percentage increase in mental health deficit (ln(GHQ deficit+1)).

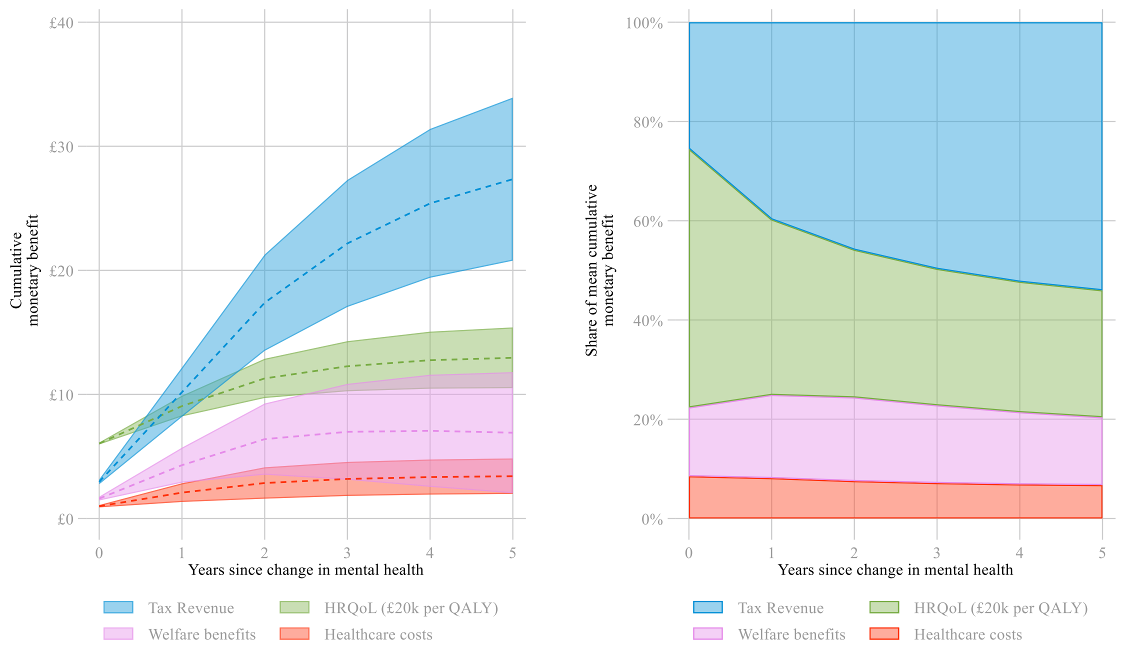


Figure Q demonstrates the potential for dynamic approaches to present the absolute and relative value of improving health across sectors without the need for complex results tables or metrics such as incremental cost-effectiveness ratios. Bootstrapping the mean cumulative monetary benefit share could also generate percentile-based estimates of the frequency of cost shares, akin to current applications of non-parametric cost-effectiveness acceptability curves.

Discussion

This study examined the relative size of multi-sectoral costs attributable to mental and physical health changes. To this end, I analysed the dynamic relationships between mental health, physical health, healthcare costs, welfare benefits, and labour tax revenue (proxying for employment capacity) variables in the UKHLS panel study. I employed methods that address several sources of endogeneity biases and capture cumulative indirect effects. The PVAR results were broadly consistent with existing literature, e.g., the significant autoregression of physical health deficit, labour tax revenue, welfare benefits, and healthcare costs, alongside the impact of mental health deficits on reduced tax revenue.[317]

The mean results of IRFs were also sensible, where increases in MCS deficits increase welfare benefit receipts over time, while physical health deficits showed minimal impact, which is in line with the diagonal accounting issues of the costs for mental health existing across other sectors.[24] Both types of health deficit reduced labour tax revenue, aligning with the established relationship between individual health and employment capacity.

Healthcare costs appeared relatively insensitive to changes in health. This might appear surprising in a dynamic model, but several plausible explanations exist. Whether healthcare or welfare, government spending will not necessarily lead to positive changes in outcomes. In many cases, such spending reduces further declines in outcomes. There could also be a lack of awareness of one's health state or available services. Mental health might be subject to stigma and act as a barrier to seeking care. Affective traits associated with mental health problems can impact data-generating mechanisms, as individuals with mental health problems are less likely to access care and are more likely to drop out of treatment.[318]

Overall, estimates were too uncertain to determine the relative value across sectors for mental or physical health changes. The IRF confidence intervals were expected to widen with each year because each repetition will inevitably compound the uncertainty. However, initial high levels of uncertainty led to extreme uncertainty in cumulative estimates over time.

Although this study satisfied the usual criteria for a GMM approach of small T and large N. Undoubtedly, certain models and specifications exist that would provide evidence that aligns with my hypotheses. However, such analyses would lack face validity, as every modelling decision in this study has followed best practices and incorporates prior beliefs where they exist. Reanalysis of the same cohort, once further waves beyond the seven analysed are published, would significantly increase the power of such analysis, e.g., because of the lags required for estimation, an additional two waves of data collection would almost double the T per panel. Such reanalysis would also benefit from FOD transformations with more extended periods to average across (it tends to first differences with only two-time points). Investigation using alternative methodologies is also essential, e.g. as discussed earlier, GMM using non-linear moments,[239] or cross-lagged ML-SEM,[240,241] despite their inability to implement IRFs to estimate the dynamic causal effects of improving health across all outcomes as a system.

Examining lifecourse spillovers across diagnoses, conditions, or treatments would be a strong avenue of research if linked to datasets such as Hospital Episode Statistics or the Mental Health Services Dataset. Findings may also differ across countries, where the interpretation of mental health and the role of healthcare and social support systems can vary significantly. Similar estimation strategies which tackle the impacts of health as a system should be considered for interindividual effects, and examination within randomised settings to derive comparable monetary benefit across sectors should be a priority.

Limitations

Estimates of changes in tax revenue and benefits at the margin indicate underlying employment and behaviours and should not be extrapolated to individual estimates because analysis averages across extreme events, e.g., job loss due to mental health decline leading to lowered tax revenue and increased use of social benefits. There may be bias in the measurement of tax revenue if people more likely to experience variance in mental health tend to exhibit less change in employment characteristics, which would minimise derived estimates of the state burden of changes in employment. The impacts of changes in health on welfare benefits are inevitably country-specific, and findings might change based on underlying policy and trends.

UKHLS does not include questions to capture resource use for programmes such as NHS Talking Therapies (formerly known as Improving Access to Psychological Therapies) and other community mental health schemes for chronic or low to medium-severity conditions, which will underestimate the healthcare costs associated with deterioration in mental health. Linear estimates of costs, when aggregated across GP, outpatient, and inpatient services, reduce zero inflation and avoid exponential increases in instruments, but they may be unable to address the underlying non-linear relationships. Furthermore, specifying a value of 11 for self-reports within the bracket of >10 outpatient attendances or GP contacts might influence estimated relationships. Because self-reported costs are accrued during the previous year, they are, on average, equidistant to collection periods *t* and *t*-1 and further compounded using forward orthogonal deviations. Adequately characterising the long-term impacts of health changes is crucial for policymaking, but the time horizon might not be sufficient to capture relevant dynamics, as found in studies of cost-offsetting and depression,[262] in particular, the length of time to observe the influence of mental health on the development of comorbidities and healthcare service use.

Self-assessed health instruments may be subject to misclassification bias,[319] and recall bias concerning resource use,[320] although self-reporting is generally considered acceptable for collecting such service use data.[321] Because of fragmentation in medical systems, self-reports may be less accurate for GP service use than their GP records but more accurate for indicating the use of health services more generally.[322] The change in HRQoL is assumed to hold for one year, and incomes are only reported for the survey month, which may under/overestimate derived annualised benefits and costs. QALYs derived using the SF-6D may differ in estimated relationships and value from one using the EQ-5D; however, mapping should only be used where necessary.[323]

Concluding remarks

Multi-sectoral spillovers represent significant value drivers in improving health. However, given the limited data collection in RCTs, examining broader spillovers in such populations is not easily undertaken, so this study hoped to shed light on this subject in a surrogate manner. Dynamic panel approaches offer many advantages over traditional longitudinal modelling, and their utility applied to UKHLS to examine multi-sectoral spillovers will greatly increase as more data become available. The challenges faced by this study in estimating health's dynamic impacts underline the utility of broader cost collection in randomised settings.

Acknowledgements

Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. The COVID-19 study (2020-2021) was funded by the Economic and Social Research Council and the Health Foundation. Serology testing was funded by the COVID-19 Longitudinal Health and Wealth – National Core Study. Fieldwork for the web survey was carried out by Ipsos MORI and for the telephone survey by Kantar.

Chapter 4. Interindividual health spillovers in economic evaluations: on mechanistic definitions

"*How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it*." Adam Smith[324]

The degree to which changes in an individual's health status impact the health of their network is critical to representing the true societal value of improving health;[171,173] targeting individual and public interventions appropriately and suitably characterising the technical and allocative efficiency of spending. A recent review found that globally, national healthcare economic evaluation guidelines frequently recommend that "the target population is the one that is most likely to receive the proposed intervention in clinical practice".[184] But overall, there is much heterogeneity, leading some to point out the lack of recommendations specifically about the unit of analysis, i.e., whom to measure (in their case regarding resource use).[134] This discussion is distinct from extra-welfarist vs welfarist approaches,[325,326] the distinction being that outcomes may lie within multiple individuals versus collective social welfare, and each can incorporate health and non-health benefits through cost-benefit or cost-effectiveness analysis.[42] Despite some laudable attempts to clarify terminologies, there has been relatively little multidisciplinary discussion of the mechanisms for health spillovers. This Chapter defines externalities and spillovers in healthcare, presents a visual framework to understand their mechanisms, and discusses their technical and normative relevance to economic evaluations.

Economics is fond of its nomenclatures; non-market interactions refer to processes such as externalities and spillovers. Although more of a technicality, the cross-discipline examination of mental health's insertion into economics poses difficulties in terminology, from the irregularity of demand and information asymmetry to externalities and heterogeneity in the concept of spillovers. This lack of clarity poses a problem for precision in research, undermining the ability of researchers to address research questions and find relevant literature despite the content aligning precisely with their interests.

In economics, heterogeneity in the concept of externalities has long been discussed.[327] Externalities are direct inter-agent effects (whether consumer or firm) and are defined as "*a cost or benefit imposed on others (without compensation) as the result of some economic activity"* [328]p.154, or more generically, where production activities or consumption behaviour affect third parties.[42] Such description is used frequently to propose the idea of internalising externalities, e.g., the tragedy of the commons and divergence between private and social costs/benefits. *The Myth of Social Cost* offers a thorough economic discussion,[329] and an insightful read is available in Arrow's *Political and economic evaluation of social effects and externalities*.[330]

Initially, externalities were also discussed in the healthcare context as arising from the health status of treated individuals,[331] therefore representing the indirect effects of treatment. However, the term spillovers, initially interchangeable with externalities, is now broadly used to differentiate such indirect effects from the direct effects of externalities. Spillovers remain a heterogeneous concept, and Muir and Keimplass offer a valuable review to inform the consideration of disparate agents in economic evaluation.[332] Briefly, their key attributes specify an initial action and targeted outcome, e.g., treatment and health, two entity involvement, e.g., direct impacts on the patient and indirect impacts on a caregiver, and a lack of intention (unintentional impact resulting from initial action). Notably, positive spillover is a burgeoning area where treatment may not explicitly intervene at those points; decision-makers can infer the spillover, e.g., in behavioural spillover literature.[333]

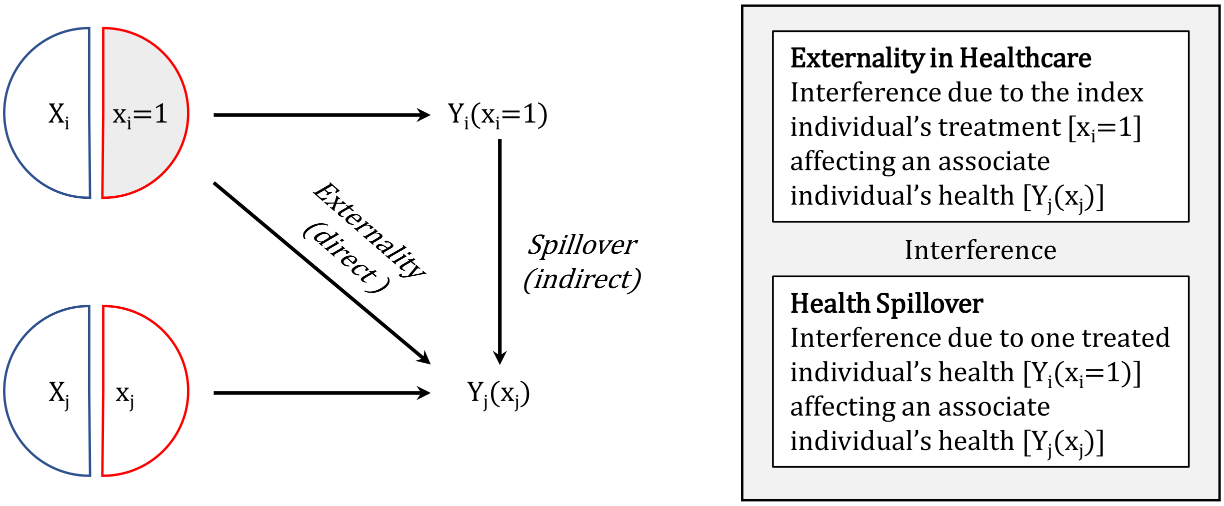
How we define spillovers is inherently linked to how we measure them. Various disciplines are starting to focus on spillovers and have made significant contributions to measurement frameworks. Benjamin-Chung et al.,[334] provide an in-depth epidemiological discussion of study design and methodological considerations. Angelucci and Di Maro do so from the lens of programme evaluation,[335] and others tackle applied topics such as estimating total treatment effects.[336] However, these works primarily focus on evaluating interventions at the group level; this Chapter seeks to disentangle spillovers at the individual level. Francetic et al. developed a framework for identifying and measuring healthcare spillover effects, dimensionally classifying them by the [un]intended effect on [non]targeted units.

Conversely, Mendoza-Jiménez, van Exel, and Brouwer's recent mapping review of spillovers in economic evaluations poses a simpler definition:

"In the context of economic evaluations of health interventions, spillovers are all impacts from an intervention on all parties or entities other than the users of the intervention under evaluation"[188]p.3

This definition removes targeting, includes both direct and indirect effects, and only focuses on users and non-users of interventions. Loosely, this can be aligned with Francetic et al. by collapsing their dimensional definitions of non-spillovers to be intended effects in a targeted unit (user), and the definition of spillovers to be all other impacts, e.g., [non]intended effects in a strictly non-targeted unit (user). However, this presents two distinct problems. 1. This presumes there is no possibility of spillover between two targeted units (users). 2. Defining spillovers through a measurement lens (a change in outcome) impinges on clarifying the mechanisms for externalities and spillovers. Instead, the mechanical definitions of externalities and spillovers (agnostic to treatment status in non-index individuals) can be combined with causal frameworks to present treatment impacts in clear visual and mathematical terms. Figure R presents a Single World Intervention Graph (SWIG) to classify treatment externalities and spillovers through the lens of statistical interference.[337,338]

Figure R. A single-world intervention graph to classify treatment externalities and spillovers.



Where the treatment assignment dummy x = 1 for the (index) individual i, and x ∈ {0,1} for the (associate) individual j.

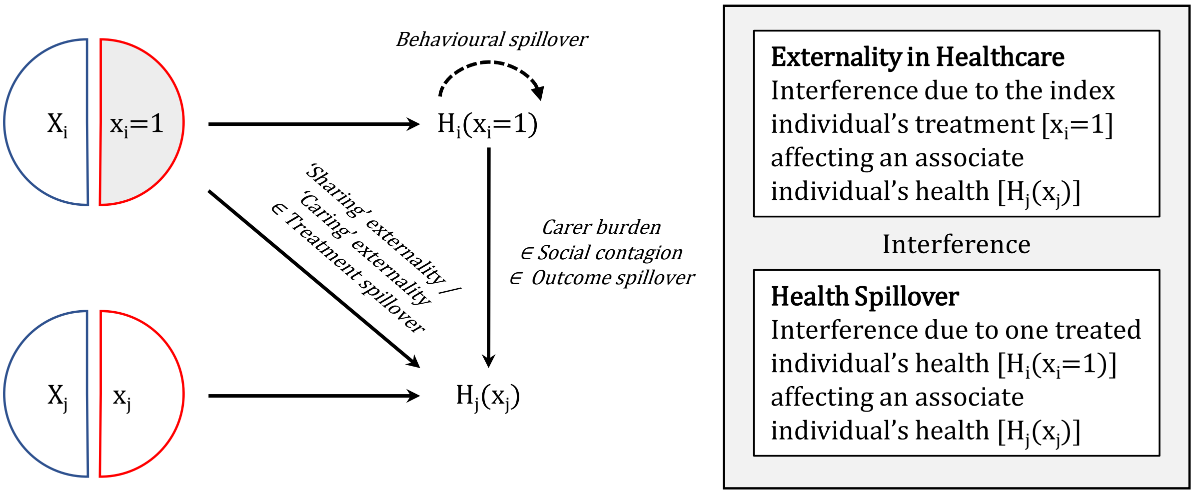
Breaking down the mechanisms for externalities and spillovers in healthcare

In health economics, Culyer posits that healthcare has two types of externalities.[339] First, the "sharing" (also termed "selfish") externality, wherein benefits accrued externally (by others), influences individual consumption in line with pure economics definitions, e.g., "*external effects exist in consumption whenever the shape or position of a man's indifference curve depends on the consumption of other men."* [326]*p.43*. For example, the vaccination of others reduces the chance of transmission to self, therefore reducing the expected benefit of self-vaccination. Interestingly, this is the inverse function of the dilemma introduced by anti-microbial resistance, wherein the treatment of others affects future health payoffs to self. Second, the subtler kind of direct interdependence between an individual's welfare and another's consumption: termed a "caring" externality (individual's utility modified by other's wellbeing or the feeling another is not getting adequate healthcare, i.e., distributional concerns, but not through their health, or concern therein, being modified).[339]

The relevance of spillovers (and their case for inclusion in economic evaluations) depends on credible mechanisms for the spillover.[340] To be clear, I will not discuss the protective effects of immunisation on other (unimmunised) family members, but this is a clear example of the sharing externality in the context of physical health. The interindividual spillover literature primarily focuses on caregiver burden, for which a wealth of discussion exists.[341,342] More recently, a case study on incorporating household spillovers classifies the different mechanisms of health spillovers, summarising theory to broadly fit into three categories: informal care burden, the mental distress of witnessing suffering, and the influences of behaviours on another.[343]

The economic literature uses umbrella terms, such as "family effects", to describe impacts beyond the informal care burden,[344] but I reiterate, as raised in Chapter 2, that being a carer is not a binary allocation. However, all such mechanisms fall within the broader remit of social contagion. Social contagion, or symptom transmission as a subset, is the classification of any effect or behaviour that can be socially transmitted.[54,345] It has been examined across economics, perhaps most famously related to medical innovation and physician prescribing practices.[346] This umbrella term has also previously been applied to transmitting emotions, feelings, or mental states through a connected network of individuals.[52] Figure S classifies treatment externalities and spillovers in healthcare and reframes the treatment impact SWIG around health change. A worked example is presented in Box 3.

Figure S. A single world intervention graph to classify treatment externalities and spillovers in healthcare.



Where the treatment assignment dummy x = 1 for the (index) individual i, and x ∈ {0,1} for the (associate) individual j.

Although this Chapter does not focus on young and adolescent populations, much of the literature examining social contagion focuses on this age bracket, which is unsurprising given periods of socio-developmental growth in young people and the impact of mental health and behavioural problems.[347] Notably, not all social interactions are opportunities for negative contagion; the social contagion literature describes the buffering capacity of social networks as an area for further investigation.[348]

Box 3. Worked example: The EMOTION randomised controlled trial (RCT).

The EMOTION RCT evaluated the effectiveness of a CBT-based transdiagnostic indicated prevention programme, "Coping Kids" Managing Anxiety and Depression," targeting schoolchildren exhibiting symptoms of anxiety and depression.[418,419] The programme significantly reduced youth-reported symptoms of anxiety and depression, as well as parent-reported depression.*

Direct effects

Treatment (H_i Index Individual)

- Health Improvement: The intervention significantly improves the mental health of participating children, reducing symptoms of anxiety and depression.

Caring Externality (H_j Associate Individual)

- Recognition and Support: Parents feel relief through the acknowledgement of their child's mental health challenges and the structured care provided by the programme. This recognition alleviates feelings of isolation and helplessness, often associated with managing childhood mental health issues.

Indirect effects

Spillover (H_j Associate Individual)

- Reduced Caregiving Burden: Improved child mental health reduces the time and resources required for support.
- Perception of Child's Health: Parents care about their child's health and wellbeing, and improvements in the child's health alleviate worry and distress – improving the parents' mental health.
- Behavioural Exposure: Reducing depressive and anxious symptoms in the child contributes to healthier family dynamics.

**Improved parental mental health may also spill over to the child, reaching some equilibrium over time.*

Long et al.[345] provide an excellent overview of social contagion-associated concepts such as the interpersonal theory of depression and explanatory behaviours such as excessive reassurance-seeking, negative feedback-seeking, and conversational self-focus. Other relevant mechanisms include homophily (individuals tending to associate with similar others),[349] or peer socialisation of problem behaviours, and newer theories with costs and benefits such as empathetic distress and co-rumination and their links to adaptive and maladaptive adjustment.[350] A systematic review and meta-analysis found that co-rumination is modestly associated with internalising problems such as depression and anxiety,[351] and a newer study on mood in social networks found no evidence for homophily (selection based on mood).[348]

In some cases, behavioural interventions at the family level are necessary for successful health intervention. For example, in cases of co-dependency—a common theoretical framework used when treating families affected by addiction disorders—addressing a family member's co-dependency is considered integral to the effective management of the addictive disorder.[352,353] Alternative non-health examples include interventions targeting the child-caregiver dyad (the pair as a unit), such as universal parenting programmes.[354]

The mechanisms for health spillovers discussed above may be relevant to physical or mental health interventions. Given the focus of this thesis on mental health, it is important to situate mental health-specific pathways of spillovers in the context of social contagion. The effects of an individual's mental health on others is not a new idea, although the knowledge base is more extensive in some areas of psychopathology than others, e.g., greater for trauma or depression than OCD or psychosis.[355] Folie à deux, where an identical or similar mental disorder affects two (or more) individuals in close association, has been reported for over a century; historically focusing upon psychotic symptoms and the transfer of delusional ideas/abnormal behaviours.[356] Such folies fall into 4 subtypes: (a) *folie imposée* (imposed) (b) *folie simultanée* (simultaneous) (c) *folie communiquée* (communicated) and (d) *folie induite* (induced).[357] Shared delusions have their categories within the DSM, which have been shown to occur within syndromes for which it is unusual to present an affective disorder (of psychotic symptoms), such as bipolar affective disorder.[358] Shared delusions can also present in carers, such as in the case of cancer,[359] or, in some cases, between fellow hospital inpatients.[360]

Neuroscientific investigation of contagion has also garnered attention over the last decade, for example, in the case of depression.[361] However, it should be noted that the discussed *folies* and social contagion are not mutually exclusive. A subset of associate patients (who have had delusional beliefs transmitted to them by the primary), may "*come to share the belief via normal processes of social contagion* [and] *only qualify as* *delusional by virtue of the abnormal persistence of their belief"*[56]p.72. Accordingly, the term *folie imposé* is typically used as the construed meaning of folie à deux, whereas social contagion represents a modernised umbrella term for both *folie simultanée and folie communiquée*. Because of the rarity of *folie imposé*, often dependent on the isolation of the family unit, it may be regarded as insignificant when compared to contemporary issues such as the interindividual transmission of addictions, self-harm, antisocial behaviour, or suicide. Therefore, social contagion is likely a suitable umbrella for all interindividual health spillovers in economic evaluations.

Chapter 5. General Discussion

This thesis investigated the links between the methods for economic evaluation and the value of improving mental health. Chapter 1 outlined the global burden of mental and addictive disorders, emphasising their growing share of disability-adjusted life years (DALYs) and the pressing need for improved mental health treatment, funding, and research. It also highlighted the need to interrogate the suitability of current methods for capturing the diverse value of mental health, laying the foundation of this thesis. Chapter 2 critically evaluated current practices for economic evaluations in health, emphasising their limitations in capturing the full spectrum of mental health outcomes, interindividual spillovers, and multi-sectoral costs.

Chapter 3 addressed the multi-sectoral cost evidence gap by applying dynamic panel methods to estimate how health improvements influence costs and revenues across healthcare, welfare benefits, and labour tax revenue. My analyses indicated that improving population mental health generates substantial economic value through increased tax revenues. However, this study also demonstrates some of the significant methodological challenges in the causal estimation of value across sectors using observational data. Chapter 4 examines interindividual spillovers and presents a detailed framework to standardise terminology and elucidate the mechanisms through which health (with an emphasis on mental health) affects families, social networks, and broader societal structures.

Overall, this thesis highlights two distinct priorities for advancing research and policy: first, improving the capture of multi-sectoral spillovers, and second, incorporating interindividual spillovers into the decision rules of economic evaluation, particularly in the context of mental health. It calls for broadening evaluative perspectives to reflect societal impacts better while highlighting the need for targeted research to quantify spillovers and assess their implications for decision-making processes. However, the inherent complexities of these issues make it challenging to draw definitive conclusions. Consequently, the following sections focus on these two themes, providing specific recommendations to enhance evidence generation, research, and policymaking, thereby aligning economic evaluations more effectively with societal value.

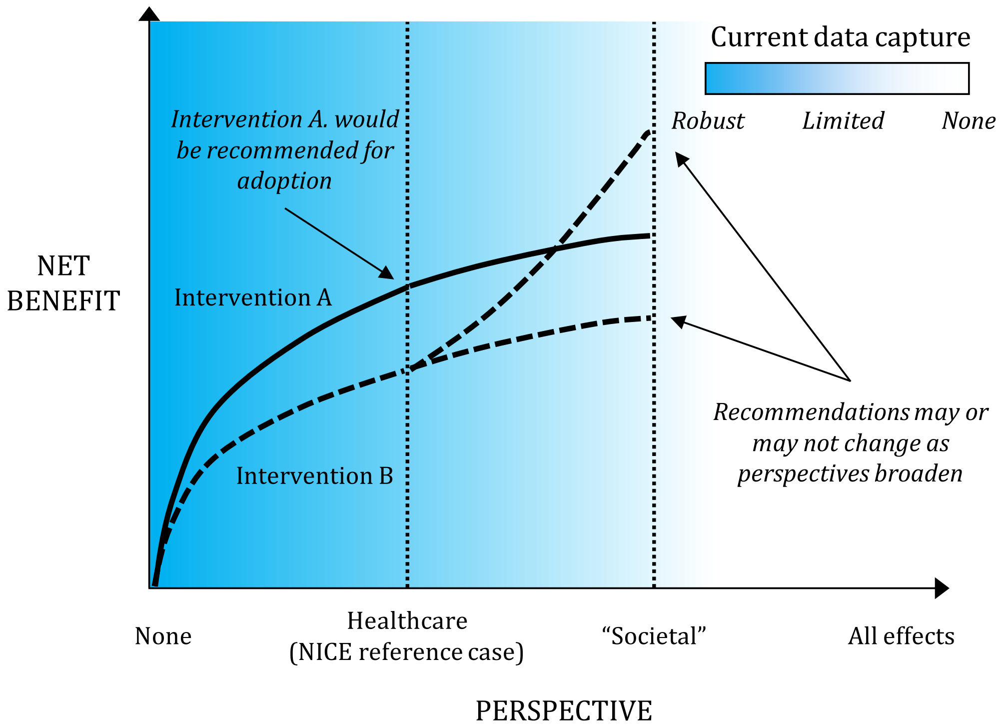
i. Measuring employment-related multi-sectoral spillovers in economic evaluations: worth doing badly

As discussed in Chapter 2, the scope of benefits and costs considered influences the conclusions drawn from economic evaluations in healthcare,[138,169–171,186] i.e., a narrow focus may not adequately represent the broader implications seen from wider perspectives. Non-healthcare sector costs are directly relevant to characterising the societal value of improving mental health; however, there is little economic evidence to demonstrate the extent this diverges from other forms of health, mainly because applied evaluations have primarily been limited to a healthcare perspective,[119] and societal perspectives to the addition of absenteeism costs.[122] Chapter 3 attempted to tackle this question using panel data but demonstrated some of the challenges of causal estimation in observational data. Therefore, one conclusion of my thesis is to advocate for enhancing the evidence base through systematically collecting data on specific multi-sectoral spillover effects in primary data-collection settings. I propose that broadening and simplifying resource use data collection could yield substantial benefits. I then briefly consider whether such spillovers should be included in the decision rule of economic evaluations in health.

Lack of empirical evidence

Including family health or societal costs in evaluations generally improves the apparent cost-effectiveness of interventions (more health produced per cost).[120,144,167,168] As noted earlier, several studies have demonstrated that decision recommendations based on a healthcare perspective can sometimes be suboptimal from a societal perspective.[138,169–171] However, these studies state that they could only test minor amendments to the cost-effectiveness calculation, such as the inclusion of absenteeism costs or social costs, because they are constrained by the data of the underlying evaluations (again, most applied evaluations take a healthcare perspective,[119] and societal perspectives usually only extend this to the addition of absenteeism costs).[122] Therefore, we need more information regarding how often decision recommendations would change under broader perspectives (see Figure T). That is not to say that there have been no evaluations under broader perspectives, e.g. costs incurred by the criminal justice system, but these represent a small fraction of the total generated evidence.

Figure T. A simplified illustration of the capture of broader perspectives in economic evaluations in health and their possible influence on decision recommendations.



We also do not know which types of health or intervention are most impacted except for in the most obvious cases, e.g., caregiver health in degenerative diseases or severe mental illnesses. The relevance of multi-sectoral costs has primarily been addressed through expert surveys,[192,193] but the scarcity of empirical evidence on the societal impacts of health interventions hinders retrospective secondary analysis and limits the prospective identification of significant cost drivers outside the healthcare domain.[188] As Brouwer,[177] and McCabe,[179] discuss in their back-and-forth commentaries on the relevance of spillovers, including spillovers in economic evaluation would greatly increase our understanding of spillovers.

In Chapter 2, I suggest that interventions that incrementally improve mental health dimensions might be significantly affected, but I note that the empirical evidence base is limited. However, such considerations are likely not isolated to mental health. Health economics should investigate groupings of disease areas with similar downstream consequences, which are more affected by current social value judgements than others, e.g., including costs beyond healthcare. This could be considered a form of equifinality, which refers to the observation that in any open system, a diversity of pathways may lead to the same outcome,[362] in this case, similarities in wider societal cost burden.

This lack of data impacts other policy-relevant research, which otherwise could demonstrate society-wide effects that could contribute to our understanding of allocative efficiency, e.g., the impact of health on the receipt of welfare benefits.[124,204] Such impacts are increasingly relevant to the current discourse on welfare conditionality.[195,205] Suppose such broader impacts of health are relevant to policymaking. In that case, a better way to tackle such issues might be to give value to welfare benefits in the healthcare decision-making process in the first place. This paucity of evidence exists in both directions, as we are also missing an opportunity to characterise the allocative or technical efficiency of spending by other governmental departments on health outcomes (despite many threshold studies in healthcare estimating the value of an *exogenous* but marginal change in expenditure).[363]

More generally, existing evidence on multi-sectoral costs is only possible by resorting to systematic review processes, as remarked in the context of quality-of-life data; standardisation in cost collection and better signposting of the costs considered in titles and abstracts would help researchers find relevant literature.[180] A notable limitation is the often insufficient funding for studies to incorporate dedicated health-economic components or to support the expertise required for such analyses. There is a clear need for the inclusion of straightforward, society-wide cost measures in research, which would facilitate a more holistic understanding of the economic implications of health interventions.

Data collection is a trade-off

Additional data collection comes at a cost, whether financial, increased participant burden, or potential compromise of clinical objectives.[22] Expanding the remit of data collection can adversely impact response quality indicators, notably the rate of missing values, reliability, and accuracy,[322,364] i.e., the intensive nature of health economics pro forma often results in high attrition rates, non-response, or significant levels of item missingness. A crucial trade-off exists between the comprehensiveness of questionnaires and the response rate at the individual level. Despite the critical importance of these issues, the literature on data missingness in health economics remains sparse.

Simplify

Adopting G.K. Chesterton's ethos, "*If a thing is worth doing, it is worth doing badly*.",[365] this commentary advocates for a pragmatic widening and simplification of resource use data collection and valuation. In striving for technical precision, we miss points of social relevance, which Olsen and Richardson previously raised in their discussion of socially relevant and socially irrelevant production gains.[158] We could lose some precision in indirect healthcare costs (such as indirect medicine dosage and timing) and increase the accuracy of societal effects, even if these are ballparks, e.g., income data can be reliably collected using a single question.[366] The first goal should be the unbiased and consistent estimation of societal impacts rather than granularity.

Although health drives costs across the welfare, criminal justice, and employment sectors,[24,192] many of these costs are not equally relevant to, nor incurred by, individuals in primary data collection settings, e.g., educational costs in adult populations, or those of the criminal justice system where healthcare intervention is relatively minimal.

Employment capacity and productivity are likely more relevant in evaluations in adult populations and of particular interest to policymakers and economists alike. Public health and healthcare are critical for economic performance and long-term economic growth.[367,368] This idea is tentatively established among health economists, yet the president of the International Health Economics Association (IHEA) recently remarked that "*it is not always clear that economic policymakers in other fields appreciate this*".[368]p.1. The Office for Budget Responsibility (OBR) in the United Kingdom notes poor [mental] health as the largest contributor to working-age employment inactivity, and describes the fiscal implications as forgone tax revenue, higher welfare spending, and higher spending on healthcare services (indeed, many healthcare costs are themselves spillovers).[369] Lord Darzi's independent review of the NHS also highlights national prosperity and getting individuals into employment as an important theme for a future 10-year health plan.[370] These findings have been followed by the Get Britain

Working White Paper, which, in the words of Liz Kendall MP, the UK Secretary of State, outlines *"how, together, we can build a healthier, wealthier nation - driving up employment and opportunity, skills and productivity – while driving down the benefit bill*."[371]

Furthermore, in the context of mental health, I raised in Chapter 2 how economic evaluations value averting productivity losses but not productivity gains, e.g., changes in employment status.[24] Some economic analyses, such as the latest *The economic and social costs of mental ill health*, apply the median net salary for economically inactive time.[224] However, neither productivity losses nor such costing methods for the economically inactive are opportunity costs in governmental spending. Such impacts could instead be captured through tax revenue measures, which have been previously argued as a value driver in health, e.g., in scaling up evidence-based psychotherapies,[94] or in understanding the cost of unpaid care by young people.[227] However, this should not necessarily extend to all tax revenue but rather the more estimable tax receipts attributable to deviations in individuals' employment capacity, e.g., income tax and national insurance contributions, representing most of the UK government's direct tax receipts.[196]

It is established that there are distributional concerns if productivity loss estimation used actual incomes, which would bias economic evaluations against improving health for individuals in lower-paid employment, such as those with serious mental illness.[117] Proponents of broader wealth effects (focusing on productivity), such as Garau et al.,[372] suggest prioritising diseases affecting individuals of working age, whereas opponents, such as Shearer, Byford, and Birch,[373] raise significant distributional concerns regarding including patient production losses in evaluations under any scenario. Replacing productivity losses with tax revenue alone could address these concerns,[373] instead of prioritising the working population, it would increase the working population through health improvement under a (perhaps strong) presumption of reinvesting increased tax yields into equitable healthcare.

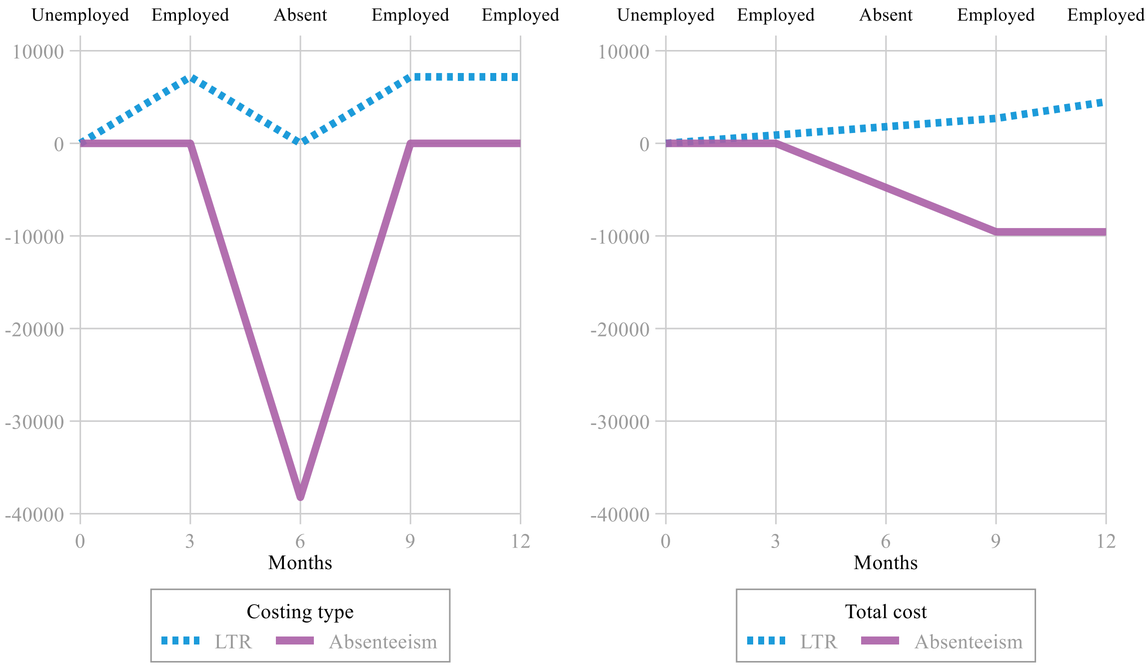
In practice, although using gross minus net income to estimate labour tax revenue would be more precise, this would again introduce distributional concerns in addition to complexity. Instead, evaluations could attribute the country-specific tax revenue of mean or median income to those in employment at observed time points and then adjust for areas under the curve to derive totals. Attributing tax revenue in this way would take the form of the inverse of a cost, and it is only through comparison to another intervention that the difference becomes an incremental cost. *The economic and social costs of mental ill health* also estimated tax revenue in addition to productivity losses;[224] however, such cases could double-count the value of improving employment capacity. Instead, to estimate absenteeism, an alternative approach could involve subtracting pro-rata tax revenue for the duration of time away from work, analogous to the impact of unemployment for that period. This approach serves two key objectives:

1. **Converting absenteeism into a tangible societal opportunity cost**
   Reflects the economic reality that individuals absent from work, particularly those on zero-hour contracts or in certain self-employed roles, may cease to generate labour tax contributions during their absence. Salaried employees' opportunity cost depends on the duration of absence, employer-provided sick pay and country-specific regulations.

2. **Maintaining consistency in decision rules**
   Mirrors the negative trend in value introduced by absenteeism in the decision rule of economic evaluations in health.

Figure U applies this methodological approach to valuing fluctuations in employment capacity and absenteeism within the simplified context of a hypothetical clinical trial. It illustrates how these changes can be translated into economic terms using tax revenue measures, such as income tax and national insurance contributions, rather than conventional productivity loss metrics.

Figure U. Costing labour tax revenue (LTR) and absenteeism of a hypothetical 12-month clinical trial in the UK.[ab]



[a] *Mean gross annual pay for all employee jobs in the UK 2024, rounded to the nearest pound: £38,224.*[374] *Mean labour tax revenue for all employee jobs calculated using the mean gross annual pay for all employee jobs minus take-home pay: £38,224 - £31,043 = £7181 (estimated using the UK government's online income tax calculator assuming no pension contributions, student loan repayments, or tax codes that impact personal allowances.*[375]

[b] *Cumulative totals estimated by area under the curve (AUC) where interval* $[t_{i-1}, t_i] = \left(\frac{£_{t_{i-1}} + £_{t_i}}{2}\right) \times \left(\frac{t_i - t_{i-1}}{12}\right)$. *The estimation of absenteeism costs using AUC (which presumes a linear relationship between time points) is for visual comparison as, in practice, participants will recall their illness-related days off work in the period since the last instance of data collection (otherwise, a binary indicator would introduce bias because participant response is likely inversely correlated with illness that leads to absence from work at the time of data collection).*

Such valuation is arguably arbitrary, and it reduces in absolute terms the value of reducing absenteeism. However, the current equivalence of one monetary unit in productivity loss to one in healthcare cost is already arbitrary. Such estimation will observe the same trend in costs to the human capital approach of productivity losses, although the friction cost approach would otherwise presume production, and so tax revenue, returns to normal. It would also assume that gains in employment, or losses in employment, attributable to the observed individuals, would not be replacing or displacing others taking up that post or that their post would get filled, i.e., the labour reserve of the non-self-employed (which the friction cost approach to productivity losses is concerned with). This paper does not seek to debate overall labour demand versus supply, or contribute to the ongoing discussion around the "lump of labor" fallacy,[376] but increasing the effective working-age population (e.g., through health intervention) feasibly influences long-run national unemployment rates,[377] and recent evidence suggests significant tax revenue losses attributable to health at the country level.[317]

Other employment-related cost drivers, such as welfare benefits, could focus on simple estimates of benefits that are more common among countries with such social safety nets, e.g., unemployment benefits, again attributing some country-specific median value. This is not to say these would reflect a perfect way to value such impacts. Instead, such methods should be evaluated against the next best alternative, i.e., current practice.[378] The combination of measures of tax revenue and welfare benefits would fully account for changes in employment status (which are not currently valued) and would be able to incorporate the costs of productivity losses, thus satisfying criteria of addressing both outcome (employment capacity) and cost (tax revenue, absenteeism, and welfare benefits).

Such simplification (employment status, days off work, and an indicator of welfare receipt) would be less intrusive, less burdensome, and more acceptable to participants. Strive for efficiency and simplicity in data collection.

Discussion

There is a broader concern that cost-effectiveness analyses do not always reflect displaced services, i.e., an intervention estimated to save healthcare costs might not save resources in practice.[379] The same can be said about tax revenue (as per earlier references to labour reserve); however, certain spillovers, such as welfare benefits through fiscal transfer, will always represent savings. Multi-sectoral costs, unlike disparate benefits, are compatible with current value frameworks.

Although not the explicit target of this commentary, it is worth acknowledging that the significance of spillovers does not always equal relevance, i.e., economic evaluations present the efficient use of healthcare resources without necessarily reflecting on where all the consequences of treatment lie.[23,340] As McCabe has also discussed, current economic evaluations in health most often inform adoption decisions around specific technologies under a fixed healthcare budget.[179] Therefore, relevance for decision-makers in these silos is dependent on the provision of cross-agency compensation arrangements.[23,380] However, the relevance for decision-makers in silos does not reflect the relevance of evaluation at the whole system level,[191,381] and I direct readers to Jönsson's editorial on the technical and normative justifications for the relevance of multi-sectoral spillovers.[191]

Crucially, although cost-effectiveness predicts most HTA decisions, such as those made by NICE,[48] decision-making committees do not use the conclusions of cost-effectiveness analyses in isolation,[382] i.e., "the HTA is not the decision—it is a tool designed to help make better decisions".[191]p.3 Many health-related decisions are also made without the systematised use of economic evaluative methods,[85] e.g., local commissioners frequently use economic evaluation data to allocate budgets. Therefore, subsuming wider economic costs into societal perspectives can inform decision-makers without impinging on the conclusions of healthcare-specific cost-effectiveness. In specific cases, requiring non-inferiority in clinical outcomes or QALYs might also be prudent, permitting broader economic costs to offset medical costs only.

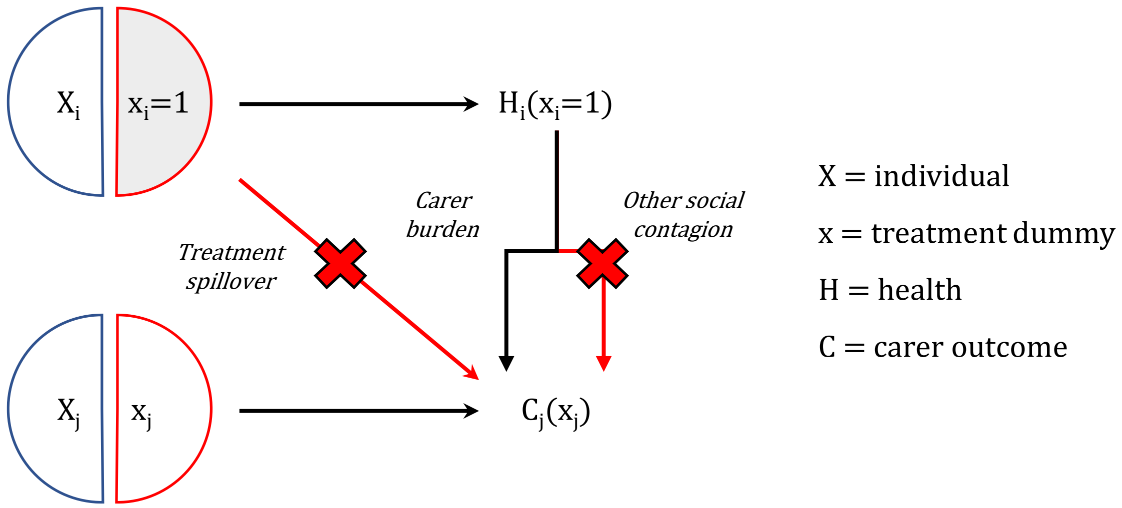
This commentary contends that by simplifying data collection and broadening the scope of costs considered, economic evaluations can better reflect the societal effects of health interventions and enhance their utility in informing decision-making at the healthcare system and policy levels. Such evidence may tend towards a step-in-the-right-direction approach.[383] Broader scopes would improve the ability of researchers to examine many relevant policy questions, such as the size and mechanisms of spillovers. Simplification would also mitigate participant burden, reduce data missingness, and minimise assumptions in analysis, yielding more accurate and actionable evaluations. Retrospective analysis (through the attachment of labour tax revenue and welfare benefits) might be possible in some instances where economic evaluations report employment status in sufficient granularity. Such secondary research could either exploit existing databases of economic evaluations or follow the practices of recent reviews that investigated how often decision recommendations change in different scenarios.[138,169–171]

In conclusion, as the landscape of health interventions becomes increasingly complex, our methods for evaluating their economic implications need not necessarily become more complicated. By embracing a more pragmatic approach to cost collection and valuation, we can ensure that economic evaluations are a robust tool for research and policy-making.

ii. Relevance of interindividual health spillovers to economic evaluations and analytical challenges

Chapter 4 describes and defines interindividual spillover effects, and here I discuss their relevance to economic evaluations. Recent resources have described how to include and value informal care in economic evaluations.[384,385] However, standard HRQoL instruments might be insensitive to the relevant impacts of informal caregiving,[386,387] resulting in several carer-specific outcome measures for health and non-health benefits.[388–390] Such instruments also disentangle caregiving and family effects.[140] For example, by using outcomes that focus on the impacts of caregiving on the caregiver, the impact of other mechanisms beyond caregiving can be removed (Figure V). Empirical evidence confirms that carer instruments measure constructs different from standard HRQoL measures such as the EQ-5D and, therefore, cannot be used interchangeably.[387]

Figure V. Single world intervention graph to untangle treatment spillover mechanisms in healthcare.



This Chapter does not address the comparison problems of using different instruments to derive QALYs between the treated and the associate individuals. Instead, the intention is to make clear that the consequences of all mechanisms for health spillovers are already included in all economic evaluations which measure carer health using standard HRQoL instruments. Therefore, there is precedence for their inclusion in evaluations.

Relevance and constitutional mandates

Despite the precedence set by previous evaluations supporting the relevance of interindividual health spillovers in evaluations, it is an insufficient condition. National Health Technology Assessment (HTA) systems, such as NICE in England and Wales,[86–88] have much say over the scope of economic evaluations in health. However, these bodies are ultimately answerable to the publicly mandated principles of their respective healthcare systems. Below, I briefly address the principles (and quotes from their subsections) of the constitution of the National Health Service (NHS) in England.[391]

The NHS provides a comprehensive service, available to all

"The service is designed to improve, prevent, diagnose and treat both physical and mental health problems with equal regard."

As described by Tilford and Payakachat, for many conditions, a broader family evaluative scope would be more precise and better aligned with constituent understanding and interests.[392] Would we ignore the interindividual effects of treating infections? The vaccine literature notes the importance of interindividual effects while acknowledging that broadening the methodology for economic evaluations might not always benefit the comparative economic case for vaccination, as all interventions should be treated equally in inclusion criteria of costs and benefits.[393]

I argue that health spillovers primarily constitute mental health impacts because physical health does not have the same capacity for contagion except through the transmission of pathogens (or their inhibition due to treatment). Emotional contagion may spread up to three degrees of separation,[143] and there is increasing evidence for the contagion, or symptom transmission, of psychopathology such as anxiety and depression,[53,355,394] and of mental disorders more generally.[395] Empirical evidence supports poor mental health as a disproportionate cause and consequence of interindividual health spillovers. Informal caregiving affects the mental functioning and psychological health of carers to a greater degree than their physical health,[135,396,397] and caregivers of people with mental illness experience a higher subjective burden than those caring for people with a somatic illness.[136] Conditions such as depression affect the health of others to a greater degree than some physical health conditions,[164,175] and health states significantly deteriorate the mental health of families and social networks;[53,141–143] comparable evidence is not found for other health dimensions.[139,164] Not including these impacts in evaluations undermines the recognition of mental health as an essential dimension of health status.[92,398]

"It has a wider social duty to promote equality through the services it provides and to pay particular attention to groups or sections of society where improvements in health and life expectancy are not keeping pace with the rest of the population."

Constraining the perspectives of evaluations to the individual (or limiting broader effects to a carer outcome) would penalise interventions that improve health domains or symptomologies that impact the health of others to a greater degree,[172–174] in this case, of mental health. Beyond parity for mental health generally, this might lead to significant distributional concerns when the burden of mental health problems is tied to economic inequality and socioeconomic disadvantage.[399,400] Indeed, evidence suggests that mental health spillovers are greater in lower-income households.[139,401] Socioeconomically disadvantaged groups are also more likely to provide informal caregiving,[402] a relationship moderated by the social capital of carers and communities.[403] Therefore, including interindividual spillovers in the decision-rule of economic evaluations might reduce inequality.

This is not to say that distributional impacts are straightforward to infer. As raised by Basu et al.,[174] there is a complex moral dilemma where treating [disease areas that affect] people with more extensive social networks is worth more.[178] Similar dilemmas are also found in the QALY trap and the more recently recognised carer QALY trap.[404] Recent work by Henry and Cullinan more thoroughly addresses the distributional impacts of spillovers in evaluations.[405] However, these are health effects in others, and valuing a greater quantity of health equal to a smaller quantity might be considered unjust.

The NHS is committed to providing best value for taxpayers' money

"It is committed to providing the most effective, fair and sustainable use of finite resources."

From a philosophical perspective, it has been argued that indirect, non-health effects relating to the satisfaction of urgent needs, such as externalities arising from caring, education, and employment, should be included.[406] It has been further argued that there is "*no morally significant difference between direct and indirect benefits when allocating scarce resources for health*"[407]p.554, although the authors note that there may be a moral distinction in the type of benefits considered. Furthermore, some have posited that the lack of consensus for including family effects in economic evaluations is partly due to under examination in the literature but also because the scope of evaluations *"normally focuses on health".*[140] I reiterate these points because there are no such quandaries regarding interindividual health spillovers.

The outcome is not in contention; it is health. Neither is the mechanism in contention; these are relevant effects attributable to changes in health status (Chapter 4 – Figure S) The question must, therefore, be reduced to quantifying the relevance in magnitude and dispersion of health spillovers across individuals; for example, there is tentative evidence that caring "about" a family member spills over as much as caring "for" them.[140] Therefore, including interindividual spillovers beyond the carer burden might meet the health maximisation goals of healthcare systems.

The NHS aspires to the highest standards of excellence and professionalism

"*Through its commitment to innovation and to the promotion, conduct and use of research to improve the current and future health and care of the population".*

As stated in Chapter 2, this thesis joins other voices in recent research that argues for including informal caregiver and family effects in economic evaluations.[120,137,140,171,174,177,178] Crucially, evidence also suggests that the UK public also views interindividual health spillovers as important.[178] This does not only mean inclusion in the decision rule of evaluations. Addressing the significant evidence gap is necessary to improve future population health.

Challenges of evidence generation

Although I argue for the relevance of including interindividual health spillovers in economic evaluations, it is difficult to demonstrate concretely the impact it would have on the conclusions of economic evaluations because of data limitations.[24] The literature on interindividual spillovers within randomised settings is scarce,[168,392] and, when considered, often focuses on costs rather than outcomes.[137,408] Furthermore, where studies report carer or family member utilities, these are frequently without any comparator, inhibiting the ability for analyses to estimate spillover effects.[409] Although there is growing attention to spillovers in general and mental health spillovers through caregiving, longitudinal pieces of evidence at the household level or non-caregiving family members are sparse and almost absent in terms of peer effects.[24] There exists no evidence in a health economics context of contagion of mental health in a longitudinal setting inclusive of a generic preference-based measure of health and disparate costs.

Some have proposed measuring patient and carer health status, then using multiplier effects for network members affected.[173] How would we estimate such multipliers pragmatically? Both economic and non-economic analyses of social interactions and networks have a long history,[410] but causal estimation of the reach of spillovers is challenging. The size of affected networks has been examined in several ways, and Christakis and Fowler provide an excellent review of methods and studies which examine dynamic social networks.[54] Compared to the challenges faced in estimating multi-sectoral cost spillovers (Chapter 3), such challenges are even greater for estimating interindividual effects in observational datasets. Individual health is dynamically related to family and social networks, requiring comprehensive data and an analytical strategy capable of examining the co-evolution of health and resource use.

Social support is key to managing individual health,[411] but mental health is also related to network composition,[412] i.e., an individual's social function can lead to changes in household composition. Households that do not break apart may be more robust to decreases or changes in mental health, and poor mental health may also lead to social isolation and decreased network sizes, [412] lowering contagion through less exposure by associate individuals.[413]

Such issues introduce many endogeneity biases in observational settings and other issues such as selection on variance in sample outcome. All methods have trade-offs; structural equation modelling could find a bigger picture of the network and associations but struggles with causal estimation of the cumulative size of such effects. However, most analyses in observational datasets cannot robustly deal with disparate nest sizes, nor changes in the size of such clusters over time, while addressing endogeneity biases. Randomised controlled trials would be an ideal setting to exploit, wherein individual and family/carer outcomes are collected concurrently in clinical samples. Recent evidence indicates that up to two degrees of separation can be induced in experimental settings.[414] For health technology assessment, most existing lifecourse modelling approaches do not allow for these complex association structures, but some newer approaches attempt to address this gap.[415,416]

The impacts of mental health on social isolation and network structure,[412,413] and the possible moderating role of social networks on carer burden raise an interesting question regarding cases where analysis may observe little interindividual benefit in marginally improving the welfare of isolated individuals. However, a more considerable improvement, which leads to reduced social isolation, would lead to exponential societal benefits over time. For example, *"happy individuals are more inclined than less happy individuals to use positive social events as cross-domain buffers against loss".*[413] Therefore, the phenomenon may show thresholds, wherein little value is observed up to a specific level or change (or can be considered exponential past a set point) and might be an interesting consideration for targeting socially isolated individuals in policy. This could be the inverse of arguments for using social network targeting to maximise population behaviour change.[417]

Future multiplier development might wish to attribute average effects instead of individual and circumstance-based estimates so that all are treated equally. Accordingly, we might need to consider an ethical rule that any adopted intervention must be non-inferior for the treated individual compared to alternative interventions. This begs the question; how would the conclusions of economic evaluations be able to change with such constraint? One avenue is that it would change the relative weight of costs in the decision rule of evaluations. For example, offsetting costs that do not produce significant QALY gains in the index individual but improve specific HRQoL dimensions, such as mental health (or unobserved dimensions) that influence the health of those around them.

Conclusions

There are justifiable calls to include carer spillovers in different contexts. I extend these arguments to include all family effects and beyond where feasible. I describe how changes in health status might impact the health of associate individuals. These health spillovers are accrued through overlapping mechanisms that extend beyond simple interpretations of informal caregiver burden. Therefore, including broader interindividual spillovers might meet the health maximisation goals of healthcare systems. Mental health is a disproportionate cause and consequence of health spillovers, and such spillovers are likely significant in all health interventions. From a distributional perspective, including these spillovers might reduce inequality. However, it is not currently possible to estimate how frequently such spillovers would change the conclusions of economic evaluations. Research into the size of such spillovers is essential but presents several methodological challenges in observational data.

References

1 Fuchs VR. The future of health economics. *J Health Econ* 2000; **19**: 141–57.

2 Rehm J, Shield KD. Global Burden of Disease and the Impact of Mental and Addictive Disorders. *Curr Psychiatry Rep* 2019; **21**. DOI:10.1007/s11920-019-0997-0.

3 Schnitzler L, Janssen LMM, Evers SMAA, *et al.* The broader societal impacts of COVID-19 and the growing importance of capturing these in health economic analyses. *Int J Technol Assess Health Care* 2021; : 17–20.

4 Mcgrath JJ, Al-hamzawi A, Alonso J, *et al.* Age of onset and cumulative risk of mental disorders: a cross-national analysis of population surveys from 29 countries. *Lancet Psychiatry* 2023; **0366**: 1–14.

5 Layard R. The economics of mental health: With modern psychological therapy , mentally ill people can become more productive and more satisfied with life. *IZA World Labor* 2017; : 1–10.

6 Wykes T, Bell A, Carr S, *et al.* Shared goals for mental health research: what, why and when for the 2020s. *J Ment Heal* 2021; **32**: 997–1005.

7 Wykes T, Haro JM, Belli SR, *et al.* Mental health research priorities for Europe. *The Lancet Psychiatry* 2015; **2**: 1036–42.

8 United Kingdom Department of Health. A Framework for mental health research. 2017; : 53.

9 Occhipinti J, Skinner A, Doraiswamy PM, *et al.* Mental health: build predictive models to steer policy. *Nature* 2021; : 633–6.

10 McGinty E, Eisenberg M. Mental Health Treatment Gap—The Implementation Problem as a Research Problem. *JAMA Psychiatry* 2022. DOI:10.1001/jamapsychiatry.2022.1468.

11 Dodge KA, Prinstein MJ, Evans AC, *et al.* Population mental health science: Guiding principles and initial agenda. *Am Psychol* 2024. DOI:https://dx.doi.org/10.1037/amp0001334.

12 Hazo JB, Gandré C, Leboyer M, *et al.* National funding for mental health research in Finland, France, Spain and the United Kingdom. *Eur Neuropsychopharmacol* 2017; **27**: 892–9.

13 Woelbert E, Lundell-Smith K, White R, Kemmer D. Accounting for mental health research funding: developing a quantitative baseline of global investments. *The*

Lancet Psychiatry 2021; **8**: 250–8.

14 Patel V. Mental health research funding: too little, too inequitable, too skewed. *The Lancet Psychiatry* 2021; **8**: 171–2.

15 Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. *World Psychiatry* 2015; **14**: 36–42.

16 Cuijpers P, Quero S, Noma H, *et al.* Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry* 2021; **20**: 283–93.

17 Williams AJ, Botanov Y, Kilshaw RE, Wong RE, Sakaluk JK. Potentially harmful therapies: A meta-scientific review of evidential value. *Clin Psychol Sci Pract* 2021; **28**: 5–18.

18 Brettschneider C, Djadran H, Härter M, Löwe B, Riedel-Heller S, König H-H. Cost-Utility Analyses of Cognitive-Behavioural Therapy of Depression: A Systematic Review. *Psychother Psychosom* 2015; **84**: 6–21.

19 Health Quality Ontario. Ontario health technology assessment series: Internet-delivered cognitive behavioural therapy for major depression and anxiety disorders: A health technology assessment. *Ont Health Technol Assess Ser* 2019; **19**: 1–199.

20 World Health Organisation. Mental health atlas 2017. 2018 https://www.who.int/publications/i/item/9789241514019.

21 Taube CA, Mechanic D, Hohmann AA. The future of mental health services research. US Department of Health and Human Services, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Division of Biometry and Applied Sciences, 1989.

22 O'Sullivan AK, Thompson D, Drummond MF. Collection of health-economic data alongside clinical trials: Is there a future for piggyback evaluations? *Value Heal* 2005; **8**: 67–79.

23 Knapp M, Wong G. Economics and mental health: the current scenario. *World Psychiatry* 2020; **19**: 3–14.

24 Lathe J, Silverwood RJ, Hughes AD, Patalay P. Examining how well economic evaluations capture the value of mental health. *The Lancet Psychiatry* 2024; **11**: 221–30.

25 Zechmeister I, Kilian R, McDaid D, *et al.* Is it worth investing in mental health

promotion and prevention of mental illness? A systematic review of the evidence from economic evaluations. *BMC Public Health* 2008; **8**: 1–11.

26 Culyer AJ, Newhouse JP, Pauly M V., McGuire TG, Barros PP. Handbook of health economics. Volume 1B. Elsevier, 2000.

27 Pauly M V., McGuire TG, Barros PP. Handbook of health economics. Volume 2. Elsevier, 2012.

28 Glied S, Smith PC. The Oxford handbook of health economics. Oxford university press, 2013.

29 Drummond, Michael F McGuire A. Economic evaluation in health care: merging theory with practice. Oxford university press, 2001.

30 Glied S, Frank R. The Evolution of Mental Health Policy and Economics. *Oxford Res Encycl Econ Financ* 2018; : 1–21.

31 Castelnuovo G, Pietrabissa G, Cattivelli R, Manzoni GM, Molinari E. Not only clinical efficacy in psychological treatments: Clinical psychology must promote cost-benefit, cost-effectiveness, and cost-utility analysis. *Front Psychol* 2016; **7**: 1–4.

32 Shearer J, Byford S. The basics of economic evaluation in mental healthcare. *BJPsych Adv* 2015; **21**: 345–53.

33 Luyten J, Knapp M. Economic evaluation of mental health promotion and mental illness prevention. In: Global Mental Health: Prevention and Promotion. 2017: 207–22.

34 Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oup Oxford, 2006.

35 Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes, 4th edn. Oxford: Oxford university press, 2015.

36 Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic evaluation in clinical trials. Oxford university press, 2014.

37 Gray A, Clarke P, Wolstenholme, JL Wordsworth S. Applied methods of cost-effectiveness analysis in health care. Oup Oxford, 2011.

38 Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats T. Cost-effectiveness in health and medicine, 2nd edn. Oxford university press, 2016.

39 Thornicroft G, Tansella M, editors. Mental health outcome measures, Third.

RCPsych Publications, 2010.

40 Razzouk D. Mental Health Economics, the Costs and Benefits of Psychiatric Care, 1st edn. Springer, Cham, 2017 DOI:https://doi.org/10.1007/978-3-319-55266-8.

41 McDaid D, Park A La, Wahlbeck K. The Economic Case for the Prevention of Mental Illness. *Annu Rev Public Health* 2019; **40**: 373–89.

42 Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford university press, 2017.

43 Brazier J. Measuring and valuing mental health for use in economic evaluation. *J Heal Serv Res Policy* 2008; **13**: 70–5.

44 Shearer J, McCrone P, Romeo R. Economic Evaluation of Mental Health Interventions: A Guide to Costing Approaches. *Pharmacoeconomics* 2016; **34**: 651–64.

45 Loewenstein G, Read D, Baumeister R. Time and decision: Economic and psychological perspectives on intertemporal choice., 1st edn. Russell Sage Foundation, 2003.

46 Glanz K, Rimer BK, Viswanath K. Health behavior: Theory, research, and practice., 5th edn. Hoboken, NJ, US: Jossey-Bass/Wiley, 2015.

47 Strulik H. An economic theory of depression and its impact on health behavior and longevity. *J Econ Behav Organ* 2019; **158**: 269–87.

48 Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D. The influence of cost-effectiveness and other factors on nice decisions. *Heal Econ (United Kingdom)* 2015; **24**: 1256–71.

49 Peirce CS. Note on the Theory of the Economy of Research. *Oper Res* 1967; **15**: 643–8.

50 Ulmer JB. Pivots and Pirouettes: Carefully Turning Traditions. *Qual Inq* 2020; **26**: 454–7.

51 Arrow KJ. Uncertainty and the Welfare Economics of Medical Care. *Am Econ Rev* 1963; **53**: 941–73.

52 Hatfield E, Cacioppo J, Rapson R. Emotional Contagion. *Curr Dir Psychol Sci* 1993; **2**: 96–100.

53 Eisenberg D, Golberstein E, Whitlock JL, Downs MF. Social contagion of mental health: Evidence from college roommates. *Heal Econ (United Kingdom)* 2013; **22**: 965–86.

54 Christakis NA, Fowler JH. Social contagion theory: Examining dynamic social networks and human behavior. *Stat Med* 2013; **32**: 556–77.

55 Cacioppo JT, Fowler JH, Christakis NA. Alone in the Crowd: The Structure and Spread of Loneliness in a Large Social Network. *J Pers Soc Psychol* 2009; **97**: 977–91.

56 Langdon R. Folie à deux and its Lessons for Two-Factor Theorists. *Mind Lang* 2013; **28**: 72–82.

57 Chi P, Du H, King RB, Zhou N, Cao H, Lin X. Well-Being Contagion in the Family: Transmission of Happiness and Distress Between Parents and Children. *Child Indic Res* 2019; **12**: 2189–202.

58 Knapp M, McDaid D, Parsonage M. Mental health promotion and mental illness prevention: the economic case. 2011. https://www.gov.uk/government/publications/mental-health-promotion-and-mental-illness-prevention-the-economic-case.

59 Frank RG, Glied SA. America's Continuing Struggle with Mental Illnesses: Economic Considerations. *J Econ Perspect* 2023; **37**: 153–78.

60 Insel T. Healing: our path from mental illness to mental health. Penguin, 2022.

61 Tian YE, Di Biase MA, Mosley PE, *et al.* Evaluation of Brain-Body Health in Individuals With Common Neuropsychiatric Disorders. *JAMA Psychiatry* 2023; : 1–10.

62 Brazier J. Is the EQ-5D fit for purpose in mental health? *Br J Psychiatry* 2010; **197**: 348–9.

63 Colson AR, Morton A, Årdal C, *et al.* Antimicrobial Resistance: Is Health Technology Assessment Part of the Solution or Part of the Problem? *Value Heal* 2021; **24**: 1828–34.

64 McGuire J, Dupret S, Plant M. Happiness for the whole family. 2022 https://www.happierlivesinstitute.org/report/happiness-for-the-whole-family/.

65 Choudry MM, Murray N, Dindinger-Hill K, *et al.* Genitourinary cancer and family: The reverberating psychological and cardiovascular effects of a genitourinary cancer diagnosis on first-degree relatives and spouses. *Cancer* 2024. DOI:10.1002/cncr.35486.

66 Ridyard CH, Hughes DA. Methods for the collection of resource use data within clinical trials: A systematic review of studies funded by the UK health technology

assessment program. *Value Heal* 2010; **13**: 867–72.

67 Thorn JC, Brookes ST, Ridyard C, *et al.* Core Items for a Standardized Resource Use Measure: Expert Delphi Consensus Survey. *Value Heal* 2018; **21**: 640–9.

68 Garfield K, Thorn JC, Noble S, Husbands S, Hollingworth W. Development of a brief, generic, modular resource-use measure (ModRUM): piloting with patients. *BMC Health Serv Res* 2023; **23**: 1–10.

69 Pokhilenko I, Janssen LMM, Paulus ATG, *et al.* Development of an Instrument for the Assessment of Health-Related Multi-sectoral Resource Use in Europe: The PECUNIA RUM. *Appl Health Econ Health Policy* 2023; **21**: 155–66.

70 Woods B, Fox A, Sculpher M, Claxton K. Estimating the shares of the value of branded pharmaceuticals accruing to manufacturers and to patients served by health systems. *Heal Econ (United Kingdom)* 2021; **30**: 2649–66.

71 Bullmore E. The inflamed mind: a radical new approach to depression. Picador, 2018.

72 Scotland G, Bryan S. Why Do Health Economists Promote Technology Adoption Rather Than the Search for Efficiency? A Proposal for a Change in Our Approach to Economic Evaluation in Health Care. *Med Decis Mak* 2016; **37**: 139–47.

73 Ekstrand J, Fattah C, Persson M, *et al.* Racemic Ketamine as an Alternative to Electroconvulsive Therapy for Unipolar Depression: A Randomized, Open-Label, Non-Inferiority Trial (KetECT). *Int J Neuropsychopharmacol* 2021; : 1–11.

74 Uthaug M V., Lancelotta R, van Oorsouw K, *et al.* A single inhalation of vapor from dried toad secretion containing 5-methoxy-N,N-dimethyltryptamine (5-MeO-DMT) in a naturalistic setting is related to sustained enhancement of satisfaction with life, mindfulness-related capacities, and a decrement of psyc. *Psychopharmacology (Berl)* 2019; **236**: 2653–66.

75 Uthaug M V., Lancelotta R, Szabo A, Davis AK, Riba J, Ramaekers JG. Prospective examination of synthetic 5-methoxy-N,N-dimethyltryptamine inhalation: effects on salivary IL-6, cortisol levels, affect, and non-judgment. *Psychopharmacology (Berl)* 2020; **237**: 773–85.

76 Marseille E, Mitchell JM, Kahn JG. Updated cost-effectiveness of MDMA-assisted therapy for the treatment of posttraumatic stress disorder in the United States: Findings from a phase 3 trial. *PLoS One* 2022; **17**: 1–12.

77 Daws RE, Timmermann C, Giribaldi B, *et al.* Increased global integration in the

brain after psilocybin therapy for depression. *Nat Med* 2022. DOI:10.1038/s41591-022-01744-z.

78 Raison CL, Sanacora G, Woolley J, *et al.* Single-Dose Psilocybin Treatment for Major Depressive Disorder A Randomized Clinical Trial. *Jama* 2023; **330**: 843–53.

79 Kaul I, Sawchak S, Correll CU, *et al.* Efficacy and safety of the muscarinic receptor agonist KarXT (xanomeline–trospium) in schizophrenia (EMERGENT-2) in the USA: results from a randomised, double-blind, placebo-controlled, flexible-dose phase 3 trial. *Lancet* 2024; **403**: 160–70.

80 Amato M, Babb R, Dixon R, *et al.* Understanding decision-making and barriers in the Mental Health R&D ecosystem. 2024 DOI:10.21955/wellcomeopenres.1115397.1.

81 Gründer G, Brand M, Mertens LJ, *et al.* Personal View Treatment with psychedelics is psychotherapy: beyond reductionism. *The Lancet Psychiatry* 2023; **0366**: 8–13.

82 Catchpole P, Barrett V. Keeping Pace with Pharmaceutical Innovation: The Importance of the NICE Methods Review. *Pharmacoeconomics* 2020; **38**: 901–3.

83 Kestel D. Transforming mental health for all: a critical role for specialists. *World Psychiatry* 2022; **21**: 333–4.

84 Mental Health Policy Group. A mentally healthier nation. 2023.

85 Howdon D, Hinde S, Lomas J, Franklin M. Economic Evaluation Evidence for Resource-Allocation Decision Making: Bridging the Gap for Local Decision Makers Using English Case Studies. *Appl Health Econ Health Policy* 2022; **20**: 783–92.

86 National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013 http://www.nice.org.uk/media/D45/1E/GuideToMethodsTechnologyAppraisal2013.pdf.

87 National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. 2022 https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation.

88 Allen N, Pichler F, Wang T, Patel S, Salek S. Development of archetypes for non-ranking classification and comparison of European National Health Technology Assessment systems. *Health Policy (New York)* 2013; **113**: 305–12.

89 Henderson N, Brassel S, O'Neill P, Allen R, Largeron N, Garau M. Do NICE's

Decision Outcomes Impact International HTA Decision-making? OHE Contract Research Report. London, 2023.

90 Wang A, Rand K, Yang Z, Brooks R, Busschbach J. The remarkably frequent use of EQ-5D in non-economic research. *Eur J Heal Econ* 2022; **23**: 1007–14.

91 Pike J, Grosse SD. Friction Cost Estimates of Productivity Costs in Cost-of-Illness Studies in Comparison with Human Capital Estimates: A Review. *Appl Health Econ Health Policy* 2018; **16**: 765–78.

92 Prince M, Patel V, Saxena S, *et al.* No health without mental health. *Lancet* 2007; **370**: 859–77.

93 Zala D, Brabban A, Stirzaker A, Kartha MR, McCrone P. The Cost-Effectiveness of the Improving Access to Psychological Therapies (IAPT) Programme in Severe Mental Illness: A Decision Analytical Model Using Routine Data. *Community Ment Health J* 2019; **55**: 873–83.

94 Layard R, Clark DM. Why more psychological therapy would cost nothing. *Front Psychol* 2015; **6**: 4–6.

95 Stevens KJ. How Well Do the Generic Multi-attribute Utility Instruments Incorporate Patient and Public Views Into Their Descriptive Systems? *Patient* 2016; **9**: 5–13.

96 Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy* 2017; **15**: 127–37.

97 Brazier J, Connell J, Papaioannou D, *et al.* A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technol Assess (Rockv)* 2014; **18**: 1–188.

98 Mulhern B, Mukuria C, Barkham M, *et al.* Using generic preference-based measures in mental health: Psychometric validity of the EQ-5D and SF-6D. *Br J Psychiatry* 2014; **205**: 236–43.

99 Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in Schizophrenia? A systematic review. *Value Heal* 2011; **14**: 907–20.

100 Barton GR, Hodgekins J, Mugford M, Jones PB, Croudace T, Fowler D. Measuring the benefits of treatment for psychosis: Validity and responsiveness of the EQ-5D. *Br J Psychiatry* 2009; **195**: 170–7.

101 National Institute for Health and Care Excellence (NICE). CHTE methods review. Health-related quality of life. Task and finish group report. 2020 https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/nice-guidance/chte-methods-consultation/Health-related-quality-of-life-task-and-finish-group-report.docx.

102 Rencz F, Janssen MF. Analyzing the Pain/Discomfort and Anxiety/Depression Composite Domains and the Meaning of Discomfort in the EQ-5D: A Mixed-Methods Study. *Value Heal* 2022; **25**: 2003–16.

103 Dawoud D, Lamb A, Moore A, *et al.* Capturing what matters: updating NICE methods guidance on measuring and valuing health. *Qual Life Res* 2022. DOI:10.1007/s11136-022-03101-6.

104 van Krugten FCW, Feskens K, Busschbach JJV, Hakkaart-van Roijen L, Brouwer WBF. Instruments to assess quality of life in people with mental health problems: a systematic review and dimension analysis of generic, domain- and disease-specific instruments. *Health Qual Life Outcomes* 2021; **19**: 1–13.

105 Keetharuth AD, Brazier J, Connell J, *et al.* Recovering Quality of Life (ReQoL): A new generic self-reported outcome measure for use with people experiencing mental health difficulties. *Br J Psychiatry* 2018; **212**: 42–9.

106 Wickramasekera N, Tubeuf S. Measuring quality of life for people with common mental health problems. *J Ment Heal* 2020; **0**: 1–8.

107 van Krugten FCW, Busschbach JJV, Versteegh MM, Hakkaart-van Roijen L, Brouwer WBF. The Mental Health Quality of Life Questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Qual Life Res* 2022; **31**: 633–43.

108 Haslam N, McGrath MJ, Viechtbauer W, Kuppens P. Dimensions over categories: A meta-analysis of taxometric research. *Psychol Med* 2020; **50**: 1418–32.

109 Layard R, De Neve J-E. Wellbeing: Science and Policy. Cambridge: Cambridge University Press, 2021 DOI:10.1017/9781009298957.

110 Frijters P, Krekel C. A handbook for wellbeing policy-making: History, theory, measurement, implementation, and examples, 1st edn. Oxford University Press, 2021.

111 Böhnke JR, Croudace TJ. Calibrating well-being, quality of life and common mental

disorder items: Psychometric epidemiology in public mental health research. *Br J Psychiatry* 2016; **209**: 162–8.

112 van IJzendoorn MH, Bakermans-Kranenburg MJ. Problematic cost–utility analysis of interventions for behavior problems in children and adolescents. *New Dir Child Adolesc Dev* 2020; **2020**: 89–102.

113 Brazier J, Tsuchiya A. Improving Cross-Sector Comparisons: Going Beyond the Health-Related QALY. *Appl Health Econ Health Policy* 2015; **13**: 557–65.

114 Brazier J, Peasgood T, Mukuria C, *et al.* The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. *Value Heal* 2022; **25**: 482–91.

115 Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. *Qual Life Res* 2012; **21**: 167–76.

116 Culyer AJ. Cost, context, and decisions in health economics and health technology assessment. *Int J Technol Assess Health Care* 2018; **34**: 434–41.

117 Sanders GD, Neumann PJ, Basu A, *et al.* Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. *JAMA - J Am Med Assoc* 2016; **316**: 1093–103.

118 Garrison LP, Pauly M V., Willke RJ, Neumann PJ. An Overview of Value, Perspective, and Decision Context—A Health Economics Approach: An ISPOR Special Task Force Report [2]. *Value Heal* 2018; **21**: 124–30.

119 Kim DD, Silver MC, Kunst N, Cohen JT, Ollendorf DA, Neumann PJ. Perspective and Costing in Cost-Effectiveness Analysis, 1974–2018. *Pharmacoeconomics* 2020; **38**: 1135–45.

120 Pennington B, Eaton J, Hatswell AJ, Taylor H. Carers' Health-Related Quality of Life in Global Health Technology Assessment: Guidance, Case Studies and Recommendations. *Pharmacoeconomics* 2022; **40**: 837–50.

121 Drost RMWA, Paulus ATG, Evers SMAA. Five pillars for societal perspective. *Int J Technol Assess Health Care* 2020; **36**: 72–4.

122 Drost RMWA, Van Der Putten IM, Ruwaard D, Evers SMAA, Paulus ATG. Conceptualizations of the societal perspective within economic evaluations: A systematic review. *Int J Technol Assess Health Care* 2017; **33**: 251–60.

123 Bokma WA, Batelaan NM, van Balkom AJLM, Penninx BWJH. Impact of Anxiety

and/or Depressive Disorders and Chronic Somatic Diseases on disability and work impairment. *J Psychosom Res* 2017; **94**: 10–6.

124 Hoedeman R. Sick on the job? Myths and realities about mental health and work. 2012 DOI:10.1007/s12498-012-0114-3.

125 Kigozi J, Jowett S, Lewis M, Barton P, Coast J. The Estimation and Inclusion of Presenteeism Costs in Applied Economic Evaluation: A Systematic Review. *Value Heal* 2017; **20**: 496–506.

126 Doran CM, Kinchin I. A review of the economic impact of mental illness. *Aust Heal Rev* 2019; **43**: 43–8.

127 Knapp M, Ardino V, Brimblecombe N, *et al.* Youth Mental Health: New Economic Evidence. 2016 www.pssru.ac.uk/publication-details.php?id=5160.

128 Pokhilenko I, Janssen LMM, Evers SMAA, Drost RMWA, Schnitzler L, Paulus ATG. Do Costs in the Education Sector Matter? A Systematic Literature Review of the Economic Impact of Psychosocial Problems on the Education Sector. *Pharmacoeconomics* 2021; **39**: 889–900.

129 Sveen CA, Pedersen G, Ulvestad DA, Zahl KE, Wilberg T, Kvarstein EH. Societal costs of personality disorders: A cross-sectional multicenter study of treatment-seeking patients in mental health services in Norway. *J Clin Psychol* 2023; : 1–18.

130 Senior M, Fazel S, Tsiachristas A. The economic impact of violence perpetration in severe mental illness: a retrospective, prevalence-based analysis in England and Wales. *Lancet Public Heal* 2020; **5**: e99–106.

131 Beecham J, Knapp M. Costing psychiatric interventions. In: Measuring Mental Health Needs, Second. Royal College of Psychiatrists, 2001: 200–24.

132 Delgado D, Breth A, Hill S, Warburton K, Stahl SM. Economics of decriminalizing mental illness: When doing the right thing costs less. *CNS Spectr* 2020; **25**: 566–70.

133 Knapp M, Wong G. Economic evaluations of mental health interventions in criminal justice. *Crim Behav Ment Heal* 2023; **33**: 139–48.

134 Janssen LMM, Drost RMWA, Paulus ATG, *et al.* Aspects and Challenges of Resource Use Measurement in Health Economics: Towards a Comprehensive Measurement Framework. *Pharmacoeconomics* 2021; **39**: 983–93.

135 Pinquart M, Sörensen S. Differences between caregivers and noncaregivers in psychological health and physical health: A meta-analysis. *Psychol Aging* 2003;

18: 250–67.

136 Hastrup LH, Van Den Berg B, Gyrd Hansen D. Do informal caregivers in mental illness feel more burdened? A comparative study of mental versus somatic illnesses. *Scand J Public Health* 2011; **39**: 598–607.

137 Goodrich K, Kaambwa B, Al-Janabi H. The inclusion of informal care in applied economic evaluation: A review. *Value Heal* 2012; **15**: 975–81.

138 Krol M, Papenburg J, van Exel J. Does Including Informal Care in Economic Evaluations Matter? A Systematic Review of Inclusion and Impact of Informal Care in Cost-Effectiveness Studies. *Pharmacoeconomics* 2015; **33**: 123–35.

139 Henry E, Cullinan J. Mental health spillovers from serious family illness: Doubly robust estimation using EQ-5D-5L population normative data. *Soc Sci Med* 2021; **279**: 113996.

140 Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. Caring for and caring about: Disentangling the caregiver effect and the family effect. *J Health Econ* 2010; **29**: 549–56.

141 Graham C, Higuera L, Lora E. Which health conditions cause the most unhappiness? *Health Econ* 2011; **20**: 1431–47.

142 Powdthavee N, Vignoles A. Mental health of parents and life satisfaction of children: A within-family analysis of intergenerational Transmission of Well-Being. *Soc Indic Res* 2008; **88**: 397–422.

143 Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ* 2008; **338**: 23–6.

144 Scope A, Bhadhuri A, Pennington B. Systematic Review of Cost-Utility Analyses That Have Included Carer and Family Member Health-Related Quality of Life. *Value Heal* 2022. DOI:10.1016/j.jval.2022.02.008.

145 National Institute for Health and Care Excellence (NICE). CHTE methods review. Costs used in Health Technology Assessment. Task and finish group report. 2020 https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/nice-guidance/chte-methods-consultation/Costs-used-in-health-technology-assessment-task-and-finish-group-report.docx.

146 Morton A, Adler AI, Bell D, *et al.* Unrelated Future Costs and Unrelated Future Benefits: Reflections on NICE Guide to the Methods of Technology Appraisal. *Heal*

Econ (United Kingdom) 2016; **25**: 933–8.

147 Soekhai V, Whichello C, Levitan B, *et al.* Methods for exploring and eliciting patient preferences in the medical product lifecycle: a literature review. *Drug Discov Today* 2019; **24**: 1324–31.

148 Roudijk B, Janssen B, Olsen JA. Value Sets for EQ-5D-5L, 1st edn. Springer, Cham, 2022 DOI:10.1007/978-3-030-89289-0.

149 Chisholm D, Healey A, Knapp M. QALYs and mental health care. *Soc Psychiatry Psychiatr Epidemiol* 1997; **32**: 68–75.

150 Brazier J, Akehurst R, Brennan A, *et al.* Should patients have a greater role in valuing health states? *Appl Health Econ Health Policy* 2005; **4**: 201–8.

151 Powell PA, Karimi M, Rowen D, Devlin N, van Hout B, Brazier JE. Hypothetical versus experienced health state valuation: a qualitative study of adult general public views and preferences. *Qual Life Res* 2022. DOI:10.1007/s11136-022-03304-x.

152 Dolan P, Metcalfe R. Valuing health: A brief report on subjective well-being versus preferences. *Med Decis Mak* 2012; **32**: 578–82.

153 Hausman DM. Valuing health properly. *Heal Econ Policy Law* 2008; **3**: 79–83.

154 Dang A-T. Amartya Sen's Capability Approach: A Framework for Well-Being Evaluation and Policy Analysis? *Rev Soc Econ* 2014; **72**: 460–84.

155 Krol M, Brouwer W. How to estimate productivity costs in economic evaluations. *Pharmacoeconomics* 2014; **32**: 335–44.

156 Brouwer W, Verbooy K, Hoefman R, van Exel J. Production Losses due to Absenteeism and Presenteeism: The Influence of Compensation Mechanisms and Multiplier Effects. *Pharmacoeconomics* 2023. DOI:10.1007/s40273-023-01253-y.

157 Pauly M V., Nicholson S, Polsky D, Berger ML, Sharda C. Valuing reductions in on-the-job illness: 'Presenteeism' from managerial and economic perspectives. *Health Econ* 2008; **17**: 469–85.

158 Olsen JA, Richardson J. Production gains from health care: What should be included in cost-effectiveness analyses? *Soc Sci Med* 1999; **49**: 17–26.

159 Bryan ML, Rice N, Roberts J, Sechel C. Mental Health and Employment: A Bounding Approach Using Panel Data*. *Oxf Bull Econ Stat* 2022; : 1–34.

160 Drake RE, Wallach MA. Employment is a critical mental health intervention. *Epidemiol Psychiatr Sci* 2020; : 10–2.

161 Williams AO, Rojanasarot S, McGovern AM, Kumar A. A systematic review of discounting in national health economic evaluation guidelines: healthcare value implications. *J Comp Eff Res* 2023; **12**. DOI:10.2217/cer-2022-0167.

162 Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. *Pharmacoeconomics* 2018; **36**: 745–58.

163 Thompson EJ, Richards M, Ploubidis GB, Fonagy P, Patalay P. Changes in the adult consequences of adolescent mental ill-health: Findings from the 1958 and 1970 British birth cohorts. *Psychol Med* 2021; **53**: 1074–83.

164 Wittenberg E, Saada A, Prosser LA. How illness affects family members: A qualitative interview survey. *Patient* 2013; **6**: 257–68.

165 Goodman A, Joyce R, Smith JP. The long shadow cast by childhood physical and mental problems on adult life. *Proc Natl Acad Sci U S A* 2011; **108**: 6032–7.

166 Favril L, Yu R, Uyar A, Sharpe M, Fazel S. Risk factors for suicide in adults: systematic review and meta-analysis of psychological autopsy studies. *Evid Based Ment Health* 2022; **25**: 148–55.

167 Wansink HJ, Drost RMWA, Paulus ATG, *et al.* Cost-effectiveness of preventive case management for parents with a mental illness: a randomized controlled trial from three economic perspectives. *BMC Health Serv Res* 2016; **16**: 1–15.

168 Pennington BM. Inclusion of Carer Health-Related Quality of Life in National Institute for Health and Care Excellence Appraisals. *Value Heal* 2020; **23**: 1349–57.

169 Peña-Longobardo LM, Rodríguez-Sánchez B, Oliva-Moreno J, Aranda-Reneo I, López-Bastida J. How relevant are social costs in economic evaluations? The case of Alzheimer's disease. *Eur J Heal Econ* 2019; **20**: 1207–36.

170 Rodriguez-Sanchez B, Aranda-Reneo I, Oliva-Moreno J, Lopez-Bastida J. Assessing the effect of including social costs in economic evaluations of diabetes-related interventions: A systematic review. *Clin Outcomes Res* 2021; **13**: 307–34.

171 Leech AA, Lin PJ, D'Cruz B, Parsons SK, Lavelle TA. Family Spillover Effects: Are Economic Evaluations Misrepresenting the Value of Healthcare Interventions to Society? *Appl Health Econ Health Policy* 2023; **21**: 5–10.

172 Wittenberg E, Prosser LA. Disutility of illness for caregivers and families: A systematic review of the literature. *Pharmacoeconomics* 2013; **31**: 489–500.

173 Al-Janabi H, Van Exel J, Brouwer W, Coast J. A framework for including family

health spillovers in economic evaluation. *Med Decis Mak* 2016; **36**: 176–86.

174 Basu A, Meltzer D. Implications of spillover effects within the family for medical cost-effectiveness analysis. *J Health Econ* 2005; **24**: 751–73.

175 Prosser LA, Lamarand K, Gebremariam A, Wittenberg E. Measuring family HRQoL spillover effects using direct health utility assessment. *Med Decis Mak* 2015; **35**: 81–93.

176 Cubi-Molla P, Buxton M, Devlin N. Allocating Public Spending Efficiently: Is There a Need for a Better Mechanism to Inform Decisions in the UK and Elsewhere? *Appl Health Econ Health Policy* 2021. DOI:10.1007/s40258-021-00648-2.

177 Brouwer WBF. The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. *Pharmacoeconomics* 2019; **37**: 451–6.

178 Henry E, Al-Janabi H, Brouwer W, *et al.* Recommendations for Emerging Good Practice and Future Research in Relation to Family and Caregiver Health Spillovers in Health Economic Evaluation: A Report of the SHEER Task Force. *Pharmacoeconomics* 2023. DOI:10.1007/s40273-023-01321-3.

179 McCabe C. Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. *Pharmacoeconomics* 2019; **37**: 457–60.

180 McCreanor V, Lum E, Graves N, Luo N, Parsonage W, Barnett A. Reducing waste in collection of quality-of-life data through better reporting: a case study. *Qual Life Res* 2022; **31**: 2931–8.

181 Sterling P. Homeostasis vs allostasis implications for brain function and mental disorders. *JAMA Psychiatry* 2014; **71**: 1192–3.

182 Eaton NR, Bringmann LF, Elmer T, *et al.* A review of approaches and models in psychopathology conceptualization research. *Nat Rev Psychol* 2023. DOI:10.1038/s44159-023-00218-4.

183 Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 2018; **18**: 104–16.

184 Sharma D, Aggarwal AK, Downey LE, Prinja S. National Healthcare Economic Evaluation Guidelines: A Cross-Country Comparison. *PharmacoEconomics - Open* 2021; **5**: 349–64.

185 Amri M, Chatur A, O'Campo P. Intersectoral and multisectoral approaches to health policy: an umbrella review protocol. *Heal Res Policy Syst* 2022; **20**: 1–5.

186 Duevel JA, Hasemann L, Peña-Longobardo LM, *et al.* Considering the societal perspective in economic evaluations: a systematic review in the case of depression. *Health Econ Rev* 2020; **10**: 1–19.

187 Catarino A, Harper S, Malcolm R, *et al.* Economic evaluation of 27,540 patients with mood and anxiety disorders and the importance of waiting time and clinical effectiveness in mental healthcare. *Nat Ment Heal* 2023; **1**. DOI:10.1038/s44220-023-00106-z.

188 Mendoza-Jiménez MJ, van Exel J, Brouwer W. On spillovers in economic evaluations: definition, mapping review and research agenda. *Eur J Heal Econ* 2024. DOI:10.1007/s10198-023-01658-8.

189 Woods B, Lomas J, Sculpher M, Weatherly H, Claxton K. Achieving dynamic efficiency in pharmaceutical innovation: Identifying the optimal share of value and payments required. *Heal Econ (United Kingdom)* 2024; : 804–19.

190 Hoedemakers M, Tsiachristas A, Rutten-van Mölken M. Moving Beyond Quality-Adjusted Life-Years in Elderly Care: How Can Multicriteria Decision Analysis Complement Cost-Effectiveness Analysis in Local-Level Decision Making. *Value Heal* 2022. DOI:10.1016/j.jval.2022.04.1728.

191 Jönsson B. Editorial: Ten arguments for a societal perspective in the economic evaluation of medical innovations. *Eur J Heal Econ* 2009; **10**: 357–9.

192 Pokhilenko I, Janssen LMM, Hiligsmann M, *et al.* The Relative Importance of Education and Criminal Justice Costs and Benefits in Economic Evaluations: A Best–Worst Scaling Experiment. *Pharmacoeconomics* 2021; **39**: 99–108.

193 Janssen LMM, Pokhilenko I, Evers SMAA, *et al.* Exploring the identification, validation, and categorization of the cost and benefits of criminal justice in mental health: The PECUNIA project. *Int J Technol Assess Health Care* 2020; **36**: 418–25.

194 Layard R, Clark D, Knapp M, Mayraz G. Cost-benefit analysis of psychological therapy. *Natl Inst Econ Rev* 2007; **202**: 90–8.

195 Dwyer P, Scullion L, Jones K, McNeill J, Stewart ABR. Work, welfare, and wellbeing: The impacts of welfare conditionality on people with mental health impairments in the UK. *Soc Policy Adm* 2020; **54**: 311–26.

196 Keep M. Tax statistics: an overview. 2024 https://commonslibrary.parliament.uk/research-briefings/cbp-8513/ (accessed April 21, 2024).

197 Pollard J, Reardon T, Williams C, *et al.* The multifaceted consequences and economic costs of child anxiety problems: A systematic review and meta-analysis. *JCPP Adv* 2023. DOI:10.1002/jcv2.12149.

198 Hernán MA, Greenland S. Why Stating Hypotheses in Grant Applications Is Unnecessary. *JAMA* 2024; **331**: 285.

199 Binder M, Coad A. An examination of the dynamics of well-being and life events using vector autoregressions. *J Econ Behav Organ* 2010; **76**: 352–71.

200 Frijters P, Johnston DW, Shields MA. THE EFFECT OF MENTAL HEALTH ON EMPLOYMENT: EVIDENCE FROM AUSTRALIAN PANEL DATA. *Health Econ* 2014; **23**: 1058–71.

201 Cornaglia F, Crivellaro E, McNally S. Mental health and education decisions. *Labour Econ* 2015; **33**: 1–12.

202 De Nardi M, Pashchenko S, Porapakkarm P. The lifetime costs of bad health. Cambridge, 2023 http://www.nber.org/papers/w23963.

203 Andersen A, Newman J. Societal and Individual Determinants of Medical Care Utilization in the United States. *Milbank Mem Fund Q Heal Soc Winter 1973;51(1)95-124* 1973; **51**: 95–124.

204 Ribanszki R, S Taylor K, Scheutzow J, Saez Fonseca JA, Ponzo S. Welfare systems and mental health in OECD and EEA countries: a scoping review. *Humanit Soc Sci Commun* 2022; **9**: 1–10.

205 Williams E. Unemployment, sanctions and mental health: The relationship between benefit sanctions and antidepressant prescribing. *J Soc Policy* 2021; **50**: 1–20.

206 Brewer M, Dang T, Tominey E. Universal Credit: Welfare Reform and Mental Health. *J Health Econ* 2024; **98**. DOI:10.2139/ssrn.4114738.

207 Katikireddi SV, Molaodi OR, Gibson M, Dundas R, Craig P. Effects of restrictions to Income Support on health of lone mothers in the UK: a natural experiment study. *Lancet Public Heal* 2018; **3**: e333–40.

208 JE F, MJ S, SA S, MG M. Effects of chronic job insecurity and change in job security on self reported health, minor psychiatric morbidity, physiological measures, and health related behaviours in British civil servants: the Whitehall II study. *J Epidemiol community Heal* 2002; **56**: 450–4.

209 Thomson RM, Igelström E, Purba AK, *et al.* How do income changes impact on

mental health and wellbeing for working-age adults? A systematic review and meta-analysis. *Lancet Public Heal* 2022; **7**: e515–28.

210 Ford E, Clark C, McManus S, *et al.* Common mental disorders, unemployment and welfare benefits in England. *Public Health* 2010; **124**: 675–81.

211 Kahneman D, Deaton A. High income improves evaluation of life but not emotional well-being. *Proc Natl Acad Sci U S A* 2010; **107**: 16489–93.

212 Killingsworth MA. Experienced well-being rises with income, even above $75,000 per year. *Proc Natl Acad Sci U S A* 2021; **118**: 1–6.

213 Mulvaney C, Coupland C, Wilson A, Hammersley V, Dyas J, Carlisle R. Does increased use of private health care reduce the demand for NHS care? A prospective survey of General Practice referrals. *J Public Health (Bangkok)* 2005; **27**: 182–8.

214 Carol P. The demand for private health care in the UK. *J Health Econ* 2000; **19**: 855–876.

215 Grossman M. On the Concept of Health Capital and the Demand for Health. *Polit Econ* 1972; **80**: 223–55.

216 Sheldon KM, Lyubomirsky S. The challenge of staying happier: Testing the hedonic adaptation prevention model. *Personal Soc Psychol Bull* 2012; **38**: 670–80.

217 Diener E, Lucas RE, Lucas RE. Beyond the hedonic treadmill : Revising the adaptation theory of well-being. *Am Psychol* 2006; **61**: 305–14.

218 Bottan NL, Perez Truglia R. Deconstructing the hedonic treadmill: Is happiness autoregressive? *J Socio Econ* 2011; **40**: 224–36.

219 Soral W, Bukowski M, Bilewicz M, *et al.* Prolonged unemployment is associated with control loss and personal as well as social disengagement. *J Pers* 2024; : 1–22.

220 Balboni CA, Bandiera O, Burgess R, Ghatak M, Heil A. Why do people stay poor? 2021 http://www.nber.org/papers/w29340.

221 Mishra GD, Kuh D, Ben-Shlomo Y. Life Course Epidemiology, Second Edi. Elsevier, 2015 DOI:10.1016/B978-0-08-097086-8.14085-1.

222 Ben-Shlomo Y. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol* 2002; **31**: 285–93.

223 Lynch J, Smith GD. A life course approach to chronic disease epidemiology. *Annu Rev Public Health* 2005; **26**: 1–35.

224 Cardoso F, McHayle Z. The economic and social costs of mental ill health: Review of methodology and update of calculations. 2024 https://www.centreformentalhealth.org.uk/publications/the-economic-and-social-costs-of-mental-ill-health/.

225 Greenhalgh T, Fisman D, Cane DJ, Oliver M, MacIntyre CR. Adapt or die: how the pandemic made the shift from EBM to EBM+ more urgent. *BMJ Evidence-Based Med* 2022; **27**: 253–60.

226 Abadie A, Cattaneo MD. Econometric Methods for Program Evaluation. *Annu Rev Econom* 2018; **10**: 465–503.

227 Brimblecombe N, Knapp M, King D, Stevens M, Cartagena Farias J. The high cost of unpaid care by young people: health and economic impacts of providing unpaid care. *BMC Public Health* 2020; **20**: 1–11.

228 Ullah S, Akhtar P, Zaefarian G. Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Ind Mark Manag* 2018; **71**: 69–78.

229 Roodman D. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata J* 2009; **9**: 86–136.

230 Kropko J, Kubinec R. Interpretation and identification of within-unit and cross-sectional variation in panel data models. *PLoS One* 2020; **15**: 1–22.

231 Arellano M, Bover O. Another look at the instrumental variable estimation of error-components models. *J Econom* 1995; **68**: 29–51.

232 Blundell R, Bond S. Initial conditions and Blundell–Bond estimators. *J Econom* 2023; **234**: 101–10.

233 Hayakawa K. First Difference or Forward Orthogonal Deviation- Which Transformation Should be Used in Dynamic Panel Data Models?: A Simulation Study. *Econ Bull* 2009; **29**: 1–10.

234 Abrigo MRM, Love I. Estimation of panel vector autoregression in Stata. *Stata J* 2016; **16**: 778–804.

235 Phillips RF. A Comparison of First-Difference and Forward Orthogonal Deviations GMM. *Econ Bull* 2020; **40**: 1–13.

236 Headey B. Hedonic Adaptation. *Encycl Qual Life Well-Being Res* 2014; **I**: 2831–3.

237 Benjamin DJ, Cooper KB, Heffetz O, Kimball MS. From Happiness Data to

Economic Conclusions. *SSRN Electron J* 2023. DOI:10.2139/ssrn.4568056.

238 Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. *BMJ* 2010; **340**: 1402–6.

239 Kripfganz S. Generalized method of moments estimation of linear dynamic panel data models. In: London Stata Conference 14. 2020: 1–128.

240 Williams R, Allison PD, Moral-Benito E. Linear dynamic panel-data estimation using maximum likelihood and structural equation modeling. *Stata J* 2018; **18**: 293–326.

241 Moral-Benito E, Allison P, Williams R. Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond. *Appl Econ* 2019; **51**: 2221–32.

242 Allison PD, Williams R, Moral-Benito E. Maximum Likelihood for Cross-lagged Panel Models with Fixed Effects. *Socius Sociol Res a Dyn World* 2017; **3**: 237802311771057.

243 Leszczensky L, Wolbring T. How to Deal With Reverse Causality Using Panel Data? Recommendations for Researchers Based on a Simulation Study. *Sociol Methods Res* 2019. DOI:10.1177/0049124119882473.

244 Moral-Benito E. Likelihood-based estimation of dynamic panels with predetermined regressors. *J Bus Econ Stat* 2013; **31**: 451–72.

245 Jorda O. Local Projections for Applied Economics. 2023 DOI:10.24148/wp2023-16.

246 Lynn P. Sample Design for Understanding Society. 2009 https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2009-01.pdf.

247 Buck N, Mcfall S. Understanding Society: design overview. *Longit Life Course Stud* 2012; **3**: 5–17.

248 Institute for Social and Economic Research U of E. Understanding Society Long Term Questionnaire Content Plan. 2019; : 15–37.

249 University of Essex I for S and ER. Understanding Society: Waves 1-13, 2009-2022 and Harmonised BHPS: Waves 1-18, 1991-2009, User Guide. 2023 https://www.understandingsociety.ac.uk/documentation/mainstage/user-guides/main-survey-user-guide/citation/.

250 University of Essex I for S and ER. Understanding Society: Waves 1-13, 2009-2022

and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 18th Edition. 2023. DOI:10.5255/UKDA-SN-6614-19.

251 Curtis S, Cunningham N, Pearce J, Congdon P, Cherrie M, Atkinson S. Trajectories in mental health and socio-spatial conditions in a time of economic recovery and austerity: A longitudinal study in England 2011–17. *Soc Sci Med* 2021; **270**: 113654.

252 Thomas GO, Poortinga W, Sautkina E. The Welsh Single-Use Carrier Bag Charge and behavioural spillover. *J Environ Psychol* 2016; **47**: 126–35.

253 Zhang Y, Bennett MR, Yeandle S. Longitudinal analysis of local government spending on adult social care and carers' subjective well-being in England. *BMJ Open* 2021; **11**. DOI:10.1136/bmjopen-2021-049652.

254 Peasgood T, Bhardwaj A, Biggs K, *et al.* The impact of ADHD on the health and well-being of ADHD children and their siblings. *Eur Child Adolesc Psychiatry* 2016; **25**: 1217–31.

255 Gush K, Scott J, Laurie H. Job loss and social capital: The role of family , friends and wider support networks. 2015 http://hdl.handle.net/10419/126464.

256 Jenkinson C, Layte R. Development and testing of the UK SF-12. *J Heal Serv Res Policy* 1997; **2**: 14–8.

257 Jenkinson C, Chandola T, Coulter A, Bruster S. An assessment of the construct validity of the SF-12 summary scores across ethnic groups. *J Public Health Med* 2001; **23**: 187–94.

258 Jenkinson C, Layte R, Jenkinson D, *et al.* A shorter form health survey: Can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Heal (United Kingdom)* 1997; **19**: 179–86.

259 Ware J, Kosinski M, Turner-Bowker D, Gandek B. How to score version 2 of the SF-12 health survey (with a supplement documenting version 1). Lincoln, RI, 2005.

260 de Beurs E, Boehnke JR, Fried EI. Common measures or common metrics? A plea to harmonize measurement results. *Clin Psychol Psychother* 2022; : 1–13.

261 Olfson M, Sing M, Schlesinger HJ. Mental health/medical care cost offsets: Opportunities for managed care. *Health Aff* 1999; **18**: 79–90.

262 Simon GE, Katzelnick DJ. Depression, use of medical services and cost-offset effects. *J Psychosom Res* 1997; **42**: 333–44.

263 Fiedler JL, Wight JB. The Medical Offset Effect and Public Health Policy: Mental

Health Industry in Transition. Praeger Publishers, 1989.

264 Chiles JA, Lambert MJ, Hatch AL. The Impact of Psychological Interventions on Medical Cost Offset: A Meta-analytic Review. *Clin Psychol Sci Pract* 1999; **6**: 204–20.

265 Ronaldson A, Elton L, Jayakumar S, Jieman A, Halvorsrud K, Bhui K. Severe mental illness and health service utilisation for nonpsychiatric medical disorders: A systematic review and meta-analysis. 2020 DOI:10.1371/journal.pmed.1003284.

266 Ride J, Kasteridis P, Gutacker N, *et al.* Impact of prevention in primary care on costs in primary and secondary care for people with serious mental illness. *Heal Econ (United Kingdom)* 2022; : 343–55.

267 Gruber J, Lordan G, Pilling S, Propper C, Saunders R. The impact of mental health support for the chronically ill on hospital utilisation: Evidence from the UK. *Soc Sci Med* 2022; **294**: 114675.

268 Curtis L, Burns A. Unit Costs of Health & Social Care 2020. Unit Costs of Health and Social Care. 2020 DOI:10.22024/UniKent/01.02.84818.

269 Office for National Statistics. Consumer Prices Index including owner occupiers' housing costs (CPIH). 2023. https://www.ons.gov.uk/datasets/cpih01/editions/time-series/versions/32 (accessed April 18, 2023).

270 NHS Improvement. Archived Reference Costs. 2017/18 Reference Costs. 2020. https://webarchive.nationalarchives.gov.uk/ukgwa/20200501111106/https://improvement.nhs.uk/resources/reference-costs/ (accessed July 26, 2022).

271 Sandmann FG, Robotham J V., Deeny SR, Edmunds WJ, Jit M. Estimating the opportunity costs of bed-days. *Heal Econ (United Kingdom)* 2018; **27**: 592–605.

272 Fisher P, Fumagalli L, Buck N, Avram S. Understanding Society and its income data. *Underst Soc Work Pap Ser* 2019.

273 Shen X, Wang C, Zhou X, *et al.* Nonlinear dynamics of multi-omics profiles during human aging. *Nat aging* 2024. DOI:10.1038/s43587-024-00692-2.

274 Thomson RM, Kopasker D, Leyland A, Pearce A, Katikireddi SV. Effects of poverty on mental health in the UK working-age population: causal analyses of the UK Household Longitudinal Study. *Int J Epidemiol* 2023; **52**: 512–22.

275 Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004; **42**: 851–9.

276 Brown CC, Tilford JM, Payakachat N, *et al.* Measuring Health Spillover Effects in Caregivers of Children with Autism Spectrum Disorder: A Comparison of the EQ-5D-3L and SF-6D. *Pharmacoeconomics* 2019; **37**: 609–20.
277 Jacobsson F, Carstensen J, Borgquist L. Caring externalities in health economic evaluation: How are they related to severity of illness? *Health Policy (New York)* 2005; **73**: 172–82.
278 Pauly M V. Valuing health care benefits in money terms. In: Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies. Cambridge University Press, 1995: 99–124.
279 Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: A framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002; **11**: 415–30.
280 Cubi-Molla P, Mott D, Henderson N, Zamora B, Grobler M, Garau M. Resource allocation in public sector programmes: does the value of a life differ between governmental departments? *Cost Eff Resour Alloc* 2023; **21**: 1–18.
281 Lomas J, Martin S, Claxton K. Estimating the Marginal Productivity of the English National Health Service From 2003 to 2012. *Value Heal* 2019; **22**: 995–1002.
282 Martin S, Lomas J, Claxton K, Longo F. How Effective is Marginal Healthcare Expenditure? New Evidence from England for 2003/04 to 2012/13. *Appl Health Econ Health Policy* 2021; **19**: 885–903.
283 Martin S, Claxton K, Lomas J, Longo F. The impact of different types of NHS expenditure on health: Marginal cost per QALY estimates for England for 2016/17. *Health Policy (New York)* 2023; **132**. DOI:10.1016/j.healthpol.2023.104800.
284 Sampson C, Cookson G. Marginal cost per QALY estimates: What are they good for? *Health Policy (New York)* 2024; **142**. DOI:10.1016/j.healthpol.2024.105036.
285 Pevalin DJ. Multiple applications of the GHQ-12 in a general population sample: An investigation of long-term retest effects. *Soc Psychiatry Psychiatr Epidemiol* 2000; **35**: 508–12.
286 O'Connor DW, Parslow RA. Mental health scales and psychiatric diagnoses: Responses to GHQ-12, K-10 and CIDI across the lifespan. *J Affect Disord* 2010; **121**: 263–7.
287 Niileksela CR, Jones NB. Measurement Equality of Frequency and Severity Item

Response Options on Depression and Generalized Anxiety Scales. *Assessment* 2022. DOI:10.1177/10731911221134599.

288 Dalgaard CJ, Strulik H. Optimal aging and death: Understanding the preston curve. *J Eur Econ Assoc* 2014; **12**: 672–701.

289 Dalgaard C-J, Hansen CW, Strulik H. Accounting for fetal origins: Health capital vs. health deficits. 2019 https://www.jstor.org/stable/90023279.

290 Dalgaard CJ, Hansen CW, Strulik H. Fetal origins—A life cycle model of health and aging from conception to death. *Heal Econ (United Kingdom)* 2021; **30**: 1276–90.

291 Strulik H. Myopic misery: Maternal depression, child investments, and the neurobiological poverty trap. *Macroecon Dyn* 2019; **23**: 522–34.

292 Lutz W, Schwartz B, Delgadillo J. Measurement-Based and Data-Informed Psychological Therapy. *Annu Rev Clin Psychol* 2022; **18**: 71–98.

293 Jorm AF. National surveys of mental disorders: Are they researching scientific facts or constructing useful myths? *Aust N Z J Psychiatry* 2006; **40**: 830–4.

294 MacNeil Vroomen J, Eekhout I, Dijkgraaf MG, *et al.* Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best? *Eur J Heal Econ* 2016; **17**: 939–50.

295 Manning WG, Mullahy J. Estimating log models: To transform or not to transform? *J Health Econ* 2001; **20**: 461–94.

296 Malehi AS, Pourmotahari F, Angali KA. Statistical models for the analysis of skewed healthcare cost data: a simulation study. *Health Econ Rev* 2015; **5**. DOI:10.1186/s13561-015-0045-7.

297 Lynn P, Borkowska M. Some Indicators of Sample Representativeness and Attrition Bias for BHPS and Understanding Society. 2018.

298 Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature. *Am J Epidemiol* 2008; **168**: 355–7.

299 Royston P, White IR. Journal of Statistical Software Multiple Imputation by Chained Equations (MICE): Implementation in Stata. *J Stat Softw* 2011; **45**: 1–20.

300 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011; **30**: 377–99.

301 van Ginkel JR, Linting M, Rippe RCA, van der Voort A. Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *J Pers Assess* 2020; **102**: 297–308.

302 Kleinke K, Reinecke J, Salfrán D, Spiess M. Multiple Imputation: New Developments. In: Kleinke K, Reinecke J, Salfrán D, Spiess M, eds. Applied Multiple Imputation. Cham: Springer International Publishing, 2020: 219–56.

303 Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Model* 2001; **8**: 430–57.

304 Lee T, Shi D. A Comparison of Full Information Maximum Likelihood and Multiple Imputation in Structural Equation Modeling With Missing Data. *Psychol Methods* 2021; **26**: 466–85.

305 Larsen R. Missing data imputation versus full information maximum likelihood with second-level dependencies. *Struct Equ Model* 2011; **18**: 649–62.

306 Liu S, Molenaar PCM. iVAR: A program for imputing missing data in multivariate time series using vector autoregressive models. *Behav Res Methods* 2014; **46**: 1138–48.

307 James ND. Respondents for Nearly Three Decades: How do loyal sample members differ from others? *Surv Res Methods* 2023; **17**: 15–36.

308 Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol* 2013; **42**: 1012–4.

309 Ebrahim S, Smith GD. Commentary: Should we always deliberately be non-representative? *Int J Epidemiol* 2013; **42**: 1022–6.

310 Holtz-Eakin D, Newey W, Rosen HS. Estimating Vector Autoregressions with Panel Data. *Econometrica* 1988; **56**: 1371.

311 StataCorp. Stata Statistical Software: Release 17. 2021.

312 Andrews DWK, Lu B. Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. *J Econom* 2001; **101**: 123–64.

313 Kiviet JF. Microeconometric dynamic panel data methods: Model specification and selection issues. *Econom Stat* 2020; **13**: 16–45.

314 Burdisso T, Sangiácomo M. Panel Time Series: Review of the Methodological Evolution. *Stata J* 2016; **16**: 424–42.

315 Wooldridge JM. Econometric analysis of cross section and panel data. MIT Press, 2010.

316 Imbens GW. Statistical significance, p-values, and the reporting of uncertainty. *J*

Econ Perspect 2021; **35**: 157–74.

317 Summers JA, Wilson N, Blakely T, Sigglekow F. Disease-Related Loss to Government Funding: Longitudinal Analysis of Individual-Level Health and Tax Data for an Entire Country. *Value Heal* 2023; **26**: 170–5.

318 Wakefield S, Kellett S, Simmonds-Buckley M, Stockton D, Bradbury A, Delgadillo J. Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A systematic review and meta-analysis of 10-years of practice-based evidence. *Br J Clin Psychol* 2021; **60**: 1–37.

319 Chen L, Clarke PM, Petrie DJ, Staub KE. The effects of self-assessed health: Dealing with and understanding misclassification bias. *J Health Econ* 2021; **78**: 102463.

320 Thorn JC, Coast J, Cohen D, *et al.* Resource-use measurement based on patient recall: Issues and challenges for economic evaluation. *Appl Health Econ Health Policy* 2013; **11**: 155–61.

321 Patel A, Rendu A, Moran P, Leese M, Mann A, Knapp M. A comparison of two methods of collecting economic data in primary care. *Fam Pract* 2005; **22**: 323–7.

322 Byford S, Leese M, Knapp M, *et al.* Comparison of alternative methods of collection of service use data for the economic evaluation health care interventions. *Health Econ* 2007; **16**: 531–6.

323 Mukuria C, Rowen D, Harnan S, *et al.* An Updated Systematic Review of Studies Mapping (or Cross-Walking) Measures of Health-Related Quality of Life to Generic Preference-Based Measures to Generate Utility Values. *Appl Health Econ Health Policy* 2019; **17**: 295–313.

324 Smith A. Of Sympathy. In: The theory of moral sentiments. HG Bohn, 1853.

325 Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. *J Health Econ* 2008; **27**: 325–38.

326 Graaff J. Theoretical welfare economics. CUP Archive, 1967.

327 Buchanan JM, Stubblebine WC. Externality. *Economica* 1962; **29**: 371.

328 Clark CW. Commons, Concept and Theory of. *Encycl Biodivers Second Ed* 2013; **2**: 149–54.

329 Cheung SNS. The Myth of Social Cost: A critique of welfare, economics and the implications of public policy. 1978; : 97.

330 Arrow K. Political and Economic Evaluation of Social Effects and Externalities. In: Margolis J, ed. The Analysis of Public Output. National Bureau of Economic

Research, 1970: 1–30.

331 Culyer AJ, Simpson H. Externality Models and Health: a Rückblick over the last Twenty Years. *Econ Rec* 1980; **56**: 222–30.

332 Muir KJ, Keim-Malpass J. Analyzing the concept of spillover effects for expanded inclusion in health economics research. *J Comp Eff Res* 2020; **9**: 755–66.

333 Galizzi MM, Whitmarsh L. How to measure behavioral spillovers: A methodological review and checklist. *Front Psychol* 2019; **10**: 1–15.

334 Benjamin-Chung J, Arnold BF, Berger D, *et al.* Spillover effects in epidemiology: Parameters, study designs and methodological considerations. *Int J Epidemiol* 2018; **47**: 332–47.

335 Angelucci M, Di Maro V. Programme evaluation and spillover effects. *J Dev Eff* 2016; **8**: 22–43.

336 Yu CL, Airoldi EM, Borgs C, Chayes JT. Estimating the total treatment effect in randomized experiments with unknown network structure. *Proc Natl Acad Sci U S A* 2022; **119**. DOI:10.1073/pnas.2208975119.

337 Richardson TS, Robins JM. Single World Intervention Graphs: A Primer. *Proc 2nd Conf Uncertain Artifical Intell* 2013; : 1–11.

338 Tchetgen EJT, Vanderweele TJ. On causal inference in the presence of interference. *Stat Methods Med Res* 2012; **21**: 55–75.

339 Culyer AJ. The normative economics of health care finance and provision. *Oxford Rev Econ Policy* 1989; **5**: 34–58.

340 Francetic I, Meacock R, Elliott J, *et al.* Framework for identification and measurement of spillover effects in policy implementation: intended non-intended targeted non-targeted spillovers (INTENTS). *Implement Sci Commun* 2022; **3**: 1–15.

341 Al-Janabi H, McLoughlin C, Oyebode J, Efstathiou N, Calvert M. Six mechanisms behind carer wellbeing effects: A qualitative study of healthcare delivery. *Soc Sci Med* 2019; **235**: 112382.

342 Dixon P, Round J. Caring for Carers: Positive and Normative Challenges for Future Research on Carer Spillover Effects in Economic Evaluation. *Value Heal* 2019; **22**: 549–54.

343 Bhadhuri A, Al-Janabi H, Jowett S, Jolly K. Incorporating Household Spillovers in Cost Utility Analysis: A Case Study Using Behavior Change in COPD. *Int J Technol*

Assess Health Care 2019; **35**: 212–20.

344 Bobinac A, Van NJA, Rutten FFH, Brouwer WBF. Health effects in significant others: Separating family and care-giving effects. *Med Decis Mak* 2011; **31**: 292–8.

345 Long TJ, Harrington R V., Rowan CB, Schwartz-Mette RA. Deleterious Friendship Behaviors. Elsevier Inc., 2021 DOI:10.1016/b978-0-12-818872-9.00043-1.

346 Burt RS. Social Contagion and Innovation: Cohesion versus Structural Equivalence. *Am J Sociol* 1987; **92**: 1287–335.

347 Dishion TJ, Tipsord JM. Peer contagion in child and adolescent social and emotional development. *Annu Rev Psychol* 2011; **62**: 189–214.

348 Block P, Burnett Heyes S. Sharing the load: Contagion and tolerance of mood in social networks. *Emotion* 2020. DOI:10.1037/emo0000952.

349 Block P, Grund T. Multidimensional homophily in friendship networks. *Netw Sci* 2014; **2**: 189–212.

350 Smith RL, Rose AJ. The ‘cost of caring’ in youths’ friendships: Considering associations among social perspective taking, co-rumination, and empathetic distress. *Dev Psychol* 2011; **47**: 1792–803.

351 Spendelow JS, Simonds LM, Avery RE. The Relationship between Co-rumination and Internalizing Problems: A Systematic Review and Meta-analysis. *Clin Psychol Psychother* 2017; **24**: 512–27.

352 Stafford LL. Is Codependency a Meaningful Concept? *Issues Ment Health Nurs* 2001; **22**: 273–86.

353 Marks ADG, Blore RL, Hine DW, Dear GE. Development and validation of a revised measure of codependency. *Aust J Psychol* 2012; **64**: 119–27.

354 Mingo MV, Goldberg J, Castro M de los A, Fillol MP, Mongillo M, Bedregal P. Supporting the caregiver-child dyad’s relationship: An evaluation of implementation quality in the Chilean Crecer Jugando program. *Eval Program Plann* 2019; **76**: 101668.

355 Horesh D, Hasson-Ohayon I, Harwood-Gross A. The Contagion of Psychopathology across Different Psychiatric Disorders: A Comparative Theoretical Analysis. *Brain Sci* 2022; **12**. DOI:10.3390/brainsci12010067.

356 Dewhurst K, Todd J. The psychosis of association—folie a deux. *J Nerv Ment Dis* 1956; **124**: 451–9.

357 Gralnick A. Folie a deux-the psychosis of association - A review of 103 cases and

the entire English literature: With case presentations. *Psychiatr Q* 1942; **16**: 230–63.

358 Patel AS, Armone D, Ryan W. Folie a deux in bipolar affective disorder: A case report. *Bipolar Disord* 2004; **6**: 162–5.

359 SanjurjoHartman T, Weitzner MA, Santana C, Devine C, Grendys E. Cancer and Folie à Deux. . *Cancer Pract* 2008; **9**: 290–4.

360 Magar B, Fahy S. Is every shared psychosis a folie à deux? *Int J Geriatr Psychiatry* 2010; **25**: 1197–8.

361 Paz LV, Viola TW, Milanesi BB, *et al.* Contagious depression: Automatic mimicry and the mirror neuron system - A review. *Neurosci Biobehav Rev* 2022; **134**. DOI:10.1016/j.neubiorev.2021.12.032.

362 Cicchetti D, Rogosch FA. Equifinality and multifinality in developmental psychopathology. *Dev Psychopathol* 1996; **8**: 597–600.

363 Claxton K, Lomas J, Longo F, Salas Ortiz A. Apples and oranges. *Health Policy (New York)* 2024; **143**. DOI:10.1016/j.healthpol.2024.105041.

364 Robling MR, Ingledew DK, Greene G, *et al.* Applying an extended theoretical framework for data collection mode to health services research. *BMC Health Serv Res* 2010; **10**. DOI:10.1186/1472-6963-10-180.

365 Chesterton G. What's wrong with the world. CreateSpace Independent Publishing Platform, 1910.

366 Micklewright J, Schnepf S. How reliable are income data collected with a single question? *J R Stat Soc Ser A (Statistics Soc* 2010; **173**: 409–29.

367 Naik Y, Baker P, Ismail SA, *et al.* Going upstream - An umbrella review of the macroeconomic determinants of health and health inequalities. *BMC Public Health* 2019; **19**: 1–19.

368 Laporte A. October Newsletter: President's Message. 2022.

369 Office for Budget Responsibility. Fiscal risks and sustainability report - July 2023. 2023 https://obr.uk/frs/fiscal-risks-and-sustainability-july-2023/.

370 Darzi A. Independent Investigation of the National Health Service in England. 2024 https://www.gov.uk/government/publications/independent-investigation-of-the-nhs-in-england.

371 Department for Work and Pensions HG. Get Britain Working White Paper. 2024 https://www.gov.uk/government/publications/get-britain-working-white-

paper/get-britain-working-white-paper.
372 Garau M, Shah KK, Sharma P, Towse A. Is the link between health and wealth considered in decision making? Results from a qualitative study. *Int J Technol Assess Health Care* 2016; **31**: 449–56.
373 Shearer J, Byford S, Birch S. Reflections on the nice decision to reject patient production losses. *Int J Technol Assess Health Care* 2017; **33**: 638–43.
374 Office for National Statistics. Table 1.7a Annual pay - Gross (£) - For all employee jobs: United Kingdom, 2024. Earn. hours Work. all employees ASHE Table 1. 2024. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/allemployeesashetable1 (accessed Nov 15, 2024).
375 Department for Work and Pensions HG. Estimate your Income Tax for the current year. https://www.gov.uk/estimate-income-tax (accessed Nov 15, 2024).
376 Walker T. Why economists dislike a lump of labor. *Rev Soc Econ* 2007; **65**: 279–91.
377 Karanassou M, Snower DJ. Unemployment invariance. *Ger Econ Rev* 2004; **5**: 297–317.
378 Sampson C. NICE and the EQ-5D-5L: Ten Years Trouble. *PharmacoEconomics - Open* 2022; **6**: 5–8.
379 Donaldson C, Currie G, Mitton C. Cost effectiveness analysis in health care: Contraindications. *Br Med J* 2002; **325**: 891–4.
380 Amri M, Chatur A, O'Campo P. An umbrella review of intersectoral and multisectoral approaches to health policy. *Soc Sci Med* 2022; **315**: 115469.
381 Edwards RT. The Inclusion of Spillover Effects in Economic Evaluation: A Public Health Economics Perspective. *Pharmacoeconomics* 2019; **37**: 621–3.
382 Boysen M, Watson I. Reflections on NICE's Uptake of New Methods: Past, Present, and the 2020 Review. *Pharmacoeconomics* 2020; **38**: 243–5.
383 Dakin H, Tsiachristas A. Rationing in an Era of Multiple Tight Constraints: Is Cost-Utility Analysis Still Fit for Purpose? *Appl Health Econ Health Policy* 2024; **25**: S7.
384 Hoefman RJ, Van Exel J, Brouwer W. How to include informal care in economic evaluations. *Pharmacoeconomics* 2013; **31**: 1105–19.
385 Koopmanschap MA, Van Exel JNA, Van Den Berg B, Brouwer WBF. An overview of methods and applications to value informal care in economic evaluations of healthcare. *Pharmacoeconomics* 2008; **10**: 99–111.

386 Al-Janabi H, Flynn TN, Coast J. QALYs and carers. *Pharmacoeconomics* 2011; **29**: 1015–23.
387 Rand S, Malley J, Vadean F, Forder J. Measuring the outcomes of long-term care for unpaid carers: Comparing the ASCOT-Carer, Carer Experience Scale and EQ-5D-3 L. *Health Qual Life Outcomes* 2019; **17**: 1–12.
388 Dow J, Robinson J, Robalino S, Finch T, McColl E, Robinson L. How best to assess quality of life in informal carers of people with dementia; A systematic review of existing outcome measures. *PLoS One* 2018; **13**: 1–18.
389 McCaffrey N, Bucholc J, Rand S, *et al.* Head-to-Head Comparison of the Psychometric Properties of 3 Carer-Related Preference-Based Instruments. *Value Heal* 2020; **23**: 1477–88.
390 Batchelder L, Malley J, Burge P, *et al.* Carer Social Care-Related Quality of Life Outcomes: Estimating English Preference Weights for the Adult Social Care Outcomes Toolkit for Carers. *Value Heal* 2019; **22**: 1427–40.
391 United Kingdom Department of Health & Social Care. The NHS Constitution for England. 2023; : 1–18.
392 Tilford JM, Payakachat N. Progress in measuring family spillover effects for economic evaluations. *Expert Rev Pharmacoeconomics Outcomes Res* 2015; **15**: 195–8.
393 Jit M, Hutubessy R. Methodological Challenges to Economic Evaluations of Vaccines: Is a Common Approach Still Possible? *Appl Health Econ Health Policy* 2016; **14**: 245–52.
394 Bastiampillai T, Allison S, Chan S. Is depression contagious? the importance of social networks and the implications of contagion theory. *Aust N Z J Psychiatry* 2013; **47**: 299–303.
395 Alho J, Gutvilig M, Niemi R, *et al.* Transmission of Mental Disorders in Adolescent Peer Networks. *JAMA Psychiatry* 2024; : 1–7.
396 Lacey RE, Xue B, Di Gessa G, Lu W, McMunn A. Mental and physical health changes around transitions into unpaid caregiving in the UK: a longitudinal, propensity score analysis. *Lancet Public Heal* 2024; **9**: e16–25.
397 Bauer JM, Sousa-Poza A. Impacts of Informal Caregiving on Caregiver Employment, Health, and Family. *J Popul Ageing* 2015; **8**: 113–45.
398 Department of Health. No health without mental health: A Cross-Government

Mental Health Outcomes Strategy for People of All Ages. 2011 https://www.gov.uk/government/publications/no-health-without-mental-health-a-cross-government-outcomes-strategy.

399 Macintyre A, Ferris D, Gonçalves B, Quinn N. What has economics got to do with it? The impact of socioeconomic factors on mental health and the case for collective action. *Palgrave Commun* 2018; **4**. DOI:10.1057/s41599-018-0063-2.

400 Kirkbride J, Anglin D, Colman I, *et al.* The social determinants of mental health and disorder: evidence, prevention, and recommendations. *World Psychiatry* 2023; : 58–90.

401 Henry E, Cullinan J. Maternal Mental Health Spillovers From Child Illness and Disability: A Dynamic Panel Analysis. *Value Heal* 2024. DOI:10.1016/j.jval.2024.09.011.

402 Quashie NT, Wagner M, Verbakel E, Deindl C. Socioeconomic differences in informal caregiving in Europe. *Eur J Ageing* 2022; **19**: 621–32.

403 Barrett P, Hale B, Butler M. Informal Care in Context: An Expression of Social Relationships. In: Family Care and Social Capital: Transitions in Informal Care. Springer, 2014. DOI:https://doi.org/10.1007/978-94-007-6872-7_2.

404 Mott D, Schirrmacher H, Al-Janabi H, *et al.* Modelling Spillover Effects on Informal Carers: The 'Carer QALY Trap'. *Pharmacoeconomics* 2023. DOI:10.1016/j.jval.2022.09.022.

405 Henry E, Cullinan J. Addressing the distributional consequences of spillovers in health economic evaluation: A prioritarian approach. *Heal Econ (United Kingdom)* 2024; : 1–15.

406 Lippert-Rasmussen K, Lauridsen S. Justice and the allocation of healthcare resources: Should indirect, non-health effects count? *Med Heal Care Philos* 2010; **13**: 237–46.

407 Du Toit J, Millum J. Are Indirect Benefits Relevant to Health Care Allocation Decisions? *J Med Philos (United Kingdom)* 2016; **41**: 540–57.

408 Lavelle TA, D'Cruz BN, Mohit B, *et al.* Family Spillover Effects in Pediatric Cost-Utility Analyses. *Appl Health Econ Health Policy* 2019; **17**: 163–74.

409 Wittenberg E, James LP, Prosser LA. Spillover Effects on Caregivers' and Family Members' Utility: A Systematic Review of the Literature. *Pharmacoeconomics* 2019; **37**: 475–99.

410 Manski CF. Economic analysis of social interactions. *J Econ Perspect* 2000; **14**: 115–36.

411 Carswell C, Brown JVE, Lister J, *et al.* The lived experience of severe mental illness and long-term conditions: a qualitative exploration of service user, carer, and healthcare professional perspectives on self-managing co-existing mental and physical conditions. *BMC Psychiatry* 2022; **22**: 1–14.

412 Pearce E, Birken M, Pais S, *et al.* Associations between constructs related to social relationships and mental health conditions and symptoms: an umbrella review. *BMC Psychiatry* 2023; **23**. DOI:10.1186/s12888-023-05069-0.

413 Sul S, Kim J, Choi I. Subjective Well-Being and Hedonic Editing: How Happy People Maximize Joint Outcomes of Loss and Gain. *J Happiness Stud* 2013; **14**: 1409–30.

414 Airoldi EM, Christakis NA. Induction of Social Contagion for Diverse Outcomes in Structured Experiments in Isolated Villages. *Science (80-)* 2024; **384**. DOI:10.1126/science.adi5147.

415 Steele F, Clarke PS, Kuha J. Modeling within-household associations in household panel studies. *Ann Appl Stat* 2019; **13**: 367–92.

416 Squires H, Kelly MP, Gilbert N, Sniehotta F, Purshouse RC. The long-term effectiveness and cost-effectiveness of public health interventions; how can we model behavior? A review. *Heal Econ (United Kingdom)* 2023; : 1–19.

417 Kim DA, Hwong AR, Stafford D, *et al.* Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *Lancet* 2015; **386**: 145–53.

418 Martinsen KD, Rasmussen LMP, Wentzel-Larsen T, *et al.* Prevention of anxiety and depression in school children: Effectiveness of the transdiagnostic EMOTION Program. *J Consult Clin Psychol* 2019; **87**: 212–9.

419 Loevaas MES, Lydersen S, Sund AM, *et al.* A 12-month follow-up of a transdiagnostic indicated prevention of internalizing symptoms in school-aged children: The results from the EMOTION study. *Child Adolesc Psychiatry Ment Health* 2020; **14**: 1–13.

Appendix

Glossary of terms and definitions

Abbreviation or term	Meaning
BHPS	British Household Panel Survey
CBA	Cost benefit analysis
CBT	Cognitive behavioural therapy
CEA	Cost effectiveness analysis
Cross-sectoral	Efforts that break down boundaries between sectors, involving shared initiatives and resource exchange
CUA	Cost utility analysis
DALY	Disability-adjusted life year
EQ-5D	Euroqol-5D (a generic five-dimensional health status instrument)
FD	First differences
FOD	Forward orthogonal deviations
GHQ-12	12-item General Health Questionnaire
GMM	Generalised method of moments
GP	General practitioner
GPBMoH	Generic preference-based measure of health
HRQoL	Health-related quality of life
HTA	Health technology assessment
IAPT	Improving Access to Psychological Therapies
Interindividual	Refers to variations, differences, or interactions between individuals
Intersectoral	Consideration of, or collaboration at, the juncture between different sectors to address complex issues
MAR	Missing at random
MHSDS	Mental Health Services Dataset
ML	Machine learning
MNAR	Missing not at random
Multi-sectoral	Involvement of multiple sectors working towards a common goal, often in parallel rather than in an integrated way
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMB	Net monetary benefit
OBS	Office for Budget Responsibility
ONS	Office for National Statistics
PVAR	Panel vector autoregression
QALY	Quality-adjusted life year

RCT	Randomised controlled trial
SF-12	Short-Form 12 Health Survey
SF-6D (SF-12)	Preference-based measure of health derived from the SF-12
Spillover	All impacts from an intervention on all parties or entities other than the users of the intervention under evaluation (e.g., carer health or multi-sectoral costs)
STM	State transition model
UKHLS	UK Household Longitudinal Study
VAR	Vector autoregression
WTA	Willingness-to-accept
WTP	Willingness-to-pay

Descriptive statistics by sample wave

Table A1 - Wave g	*n*	Mean	SD	Min	p50	Max
Age	27185	45.11	10.98	25	46	64
Total GHQ	24392	11.22	5.66	0	10	36
MCS	24276	48.25	10.31	0	50.6	75.51
PCS	24276	50.66	10.29	4.85	54.26	73.59
SF-6D	24362	0.79	0.14	0.34	0.8	1
GP Visits (Last 12-months)	25383	2.71	3.13	0	1.5	11
Outpatient Visits (Last 12-months)	25394	1.39	2.52	0	0	11
Inpatient Days (Last 12-months)	25421	0.53	4.76	0	0	238
Welfare Benefits (Monthly)	27184	268.82	520.1	0	0	5381.39
Gross Labour Income (Monthly)	27184	1881.38	1965.18	-48000	1521.46	18875.91
Net Labour Income (Monthly)	27184	1414.2	1344.64	-48000	1289.28	12232.08
Annualised Healthcare Costs	27185	864.93	1514.16	0	298.92	5514.49
Annualised Welfare Benefits	27185	3145.57	5850.43	0	0	26170.68
Annualised Labour Tax Revenue	27185	5533.52	7912.26	0	2718.24	38553.36

Table A2 - Wave h	*n*	Mean	SD	Min	p50	Max
Age	25087	45.42	10.96	25	46	64
Total GHQ	23155	11.38	5.71	0	10	36
MCS	23064	47.99	10.37	0	50.17	75.2
PCS	23064	50.72	10.24	5.83	54.3	74.09
SF-6D	23131	0.78	0.14	0.34	0.8	1
GP Visits (Last 12-months)	24031	2.88	3.09	0	1.5	11
Outpatient Visits (Last 12-months)	24032	1.41	2.48	0	0	11
Inpatient Days (Last 12-months)	24064	0.55	4.9	0	0	270
Welfare Benefits (Monthly)	25086	251.54	499.7	0	0	5289.93
Gross Labour Income (Monthly)	25086	1906.37	1990.59	-62000	1571.5	18707.59
Net Labour Income (Monthly)	25086	1428.39	1373.77	-62000	1327.92	12123
Annualised Healthcare Costs	25087	763.77	1332.63	0	298.92	5514.49
Annualised Welfare Benefits	25087	2950.15	5656.89	0	0	26170.68
Annualised Labour Tax Revenue	25087	5673.03	8037.52	0	2694	38553.36

Table A3 - Wave i	*n*	Mean	SD	Min	p50	Max
Age	22835	45.76	11.06	25	47	64
Total GHQ	21286	11.59	5.79	0	11	36
MCS	21142	47.79	10.46	0	49.89	75.54
PCS	21142	50.58	10.33	5.98	54.03	73.59
SF-6D	21204	0.78	0.14	0.34	0.8	1
GP Visits (Last 12-months)	22155	2.75	2.95	0	1.5	11
Outpatient Visits (Last 12-months)	22162	1.45	2.48	0	0	11
Inpatient Days (Last 12-months)	22177	0.53	4.63	0	0	250
Welfare Benefits (Monthly)	22834	232.33	479.86	0	0	4797.44
Gross Labour Income (Monthly)	22834	1869.14	1887.83	-1090.2	1544.45	17806.24
Net Labour Income (Monthly)	22834	1406.74	1269.89	-1090.2	1308.24	10683.96
Annualised Healthcare Costs	22835	697.65	1224.35	0	298.92	5514.49
Annualised Welfare Benefits	22835	2725.43	5424.07	0	0	26170.68
Annualised Labour Tax Revenue	22835	5507.27	7855.09	0	2616.48	38553.36

Table A4 - Wave j	*n*	Mean	SD	Min	p50	Max
Age	21538	45.89	11.09	25	47	64
Total GHQ	20251	11.7	5.79	0	11	36
MCS	20067	47.11	10.58	0	49.04	72.86
PCS	20067	50.5	10.03	6.12	53.7	70.9
SF-6D	20175	0.77	0.13	0.34	0.8	1
GP Visits (Last 12-months)	21022	2.31	2.79	0	1.5	11
Outpatient Visits (Last 12-months)	21022	1.32	2.37	0	0	11
Inpatient Days (Last 12-months)	21037	0.42	3.93	0	0	300
Welfare Benefits (Monthly)	21537	213.69	458.73	0	0	5598.18
Gross Labour Income (Monthly)	21537	1876.9	1902.99	-3549	1569.37	17744.29
Net Labour Income (Monthly)	21537	1408.38	1284.82	-3549	1304.26	17567.55
Annualised Healthcare Costs	21538	613.41	1147.42	0	220.06	5514.49
Annualised Welfare Benefits	21538	2516.31	5235.08	0	0	26170.68
Annualised Labour Tax Revenue	21538	5584.65	7934.25	0	2555.28	38553.36

Table A5 - Wave k	*n*	Mean	SD	Min	p50	Max
Age	19981	46.12	11.16	25	47	64
Total GHQ	19105	11.99	5.84	0	11	36
MCS	18939	46.84	10.65	0.6	48.87	76.25
PCS	18939	50.71	9.91	6.27	53.76	73.8
SF-6D	19038	0.77	0.13	0.34	0.8	1
GP Visits (Last 12-months)	19668	2.6	2.89	0	1.5	11
Outpatient Visits (Last 12-months)	19666	1.33	2.36	0	0	11
Inpatient Days (Last 12-months)	19698	0.48	4.24	0	0	250
Welfare Benefits (Monthly)	19979	208.06	461.67	0	0	5042.93
Gross Labour Income (Monthly)	19979	1904.12	1921.26	-5034.56	1573.35	17480.97
Net Labour Income (Monthly)	19979	1424.06	1293.21	-5034.56	1340.87	11328.12
Annualised Healthcare Costs	19981	594.46	1071.26	0	298.92	5514.49
Annualised Welfare Benefits	19981	2438.82	5208.23	0	0	26170.68
Annualised Labour Tax Revenue	19981	5721.09	8029.06	0	2668.44	38553.36

Table A6 - Wave l	*n*	Mean	SD	Min	p50	Max
Age	18307	46.27	11.26	25	47	64
Total GHQ	17821	12.24	5.93	0	11	36
MCS	17728	46.47	10.76	0	48.76	74.37
PCS	17728	51.06	9.89	7.87	54.26	75.2
SF-6D	17775	0.77	0.13	0.34	0.8	1
GP Visits (Last 12-months)	18123	2.14	2.6	0	1.5	11
Outpatient Visits (Last 12-months)	18121	1.08	2.06	0	0	11
Inpatient Days (Last 12-months)	18165	0.42	3.93	0	0	180
Welfare Benefits (Monthly)	18304	198.02	449.32	0	0	5734.92
Gross Labour Income (Monthly)	18304	1943.05	1939.85	-5819.78	1648.56	16068.84
Net Labour Income (Monthly)	18304	1451.02	1306.34	-5819.78	1355.22	9995.52
Annualised Healthcare Costs	18307	486.92	966.75	0	112.1	5514.49
Annualised Welfare Benefits	18307	2328.75	5098.08	0	0	26170.68
Annualised Labour Tax Revenue	18307	5866.85	8113.8	0	2748.72	38553.36

Table A7 - Wave m	*n*	Mean	SD	Min	p50	Max
Age	17515	46.44	11.3	25	48	64
Total GHQ	16975	12.09	5.95	0	11	36
MCS	16905	46.61	10.95	0	48.87	73.98
PCS	16905	50.68	9.87	5.68	53.71	74.17
SF-6D	16961	0.77	0.14	0.34	0.8	1
GP Visits (Last 12-months)	17368	1.76	2.47	0	1.5	11
Outpatient Visits (Last 12-months)	17389	0.96	1.98	0	0	11
Inpatient Days (Last 12-months)	17442	0.42	4.38	0	0	200
Welfare Benefits (Monthly)	17515	190.68	438.47	0	0	4479.83
Gross Labour Income (Monthly)	17515	1886.58	1925.51	-11000	1600	16666
Net Labour Income (Monthly)	17515	1403.24	1294.3	-11000	1325.21	10800
Annualised Healthcare Costs	17515	425.3	917.94	0	112.1	5514.49
Annualised Welfare Benefits	17515	2245.03	5017.63	0	0	26170.68
Annualised Labour Tax Revenue	17515	5759.95	8038.8	0	2623.08	38553.36

Secondary analyses

Table A8. Results of a just-identified second-order panel vector autoregression (two-step GMM).

	ΔHRQoL deficit	Δ(ln)GHQ	ΔTax revenue	ΔWelfare benefits	ΔHealthcare costs
n	40,266	40,266	40,266	40,266	40,266
ΔHRQoL deficit					
t - 1	**0.094**	0.0213	1,186	**-3,878**	81.1
	0.0475, 0.140	-0.151, 0.193	-826.1, 3,199	-5,798, -1,957	-1,454, 1,616
t - 2	**0.044**	-0.0146	179.3	**-1,889**	25.67
	0.0174, 0.0705	-0.113, 0.0836	-927.2, 1,286	-2,940, -837.5	-761.7, 813.0
Δ(ln)GHQ					
t - 1	**0.062**	**0.142**	**-2,026**	**2,459**	439.8
	0.0325, 0.0913	0.0368, 0.248	-3,246, -806.3	1,534, 3,384	-273.0, 1,153
t - 2	**0.027**	0.0432	**-933.3**	**1,238**	188.5
	0.0122, 0.0427	-0.0126, 0.0990	-1,570, -297.0	757.6, 1,719	-187.0, 564.0
ΔTax revenue					
t - 1	8.14E-07	-1.01E-07	**0.331**	**0.055**	0.0075
	-5.81e-08, 1.69e-06	-3.50e-06, 3.30e-06	0.237, 0.425	0.0318, 0.0779	-0.0093, 0.0244
t - 2	2.83E-07	-4.70E-07	**0.128**	**0.022**	0.0021
	-1.55e-07, 7.21e-07	-2.20e-06, 1.26e-06	0.0797, 0.176	0.0112, 0.0337	-0.0056, 0.0098
ΔWelfare benefits					
t - 1	-6.41E-07	-8.32E-07	**-0.088**	**0.251**	-0.0131
	-2.00e-06, 7.18e-07	-5.98e-06, 4.31e-06	-0.130, -0.0457	0.161, 0.340	-0.0568, 0.0305
t - 2	-2.84E-07	-9.03E-07	**-0.039**	**0.103**	-0.0123
	-9.41e-07, 3.73e-07	-3.27e-06, 1.47e-06	-0.0596, -0.0174	0.0591, 0.146	-0.0321, 0.00752
ΔHealthcare costs					
t - 1	1.32E-07	7.95E-07	0.0066	**-0.067**	0.0015
	-7.63e-07, 1.03e-06	-2.59e-06, 4.18e-06	-0.019, 0.032	-0.0946, -0.0391	-0.118, 0.121
t - 2	1.21E-07	1.70E-06	-0.0026	**-0.046**	-0.039
	-3.24e-07, 5.66e-07	-1.16e-07, 3.52e-06	-0.012, 0.0075	-0.0718, -0.0201	-0.108, 0.0307
Age	**0.005**	**0.017**	-27.19	**171.8**	17.76
	0.00410, 0.00620	0.0137, 0.0209	-56.65, 2.272	125.7, 217.9	-15.18, 50.70

For a 1 percentage point change in (log) GHQ exposure, the change in any non-logged outcome is equal to $\hat{\beta}\ln(1.01)$ or approximated by $\hat{\beta}/100$ for small coefficients. Independent variables not in natural log form can be transformed to the percentage change on outcome (log) GHQ using $100\left(e^{\hat{\beta}}-1\right)$ or approximated by $100\hat{\beta}$. Change in (log) GHQ exposure on (log) GHQ dependent variable is equal to $100\left(1.01^{\hat{\beta}}-1\right)$, or be approximated as $\hat{\beta}$ = percentage change.