UNIVERSITY OF LEEDS

Modelling the Social Determinants of Health at the Individual and Neighbourhood Level

Robert Clay

Submitted in accordance with the requirements for the degree of Doctor of Philosophy in Geography

The University of Leeds

Faculty of Earth and Environment

School of Geography

30/10/24

Intellectual Property and Included Publications

Copyright

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Author CRediT Statement

The CRediT (Contributor Roles Taxonomy) outlines individuals who have authored and co-authored all work contained in this thesis. These authors have the following names and initials; Robert Clay (RC), Nik Lomax (NL), Alison Heppenstall (AH), Patrica Ternes D'Allagnollo (PTDA), Luke Archer (LA), Hugh Rice (HR), Petra Meier (PM), and Robin Purshouse (RP).

Chapter 2 is based on a paper submitted to the International Journal of Microsimulation Pending Peer review.

> Towards the Accessibility of Transition Probability Development for Dynamic Microsimulation: A review of the current literature

RC carried out literature searches and authored the manuscript. NL, AH, and PTDA edited the final draft and supervised the work.

Chapter 3 is based on a paper submitted to the International Journal of Microsimulation Pending Peer review.

> An Example Case Study on the Application of Transition Dynamics for Discrete Time Dynamic Microsimulation

RC carried out literature searches, authored the manuscript, and wrote all required code. NL, AH, and PTDA edited the final draft and supervised the work.

Chapter 4 is based on a paper submitted to the Journal of Artificial Societies and Social Simulation (JASSS) pending peer review.

> Estimating the Effects of Income Support Policies on the Mental Well-Being of the UK Population Using a Microsimulation Model

RC carried out literature searches and authored the manuscript. LA and HR wrote all required code. NL, AH, PM, RP, and PTDA edited the final draft and supervised the work.

Chapter 5 is based on a paper submitted to the Nature Energy Justice Special Collection pending peer review.

> Evaluating the Impact of Household Energy Policy on Adult Mental and Physical Health in the United Kingdom

RC carried out literature searches and authored the manuscript. LA and HR wrote all required code. NL, AH, and PTDA edited the final draft and supervised the work.

Chapter 6 is based on a paper that has not yet been submitted. The intention is to submit this paper to a journal focused on spatial health inequalities and energy justice research.

> Analysis of MINOS Dynamic Microsimulation Output Data Exploring Energy Poverty Efficacy Heterogeneity Over Space and Vulnerable Sub-Populations

RC carried out literature searches and authored the manuscript. LA and HR wrote all required code. NL edited the final draft and AH supervised the work.

Acknowledgements

First, thank you to my supervisors Professor Nikolas Lomax, Professor Alison Heppenstall, and Dr. Patricia Ternes D'Allagnollo. Your support in my professional development, opportunities to train and present at a wide range of conferences, written and verbal feedback, and four years of your time have been indispensable to authoring this thesis.

Thank you to the SIPHER research group who funded this research. They have provided many opportunities to meet and present with a broad multidisciplinary range of individuals. Iterative feedback on this thesis and connection with policy partners have been invaluable. Particular thanks go to SIPHER Workstrand 5's Luke Archer, and our many arguments, who has contributed substantially to the development of software used in this thesis, keeping me sane, and beyond.

Additional gratitude is given to Rachel Furniss and Nicola Bhuller within the School of Geography Postgraduate administration team who have provided consistent support to myself and many others particularly at the start of this thesis during pandemic lockdown(s).

To my family, thank you all for your individual contributions. To my parents Jacqueline and Mark, thank you for keeping me fed, watered, taxiing me around, emotional support to persevere, and keeping a roof over my head particularly during these interesting times. Thank you to my sister Caz and husband Nick for your constant support and facilitating writing retreats; even though you locked me in a room for several days. Thank you to my brother Tim for simply asking how I am every week as well as his wife Fiona and their children, Monty and Ellis, whose love of slides cannot be overstated. As you both start school, I hope this thesis can inspire your own academic pursuits as my Father and Uncles did.

Finally, appreciation is given to all examiners involved in the marking of this thesis throughout both interim reports and final examination whose time and critical review are essential to producing the best thesis possible.

Abstract

The United Kingdom is seeing a substantial rise in incidences of non-communicable disease including asthma and anxiety disorders. Implementing policy to counteract these conditions is crucial, but many policies implemented during the coronavirus and energy crises in the UK have been unable to help those most in need of assistance due to limited evidence.

One approach to evidence the effect of policy on public health is utilising the Complex Systems Modelling methodology. A candidate policy is parameterised as a causal loop diagram that is then emulated using individual-level modelling providing tools for policy makers to generate evidence identifying subgroups and sub-geographies of the population that may have been overlooked allowing for personalised, equitable, and efficient policy.

This thesis has developed the Microsimulation for Interrogation of Social Science Systems (MINOS) dynamic microsimulation model to estimate causal pathways between household income and health. Several policies are implemented to address child and energy poverty in the UK measuring health outcomes using tangible SF-12 scores, Quality Adjusted Life Years, and Incremental Cost-Effectiveness Ratios. Evidence to increase the Scottish Child Payment policy has been submitted to the Scottish Government suggesting increasing the payment will be cost-effective for mental health. Application of the existing Energy Price Cap Guarantee and novel Great British Insulation scheme policies to address energy poverty showed that while both policies are expensive if insulation was retroactively fitted in 2020 it would be cost-effective by as early as 2024.

These findings have strong implications for future government strategy to reduce household energy bills, reach net zero targets and improve mental well-being. Ultimately, the open source and flexible MINOS model framework has shown success estimating public health response to real government policies and could be expanded to further more complex scenarios reducing incidence of non-communicable disease and government costs.

Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Background

Life expectancy in the United Kingdom is stagnating for the first time in over 180 years (Raleigh, 2021; Broadbent et al., 2023) largely due to increasing incidence of non-communicable diseases (NCDs) including asthma (National Health Service (NHS) Confederation, 2022) and anxiety (Appleyard et al., 2023). A series of socioeconomic crises including austerity, the coronavirus pandemic, and rising fossil fuel prices (triggered by war in Ukraine) have reduced UK household spending power by 7.1% in 2022 – 2024 (Office for Budget Responsibility, 2024). This is the largest biennial decrease since records began in 1956, resulting in a rise in households being driven into poverty and unable to afford essential amenities including heating and good quality food (Broadbent et al., 2023). Houses that cannot afford heating quickly deteriorate, resulting in mold and damp causing respiratory issues including asthma and COPD (National Health Service (NHS) Confederation, 2022). Individuals who cannot afford household bills are seeing sharp increases in incidence of anxiety disorders from financial stress as they struggle to pay back short term loans (Appleyard et al., 2023) and face social stigma (Purdam et al., 2016) from using welfare resources such as food banks.

The underlying causes of NCDs are highly complex with a broad range of determinants including socioeconomic background, environmental factors, individual and regional social behaviour, and external geopolitics (Greszczuk, 2019; Meier et al., 2019). Individual households have highly heterogeneous needs and their response to socioeconomic crises can vary substantially (Peersman and Wauters, 2024). For example, as fossil fuel prices increase, some households have seen disproportionate increases (such as those in rural areas and ethnic minorities) resulting in substantially increased energy bills (UK Fuel Poverty Monitor, 2022). Another group that have been hit financially hard are workers in specific sectors such as tourism and entertainment who were made redundant nearly overnight during the Covid Pandemic and are still feeling the impacts of this (Kopasker et al., 2024). Incidences of NCDs can vary substantially across subsections of a population including low income households and rural areas. Addressing NCDs requires consideration of all possible influencing factors on health and application of flexible, equitable healthcare according to individual needs (Cairney, 2021; Meier et al., 2019).

Policy Analysis is defined as 'client-orientated advice relevant to public decision and informed by social values' (Weimer and Vining, 2017). When new issues emerge within a population, policy analysis identifies potential actions a government body may take in order to enact positive change (Weimer and Vining, 2017). Candidate policies are optimised with respect to equity, financial cost, and political feasibility in order to advise best course of action for policy makers (Weimer and Vining, 2017). Ideally, policy analysis is 'evidence-based' (Cairney and Oliver, 2017; Cairney et al., 2021) such that a diverse consensus of empirical evidence, expert opinion, and more recently statistical and computational modelling (Greszczuk, 2019; Meier et al., 2019) are utilised to inform decision making by exhaustively exploring all possible determinants of change. The intention is to increase synergy within policy design by considering all available policy scenarios as well as estimating potential policy consequences within vulnerable individuals and outcomes across seemingly unrelated government sectors (Meier et al., 2019; Greszczuk, 2019). In the health economics literature this is commonly referred to as a 'Health in All Policies' approach (Broadbent et al., 2023; Cairney et al., 2021; Greszczuk, 2019) such that all government policy across all sectors must now be considerate of the determinants of NCDs to preserve public health.

Practically, evidence-based policy has proven difficult to implement at different levels of government in the United Kingdom (Cairney and Oliver, 2017; Cairney, 2021; Scott and Gong, 2021). Government bodies work in silos solving their own internal issues while ignoring other sectors and external advice producing disjointed, black-box, and confusing policy that can vary substantially by departments and spatial areas (Ham, 2021; Cairney, 2019). Many developing health crises are completely unprecedented and have no empirical evidence available to advise policy requiring over-reliance on simulated evidence with potential strong underlying mathematical assumptions (Pawson, 2021; Saltelli and Giampietro, 2017; Cairney, 2021). Evidence is also not the only influence on policy making as Governments may simply outright ignore evidence if it does not align with party political values, long term plans, or public opinion (Cairney, 2016). These issues are perfectly encapsulated by the UK policy response to the coronavirus pandemic. Lockdown policy was, necessarily, primarily focused on disease transmission rates and did not consider health ramifications due to financial stress from mass unemployment particularly within the entertainment and tourism industries (Collard et al., 2021; Hancock and Tyler, 2024). Policy varied substantially by Local Authority area with some regions locked down and others not. This resulted in major confusion particularly for individuals living near 'unfair' (Cairney et al., 2021; Cushion et al., 2020) regional borders who often outright ignored Government guidelines and suffered different employment and other economic outcomes (Collard et al., 2021). Initial lockdown measures were based almost entirely on forecasts from the now infamous 'Ferguson Imperial Model' (Ham, 2021; Magness, 2021) based on 10 year old proprietary code that could not be reviewed. This model has since received criticism due to a poor theoretical background and model assumptions that grossly overestimated excess mortality and transmission rates (Ham, 2021). The UK Government implemented deeply unpopular 'draconian' lockdown measures (Ioannidis et al., 2022), based almost entirely on these exaggerated figures. This had the effect of drastically reducing public trust in forecasting models(Ioannidis et al., 2022).

One suggestion to facilitate trust in simulated evidence for policy analysis is ensuring that modelling efforts adhere to four 'ATOMIC' principles (Wasserstein et al., 2019). Accepting model uncertainty requires considering all potential sources of error in model inputs, clear statement of underlying model assumptions and potential caveats, and reliance on probability distributions rather than point estimates (Ioannidis et al., 2022). Being modest and clear with where a model could deviate from the truth tempers expectation for what policies can be assessed and allows for governments to carefully consider the feasibility of model forecasts as part of a larger suite of predictive tools (Cairney, 2021). Thoughtful research is considerate of previous literature and potential model users. Ideally model results are quantifiable as simple interpretable (Tomova et al., 2022) stories explaining the systems of change behind a policy (Cairney, 2021) that can be understood by non-technical audiences. Open research facilitates this requiring transparency such that all theoretical background, model assumptions, and software are clearly available in long term repositories to facilitate critical review and reproducibility (O'Donoghue and Dekkers, 2018; Wasserstein et al., 2019).

Complex systems methodology (Meier et al., 2019; Haraldsson, 2004) is one approach to generate ATOMIC evidence for policy analysis. Consultation with literature, community panels, government bodies, and academic experts explores available candidate policy to address NCDs in an iterative process (Meier et al., 2019). For each candidate policy causal systems maps (Barbrook-Johnson and Penn, 2022; Haraldsson, 2004) are constructed using the same iterative process in order to identify which characteristics of a population and individual households would change under this policy. These causal systems maps must then be operationalised using individual-level models (Höhn et al., 2023) that estimate each individual entity from a population, whether a person or household, separately. These models estimate the future state of the population under a candidate policy providing evidence for how the population changes over desirable health outcomes (Andreassen et al., 2020). A population can be analysed both at national aggregates and exploring smaller sub-populations and sub-geographies of interest allowing for optimisation of precise policy that can cost-effectively fight NCDs while preserving overall public health (Spielauer et al., 2020; Burgard et al., 2021). Policy evidence can be simulated quickly providing governments with evidence to response to urgent crises that can explore a wide range of potentially extreme policy on a national scale that is not possible with traditional randomised controlled trials and natural experiments (Smith et al., 2021).

Dynamic microsimulation is the most commonly chosen individual-level modelling technique used to simulate policy evidence (Spielauer et al., 2020; Burgard et al., 2021). A population of individual units such as households is constructed from empirical administrative data and projected forwards in time under intervention using transition dynamics models. Dynamic microsimulation is considered to be more robust than similar techniques such as agent-based modelling due to use of empirical data rather than a 'bottom up' theoretical approach (Arnold et al., 2019) used to better estimate social behaviour. Output data is typically produced in 'cross sections' demonstrating the state of the entire population simultaneously at a given point in time which is particularly useful to policy makers (Andreassen et al., 2020) interested in annual budgets.

Dynamic microsimulation has previously been used for application over economic policy such as in the Department of Work and Pensions (Li and O'Donoghue, 2013) to estimate pension contribution and aggregate economic measures such as GDP (Li and O'Donoghue, 2013). Increasingly synthetic evidence is also being used in health scenarios particularly in hospitals in order to test treatment efficacy, long term health outcomes, and resource allocation optimisation (Rutter et al., 2011). Application to health economics is a recently new application due to new available data and the health in all policies approach that is now trying to provide simulated evidence for the effect of policy to prevent NCDs (Cairney et al., 2021; Katikireddi et al., 2022; Broadbent et al., 2023; Kopasker et al., 2024). These policies typically intervene either on household income or household quality itself. Policy affecting household income are very diverse either directly providing households further income through schemes such as universal basic income (Eichhorst et al., 2010) or indirectly saving households money such as by reducing their energy bills (UK Fuel Poverty Monitor, 2022). Similarly housing support policy either directly improves individual living conditions through measures such as insulation retrofitting (Regan, n.d.) or indirectly by improving local amenities and employment opportunities (Pemberton and Humphris, 2016; Eichhorst et al., 2010).

1.2 Research Gaps

Dynamic microsimulation models are being developed by a number of research groups to generate synthetic evidence for the change in public health effect for income and housing support policy (Kopasker et al., 2024; Katikireddi et al., 2022; Meier et al., 2019; Archer et al., 2021; van de Ven et al., 2022). The main objective of this research is to formalise the implementation of causal systems modelling in policy analysis as a full life-cycle from initial public issues to development of causal loop diagrams and microsimulation to finally influence government decisions and improve public health. This thesis contributes towards the development of such a model for the Systems Science in Public Health and Health Economics Research (SIPHER) consortium (Meier et al., 2019) aiming to address three gaps in research.

1.2.1 Literature on the Best Practice of Transition Dynamics Development for Dynamic Microsimulation

Recent reviews of the state of the art for dynamic microsimulation have highlighted limited application of microsimulation to new areas of research as development methodology is tacitly hidden in out of print or proprietary research (O'Donoghue and Dekkers, 2018). Increasing accessibility of microsimulation to new model users requires standardisation of methodologies using tutorial literature providing guidelines on development best practice outlining common mistakes to avoid and improving model prediction quality (O'Donoghue and Dekkers, 2018). Tutorial papers are being released covering a areas of microsimulation development methodology including development of spatial input populations (Lomax and Norman, 2016; Lovelace, 2014), population weighting (Dekkers and Cumpston, 2012), missing data imputation (Frick and Grabka*, 2005; Kum, 2010), application in specific programming languages (Krijkamp et al., 2018), alignment methods (Li and O'Donoghue, 2013), and outcome visualisation methods (Salonen et al., 2019). There is currently limited literature available on the methodologies for the best practice of the development of transition probability models for dynamic microsimulation.

Providing research on common methods to improve transition probability models including better presentation (Lomax and Norman, 2016; Charlton and Sana, 2023; Salonen et al., 2019), uncertainty quantification using sensitivity analysis (Rutter et al., 2011; Petrik et al., 2020; Sharif et al., 2012; Burgard et al., 2019), and methods incorporating individual heterogeneity (Richiardi et al., 2014; Burgard et al., 2019) allowing for better longitudinal prediction are all common but not presently collated in one place. Demonstrating best practice using example transition models facilitates the implementation of ATOMIC practice in both general microsimulation and policy analysis research. This knowledge contributes to the remainder of the thesis and construction of full dynamic microsimulation models.

1.2.2 Application of Dynamic Microsimulation to Income and Housing Support Policy Within the United Kingdom.

Income support policies are one proposed family of policies designed to alleviate the effects of NCDs by protecting household disposable income (Kopasker et al., 2024). These policies can either directly intervene upon household income by providing additional funds or indirectly save a household money by providing them with resources including free bus passes, benefit schemes, or higher quality employment opportunities (Harrison et al., 2022; Meier et al., 2019). Housing support policy is a further subcategory of indirect income support policy (Regan, n.d.) providing household improvements including insulation and electrical heating system subsidies. Households equipped with these measures spend less on household energy bills (Regan, n.d.) and have 'freed' up income to spend elsewhere as required.

Both income (Jensen and Blundell, 2024) and housing (UK Fuel Poverty Monitor, 2022; Regan, n.d.) support policies have faced criticism as they are often expensive and have limited evidence of their cost-effectiveness. Subsidy schemes often require additional private capital to be completely funded, the general public is reluctant to provide this without clear evidence (Regan, n.d.). Interestingly recent surveys suggest many members of the general public are more interested in long term benefits of support policies rather than immediate changes in household bills (Regan, n.d.). Combined with a health in all policies approach this suggests a need for evidencing the health effects of these income and housing support policies to support their adoption.

While dynamic microsimulation has large historical applications in both income and health policy (Li and O'Donoghue, 2013; Rutter et al., 2011) scenarios, there is limited literature available estimating health outcomes for these policies (Kopasker et al., 2024). Several research groups (Meier et al., 2019; Katikireddi et al., 2022; Kopasker et al., 2024; van de Ven et al., 2022) are proposing construction of dynamic microsimulation models emphasising the potential health benefits and application to future policy. This is particularly important as income and housing support measures become more commonplace. This thesis contributes to research from the Systems Science in Public Health and Health Economics Research (SIPHER) consortium (Meier et al., 2019) aiming to construct such a dynamic microsimulation estimating the effect on mental and physical health of income and housing support policies in Sheffield, Greater Manchester and Scotland. Externally this model can be used beyond this thesis to explore new policies and to corroborate with other simulation of health policy using other modelling methods such as static microsimulation.

1.2.3 Evidence comparing the Cost-Effectiveness and Equity of Potential Income and Housing Support Policies.

UK governments are struggling to identify households who are most in need of income (Child Poverty Action Group, 2024) and housing support (Regan, n.d.). Ideally support can be targeted equitably across households who need assistance in order to ensure a 'minimum' public health (Hjelmskog, 2022). Most policies are instead applied uniformly to broad subgroups of individuals according to national-level statistics. One example is the Energy Price Cap Guarantee (EPCG) (UK Fuel Poverty Monitor, 2022) which provided households with income to cover their energy bills according to overall energy consumption. Overall energy consumption is a poor measure of overall energy poverty and many households including ethnic minorities and those in rural areas were not given enough support to fund required energy consumption (UK Fuel Poverty Monitor, 2022). Not meeting individual household demands causes large inequalities in health outcomes and rates of NCDs (Purdam et al., 2016). This can also been seen for income support policies such as the Scottish Child Payment as the Scottish Government continues to explore which subgroups of the population such as single mothers receive the most benefit from this intervention or could see further uptake (Statham et al., 2022). There is a desire to fine tune this policy ensuring uptake and potentially providing more income to these vulnerable households as required.

Another issue with these support policies is cost effectiveness. The EPCG policy discussed above is by far the most expensive income support policy ever implemented in the UK but is considered to be largely ineffective (UK Fuel Poverty Monitor, 2022). Evidence is sorely needed comparing the cost-effectiveness of competing policies to improve NCDs in the UK estimating overall intervention cost and applying health economics measures allowing for direct comparison (Chen et al., 2023).

Dynamic microsimulation for these policies may provide evidence required to explore both cost-effectiveness and application of dynamic equitable policy. Individual households and extreme distribution values can be highlighted and examined to determine what further measures are needed. Household intervention costs can be easily scaled up to the national-level to provide cost-effectiveness evidence and overall provide more optimised NCD policy in the future.

1.3 Thesis Research Questions

In order to address these gaps in literature the primary aim of this thesis is the development of a dynamic microsimulation model emulating causal pathways between income and health to robustly evidence and optimise the effect of income and housing support policy in the United Kingdom at individual and neighbourhood levels.

To achieve this aim, this thesis will address three key objectives formulated as research questions:

Can I contribute to literature codifying state of the art best practice for developing transition probability models in dynamic microsimulation?

Addressing this question requires an in depth literature review on best practice of dynamic microsimulation transition probability models exploring common pitfalls, methods to address these issues, and potential future innovations. Collecting these methods together contributes to best practice literature increasing accessibility to new users by having information in one place increasing the impact of dynamic microsimulation while also facilitating creation of reproducible, robust simulated policy evidence. This review is supplemented with a case study example developing transition probability models for use in income support policy microsimulation models. Tangible application of transition probability best practice further improves accessibility and supplements personal background knowledge required for the development of full dynamic microsimulation models throughout the remainder of this thesis.

How can dynamic microsimulation be used to operationalise causal loop diagrams as part of a complex systems modelling framework to generate synthetic evidence for policy effect on change in health outcomes?

As part of SIPHER consortium, the main objective of this thesis is the development of a dynamic microsimulation model. This model will be utilised as part of a complex systems methodology framework operationalising causal systems diagrams for income and housing support policy. This model will be fully open source using publicly available datasets as well as transparent, online documentation for all model inputs. This model will be intended for use on high performance computing systems allowing for large scale policy optimisation considering many possible policies and model runs. Initial simple household income support policies will be implemented as a test case that can be expanded upon in further SIPHER work.

How can evidence from dynamic microsimulation be used to compare the cost-effectiveness of competing policies within different priority subgroups and spatial regions?

Using this dynamic microsimulation framework more elaborate policy scenarios can the be explored. Given SIPHER interests, available data, and political relevance the remainder of this project intends to focus on child poverty policy, with energy poverty policy explored later. This requires expansion of the initial microsimulation to include further energy poverty variables and transition dynamics. Inclusion of further health measures including physical health and

Quality Adjusted Life Years (QALYs) as well as intervention cost will allow for more detailed comparison of policy. Comparison of these policies will then provide a more tangible use case for application both within the SIPHER group and potentially for wider UK Government bodies.

1.4 Methods and Data Summary

This section summarises the methods and data sources used throughout this thesis.

The Understanding Society (UKHLS) (Fumagalli et al., 2017) dataset is used throughout this thesis requiring all 21 existing waves of data to construct transition dynamics and estimate a synthetic population of the UK for use in dynamic microsimulation. From chapter 4 onwards a number of auxiliary datasets are also used summarised in Table 1.1 including midyear NOMIS (Rees et al., 2017) fertility and mortality estimates, NEWETHPOP project age, sex, and ethnicity demographic forecasts, conversion from council tax bands to monetary amounts paid by LSOA area, and CPI inflation historical data and forecasts. Chapter 5 onwards then utilises the SIPHER synthetic spatial population (Höhn et al., 2024) attaching spatial information to UKHLS data as well as several Office for National Statistics geoportal datasets assigning spatial information by LSOA area including map boundaries (Office for National Statistics, 2024), rural urban classifiers (Gale et al., 2016), index of multiple deprivation (Payne and Abel, 2012), and SIPHER inclusive economies indicators (Lupton et al., 2023).

Chapters 2 and 3 review and apply methods for ensuring interpretability and goodness of fit in transition dynamics. causal systems maps and application based on available data. Initially ordinary linear regression modelling (Harrell et al., 2001) to estimate SF-12 MCS state. Multiple Imputation Chained Equations (MICE) algorithm (Sterne et al., 2009; Krijkamp et al., 2018) is used to impute missing observations and variables not recorded in every wave such as loneliness. Several methods are used to quantify uncertainty in model inputs (Sharif et al., 2012). These include variable selection techniques such as shrinkage methods (Tibshirani, 1996) to test sensitivity of model structure. Ensemble cross validation is used to estimate uncertainty in sample data and generalised linear mixed modelling is used to incorporate individual long term well-being states using random intercepts scores (Breslow and Clayton, 1993).

Chapter 4 applies these transition dynamics methods to construction of a full dynamic microsimulation for estimation of health economics outcomes for income support policy(Katikireddi et al., 2022; Kopasker et al., 2024; van de Ven et al., 2022). The Scottish Child Payment (Statham et al., 2022), Living Wage foundation 'real living wage' (Cominetti and Murphy, 2022), and the energy price cap guarantee (UK Fuel Poverty Monitor, 2022) policy scenarios are all implemented demonstrating expected mental well-being improvement as households receive more disposable income. A modular discrete-time dynamic microsimulation framework 'vivarium' (Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA, 2023) is used to developed the model (Spielauer et al., 2020; Burgard et al., 2021). Replenishing population methodology (Archer et al., 2021) and cell based reweighting (Tysinger, 2021) are used for population demographics alignment over time. Treatment on the treated visualisation methods

(Archer et al., 2021) are used to identify how policy effects the health of different population subgroups.

Chapters 5 and 6 further apply this dynamic microsimulation model to compare the health effects of two housing support policies. The microsimulation employs a larger set of transition dynamics models including energy poverty pathways and utilises the SIPHER synthetic input population (Höhn et al., 2024) generated using the simulated annealing algorithm. This provides a full scale spatially representative population for the Greater Manchester area. Further health outcomes are also applied using physical health score (Lawrence and Fleishman, 2004), quality adjusted life years (QALYs) (Lawrence and Fleishman, 2004), and finally increment cost effectiveness ratios (ICERs) (Gafni and Birch, 2006) to further examine policy cost-effectiveness of two mutually exclusive policies. The EPCG is used again as well as the Great British Insulation Scheme (GBIS/ECO4) (Regan, n.d.) emulating how household characteristics would change under these policies.

1.5 Thesis Structure

This thesis is comprised of 8 individual chapters. Chapters 2 to 5 are analogous to 4 submitted research articles under review or in preparation (as of September 2024) in the International Journal of Microsimulation, the Journal of Artificial Societies and Social Simulation, and a Nature Energy Justice Special Collection. Each chapter is supplemented with brief connecting material summarising findings and establishing a single narrative between subsequent chapters as part of the wider thesis. Chapter 1 contains introductory material as well as an overview of the data and methods used for the remainder of the thesis. Chapter 7 provides discussion reviewing research objectives and key findings concluding with project limitations and potential future work and applications.

Chapters 2 and 3 address the first thesis objective by providing a literature review of the best practice for application of transition probability dynamics for dynamic microsimulation. Several areas where transition models can be improved are discussed in detail highlighting why they can result in poor prediction of behaviour as well as common methods used and potential future innovation. This literature review is complemented with a case study example estimating transition dynamics for mental well-being state using Short Form 12 Mental Component Score using UKHLS data. Applying discussed methods to improve transition model fit shows a significant improvement in prediction quality as well as more interpretable model that consider missing data and uncertainty.

Chapter 4 addresses the second thesis objective by constructing a dynamic microsimulation model MINOS estimating change in mental well-being score under several diverse income support policies. This chapter explains how public administrative data and SIPHER causal systems maps have been developed into a full dynamic microsimulation providing extensive technical appendices describe how the model operates as well as complete transition model formulae and coefficients. The Scottish Child Payment, Living Wage foundation 'real living wage', and the

Data source	**Chapter(s)**	**Data cleaning/ Description**
Understanding Society Data	2-6	Administrative data containing yearly waves of individual household surveys
NEWETHPOP Projections	2-6	Projections of UK fertility and mortality stratified by age, sex, and ethnicity up to 2061
Council Tax Bands	4-6	Data converting council tax band letters $A, B, C, \ldots$ into continuous intervals e.g [£1200,£1400] by Local Authority.
Quarterly Inflation CPI	4-6	Historical data and forecasts for quarterly Consumer Price Index (CPI) inflation from the Office for National Statistics (ONS) and Budget Responsibility (OBR) respectively.
GB GeoJSON Map Tiles	5-6	Geojson map tiles for Great Britain from the ONS geoportal.
SIPHER Synthetic Population for Individuals in Great Britain, 2019-2021	5-6	A synthetic full scale, spatially representative population of Great Britain from the SIPHER consortium (Höhn et al., 2024).
Quarterly Wholesale Fossil Fuel Prices	5-6	Quarterly wholesale fossil fuel price histories for gas, electricity, liquid, and solid fuels from the ONS.

Table 1.1: Data sources used within the thesis and corresponding chapters.

energy price cap guarantee policy scenarios are all implemented demonstrating expected mental well-being improvement as households receive more disposable income.

Chapters 5 and 6 address the third thesis objective by comparing two housing support policies to reduce household energy bills from 2020 onwards. The initial MINOS model is expanded to include further energy poverty pathways, a synthetic input population, and new cost-effectiveness measures of public health. The Energy Price Cap Guarantee and Great British Insulation Scheme policies are directly compared exploring their overall cost-effectiveness as well as by subgroups including income quintiles and across spatial regions. Results suggests insulation schemes have very large initial capital required but are more cost effective than capping energy prices. This has strong implications for future UK government strategy adjusting housing quality to reach net zero targets and improve public health.

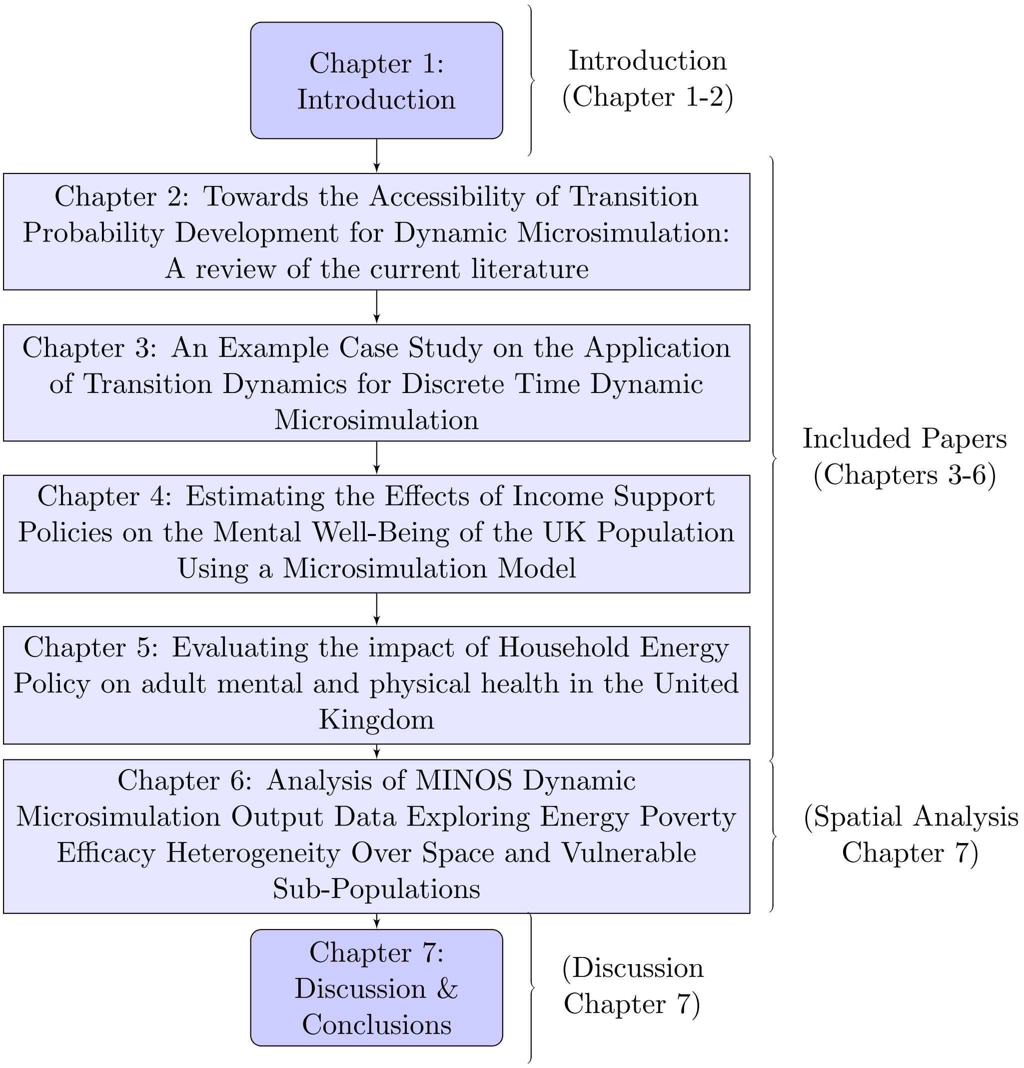


Figure 1.1: A diagram outlining the structure of this thesis submission.

Chapter 2: Towards the Accessibility of Transition Probability Development for Dynamic Microsimulation: A Review of the Current Literature

2.1 Abstract

Transition dynamics are a key component of dynamic microsimulation projecting future states of individual units forwards in time using predictive statistical modelling. Well implemented transition probabilities are critical for accurate estimation of individual behaviour and producing accurate, trustworthy microsimulation results that can be used to inform real decision making. In an effort to increase the accessibility of dynamic microsimulation to new users and fields of research review of the state of the art has proposed codification of tacit microsimulation methodological knowledge that is largely hidden in proprietary or out of print texts. Several literature reviews covering the complete microsimulation development cycle and specific methods including alignment and synthetic population generation have been published outlining typical methods used, common issues and pitfalls, speculation on future innovation, and overall guidance on best practice. This article addresses a gap in literature providing a methodological review for transition probability model development in dynamic microsimulation. Common features of transition probability models are described as well as five key areas in which prediction robustness and accuracy can be improved. A final brief overview is provided of these key issues as well as a series of questions that should be addressed in development to prevent misspecification and ensure goodness of fit when projecting out of sample forward in time and under counterfactual scenarios.

2.2 Introduction

Dynamic microsimulation (Spielauer et al., 2020; Burgard et al., 2020; Dekkers, 2015) is a modelling technique often used to estimate the effect of hypothetical 'what-if' scenarios on a population. Initially microsimulation was developed to estimate the effect of socioeconomic policy on income and demographic states in the USA where macro-level models were insufficient to capture individual long-term histories (Orcutt, 1961; Li and O'Donoghue, 2013; Zaidi and Rake, 2001). With advances in data availability and computing power, dynamic microsimulation is increasingly being applied to new areas of research including the COVID-19 pandemic

(O'Donoghue et al., 2021) and recent cost of living crises (Broadbent et al., 2023).

Dynamic microsimulation models feature two key components: a synthetic input population; and transition probability mechanics. An input dataset of individual entities such as households, people, or businesses is generated from survey data to be statistically representative of some population of interest (Lovelace and Dumont, 2017; Burgard et al., 2021; Lomax and Smith, 2017b). Transition probabilities govern how individuals evolve between states as they are projected forward in time within a dynamic microsimulation (Burgard et al., 2021). Each individual randomly draws their next state from a transition probability distribution conditional on current attributes (Burgard et al., 2021). Examples of these transitions are from alive to dead, calculating new household income, or a change in mental health state. Construction of these transition probabilities is highly flexible such that any methodology that produces a probability distribution of next state can be applied, ranging from empirical counts to complex multi-modular systems (Li and O'Donoghue, 2013; Spielauer et al., 2020; Skarda et al., 2021). Transitions can capture complex non-linear heterogeneous dynamics such as long term spatio-temporal histories and interactions between individual units (Richiardi et al., 2014; Burgard et al., 2021).

Well calibrated transition probabilities are essential to the functionality of dynamic microsimulation. Bias, projected from misspecified behaviour, result in model output that is difficult for users to trust and limits their application (Zaidi and Rake, 2001; Harding, 2007; Burgard et al., 2021). One major criticism levelled against microsimulation as a whole is that best practice on development of transition probabilities, and other methodologies, is tacit knowledge hidden in out of print and proprietary research (O'Donoghue and Dekkers, 2018; Burgard et al., 2020), limiting accessibility of microsimulation to new researchers and application to new fields of research. There have been a number of literature reviews for dynamic microsimulation methodology covering a broad range of topics including the overall state of the art (Spielauer et al., 2007; Li and O'Donoghue, 2013; O'Donoghue and Dekkers, 2018; Zaidi and Rake, 2001), application to specific areas such as health and income policy (Li and O'Donoghue, 2013; Rutter et al., 2011), microsimulation design requirements (Spielauer et al., 2020; Burgard et al., 2020) and tutorial papers that implement common methodological techniques such as alignment (Li et al., 2014; Alarid-Escudero et al., 2020) and input population generation (Li and O'Donoghue, 2016; Burgard et al., 2021; Müller and Axhausen, 2010; Lomax and Smith, 2017a). However, to the authors knowledge there is no such review for transition probability methods highlighting common issues and pitfalls, model selection criteria, and overall guidelines on best practice.

This paper aims to fill this gap by providing a literature review of key issues and best practice in the development of transition probabilities for dynamic microsimulation. These are divided into five key areas: (1) presentation (where extensive publication/communication of transition probability methods can be difficult given potential complexity and space constraints); (2) missingness (where longitudinal data offers specific challenges); (3) validation (where microsimulation output is compared against other data to determine sensible and reliable projection); (4) uncertainty quantification; and (5) heterogeneity (whereby microsimulation models need to meet their full potential in estimating behaviour across a population of individuals). For each area we explore motivation for why these problems must be addressed, popular and novel methodologies, critique, and discussion of potential future innovation. This review, focused on transition probabilities, will align with other existing microsimulation methodological reviews to provide accessible information for new researchers to the field and collate tacit knowledge. To complement this review, A second separate paper has also been provided demonstrating example application of methods to address the above issues and how they improve fit for a transition model estimating mental well-being state using linear regression methods.

The structure of this paper is as follows: Section 2.2 defines transition probabilities and classifies common methods; Sections 2.4 - 2.8 defines the above five key issues reviewing common methods and speculating on potential future innovation; and finally Section 2.9 summarises key points to address while developing transition probabilities and outlines future work.

2.3 Background

2.3.1 Dynamic Microsimulation

Dynamic microsimulation (Orcutt, 1961; Li et al., 2014; Burgard et al., 2021) is an individual-level modelling technique developed to estimate change in longitudinal household income due to hypothetical pension and general economic policy. A population of individual entities, such as households or individuals, is created to be statistically identical to some real population and evolved forwards in time using transition dynamics under varying hypothetical counterfactual scenarios. Any change in this artificial individual population should then approximate real change that would occur in a population and can be used to advise and optimise decision making (Li et al., 2014; Rutter et al., 2011).

Estimating household longitudinal behaviour was simply not possible with existing macro-economic modelling and static population microsimulation models that estimated change in population proportion of fixed demographic blocks rather than true individual entities (Orcutt, 1961; Li et al., 2014). In the early 1990s, further administrative data and computational power facilitated the development of initial dynamic microsimulation models across multiple countries such as the United Kingdom estimating pension policy outcomes using the DYNASIM and PENSIM2 models as well as more general economic and demographic models such as SAGE (O'Donoghue and Dekkers, 2018; Li et al., 2014).

After the initial success of pension policy dynamic microsimulation has facilitated expansion to several new areas of research (Rutter et al., 2011). Estimation of health and health economics policy is frequently utilised to estimate health outcomes of existing non-health policy such as pension schemes as well as epidemiological research estimating treatment cost-effectiveness and resource allocation (Rutter et al., 2009; Spielauer et al., 2007). Inclusion of non-human populations such as traffic management (Odiari and Birkin, 2022) and businesses (Ballas et al., 2013) facilitates novel application and comparison with similar individual models such as agent-based modelling (Birkin, 2021). Development of further statistical methods has given rise to new areas of research including spatial dynamic microsimulation including initial methodological papers developments and how-to guidance (Lomax and Smith, 2017b; O'Donoghue et al., 2013; Burgard et al., 2021; Zhou et al., 2022; Tanton et al., 2014) as well as case-study application (Ballas, Clarke, Dorling, Eyre, Thomas and Rossiter, 2005; Odiari and Birkin, 2022).

Modern dynamic microsimulation modelling has seen limited new areas of application beyond health and economics outcomes (O'Donoghue and Dekkers, 2018) but methods and technologies continue to improve. Open source repositories allow for full presentation of microsimulation model structure and documentation (Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023; Kopasker et al., 2023). Similarly, open source programming frameworks are continuing to standardise code structure facilitating independent reviews and faster development (De Menten et al., 2014; Li et al., 2014). Inclusion of synthetic population data allows for detailed estimation of individual units at small areas and timescales and, merging multiple datasets to include all required variables to fully describe the population and effect of interventions (Lomax and Smith, 2017b; Burgard and Schmaus, 2019; Davis and Lay-Yee, 2019). Increasingly complex predictive methods are used to incorporate further information about individual structure including longer spatio-temporal histories and interactions between proximal units in households or neighbourhoods (Richiardi et al., 2014; Burgard et al., 2021). As data availability, prediction methods, and computing power increases, dynamic microsimulation continues to see further application in new countries with more limited data resources (O'Donoghue and Dekkers, 2018) as well as application informing real government decision making in new areas such as mental well-being in countries with existing microsimulation modelling projects (Skarda et al., 2021). To facilitate this development, reviews into dynamic microsimulation have highlighted the need increase modelling accessibility by codifying common methods such as alignment (O'Donoghue and Dekkers, 2018). The remainder of this article defines transition probability methods used in dynamic microsimulation including common problems in development and overall recommended- best practice guidelines.

2.3.2 Transition Probabilities

In dynamic microsimulation an individual entity is represented as some state vector X_t at some point in time t. This state vector contains all information required to estimate future states and evaluate model output (Dekkers, 2015). Transition probability mechanics are defined as some function $x_s = F(x_t)$ that estimates the state vector for some future time $s > t$ (Dekkers, 2015). The function F consists of a series of transition probability functions $F = (f_1, \ldots, f_n)$ that each independently estimate some subset of the future state. In theory, any predictive methodology can be used for these functions resulting in a very large pool of available methods (Spielauer et al., 2020; Dekkers and Van Den Bosch, 2016). These methods can then be further divided dependent on several key features.

Time Increments

In discrete time microsimulation, prediction of new states occurs at uniform time increments such as years (Li and O'Donoghue, 2013). The change in individual state may have occurred at any point over this time period. This allows for estimation of synchronised 'cross-sections' of a population often desirable for estimations of aggregate-level change in areas such as policy analysis (Andreassen et al., 2020; Li and O'Donoghue, 2013). On the other hand continuous time microsimulation estimates both the future state and the sojourn time until this state occurs (Dekkers and Van Den Bosch, 2016; Li and O'Donoghue, 2013). This allows for estimation of the exact moment a person changes state and is useful in estimating precise longitudinal life paths common in areas such as health data for disease progression and mental well-being change due to short term issues such as unemployment (Dekkers and Van Den Bosch, 2016; Li and O'Donoghue, 2013; Archer et al., 2021). It is often desirable to estimate this short term behaviour, but it can then be difficult to produce accurate cross-sections of the whole population for a given time point(Dekkers and Van Den Bosch, 2016; Zucchelli et al., 2010; Smith et al., 2021). In general, continuous time transition models are desirable but are often not utilised due to strong data requirements and difficult implementation (Li and O'Donoghue, 2013).

Pseudo-continuous time microsimulation combines discrete and continuous time processes together (Spielauer et al., 2007). Longitudinal projections of individual behaviour are updated sequentially using data assimilation methods incorporating information from discrete time yearly cohort modelling. These transition models are used particularly in health microsimulation models (Spielauer et al., 2007; Zaidi and Rake, 2001) to estimate disease progression. For example, individual progression through a disease such as cancer may be generated for their entire lifetime estimating time points for transition between disease stages and ultimately death. These disease progression times are then updated yearly cross-section data such as tobacco and alcohol consumption. If an individual drinks a lot of alcohol their disease progression times would then move forwards and vice versa. Ideally, pseudo-continuous transition provide the best of both worlds being able to incorporate short time scale continuous processes only as needed. They have fewer data requirements that full continuous time models but still have large development time and require synchronisation to single points in time if cross-sectional examination is required (Li and O'Donoghue, 2013).

This review will focus on transition probabilities applied within discrete-time frameworks. They are the most commonly implemented form of transition probabilities due to simplicity and data availability. However, much of the discussion in this paper is also applicable to pseudo-continuous and continuous time transition probability models.

Available Calibration Data

The choice of transition model is largely dependent upon the calibration data that is available. Original transition probabilities used time-series for a single discrete variable with some k states (Orcutt, 1961) estimated using multi-state Markov modelling. The probability of transition between each pair of states is estimated from empirical data and used to inform the behaviour of all individuals (Williams et al., 2017; Li and O'Donoghue, 2013). Improved transition matrices stratify transition probabilities by demographic attributes such as age, sex, or ethnicity (Stillwell and Duke-Williams, 2003) introducing heterogeneity in individual behaviour, albeit with exponentially increased computing time for each variable added and sparsity issues (Leknes and Løkken, 2021; Stillwell and Duke-Williams, 2003; Lomax and Smith, 2017b). For modern survey data these rate tables can be difficult to implement for a large number of predictor variables. In this case it is common to utilise linear regression techniques, particularly ordinary least square and logistic regression, to estimate future state conditional on other demographic information (Burgard et al., 2019; Marois and Aktas, 2021; Zhou et al., 2021). As modern datasets include long term individual histories, spatial information, and interactions between individual units such as between households and neighbourhoods further advanced modelling techniques are implemented to capture detailed individual behaviour (DeYoreo et al., 2022; Richiardi et al., 2014). Alternatively, there may be no empirical data available at all and behaviour must be simulated using a 'bottom up' approach with behaviour governed by social theory (Arnold et al., 2019; Klevmarken, 2022). This approach is more commonly utilised in agent-based modelling but increasingly available social behaviour data and synthetic populations are increasing utilisation of agent-based modelling and microsimulation hybrid approaches combining empirical and theoretical transition modelling together (Birkin, 2021; Arnold et al., 2019; Klevmarken, 2022).

This review will focus on application of modern empirical survey data estimating transition probabilities using linear regression methods. These methods are most suitable for available yearly discrete time Understanding Society survey data and see increasing application in microsimulation literature. Again, methods discussed here extend to simpler transition matrices as well as more complex systems.

Dependent Variable Type

The choice of transition model is also dependent on response variable type. In the case of a large body of extant microsimulation literature, binary variables dictate the use of logistic regression and continuous variables result in the use of ordinary linear regression. However, while more complex dependent variables such as censored or zero inflated data are increasingly present in survey data, there is limited formal application and model assumption testing within microsimulation frameworks (McLay et al., 2015). Examples of more elaborate transition probabilities for available data and complex response variables are discussed further in Section 2.8. This study aims to incorporate transition probability models estimating continuous and discrete outcome variables. Best practice methods are largely identical but can use different algorithms and visualisation.

Programming Language

Several programming languages are utilised to implemented transition probability models. Numerous microsimulation reviews (O'Donoghue and Dekkers, 2018; Spielauer et al., 2020) have highlighted the need to use common programming languages to implement microsimulation to ensure readibility and standardisation of code. R., python, java, and C++ are the most commonly used languages dependent on user requirements including readibility, computational speed, accessibility to new programmers, and intergration into larger microsimulation frameworks such as MODGEN and LIAM2 (O'Donoghue and Dekkers, 2018). This paper will implement all transition probabilities in the R. programming language due to open source well documented statistical packages that can be implemented in python, C++, and Java microsimulation frameworks.

These criteria describe the primary classifications of transition probability models by time, available data, outcome variable type, and programming language. The remainder of this paper will discuss methods to improve development of linear regression based discrete time transition probability models in the R. programming language. While specific methods are used overall discussion on best practice intends to be transferable to other classification of transition probability models and review of specific methods for categories such as continuous time models are omitted for future work.

2.4 Presentation

A full account of transition probabilities used in a given microsimulation model can be difficult to publish, given they can contain numerous transition probability models, each with background, formulae, large tables of coefficients, and diagnostics (Rutter et al., 2011). Technical documentation is often impractical to publish due to article page constraints (Rutter et al., 2011), while complex models such as neural networks with very large numbers of coefficients are difficult to interpret, or contain sensitive, proprietary data that cannot be released to the public (O'Donoghue and Dekkers, 2018; Andreassen et al., 2020). As a result, formal specification of transition probabilities is often omitted, or they are sourced from other literature producing 'black-box' microsimulations that can be impossible to fully critique or reproduce(Rutter et al., 2011; Zhou et al., 2022). Things are changing - as public online repositories and journals supplementary material become more widely available, microsimulation literature is increasingly providing fuller documentation for transition probability models (Zaidi and Rake, 2001). However, even well documented transition models can still be difficult to interpret. Increasingly a middle ground of 'grey-box' interpretable models (Tomova et al., 2022; Spielauer et al., 2020; Zaidi and Rake, 2001) is being adopted, which provide all required data to reproduce a microsimulation but with standardised presentation of model design and visualisation to facilitate critical review and guide users through results.

2.4.1 Planning

The first stage of transition probability development is planning. An analysis plan (Harding, 2007) is written beforehand which describes expected model implementation. Transition model planning is generally integrated into overall microsimulation design. Researchers consider previous literature evidence, expert opinion, and other sources of information to identify pathways and variables that inform and predict future behaviour, often producing systems maps for the whole microsimulation (Skarda et al., 2021; Tikka et al., 2021; Katikireddi et al., 2022; Zhou et al., 2021). A transition probability is then designed to represent some subsection of this map. Review of previous literature ensures consideration of user requirements and exploration of available methods and data (Burgard et al., 2020). Comparison of similar existing models allows for alignment with other (microsimulation) scenario results allowing for external validation in meta-analyses (Wasserstein et al., 2019). Predetermined analysis prevent overfitting and exploratory analysis (Harding, 2007; Wasserstein et al., 2019) providing reproducible models that are transferable to other contexts with more limited data resources (Allendes et al., 2021; Harding, 2007).

Assumptions of statistical methodologies used should be clearly stated (McLay et al., 2015) in order to account for limitations and prevent model misspecification in results, ensuring modesty in findings that tempers user expectations. Any transition methodology will have underlying mathematical assumptions that are rarely considered (McLay et al., 2015). Diagnostic testing is increasingly used to justify the use of transition methods to reviewers. These diagnostics are highly bespoke depending on methods used with a number of how-to papers arising describing how to implement testing of assumptions (McLay et al., 2015; Lomax and Norman, 2016; Burgard et al., 2021). Model structure is also increasingly considered in transition probability design. Microsimulations are generally modular (Burgard and Schmaus, 2019; Zhou et al., 2021; Burgard et al., 2020; Harding, 2007) such that related sets of variables transition in independent groups. The order of these transition modules as well as their inputs and outputs must be considered (Li and O'Donoghue, 2013). Incorrect order can induce bias due to endogeneity, 'cyclical' transitions that are not Markov and predict themselves using current time information, and tacit assumptions of the direction of causal effect (Scott et al., 2003; Zucchelli et al., 2010). Moreover, modules typically only transition some small subset of individual attributes sequentially, ignoring intersectionality between variables and transition processes (Harding et al., 2010a; Zaidi and Rake, 2001). One common example in microsimulation is cohabitation, whereby when an individual marries they often move into a household with their partner (Zaidi and Rake, 2001). Interaction between two marriage and migration modules is essential to correctly simulate a marriage event.

Another planning consideration is the programming language for transition models. A number of Microsimulations are being developed as well documented, open source frameworks (O'Donoghue and Dekkers, 2018; Tikka et al., 2021), using widely used programming languages (O'Donoghue and Dekkers, 2018), often alongside open source visualisation tool-kits (Salonen et al., 2020; Charlton and Sana, 2023). Many microsimulations are written in common modular syntax allowing for easier understanding and reuse of transition model code and results from one project to another (O'Donoghue and Dekkers, 2018; Li and O'Donoghue, 2013). Common frameworks including LIAM2, MODGEN, M++, and EUROMOD (O'Donoghue and Dekkers, 2018; Spielauer et al., 2020; Broadbent et al., 2023; De Menten et al., 2014) are all now fully open source allowing for maintenance beyond project funding and have seen desired extended application (Spielauer et al., 2020) to further projects beyond their initial use. Discussion on differences between these frameworks including computational speed, readibility, user requirements, and accessibility to new users (De Menten et al., 2014; Li and O'Donoghue, 2013).

2.4.2 Model Structure

Description of transition probability model structure is increasingly recommended to facilitate interpretable microsimulation (Burgard et al., 2020; Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023). For a single transition probability model use of diagrams and model formulae are usually sufficient but this is not unique to microsimulation (Krijkamp et al., 2018). For multi-state Markov models, diagrams indicating possible state transitions combined with transition matrices are enough to interpret how behaviour changes over time (Krijkamp et al., 2018). Similarly regression methods have simple model structure requiring publication of just model formulae (McLay et al., 2015; Rutter et al., 2009). More complex transition models including neural networks (Essien et al., 2019) and MCMC methods (DeYoreo et al., 2022) require more elaborate model architecture graphs.

Unique presentation of model structure occurs when presenting multiple transition probability models as part of a larger dynamic microsimulation (Burgard et al., 2020). Typically a series of transition probability models are run in a cycle for each discrete time step whose structure can greatly influence microsimulation behaviour (Li and O'Donoghue, 2013). The biggest consideration is the order in which transition probability models are run in this cycle (Li and O'Donoghue, 2013; Burgard et al., 2020; Harding et al., 2010a,b). If a microsimulation contains a mortality module using this module first, such that individuals die immediately, can produce greatly different results if this module is used last and continue to operate within the population. This is particularly influential in discrete time microsimulation (Li and O'Donoghue, 2013). If an individual dies at the start vs the end of a discrete time interval this can produce different behaviour economically through pension or employment contribution.

The order of transitions also influences which predictor variables and at which time points are included in prediction of behaviour (Li and O'Donoghue, 2013). Typically microsimulation models try and implement Markov behaviour such that future time state is only estimated using current time information (Burgard et al., 2021). In some cases, prediction of future state may be performed using future time information. For example, the prediction of future mental well-being may be predicted using future estimates of household income from a transition model earlier in the microsimulation cycle. Mental well-being is then a latent variable measured as a function of household income that is controversial (Li and O'Donoghue, 2013) as it ignores how individuals trajectories between discrete time points but allows for immediate response of population health to change in income.

This also highlights the importance of interaction between transition probability modules in microsimulation (Li and O'Donoghue, 2013; Harding et al., 2010b) multiple transition models may be required to estimate a single variable. A common example is household disposable income which first requires calculation of both household net income and overheads including rent (Frick et al., 2009). This application of successive models can produce complex behaviour and logical inconsistencies not seen in individual components (Li and O'Donoghue, 2013; Harding et al., 2010b). Additionally some microsimulation models are using ensembles of transition probability models estimate the same subspace of the future state space (Parham and Hughes, 2013). Explanation is needed for how these ensembles interact demonstrating under which conditions each model best estimates future state.

Strong presentation of this transition probability model cycle is then essential for an interpretable microsimulation (Burgard et al., 2020). Most published microsimulation literature includes a visual representation of this cycle by grouping transition probability models into modules that each estimate some related subset of the state space (Li and O'Donoghue, 2013). The order in which these modules are run is then published in simple directed graphs. There are many examples of this of varying complexity ranging from simple chains of modules (Zhou et al., 2022; Burgard and Schmaus, 2019) up to very complex systems (Skarda et al., 2021; Murray et al., 2017; Broadbent et al., 2023; Katikireddi et al., 2022; Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023) . These diagrams are often supplemented with protocols (Zhou et al., 2022; Burgard and Schmaus, 2019) specifying clear instructions on the order transition probability models are run. Overall, this improved presentation of structure is popular with both model developers and users to present complex information in a non-technical way and increase microsimulation impact. One notable example is application in policy analysis allowing for easier operationalisation of complex theory of change causal diagrams into microsimulation that can be understood by policy partners in simple, interpretable stories (Meier et al., 2019).

2.4.3 Visualisation

Full transparent publication of transition probability models including technical documentation, coefficients, and model structure is ideal (Rutter et al., 2011) but can be difficult for model users to interpret (Spielauer et al., 2020). These 'white box' models are modest and open source but can be difficult to understand when identifying drivers of change due to counterfactual scenarios (Spielauer et al., 2020; Tomova et al., 2022). A 'grey-box' model compromise then provides transparency as necessary, i.e. partially omitted coefficients (Essien et al., 2019), to facilitate interpretable stories in literature and analysis with full technical documentation available elsewhere. Interpretability can be achieved by either simplification of transition models albeit with poorer prediction (Spielauer et al., 2020) and with improved visualisation exploring specifically why a population changes (Pintelas et al., 2020; Boumans, 2009).

Visualisation of microsimulation model output is highly bespoke (Zinn et al., 2014) according to user needs. Common visualisation techniques are used (Zinn et al., 2014) to demonstrate change in population distributions over time including histograms, bar plots, and boxplots (Zinn et al., 2014; Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023; Archer et al., 2021).Further visualisation largely depends on user requirements including mapping for spatial dynamic microsimulation (Lovelace and Dumont, 2017; Zhou et al., 2022) using real maps for application in agriculture and traffic as well as change in longitudinal trajectories in policy analysis (Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023; Salonen et al., 2020; Burgard et al., 2021). Visualisation for validation of microsimulation is also improving with methods such as handover plots, ROC curves (Archer et al., 2021) and bespoke trajectory analysis toolkits (Salonen et al., 2020; Spielauer et al., 2020; Harding et al., 2010a) becoming increasingly common. There is still limited standardisation of the visualisation of transition probability models relative to other components of microsimulation. Output visualisation for transition probabilities is still highly bespoke and rarely published. Approaches to speed up visualisation such as standard diagnostic plots and guidance on how to interpret these, as well as easy online documentation may substantially increase microsimulation impact(Li and O'Donoghue, 2013; O'Donoghue and Dekkers, 2018).

2.5 Missingness

Missingness is common in datasets used to calibrate transition probability models in dynamic microsimulation (Gilbert and Troitzsch, 2005; Mirzaei et al., 2022; Burgard et al., 2020). Observations may not be recorded for a multitude of reasons including non-response due to absence or non-applicability, changes in data collection such as questions no longer being asked in a longitudinal survey, or clerical errors (Frick et al., 2009; Mirzaei et al., 2022). Moreover, critical variables required for calibration may simply not be recorded in a dataset (Klevmarken, 2022). Missing data correction is defined as the adjustment of a sample population to prevent bias in model coefficients by accounting for missing values (Mirzaei et al., 2022). Often complete-case analysis is used which simply omits all observations with missing values from sample data (Gilbert and Troitzsch, 2005; Frick et al., 2009). However, complete case analysis assumes that data are missing unconditionally (missing completely at random) (Mirzaei et al., 2022) which is a strong mathematical assumption that is difficult to test and rarely holds in practice (Kum, 2010; Mirzaei et al., 2022). Missingness is often dependent on other variables, such as individuals dropping out a survey due to severe illness (Mirzaei et al., 2022), and ignoring missing values distorts the sample population, inducing selection bias in transition models (Kum, 2010). For observations that are not recorded at all, complete case analysis cannot be used, as there would be no data left, and transition models must simply omit terms that can again result in bias due to omitted variables (Skarda et al., 2021). This section reviews missing data correction techniques employed in microsimulation beyond complete case analysis in order to utilise all information and reduce bias.

Choice of missing data correction methods depends primarily on the type of variable and missingness structure (Mirzaei et al., 2022; Frick et al., 2009). Variable type, (e.g. continuous or discrete data), should be determined beforehand outlining which methods are available (Mirzaei et al., 2022). Missingness structure is more complex in longitudinal survey data (Mirzaei et al., 2022). The simplest structure contains uniformly missing variables such that an observation is either completely observed in data or completely missing (Li et al., 2013). This is rarely true in practice with many non-uniform partially observed missing data structures. Observations can be intermittently recorded over time or missing some members of a household making missing values heterogeneous and strongly conditional on available partially complete information (Richiardi and Poggi, 2012; Mirzaei et al., 2022; Frick et al., 2009). Variables may also not be recorded at all within a dataset and must be simulated, either by merging multiple datasets or creating entirely synthetic populations with all required variables together (Klevmarken, 2022; Richiardi et al., 2014; Li and O'Donoghue, 2013).

2.5.1 Imputation and Data Matching

Imputation techniques are commonly used in microsimulation to estimate missing values, using predictive models calibrated on complete information (Kum, 2010). Single (or deterministic) imputation techniques are the simplest to implement, estimating missing values dependent on some explicit formula (Kum, 2010; Haslett et al., 2010). This can be due to derived logical structure, such as an unemployed individual having no income or their net income being a function of gross income and taxes, that is easy to correct in data preprocessing (Frick et al., 2009; Kopec et al., 2010). If missing values cannot be directly derived, single imputation must then predict missing values (Kum, 2010). This can be as simple as mean imputation that assigns all missing values the mean value of complete observations. More elaborate prediction such as knn imputation or even linear regression estimates missing values as a function of other associative variables (Waddell, 2012). Data matching is much more common in microsimulation due to its ability to simulate entire missing individuals as well as missing variables. It is often used in open population microsimulation models to generate new individuals such as immigrants or spouses that enter a model on demand (Cumpston et al., 2011; Spielauer et al., 2020). Data matching (or hot-deck imputation) instead borrows information verbatim from the 'closest' available individual according to some distance metric (Kum, 2010; Cumpston et al., 2011; Haslett et al., 2010). There are many choices of distance metrics including Gower and Mahalanobis functions (Urban and Shrestha, 2023; Birkin, Morris, Birkin and Lovelace, 2017) as well as propensity score matching (Kum, 2010; Cumpston et al., 2011) calculating the probability that two observations match. Single imputation approaches are computationally inexpensive but are criticised due to lack of heterogeneity in prediction (Kum, 2010; Richiardi and Poggi, 2012; Haslett et al., 2010) and have fallen out of favour (Haslett et al., 2010). Multiple (or stochastic) imputation techniques randomly draw imputed values according to some probability distribution. This is an extension of single imputation techniques that typically provides an ensemble of imputed values pooled together to quantify uncertainty in final transition model parameter estimates (Marois et al., 2023; Schreuder et al., 2021). There are many stochastic analogues of deterministic imputation, particularly for data matching techniques, choosing values from a closest available pool of applicants rather than the single most likely value (Cumpston et al., 2011; Haslett et al., 2010). Multiple Imputation by Chained Equations imputes missing values using iterative linear regression (Jakobsen et al., 2017) and is commonly used in microsimulation due to availability in programming languages and ease of use (Marois and Aktas, 2021; Marois et al., 2023; Schreuder et al., 2021).

Many multiple imputation algorithms have strong mathematical assumptions that must be relaxed in order to estimate non-uniform missing data (Huque et al., 2018). For partially observed longitudinal data there are extensions of the MICE algorithm using linear mixed models and generalised estimating equations (Huque et al., 2018; Marois and Aktas, 2021). Likewise, MICE can only handle Normal variables and further extensions are needed for categorical and non-Normal data (Huque et al., 2018). Bayesian methods including the Expectation Maximisation (EM) algorithm are used to impute right censored data and provide consistent trajectory dynamics over time (Alsefri et al., 2020; Salonen et al., 2020; Burgard et al., 2019). Markov Chain Monte Carlo (MCMC) methods such as approximate Approximate Bayesian Computation (DeYoreo et al., 2022; Shewmaker et al., 2022) use multi-level hierarchical structures, estimating imputed missing values themselves or model parameters such as regression coefficients (DeYoreo et al., 2022; Rutter et al., 2009; Alsefri et al., 2020) conditional on each individual's available complete information.

Missing data can also be influenced by interactions between other proximal individual units (Tanton et al., 2014; Frick et al., 2009). If one individual is missing from a household then total household disposable income is not known and must estimated given information from remaining complete individuals (Frick et al., 2009; Richiardi and Poggi, 2012). Individuals within close spatial proximity can see spillover effects in response to neighbourhood crime (Ballas, Clarke, Dorling, Eyre, Thomas and Rossiter, 2005), labour market demand (Richiardi et al., 2014) or disease progression (Burgard et al., 2021; O'Donoghue et al., 2021). Imputing each member of the household individually often does not work due to illogical values being generated such as a household with no income (Frick et al., 2009; Richiardi and Poggi, 2012). Multi-level imputation methods are incorporated to capture spatial networks between multiple levels of individual units including households (Richiardi et al., 2014), commercial entities such as business and hospitals as well as sparsely observed small-scale spatial areas (Burgard et al., 2021; Hartnett et al., 2020).

2.5.2 Re-weighting and Synthetic Populations

Complete case analysis should only be applied when observations are missing completely at random (MCAR) independent of any other variables in a dataset (Mirzaei et al., 2022). However, this is rare in practice, impossible to test, and almost never recommended (Jakobsen et al., 2017). Many survey datasets now contain sample weights such that simply removing missing values is unlikely to preserve representativeness of the sample population (Mirzaei et al., 2022; Li et al., 2013). To preserve sample data, complete-case methods remove partially missing values but create a new synthetic population based on re-weighted complete values (Frick et al., 2009; Li et al., 2013; Haslett et al., 2010). Approaches such as inverse probability weights methods, iterative proportional fitting, and simulated annealing are all utilised to create a population of complete values that matches regional-level aggregates (Haslett et al., 2010; Lomax and Norman, 2016; Lomax and Smith, 2017b). While this creates a complete population of observations there is high uncertainty associated with calibrating transitions using simulated data as well as continuing to ignore any distributional effect of missing values (Li and O'Donoghue, 2013). Synthetic populations are also used to simulate non-uniform missing data, including dynamically consistent longitudinal trajectories (Salonen et al., 2020; McLay et al., 2015) as well as estimating small scale spatial areas for which data are not recorded (Burgard et al., 2021; Zhou et al., 2022)

2.5.3 Presentation of Missing Data Correction

Presentation of missing data correction methods is largely underutilised (Jakobsen et al., 2017). Authors (Jakobsen et al., 2017; Davis and Lay-Yee, 2019) have suggested a number of strategies including presentation of missingness structure, exploration of distributional differences before and after missing data correction, clear planning of correction strategy, comparison of transition probability models fitted with and without imputed data to determine goodness of fit and any differences in coefficients. Inclusion of analysis plans for imputation of missing in data in microsimulation assures consideration of previous literature, model assumptions, and interpretability discussed in Section 2.4 (Burgard et al., 2020). Novel visualisation tools for interpretation of more complex missingness structures are also being applied to microsimulation including longitudinal trajectory analyses and spatial mapping (Salonen et al., 2020; Lovelace and Dumont, 2017). As with general presentation, microsimulation would benefit from standardisation of common missing data correction algorithms in order to test model assumptions, validity, and improve development time. Especially as synthetic populations, and their black box generation methods, become increasingly common discussing how the imputed data differs from reality and any effect on transition calibration is crucial.

2.6 Validation

Validation of dynamic microsimulation is broadly defined as ensuring projected model output at individual and population level is 'reasonable and credible' to facilitate trustworthy and reproducible evidence that can be used in real decision making (Harding et al., 2010a; Dekkers, 2015; Alarid-Escudero et al., 2020). A combination of statistical testing and visualisation is used to determine if microsimulation output is not statistically distinct from sample data over aggregate-level distributions and individual trajectories (Harding et al., 2010a; Archer et al., 2021; Burgard et al., 2020; McLay et al., 2015). Validation is arguably the most important and time-consuming component of microsimulation taking up to 90% of overall development (O'Donoghue and Dekkers, 2018) time due to debugging of computationally expensive model runs and complex microsimulation structure. A large component of validation is in fact validation of transition probability models ensuring correct projection of behaviour over time (Spielauer et al., 2020; Burgard et al., 2020; Zaidi and Rake, 2001; Harding et al., 2010a; Harding, 2007). This section reviews common methods to validate transition probability models and discussion on existing literature gaps.

2.6.1 Internal Validation

Internal validation compares microsimulation projections against datasets used for calibration of the dynamic microsimulation model (Li and O'Donoghue, 2013). Initially this requires testing of underlying background model assumptions for each transition probability model (McLay et al., 2015; Goedemé et al., 2013; Harding et al., 2010a). Ensuring model assumption are met prevents biased prediction due to model misspecification and clarifies to model users prediction scope and where estimates may fail under counterfactual scenarios (McLay et al., 2015). Assumption testing applies to all statistical modelling with some unique considerations in microsimulation including a balance between strong mathematical assumptions and development time (McLay et al., 2015). Generally model assumptions are bespoke to each transition probability model. A common example application uses Ordinary Least Squares (OLS) linear regression (McLay et al., 2015). Consideration of background assumptions including independent, identically Normal distributed errors are commonly violated due to heterogeneity in survey data (McLay et al., 2015) and must be tested and often relaxed using longitudinal modelling such as mixed effects models. Additionally the OLS model will always underestimate population variance over time (Harrell Jr, 2015) and this must be accounted for to accurate estimate population distributions over time. Some reviews for common microsimulation methods are available including Multi-State Markov models outlining guidelines for testing model assumptions such as the Markov assumption (Krijkamp et al., 2018), where but further guidance has been repeatedly recommended (Harding et al., 2010a; Goedemé et al., 2013).

Remaining internal validation assess transition model performance by comparing observed and predicted data. As there is no available real data into the future and under counterfactual scenarios internal validation often requires 'nowcasting' (O'Donoghue and Loughrey, 2014) whereby a microsimulation starts in the past and prediction runs until the present day where data are available. Comparison with real data is performed using both visualisation and statistical testing.

Visualisation methods provide interpretable "eye tests" (Archer et al., 2021) making it clear to model users that microsimulation prediction preserves linear trends in population statistical moments and individual behaviour (McLay et al., 2015). Common visualisation methods (Zinn et al., 2014) including bar plots and histograms (Archer et al., 2021) are used to demonstrate differences in population distributions. Comparison between different counterfactual scenarios is also often presented again using common visualisation including kernel density estimate curves and receiver operating characteristics (ROC) curves (Archer et al., 2021). More recently the behaviour of individual trajectories are also being considered using trajectory analysis (Salonen et al., 2020) and spaghetti plots (Crooks et al., 2018) to assess if individual progression between states over time matches real data. Much of this validation is also performed longitudinally using specialised versions of common visualisation such as ridgeline and handover plots (Lang et al., 2020; Archer et al., 2021) to address issues including failure to estimate mean and variance over time particularly using regression models that underestimated observation variance over time (Harrell Jr, 2015).

Formal statistical testing is also used to determine if observed and predicted populations are statistically distinct. This can range from simple t-tests comparing population distributions and other statistics such as PRESS testing for longitudinal dynamism (McLay et al., 2015). Out of sample performance frameworks divide calibration data into training and test sets using methods such as cross-validation and common goodness of fit measures including root mean squared error (McLay et al., 2015; Archer et al., 2021; Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023).

Predicted output from transition probability models, particularly more complex models, do not accurately estimate population mean and variance over time (Li and O'Donoghue, 2012; Richiardi et al., 2014). In this case alignment methods (Alarid-Escudero et al., 2020; Li and O'Donoghue, 2012, 2013) are used to adjust the predicted population to match exogenous aggregates such as mean, variance, and percentage population belonging to each discrete state (Richiardi et al., 2014; Li and O'Donoghue, 2012; Dekkers and Cumpston, 2012). There are many available alignment methods dependent on variable type and whether alignment is done during or after a microsimulation model run (Li and O'Donoghue, 2012; Dekkers and Cumpston, 2012; O'Donoghue and Dekkers, 2018). Methods for discrete binary and ordinal variables are well described elsewhere (Li and O'Donoghue, 2012; Burgard et al., 2021) adjusting transition probabilities and moving individuals between states in order to match aggregate data. Alignment methods for continuous variables have limited formal literature review (O'Donoghue and Dekkers, 2018) but are prevalent in most microsimulation models. Methods such as multiplicative scaling (Dekkers and Cumpston, 2012) are applied again adjusting statistical moments and perturbing individual observations to match aggregates. Other alignment methods are used elsewhere including adjustment of microsimulation population weights (Dekkers and Cumpston,

2012) that are rarely used to recalibrate transition models as well.

2.6.2 External Validation

External validation instead compares microsimulation output against datasets not used in model calibration (Rutter et al., 2011). There is substantial overlap with methods used in internal validation using visualisation and statistical testing. Visualisation of distributional differences used in nowcasting is extended to application forwards in time and under counterfactual scenarios (Archer et al., 2021). Direct comparison with real data cannot be made but these plots serve as useful 'eye tests' ensuring continuation of population trends and individual behaviour.

Alignment methods adjust microsimulation output to match simulated aggregates where real data are not available (Rutter et al., 2011; Li and O'Donoghue, 2013, 2012). Prediction where real data is not available is instead aligned against either more robust macro-level projections (Li and O'Donoghue, 2013), other individual-level simulations (Kelly and Percival, 2009), or using hybrid macro-micro simulation models that inform each other over time producing dynamic aggregates (Cockburn et al., 2014). Macro-micro models align each other either using simple scaling based on mean and variance or more complex methods including Approximate Bayesian Computation (Shewmaker et al., 2022; Asher et al., 2023) and potentially data assimilation (Birkin, 2021).

Ideally, external validation allows for analysis of predicted behaviour from multiple simulation models using meta analyses but this is rare requiring multiple models all simulating the same counterfactual scenarios (Rutter et al., 2011; O'Donoghue and Dekkers, 2018; McLay et al., 2015). Comparison can be made using transition probability models used to estimate the same processes from different microsimulation. Combination of several different transition probability models can also be used in meta-modelling choosing the best fitting model out of an available pool dependent on goodness of fit criteria (Zhong et al., 2022).

Validation is, due to its importance, one of the best documented aspects of microsimulation and transition probability development (O'Donoghue and Dekkers, 2018). A variety of methods are available to demonstrate projections match real data and are reliable enough to use when forecasting forwards in time. As microsimulation models become more complex further validation techniques are being developed to visualise and align individual-level data over more detailed aggregates including small, sparse spatial areas and subgroups. Validation is also being developed to ensure correct behaviour of individual-level trajectories ensuring suitable dynamism and consideration of long term histories. This is particularly useful as dynamic microsimulation continues to grow allowing comparison between output from two microsimulation models in external validation. There is also a push for further tutorial literature in dynamic microsimulation validation with gaps in research including continuous variables (O'Donoghue and Dekkers, 2018) and validation presentation (Salonen et al., 2020) that would further codify standardised methods and address criticism on reproducibility.

2.7 Uncertainty/Sensitivity Analysis

Another major component of microsimulation validation is sensitivity analysis (Burgard and Schmaus, 2019). Testing how much change in model input parameters changes overall microsimulation output is used to determine projection robustness. For example, if projections are sensitive to interest rates or energy pricing (i.e. highly volatile processes), then inferences drawn from a microsimulation can be hard to trust due to high variance (Burgard and Schmaus, 2019). Conversely, if model output does not change as input parameters vary it suggests these events are likely to occur (Sharif et al., 2012). Sensitivity analysis requires quantification of model uncertainty, exploring which input parameters should be varied and by how much to reflect uncertainty in real life observations and future projections. As microsimulations are computationally expensive, experimental design allows for efficient exploration of the input parameter space (Petrik et al., 2020; Sharif et al., 2012) in order to sufficiently test model robustness. While sensitivity analysis is largely performed for microsimulation outputs as a whole it should also be performed for all transitions models with respect to the following sources of uncertainty (Harding et al., 2010a; Burgard and Schmaus, 2019).

2.7.1 Uncertainty in Calibration Data

As discussed in Section 2.5 survey data contains a sample that may not be representative of the full population. This is particularly true when extrapolating data to new contexts (Allendes et al., 2021; Essa and Sayed, 2015). Model parameters and datasets are commonly sourced from other simulation literature to increase development speed (Kopec et al., 2010). For example, demographic rate tables for fertility and mortality are available from Government Statistical Agencies and can be used to estimate the dynamics of these processes within a model. Microsimulation for large geographic areas such as the EU or USA may borrow model parameters from 'similar' socioeconomic countries or states (Sutherland, 2018). Population demographics and behaviours can change substantially over a microsimulation's time horizon (Lomax and Smith, 2017b; Allendes et al., 2021; Alsefri et al., 2020). The transferability of a dataset from one context to another must be considered when validating a microsimulation (Allendes et al., 2021; Kopec et al., 2010; Sharif et al., 2012). Data sourced from other literature must be justified as suitable for use (Kopec et al., 2010). Considering whether data comes from a reputable journal, any modelling assumptions made to generate data and parameters, and any further missing data that must be estimated such as model parameters for a new subgroup of the population (Kopec et al., 2010) or missing data (Klevmarken, 2022). Ideally data used to calibrate borrowed transition probabilities is available, allowing for direct comparison between datasets to determine distributional differences and any potential biases (Kopec et al., 2010; Sutherland, 2018). Survey datasets typically contain large outliers (Lappo et al., 2015; Wolf, 2001) that can cause large changes in model parameters, particularly for small datasets. For example, individuals with very high incomes may exhibit very different behaviour from the rest of the general population. Addressing the effect of outliers can be done in a number of different ways. For the baseline OLS model, there are a number common tests for outliers including leveraging and

Cook's distance (Harrell Jr, 2015). More generally, cross validation methods are used to assess for outliers by partitioning calibration data into multiple pieces. Transition models are fitted to multiple datasets with each piece removed to determine any change in model coefficients (McLay et al., 2015). Cross validation and bootstrapping methods are commonly used as they are efficient and effective but higher order folds such as leave one out cross validation are used with large numbers of outliers (McLay et al., 2015). Cross validated models are then pooled together to produce final parameter estimates. This approach is robust against outliers, such as individuals with very high income seen in Section 2.4, but requires large sample sizes.

Dynamic microsimulations also extrapolate forwards in time and under different hypothetical scenarios. Many of the methods discussed in Section 2.5 to interpolate missing longitudinal data can also be used to extrapolate calibration data into the future and inform adaptive transition models. Alignment methods adjust the microsimulation population according to projected aggregated demographic information from robust and well-validated macro-models (Dekkers, 2015; Cockburn et al., 2014). Sample weights can be updated as if they were another variable in a microsimulation and used to perform post-hoc weighted aggregation (Dekkers and Cumpston, 2012). Microsimulation populations frequently add replenishing cohorts of individuals to population data in order to align with migration and demographic change (Archer et al., 2021). Using simulated data can reduce bias in model coefficients, as the population is changed via microsimulation dynamics to reflect change over time and counterfactuals, albeit with much higher uncertainty that is undesirable (Arnold et al., 2019; Li and O'Donoghue, 2013) to some model users and can required significantly more microsimulation iterations and computational power. As a compromise, integration of data assimilation into microsimulation has been suggested (Spielauer et al., 2007; Birkin, 2021; Asher et al., 2023). Transition models are updated as new survey data becomes available, allowing for rapid updating of model parameters.

Synthetic input data has a number of benefits (Klevmarken, 2022; Li and O'Donoghue, 2013; Richiardi et al., 2014) but introduces further uncertainty (Lomax and Norman, 2016). Generation of synthetic data can be sensitive to the choice of real input datasets, data generation algorithms, and validation methods (Müller and Axhausen, 2010; Lomax et al., 2024). Choice of real datasets used in data generation can produce substantially different results particularly if any datasets suffer from bias or low sample size. There are a number of available algorithms used to generate synthetic data including robust deterministic methods such as iterative proportional fitting and simulated annealing (Lomax and Smith, 2017b; Lomax and Norman, 2016; Zhou et al., 2022) and more recent stochastic Bayesian methods (Burgard et al., 2021; Zhou et al., 2022). Deterministic methods generate the same input population every time conditional on input parameters that should be tested using sensitivity analyses (Burgard et al., 2021). Stochastic methods also require testing of input parameters but produce different synthetic populations every time. Use of an ensemble of synthetic populations is required for calibrating transition probability allowing for inclusion of Monte Carlo uncertainty in population generation (Burgard et al., 2021). Synthetic data are validated against national-level or smaller aggregates on desired variables of interest such as household income distribution (Lomax and

Norman, 2016; Lovelace and Dumont, 2017). While synthetic data may match these aggregates it is not guaranteed to produce sensible outputs elsewhere. Transparency in which variables are validated against as well as clear descriptive statistics of synthetic data ensures datasets are representative demonstrating potentially uncertain policy outcomes(Lomax and Smith, 2017b; Lomax et al., 2024).

Synthetic data may also be utilised to simulate microsimulation data at varying time and spatial scales. Estimating small-scale processes such as unemployment or disease progression which can occur over shorter time-scales (such as days and weeks) (Archer et al., 2021) cannot be sufficiently captured using yearly transition modelling. Pseudo-continuous microsimulation modelling has been utilised in health microsimulation to combine short and long term processes together allowing for more detailed health progression (Spielauer et al., 2007). These methods have seen overall limited application outside of health microsimulation due to the difficulty fusing and aligning different sources of data into desired cross-sections and limited longitudinal data for the same individuals. New proposed methods including sensor fusion and data assimilation (Birkin, 2021) have potential to further utilise synthetic population data and provide increased policy detail on desired outcomes.

2.7.2 Uncertainty in Model Parameter Values and Behaviour

All estimated parameters in transition models have associated uncertainty (Sharif et al., 2012; Kopec et al., 2010). In the baseline OLS model, all model coefficients are Normal random variables. Different values these random variables can take may produce significantly different microsimulation results (Rutter et al., 2011). (Probabilistic) Sensitivity Analysis methods are used to (Spielauer et al., 2020; Rutter et al., 2011; Manoukian et al., 2022) perturbate input transition model parameters to determine effect on model output. These perturbations are implemented either as randomisation due to uncertainty in model parameters or change in coefficient due to hypothetical shocks (Rutter et al., 2011; Creedy et al., 2007). Sampling these coefficients is dependent on the probability distribution of model coefficients chosen by the user (Petrik et al., 2020; Sharif et al., 2012; Rutter et al., 2011). The simplest approach involves complete factorial design (Sharif et al., 2012; Petrik et al., 2020) but this often requires a very large number of microsimulation iterations to test all possible combinations. Random sampling from probability distributions is more commonly used (Petrik et al., 2020) whether this is from explicit distribution of parameters assumed in models such as OLS (Petrik et al., 2020), posterior distributions from Bayesian inference (Rutter et al., 2011) and more targeted importance sampling (Asher et al., 2023) that intelligently sample possible model parameters reducing required microsimulation iterations. Sensitivity analysis under hypothetical shocks is largely discretionary depending on which scenarios modellers think are most important to ensure robustness against (Rutter et al., 2009) such as worst case traffic or disease progression scenarios.

Additionally model outputs themselves may be considered as random variables. In the OLS example predicted values are not a point estimate but are also Normal distributed. Microsimulation commonly adds noise to predicted variables called Monte Carlo error (Sharif et al., 2012). Simple practical examples involve simply adding Normal noise to prediction of a continuous variable (Krijkamp et al., 2018; Petrik et al., 2020). Recently probabilistic methods are being implemented to estimate both mean and variance of individual state using Stochastic Differential Equations (Iskandar, 2019), Bayesian methods (Rutter et al., 2009) and probabilistic programming (Manoukian et al., 2022) that create individual units that exist in superposition albeit with small numbers of variables.

2.7.3 Uncertainty in Model Structure

There is uncertainty in the choice of model structure itself (Rutter et al., 2009; Sharif et al., 2012; Petrik et al., 2020) including both the choice of coefficients and statistical methods used. Variables that are predictive of next state but not included in transition models can result in omitted variable bias (Skarda et al., 2021). Clear understanding of transition probability methodology is required in order to determine any possible model misspecification or uncertainty involved in choice of model parameters (McLay et al., 2015). Significant model coefficients can vary over time and across strata requiring different transition models (Skarda et al., 2021). To ensure selection of all significant transition probability models microsimulation work suggests use of variable selection techniques. These methods sweep the space of available model coefficient choices, searching for best fitting models according to some information criterion (Sharif et al., 2012). There are a number of more elaborate variable selection techniques such as the fused LASSO (Adhikari et al., 2019) that are well suited for estimation of consistent trends in change in model coefficients over time but not yet used in microsimulation. Random samples of transition models are also implemented, each using some randomly drawn set of coefficients in sensitivity analyses (Petrik et al., 2020). This can be as simple as randomly choosing whether to include variables or not but Bayesian methods are available that can choose and switch between choices of model coefficients over time, again dependent on goodness of fit (DeYoreo et al., 2022; Shewmaker et al., 2022; Asher et al., 2023; Alsefri et al., 2020).

2.7.4 Monte Carlo Random Noise

Many transition probability methods are deterministic always predicting the same state for an individual unit (Krijkamp et al., 2018). The predicted state is taken as the mean point estimate when the true state belongs to a probability distribution defined by confidence intervals. Different potential state values cause small perturbations that can evolve into large differences in state over time across repeated microsimulation model runs (Sharif et al., 2012; Krijkamp et al., 2018). Sensitivity analysis must then quantify how much these perturbations effect prediction for individual units over time (Sharif et al., 2012; Krijkamp et al., 2018).

The simplest approach is to add random noise to predictions (Krijkamp et al., 2018) according to confidence intervals. Continuous variables often apply Normal noise or other specialised noise distributions (Krijkamp et al., 2018) while discrete variables apply adjustments to multinomial probability vectors. More elaborate methods generate and sample from posterior distributions using MCMC methods, purely probabilistic microsimulation, and data assimilation (DeYoreo et al., 2022; Birkin, 2021). These methods better estimate heterogeneity in individual probability distributions but are much more bespoke and expensive to implement.

2.8 Heterogeneity

Early microsimulation utilised low complexity transition probability models, these transition models were fitted to small survey datasets with limited covariates. Low complexity prediction is reliable and interpretable with low uncertainty providing strong estimation of aggregate cross-sectional population behaviour useful in certain applications such as traffic and policy analysis (Rutter et al., 2011; Li and O'Donoghue, 2013). The drawback is that individual behaviour can be poorly estimated, especially for individuals with extreme / outlying values. Individual heterogeneity is largely ignored from sources including longitudinal histories, spatial information and interaction with other units, and unexplained variance in behaviour (Richiardi and Poggi, 2012; Klevmarken, 2022; Li and O'Donoghue, 2013; Burgard et al., 2021; Rutter et al., 2009). As a result, the dynamics of individual trajectories can be highly biased due to significant under-dispersion and regression to group means (McLay et al., 2015; Salonen et al., 2020). As larger datasets and richer predictive methodologies become available, dynamic microsimulation is increasingly incorporating this heterogeneity to utilise all available information and better predict both individual and population behaviour simultaneously (O'Donoghue and Dekkers, 2018; Burgard et al., 2021). This section reviews these common sources of heterogeneity and methods used to address them in microsimulation.

2.8.1 Longitudinal Heterogeneity

Microsimulation was initially developed to incorporate long-term individual histories not possible with traditional macroeconomic models (Orcutt, 1961). Many microsimulation use first-order Markov transition models such that prediction of next state is performed using only current time information (McLay et al., 2015; Li and O'Donoghue, 2013). Methods such as linear regression use lagged dependent variables using previous state to inform next prediction (McLay et al., 2015; Burgard et al., 2021). Microsimulations then track variables over time accumulating desired population statistics such as pension contribution (Orcutt, 1961; Salonen et al., 2020). Increasingly higher order Markov prediction methods are being used to incorporate longer individual trajectories (Marois and Aktas, 2021; Li and O'Donoghue, 2013). Mixed effects models including Generalised Estimating Equations (Marois and Aktas, 2021; McLay et al., 2015; Zhou et al., 2021), higher order multi-state Markov models (Liao et al., 2021), and convolutional neural networks (Essien et al., 2019) all utilise information from further into the past. These methods are significantly more complex than first order Markov models requiring more development time and computational power (Alsefri et al., 2020). Accessibility is limited and further how-to papers may greatly improve the uptake of these methods.

2.8.2 Spatial Heterogeneity

Processes such as migration, disease progression and labour market supply are highly conditional on spatial orientation (Tanton et al., 2014). Generation of small-area synthetic population data is a rich field of research in dynamic microsimulation, aligning small-scale information against regional aggregates (Ballas, Clarke, Dorling, Eyre, Thomas and Rossiter, 2005; Lomax and

Smith, 2017b; Tanton et al., 2010; Odiari and Birkin, 2022; Burgard et al., 2021; Zhou et al., 2022) using methods such as simulated annealing and iterative proportional fitting (Lomax and Smith, 2017b). These spatial data can be projected forwards in time and including in predictive modelling allowing for spatial analysis within novel applications including disease progression (O'Donoghue et al., 2021), health economics (Smith et al., 2021) and agriculture (O'Donoghue et al., 2017). While these populations contain spatial information, transition probabilities conditional on spatial information are less widespread (Burgard et al., 2021). Spatial prediction models including spatial regression (Tanton et al., 2014), spatial interactions (Tanton et al., 2014; Ballas et al., 2013), spatial reweighting (Bhattacharjee et al., 2023), spatial linear regression (Tanton et al., 2014; Ballas, Clarke, Dorling, Eyre, Thomas and Rossiter, 2005), kriging (Selby, 2011), and empirical origin-destination data (Marois and Aktas, 2021) are all utilised to include this spatial information. Bayesian methods and random effects models are proposed, assuming a super-population of regional effects distributed according to some posterior distribution (Haslett et al., 2010; Arnold et al., 2019). This is particularly useful when estimating effects for areas with small populations due to being small-scale or rural (Burgard et al., 2021).

However, even if data for the full population are available, spatial effects may not be estimable due to sparsity and low sample size (Burgard et al., 2021; Leknes and Løkken, 2021). Instead, transition probabilities can be aligned against available region-level aggregates (Burgard et al., 2021) by including logistic scaling methods, adjusting transition parameters using iterative proportional fitting as well as maximum likelihood estimates constraining transition model parameters using quadratic programming (Burgard et al., 2021). These methods have seen substantial application to the generation of spatially representative input microsimulation data (Lovelace and Dumont, 2017; Lomax and Norman, 2016) and areas where policy varies over space including farming (O'Donoghue et al., 2017), the effect of the weather on general mood and energy spending (Ballas, 2020), urban development (Ballas, Clarke, Dorling, Eyre, Thomas and Rossiter, 2005), and energy poverty (Neto-Bradley et al., 2023).

2.8.3 Causal Effects

Transition models in dynamic microsimulation usually apply associative effects (O'Donoghue and Dekkers, 2018). Attributes in a dataset may be predictive of an outcome due to high correlation but not actually cause an individual to change state (Arnold et al., 2019). Being bald is a strong indicator of being male and hence certain diseases such as prostate cancer, but is unlikely to be an underlying cause. Causal inference is rarely applicable in policy analysis. Analysis is performed using retrospective datasets that rarely apply policy exclusively to two randomised groups as in a natural experiment (Marois and Aktas, 2021). Microsimulation provides a possible alternative with multiple copies of the same population available. Ideally two populations can be generated using microsimulation and compared using causal analysis (Kouser et al., 2021; Katikireddi et al., 2022). Use of structures such as directed acyclic graphs (DAGs) (Arnold et al., 2019) allows for clear presentation of decision pathways and aid understanding of why an individual moves state. Application of g-formula methods (Murray et al.,

2017) incorporate time varying and bidirectional causal effects which are difficult to estimate with associative methods. However, causal effects are difficult to implement in practice. Data requirements are very large needing information on critical life events in near continuous time that are rarely available. Microsimulation can simulate these data but any inferences made are unreliable (Ackley et al., 2022). Policies that can be tested must be carefully chosen due to strong mathematical assumptions that can limit utility of any microsimulation (Kouser et al., 2021; O'Donoghue and Dekkers, 2018). Significant methodological improvements are being made to test when causal inference is applicable in microsimulation but there are still limited successful applications (Skarda et al., 2021; Murray et al., 2017; Kouser et al., 2021).

2.8.4 Interaction Effects

Prediction of individual next state is often conditional on other nearby individuals(Dekkers and Cumpston, 2012; Burgard et al., 2021; Zinn et al., 2014). There can be multiple types of individuals, such as households, people, and business, all with different variables, interaction mechanisms, and complex social network structures (Bae et al., 2016; Axtell and Farmer, 2022; Hartnett et al., 2020). These interactions can result in non-linear emergent behaviour over time due to spillover effects and feedback loops (Arnold et al., 2019) that particularly influence social behaviour and decision making (Klevmarken, 2022).

These kind of social interactions are historically dominated by 'bottom up' individual-level modelling techniques such as agent based modelling (ABMs) that employ toy (simple) datasets favouring analysis of theory over empirical basis (Birkin, 2021; Bae et al., 2016). As data becomes available hybrid ABM-microsimulation methods are increasingly adopted to combine the benefits of empirically driven microsimulation and complex social interactions (Birkin, 2021; Bae et al., 2016; Axtell and Farmer, 2022).

These hybrid methods are primarily used in socioeconomic applications (Axtell and Farmer, 2022). Estimation of labour supply markets is used to drive individual migration and housing behaviour as they search for the best available work (Axtell and Farmer, 2022). Stochastic social behaviour is employed determining change in investment habits in response to government policies(Klevmarken, 2022). More recently the coronavirus pandemic has inspired a surge in health based interaction behaviour(O'Donoghue et al., 2021; Rutter et al., 2011; Zucchelli et al., 2010; Smith et al., 2021). Contagious disease are much more likely to spread between individuals who share close contacts whether in households or airports (O'Donoghue et al., 2021). Future expansion of ABM techniques are aiming to include further stochastic behavioural response in order to better estimate individual behaviour and overall aggregate microsimulation output (Klevmarken, 2022; Axtell and Farmer, 2022)

Interactions also occur within individuals whereby change in one variable influences change in another. These changes can occur simultaneously (Birkin, 2021; Frick et al., 2009), such as an individual losing their job and immediately losing their employment income, or subsequently such as having a child and moving to a larger household or area with good schools. These interactions can either using several transition models in sequence such as births module

(Tikka et al., 2021; Li and O'Donoghue, 2013) and then a housing module where a new birth conditions moving behaviour, or use joint modelling estimating multiple variables in one model (Klevmarken, 2022) using methods such as multiple regression. Use of joint models that capture endogeneity between variables is ideal but rarely practical (Klevmarken, 2022). Additionally, careful microsimulation design is required to ensure that transition modules progress in correct order capture concurrent interactions and preventing illogical new states (Li and O'Donoghue, 2013; Müller and Axhausen, 2010; Andreassen et al., 2020). Joint variable modelling is fairly uncommon in microsimulation using methods such as multivariate regression and hazard ratio models (Zaidi and Rake, 2001) favouring simpler univariate techniques that are generally easier to implement and interpret.

2.8.5 Social Heterogeneity

Behaviour in a human population can vary due to social heterogeneity (Reeves et al., 2022; Klevmarken, 2022). Individual people with identical objective socioeconomic characteristics have different perceptions and behaviour in response to new government policy and socioeconomic events (Reeves et al., 2022). One well discussed example is incorporation of behavioural response to COVID-19 in individual-level models trying to identify individuals who adhered to government guidelines including mask use and staying indoors (Reeves et al., 2022). Microsimulation historically does not estimate social behaviour due to no available variables in administrative datasets (Arnold et al., 2019; Klevmarken, 2022). Modelling of social behaviour was reserved for agent-based modelling that emulates social theory using a bottom up approach (Arnold et al., 2019).

Modern administrative data is beginning to include subjective variables describing individual opinions such as financial stability and thermal comfort. These variables are being incorporated into transition probabilities to more accurately estimate personal decisions and move away from 'average behaviour with random deviation' (Klevmarken, 2022).

Another approach is incorporation of transition dynamics from agent-based models in a hybrid approach (Birkin and Wu, 2012). Calibration of social science theory against empirical data has shown to able to better capture adaptive and emergent social behaviour in several microsimulation models (Birkin and Wu, 2012; Bae et al., 2016) for estimation of economic (Bae et al., 2016) and health policy (Asher et al., 2023). Substantial areas of future work are available including better calibration of hybrid models using data assimilation approaches (Birkin, 2021) and application to new areas of research including housing and energy policy (Khalil and Fatmi, 2022). Social behaviour intersect strongly with other heterogeneity discussed here including spatial modelling of phenomena including social weather (Ballas, 2020) and neighbourhood interactions (Birkin, 2021).

Incorporation of heterogeneity into transition probability models allows for better prediction of behaviour by inclusion of further information on differences in individual behaviour beyond simple demographic strata. Methods are showing improvement in prediction quality of individual trajectories across multiple applications. While it is likely impossible to fully capture the complexity of human and other individual behaviour higher accessibility of these more advanced methods through tutorial papers outlining model implementation through assumptions and diagnostics would utilise modern rich datasets and further improve long term prediction quality of microsimulation models.

2.9 Discussion

Dynamic microsimulation models estimate the future state of a population of individual units using transition probability models that must be well-calibrated to ensure accurate, robust prediction that can be trusted to affect real-life decision making. Development of these transition probability models can be difficult due to the complex nature of human behavioural systems including interacting causal mechanisms and large heterogeneity. Careful selection of modelling technique is required ensuring background mathematical assumptions, justifying sensible choices of predictor variables, and extensive validation to ensure sensible behaviour over time and under counterfactual interventions.

Broader review of the state of dynamic microsimulation (O'Donoghue and Dekkers, 2018) has highlighted that knowledge of best practice and common pitfalls for several methods including transition probability development are often hidden in proprietary research and out of print journals. Microsimulation methods must then be codified and standardised in open source literature to ensure transmission of tacit knowledge to new users and further application of microsimulation to new areas to match sister techniques such as Agent-Based Modelling (O'Donoghue and Dekkers, 2018).

To counteract this issue a number of tutorial literature and software packages are being released in order to guide new users and providing living, open-source documentation of microsimulation methods including alignment and input population generation (Lomax and Smith, 2017b; Burgard et al., 2020). This paper provides a review into the current state of transition probability methodology, outlining common issues in development and publication, proposed methods to address them, and potential future innovation in four key areas:

1. Presentation - Better presentation of the development cycle of transition probabilities improves interpretability and ability to critique current models. High accessibility to new users increases rate of expansion to new fields and new ideas.

2. Missingness - Handling missing data increases the sample size of transition models improving statistical power and preventing bias due to attrition. Missing data handling techniques are very well implemented in microsimulations but could be expanded to more complex missing types to even further accommodate available information.

3. Validation and Sensitivity Analyses - Generation of spatial data as well as projection into the future and under counterfactual scenarios in microsimulation all have associated uncertainty. Quantifying uncertainty in transition probability models as well as microsimulation as a whole is essential to ensure robustness to model misspecification and potential sensitivity in model coefficients and from future events.

4. Heterogeneity - Transition probability models are often oversimplified and do not include individual effects. A number of methodologies are available dependent on context to increase the complexity of transitions to better predict individual behaviour while retaining parsimony to keep simplicity and easier interpretation.

Methods to address these issues have improved substantially over the last 30 years as computation power and available data increases. However there are still a number of key criticisms levelled against current implementation of transition probabilities.

Transition probability models have low interpretability. It is difficult to determine 'why' certain future states are predicted as well as how these future states interact with other transition models and counterfactual scenarios. There is limited testing of model assumptions for each transition model. Testing background assumptions should be considered essential (McLay et al., 2015) in order to prevent model misspecification and elucidate potential bias. Transition probabilities are under-reported and not as standardised as other areas of microsimulation development (Burgard et al., 2020). Methods used are typically omitted from research literature making it difficult to review and replicate new methods (Richiardi et al., 2014) slowing down uptake of new methods and accessibility to new fields.

One proposition to address these issues is to further standardise the implementation of transition probability methods. Set guidelines for statistical testing, sensitivity analyses and visualisation would allow for faster development time that allows for assumption testing and improved interpretability. Further papers reviewing specific areas of transition development in detail combined with tangible case studies would further increase accessibility of dynamic microsimulation. Additional technologies allowing for conversion of transition code and notebooks into online documentation and repositories would make publication substantially less time consuming, open source, and easier to maintain beyond project funding.

2.9.1 Best Practice Microsimulation Development

This section provides a brief overview of the development cycle of transition probabilities identifying an overall structure as well as key considerations to ensure best practice. Examples from published literature that apply these principles are provided.

Planning

As with general microsimulation guidelines (Burgard et al., 2020) the majority of development time should ideally be spent on planning. Justification of variables, model type, model structure, and sources of uncertainty should be performed before any implementation to ensure consideration of previous literature and all available options that best fit data, prevent over-fitting, and produce 'thoughtful' (Wasserstein et al., 2019) considerate of previous and future application in other microsimulation literature.

- What variables should be included in prediction according to existing literature?
- How are transition models for a given process implemented in other microsimulation literature?
- What sources of uncertainty must be accounted for?

- What programming language and microsimulation framework will be used? Ideally existing model frameworks are used to promote readability.
- What datasets are used for calibration?
- What is the structure of missing data and what correction methods are used?
- What modelling methodology and structure should be used?
- What are the underlying transition probability model assumptions?
- How will a transition model be assessed?
- If there are several transition probability models, which order do they occur and how do they interact?
- How will counterfactual scenarios affect transition models?

Ideal model structure design is provided in (Skarda et al., 2021). Highly detailed model structures are provided including justification of predictor variables based on available causal evidence. When testing transition probability model assumptions (McLay et al., 2015) compares several potentially transition probability models outlining their different background assumptions including whether they are met and whether they influence prediction. Description of missing data correction in (Davis and Lay-Yee, 2019) provides a full imputation methodology in line with wider practice as well as detailing how uncertainty may affect prediction. Many modern microsimulation papers provide detailed open source code based on standardised frameworks (Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023) that are considered in wider planning. Further overall improvement may be provided using standardised design documents and supplementary material that is flexible enough to account for highly bespoke microsimulation models.

2.9.2 Validation

Practically, the majority of development time will be spent validating transition probability models. Extensive testing and clarity in model background assumptions prevents misspecification and allows for clear demonstration of where these assumptions may fail. Testing model outputs against out of sample internal data and external independent model results using statistical testing and visualisation ensures that predicted outputs are sensible and match reality. Alignment methods are used to adjust these outputs in order to match statistical moments (mean and variance) over time and under hypothetical scenarios. Finally sensitivity analysis is used to test the robustness of transition models to change in input parameters and determine where overall microsimulation output is vulnerable to changes in model assumptions.

- Are transition model assumptions satisfied? If not, how could this bias prediction?
- How is this model selection justified against other available predictive methods?

- How do model outputs compare again in sample data and out of sample reserved data? Do both statistical testing and 'eye test' visualisation agree results are sensible?
- What predictive variables are included in this model? Are they statistically significant? If not, why are they included otherwise?
- Does validation of model predictions against internal and external data produces sensible and credible output?
- How is a transition model aligned to better estimate real data?
- How sensitive is prediction to sources of uncertainty from input data, model coefficients and structure, and Monte Carlo error?

Strong internal validation is provided in (Skarda et al., 2021) comparing a simulated transition model and wider microsimulation output against available real data in nowcasting. Similar strong external validation (Skarda et al., 2021; Lomax and Smith, 2017b) is also provided comparing microsimulation output against external real data. Further external validation into the future and under counterfactual scenarios is highly bespoke but several papers have demonstrated application against other real data (Lomax and Smith, 2017b; Skarda et al., 2021) as well as against other projected output in meta-analyses (Archer et al., 2021). However, this is still limited due to few microsimulation models estimating the same scenarios. Validation of model background assumptions provided in (McLay et al., 2015) tests a suite of potential models identifying where model assumption fail and how this affects predicted output. Many articles show strong incorporation of uncertainty including outlining overall sources and methods (Sharif et al., 2012) as well as real application (Petrik et al., 2020) in overall model development and sensitivity analyses (Burgard and Schmaus, 2019). Sensitivity analysis estimating interactions between transition probability models (Harding, 2017) is still largely under-reported. Inclusion of sensitivity analysis exploring the 'ordering of modules' (Münnich et al., 2021) and interaction with policy interventions could be used to highlight why interventions enact individual change. These areas, combined with exploration of model structure uncertainty using variable selection methods (de Oliveira et al., 2024), would further increase robustness of microsimulation prediction.

2.9.3 Interpretation

Interpretation should ultimately make it clear to microsimulation users how and why change in behaviour occurs due to a specific transition model. This is the vaguest aspect of transition development due to being highly complex and bespoke to each microsimulation but with a number of common themes. There is significant overlap with clear presentation of model validation above with some additional considerations. Further visualisation methods are being developed in order to assess how individual trajectories and population distributions change into the future and under counterfactual scenarios for individual variables. This is particularly import for the assessment of individual longitudinal trajectories to ensure suitable dynamicism as well as consistency (McLay et al., 2015; Salonen et al., 2020). Being able to trace direct variable change due to counterfactual scenarios as well as interactions with other transition models allows for clear narrative of how change propagates through a microsimulation highlights key systems and potential further counterfactual scenarios. Technologies such as GitHub pages make generation of online documentation much more feasible but standardisation of documentation and methods to quickly convert development code such as in R. notebooks to online websites may further increase uptake.

- Are the above planning and validation well presented in literature or supplementary material?
- Is supplementary material fully describing model development available in open source technical documentation either in literature or on website such as GitHub pages?
- What visualisation methods are used to indicate how individual and global population states change over time? Are these common methods or should custom visualisation be used?
- Do these plots clearly explain how each transition probability model influences change in key model results? Can these changes be represented as simple stories that could be understood by the wider public?
- Can visualisation be aligned to existing literature to ensure easier comparison and review?
- Can model output be disaggregated to demonstrate how different subgroups of the population respond to microsimulation scenarios?
- Are provided results modest highlighting uncertainty and where prediction may deviate from reality over time?

Many papers publish full documentation (Richiardi, Bronka, van de Ven, Kopasker and Katikireddi, 2023; Kopasker et al., 2023) presenting model formulae, coefficients, larger model order of operations, and background code. Further presentation of transition models may seek to include more of the full development process from initial policy ideas and literature evidence using approaches such as systems thinking (Meier et al., 2019). Combined with largely standardised visualisation of outcome variable using standard bar plots and confidence intervals (Zinn et al., 2014) it is straightforward to identify variables that determine future individual state and cross-sectional behaviour over time. While transparency has greatly improved interpretability further innovation continues to present microsimulation output in terms of non-technical narratives. Several papers highlight detailed change in populations over time and under interventions using disaggregation by subgroups (Archer et al., 2021), longitudinal (Salonen et al., 2020) and spatial (Lovelace and Dumont, 2017) histories, as well as interactions between individual units (Burgard and Schmaus, 2019). Further innovation continuing to display detailed individual life paths and change due to interventions can further help to develop strong, interpretable narratives of intervention effect caused by interactions between multiple microsimulation transition modules and other components- (Morrison, 2008).

Application of dynamic microsimulation is increasingly rapidly as access to open source individual-level data, coding frameworks, technologies for faster development, and demand for simulated evidence increases. Strong standardisation of methodological background is essential to ensure new research teams can adopt best practice of transition models allowing easy access to existing common methods and integration of more complex models as they arise.

Inclusion of new predictive methods including artificial intelligence to estimate non-linear behaviour, inclusion of finer granularity data over short time points and spatial scales, and estimation of social behaviour in hybrid-abm microsimulation. These methods all have immediate potential to be integrated into microsimulation relative to key issues and development questions defined above increasing impact through expansion to new sectors and research questions beyond traditional policy analyses.

2.10 Connection between Chapter 2 and Chapter 3.

In order to increase the impact of dynamic microsimulation modelling this chapter contributes to the literature codifying best practice of transition probability model development for dynamic microsimulation. Five key categories of methodological improvements have been identified that contribute to both the application of ATOMIC principles, and improvement in goodness-of-fit predicting individual future states.

Good presentation of transition models should include full publication of model structure, diagnostic testing, justification for predictive variables used, and complete model coefficient tables. Transparent publication of all available material ensures open, reproducible research. Sufficient testing of model assumptions allows sufficient, modest review of potential model misspecification. A clear model structure with diagrams will improve interpretability of complex systems of interacting transition probability models and ensure thoughtful consideration of previous simulation modelling, evidence from literature and on pathways effected by a policy.

Consideration of missing data, uncertainty due to input parameters, and methods of validation constitute the majority of transition model development time. Ensuring that model output is sensible preserving overall statistical moments as well as individual dynamics in confidence intervals is essential for reliable prediction that is accepting of uncertainty and modest to large deviations in prediction when predicting far into the future. Combining these methods again with strong presentation again ensures transparent modelling that can be aggressively reviewed by Government bodies before use as formal evidence. Microsimulation provides unique application of a number of visualisation techniques to ensure predictive quality of transition model output over time, handover plots, trajectory analyses and spatial mapping.

While these methods improve the reliability of microsimulation modelling they typically do not improve prediction quality. Better estimation of individual future state largely relies on heterogeneous modelling methods able to include more variables and more data to account for the high complexity of individual trajectories. The most common example is application of longitudinal modelling due to repeated wave information from administrative data. As datasets continue to improve, the robustness of prediction of policy effects will also increase in quality.

This literature review provides a useful general summary of available methodology that can be extended through further papers exploring individual methods and best practice in case studies with tangible output in larger microsimulation models and reusable software for creation of online documentation using common programming languages for transition modelling such as R.. The following chapter aims to apply these methods to estimate Short Form 12 Mental Component Score (SF12-MCS) using the Understanding Society data exploring how a simple linear regression model can be extended to considering model diagnostics, out of sample model performance, uncertainty quantification, and model selection criteria. Estimating SF-12 MCS is then required as part of a larger microsimulation model estimating change in health due to income support policy.

Chapter 3: An Example Case Study on the Application of Transition Dynamics for Discrete Time Dynamic Microsimulation

3.1 Abstract

Transition dynamics are used in microsimulation to evolve the state of individual entities over time. A random draw from a probability distribution predicts an individual's next state conditional on current information. Good calibration of transition probabilities is critical for sensible projection of individual and population level behaviour reducing bias in any projection results. In order to facilitate best practice and accessibility of microsimulation to new users a series of 'how-to' tutorial papers have been published demonstrating application of microsimulation methods such as synthetic population generation.

This paper provides such a tutorial case study constructing a discrete-time transition probability model to estimate Short Form 12 mental component score (SF-12 MCS, a summary metric of individual's mental well-being) for individuals for the United Kingdom using the Understanding Society dataset. An initial Ordinary Least Squares model is applied to SF-12 MCS score and then improved upon by considering improved presentation of results, missing data, uncertainty, and heterogeneity due to repeated household observations. A final Generalised Linear Mixed Gamma-Normal model provides a substantially better fit than the baseline linear model showing improvement in several diagnostic tests and a reduction 5-fold cross validation root mean squared error from 0.223 ± 0.015 to $0.152 \pm 6 \times 10^{-3}$.

3.2 Introduction

Transition dynamics are a key component of dynamic microsimulation models (Spielauer et al., 2020; Burgard et al., 2020). A population of individual entities such as households or humans is generated and projected forwards in time to estimate the effect of some hypothetical counterfactual scenario. This projection is typically performed using a system of transition probability models each predicting some subset of an individual's future state space condition on current time information (Harding, 2007; Rutter et al., 2011). In theory, any predictive methodology can be used as a transition probability model as long as it estimates future state (Zaidi and Rake, 2001). Simplistic transition models such as logistic regression (Andreassen et al., 2020) have seen great success in microsimulation modelling as they are easy to implement, diagnose and interpret. However, the processes that govern individual populations, particularly human populations, can be highly complex, non-linear, and contain multiple competing socioeconomic forces (Marois and Aktas, 2021; Richiardi et al., 2014). Multiple interacting causal mechanisms (Skarda et al., 2021) and heterogeneity across socioeconomic background (Marois and Aktas, 2021), environment (Burgard et al., 2021), and individual behaviour (Klevmarken, 2022) can all influence future state. As a result, increasingly complex transition probability models are being used in microsimulation to capture individual dynamicism (McLay et al., 2015). Mixed and random effects models are utilised to account for repeated observations over time from the same individuals and the same spatial areas (Marois and Aktas, 2021). Neural network models (DeYoreo et al., 2022) are used to also capture individual histories and highly non-linear behaviour. Multi-level and hierarchical models including Markov Chain Monte Carlo (Serena et al., 2023; DeYoreo et al., 2022; Leknes and Løkken, 2021) are utilised to estimate behaviour in sparse conditions with limited data and account for interaction between multiple levels of individual within households or neighbourhoods.

As transition probability models become increasingly complex they are more difficult to validate and interpret (Li and O'Donoghue, 2013). Model selection must be considerate of traditional goodness of fit metrics and background mathematical assumptions (McLay et al., 2015) as well as unique microsimulation considerations including interaction between transition probability modules (Harding et al., 2010a), the order in which these transitions occur (Li and O'Donoghue, 2013), and repeated application of statistical models (Richiardi et al., 2014). Ideally, description of the full development of transition probability models is included in microsimulation literature to facilitate accessibility of more complex transition methods to new users and critical review. However, this information is typically relegated to supplementary material or outright omitted due to page constraints, results focused publication pressure, and expenses associated with the development and maintenance of full technical documentation (Richiardi et al., 2014; Goedemé et al., 2013; Davis and Lay-Yee, 2019; O'Donoghue and Dekkers, 2018).

In order to increase the impact of dynamic microsimulation, research suggests that standardisation of common microsimulation methodological techniques into open source best practice will reduce tacit knowledge and prevent new users repeating common mistakes (O'Donoghue and Dekkers, 2018). There is already a set of published articles covering several microsimulation methods including population generation (Lomax and Smith, 2017b; Lovelace and Dumont, 2017), alignment (Li and O'Donoghue, 2016), validation methods (Li and O'Donoghue, 2013), and the general state of the art (Li and O'Donoghue, 2013; O'Donoghue and Dekkers, 2018). However, there is limited evidence available on best practice of transition probability development for dynamic microsimulation outlining common methods, model selection and assessment of model fit in and out of sample performance and validation.

The paper provides case study uses the UK Understanding Society annual survey data to estimate change in mental well-being using Short Form 12 Mental Component Score (SF-12 MCS) as part of the SIPHER Microsimulation for Interrogation of Social and Health Systems (MINOS) model. While this paper can be read as a stand-alone article, readers may also be interested in a supplementary literature review [CITE AFTER RELEASE] that provides an overall review of methods used to interpret and improve transition probability models. An initial Ordinary Least Squares model is applied to estimate SF-12 MCS and shows overall poor prediction of future state violating of several underlying model assumptions and failing to preserve trends in SF-12 MCS mean and variance trends over time. Better presentation of the OLS model highlights violation of these background statistical assumptions and highlights potential higher complexity models that could better estimate SF-12 score. Sensitivity analysis of the baseline model is also performed testing vulnerability to potential sources of uncertainty. Application of missing data imputation using the Multiple Imputation Chained Equations (MICE) algorithm, randomisation of model coefficients, and application of the LASSO variable selection algorithm determined the baseline model is not sensitive to model structure and coefficients but prediction is significantly affected for years in which variables such as loneliness are missing in UKHLS data. Results suggested use of Generalised Linear Mixed modelling and lagged dependent variables to account for heterogeneity over time due to inconsistent measurements and repeated observations from the same individuals showed improved model fit better for both model assumptions and goodness of fit metrics.

3.3 Baseline SF-12 MCS Transition Probability Model

3.3.1 The MINOS Dynamic Microsimulation framework

Within a UK context, initial work on dynamic microsimulation applications included models such as PENSIM2 and DYNASIM (Li and O'Donoghue, 2013). These microsimulations estimated long term change in pension contribution and other economic outcomes for United Kingdom households using simplistic but highly reliable transition dynamics. Further work extended these applications to health outcomes exploring the change in further health economics outcomes (Skarda et al., 2021; Spielauer et al., 2007), as well as hospital clinical trials simulating disease prognoses and resource management (Rutter et al., 2011). These applications highlight the potential for novel applications of dynamic microsimulation in the UK and other international contexts.

As data becomes available there is now increasing interest applying dynamic microsimulation to estimate both mental and physical well-being outcomes (Kopasker et al., 2023; de Oliveira et al., 2024). Recent literature (de Oliveira et al., 2024; Kopasker et al., 2023; Katikireddi et al., 2022; Broadbent et al., 2023) is increasingly estimating mental well-being outcomes for the United Kingdom population in response to multiple crises including the coronavirus pandemic, high energy prices and general cost of living, and transition to carbon neutrality.

There is strong potential for microsimulation to address these mental well-being issues as part of wider complex systems thinking framework (Meier et al., 2019; Schünemann et al., 2024) that requires holistic consideration of seemingly unrelated aspects of human population behaviour when implementing policy. For example, this approach requires consideration of health outcomes for non-health policies in other sectors such as employment and transport (Meier et al., 2019). Exhaustive exploration of all possible variables that could influence behaviour requires consultation with domain experts and literature evidence combining this data into causal loop diagrams (Meier et al., 2019; Schünemann et al., 2024). These diagrams are then emulated using interpretable microsimulation models (Zaidi and Rake, 2001) creating narratives (Meier et al., 2019; Schünemann et al., 2024) describing why policies enact change. These narratives can be shared across sectors and to the general public, increasing trust in these simulations and the results that they produce. The MINOS model contributes towards interpretable microsimulation as part of a complex systems thinking framework using transparent open source data and code as well as a flexible modular design facilitating review of model structure and easier application to new scenarios. The MINOS model can then be used to generate evidence evidence for these narratives using transition models demonstrated within this paper.

The Microsimulation for Interrogation of Social and Health Science Systems (MINOS (Clay et al., 2023; Lomax et al., 2023)) model has been developed by the SIPHER consortium. The MINOS model is fully open source, written in the R and Python languages, and utilising data that are readily available to academic research teams. MINOS estimates the health effect of non-health policies on the wider population and specific vulnerable subgroups as part of wider

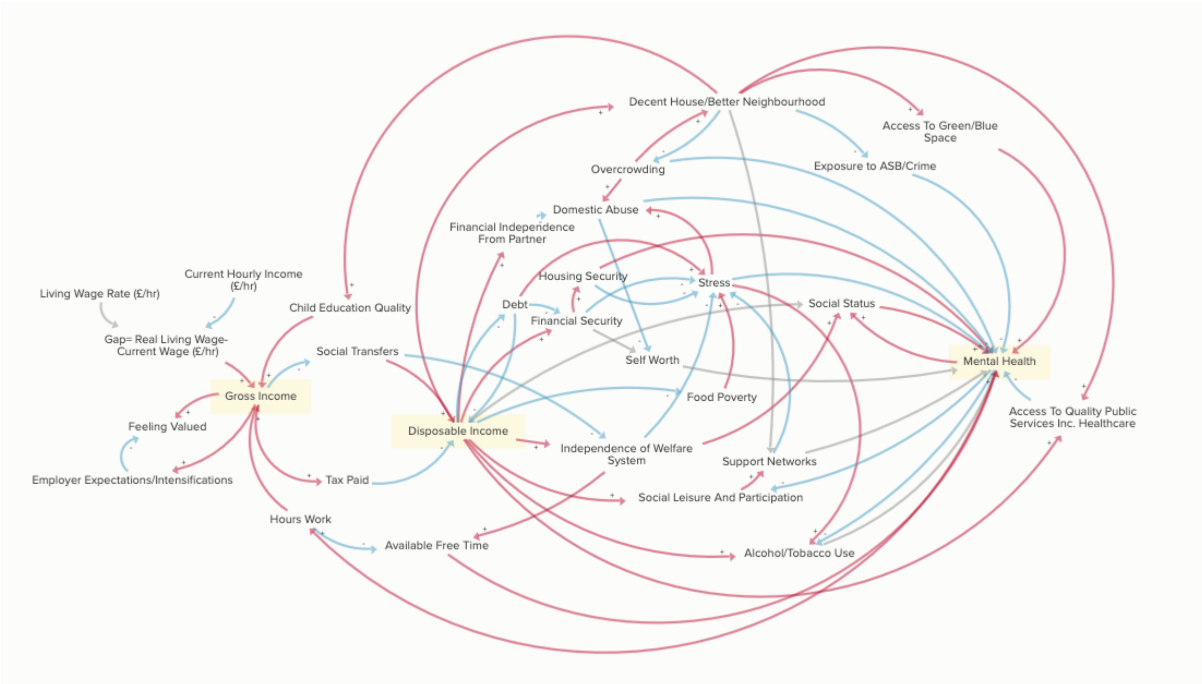


Figure 3.1: The SIPHER (Consortium, 2023) causal loop diagram between household disposable income and health used to inform construction of MINOS transition model pathways.

systems thinking (Meier et al., 2019). A simplified structure of the transition probability models used in the MINOS model is provided in Figure 3.2. This demonstrates how the microsimulation applies policy interventions to household disposable income before a final estimation of mental well-being outcomes using SF-12 MCS score (Lawrence and Fleishman, 2004). The remainder of this paper sets out how the Mental Health module was developed for MINOS starting with initial theory driven specification, parameterisation of causal loop diagrams conditional on available data, and optimisation of an initial simple linear regression model predicting future SF-12 MCS state.

3.3.2 Causal Loop Diagram Refinement into an Associative Pathways for an SF-12 MCS Transition Probability Model

Initial development of an SF-12 MCS transition probability module began with a detailed disposable income to health causal loop diagram (CLD) developed by a multidiciplinary team of academic and policy partners who form the SIPHER consortium (Hill O'Connor et al., 2023; Campbell et al., 2023; Consortium, 2025) in Figure 3.1. The process used to construct this CLD is described elsewhere (Campbell et al., 2023; SIPHER Consortium, 2025; Consortium, 2025) outlining how causal literature evidence and opinion from community member and academic expert panels describes causal relationships between individual, household, and neighbourhood characteristics. This diagram describes the pathways between changes in household disposable income and health in the United Kingdom population and can be used to construct narratives highlighting which nodes a policy intervenes upon, and propagation through causal pathways to desired health outcomes.

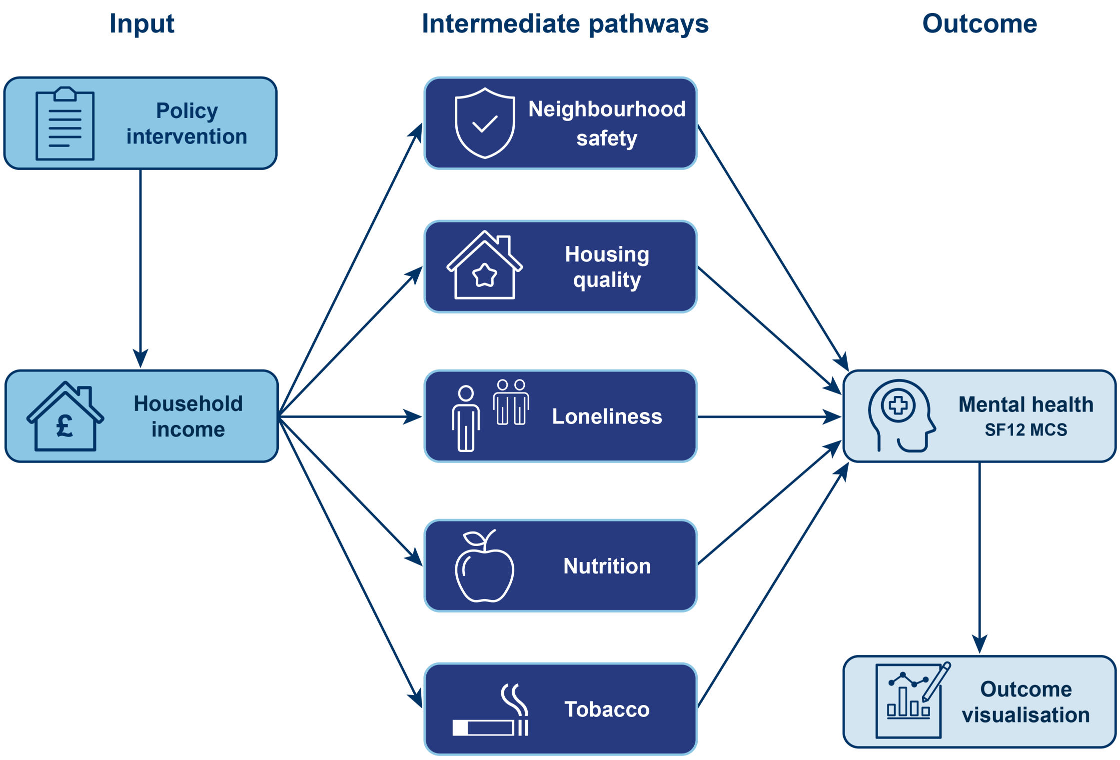


Figure 3.2: High Level pathways derived from SIPHER consortium causal systems maps. These pathways indicate key variables that are influenced by change in household disposable income and go on to influence change in SF-12 MCS score. Each pathway module contains a series of related transition probability models.

The MINOS model is used to emulate these causal relationships. Following the wider microsimulation literature, this causal loop diagram is partitioned into a series of transition probability modules each predicting future state for some subset of nodes (Arnold et al., 2019). A transition model defines some dependent variables that represent each node as well as a model structure and independent predictor variables estimating all incoming causal relationships from other nodes. Variable choices are highly dependent on available data. The choice of model is similarly driven by variable availability and the need to deliver efficient performance within the dynamic microsimulation framework (Andreassen et al., 2020; Zinn et al., 2014; Zaidi and Rake, 2001). Application of true causal modelling is rare in microsimulation due to the event-based data required that is not currently present in administrative cross-sectional data (Skarda et al., 2021; Kouser et al., 2021; Katikireddi et al., 2022). Associative modelling techniques are typically used to approximate future state rather than fully emulate a causal sequence of life events (Zaidi and Rake, 2001). For the MINOS SF-12 MCS model, the choice of predictor variables and model structure are defined in more detail in the following section.

Fitted SF-12 MCS transition model discussed throughout the remainder of this article underwent several internal SIPHER reviews with domain experts assessing the initial income to health CLD and choice of predictor variables and model structure. Each predictor variable is examined to ensure it sufficiently represents each CLD node and connecting causal pathways. Any potential pathways that cannot be estimated by a transition model must be justified with particular consideration on how this assumption may bias future prediction. Conversely, the choice of predictors can highlight any potential gaps in the CLD due to insignificant variables or poor model fit. This review process is performed iteratively updating the CLD, predictor variables, and output interpretation in sequence.

For the MINOS microsimulation this review was performed by compiling transition model output into R. notebooks shared internally within the SIPHER consortium. Significant refinement based on available data and desired modelling using simple linear regression resulted in a final SF-12 MCS model structure as well as the full series of modules provided in Figure 3.2.

3.3.3 Data

The United Kingdom Household Longitudinal Survey (UKHLS) (Benzeval et al., 2020) consists of 13 waves (2009 − 2021) with a sample in 2009 of circa 40,000 households, designed to be representative of the UK population. At time of writing wave 13 had just been released, our results use waves 1−12. The UKHLS contains hundreds of attributes across numerous categories including demographics, labour state, education and health in both household and individual respondent files. SF-12 MCS score (Lawrence and Fleishman, 2004) is our outcome variable of interest. SF-12 MCS is available in all 12 waves of data. SF-12 MCS provides a continuous score of mental well-being derived from 12 binary questions using principal component analyses (Lawrence and Fleishman, 2004). This score is normalised to approximate a Normal distribution with mean 50 and variance 10^2, providing a single continuous mental functioning score between 0 (low functioning) and 100 (high functioning) (Lawrence and Fleishman, 2004). Use of a Normal variable is favourable due to its extensive use in microsimulation, availability in a wide suite of modelling techniques, goodness of fit diagnostics, and implementation in most programming languages (Harrell Jr, 2015).

Explanatory variables used to to predict SF-12 MCS score are provided in Table 3.1, all of which have been shown to influence mental well-being or have been identified in previous evidence synthesis and exploration of available UKHLS variables (Hill O'Connor et al., 2023; Tsuchiya et al., 2022). These include demographic variables (age, sex, and ethnicity (Burdine et al., 2000)), education state (Jones, 2017), net household income (Jones, 2017), labour state (Burdine et al., 2000), and NS-SEC socioeconomic group (Chandola and Jenkinson, 2000). Five further key variables for the prediction of SF-12 MCS (Hill O'Connor et al., 2023) including loneliness, smoking via weekly cigarette consumption, nutrition quality, housing quality, and neighbourhood safety are omitted from the baseline model. Loneliness and cigarette consumption are not available in any waves from $2011 - 2013$. Housing quality, nutrition quality, and neighbourhood safety are proxy variables not directly present in UKHLS data and must be derived using formulae (Tsuchiya et al., 2022) derived elsewhere. These proxy variables are also only available every three years and application to all waves of UKHLS data requires missing data correction provided in Section 3.5. An individual identifier (pip) is included to track which repeated observations belong to which individual for use in longitudinal modelling. This is an intentionally small subset of variables to allow application of methodology to other datasets with more limited resources and encourage parsimonious modelling. If further mutable variables are included they require their own transition models that can exponentially increase microsimulation complexity. There are many other strong predictors of SF-12 MCS that are not included such as loneliness and housing quality because of limited availability across waves, high data missingness, or difficulty in applying a transition probability model.

A subset of three years of UKHLS data (2011-2013) is used in this study, providing $n = 139813$ observations from $k = 48111$ unique individuals. This represents a period where data availability for variables in Table 3.1 is as good as possible (see imputation below). Further data up to 2020 are reserved for model validation.

Preprocessing UKHLS data is done in two stages. The first stage is variable encoding. Many categorical variables are redefined from integer to string values. This is mainly for readability as it is easier to read 'employed' than 'labour state 2'. Categorical variables that are re-encoded are specified in Table 3.1 The second stage is missing data correction. Only 41% of all observations are complete. Missing values in the NS-SEC, SF-12 MCS, education and ethnicity variables account for over 99% of all missing data. The baseline model only uses complete cases such that any individual observation with missing values for variables in Table 3.1 are removed. This is also applied to SF-12 MCS observations at the next point in time. For example, an observation in 2011 that does not have an SF-12 MCS value in 2012 is removed. It is impossible to estimate a transition to the next state with no data. After complete case correction there are $n = 57195$ remaining observations.

Variable Name	Description	Units
SF-12 MCS	SF12 Mental Component Score	Continuous (0-100)
pidp	Personal identifier	Integers > 0
time	Interview Year	(2011, 2012, 2013)
sex	Subject Biological sex	Categorical M/F
age	Subject Age	Years
ethnicity	Subject Ethnicity	Categorical (reference White British)
education_state	Highest education qualification achieved	Categorical (reference GCSE)
labour_state	Current labour state	Categorical (reference Employed)
region	Administrative Region	Categorical (reference London)
net_hh_income	Net Household Income	Continuous
NSSEC	NS_SEC socioeconomic class	Categorical (reference unemployed)
ncigs	Weekly Cigarette consumption	Counts (>= 0)
housing_quality	Housing Quality	Ordinal (1-3)
neighbourhood_safety	Neighbourhood Safety	Ordinal (1-3)

Table 3.1: Variables from the Understanding Society (UKHLS) dataset used to predict SF-12 MCS. Full discrete variable encodings are provided in online supplementary material https://leeds-mrg.github.io/Minos/.

3.3.4 MINOS SF-12 MCS Module

Baseline prediction of SF-12 MCS is achieved using ordinary least squares (OLS) regression (Harrell Jr, 2015). SF-12 MCS score at the next point in time $t + 1$ is estimated as a linear function of the variables specified in Table 3.1 at time t. Next the SF-12 MCS state is estimated using a linear equation given as

$$\begin{aligned}\log(\text{SF-12\_MCS+10}) = \beta_0 &+ (\text{sex} \times \beta_1) + (\text{age} \times \beta_2) + (\text{ethnicity} \times \beta_3) + (\text{education\_state} \times \beta_4)\\ &+ (\text{labour\_state} \times \beta_5) + (\text{region} \times \beta_6) + (\text{net\_household\_income} \times \beta_7)\\ &+ (\text{NSSEC} \times \beta_8)\end{aligned}$$

The OLS is straightforward to implement using the R (R. Core Team, 2021) 'lm' function. Model fit is assessed using model coefficients provided in Table 3.2. There are a number of significant coefficients that have established association with mental well-being. Being male, higher gross household income, higher education, higher job quality (NS-SEC), and being outside of the London region all have a positive association with SF-12 MCS. Most minority ethnic groups and those in vulnerable labour states such as Sick/Disabled or Maternity Leave indicate lower values of SF-12 MCS. Many attributes such as Government Training and the Other Black (OBL) ethnic group have small sample sizes and large confidence intervals suggesting it is difficult to infer anything meaningful. An adjusted R^2 (Harrell Jr, 2015) score of 0.034 suggests this OLS model is a poor fit for estimation of SF-12 MCS.

Out of sample performance is also assessed using 5-fold cross validation using the R. caret package. The baseline model is assessed using Root Mean Squared Error (RMSE) (Harrell Jr, 2015; McLay et al., 2015) score comparing reserved test data against predictions for each of the 5 folds. The mean and variance of these 5 RMSES scores is $RMSE = 0.223 \pm 0.015$ which is used for comparison against further modelling in this paper.

	Baseline SF_12 OLS
(Intercept)	4.0073 (0.0036)***
factor(sex)Male	0.0353 (0.0013)***
relevel(factor(ethnicity), ref = "WBI")BAN	0.0008 (0.0054)
relevel(factor(ethnicity), ref = "WBI")BLA	0.0319 (0.0046)***
relevel(factor(ethnicity), ref = "WBI")BLC	0.0009 (0.0048)
relevel(factor(ethnicity), ref = "WBI")CHI	0.0141 (0.0093)
relevel(factor(ethnicity), ref = "WBI")IND	0.0113 (0.0037)**
relevel(factor(ethnicity), ref = "WBI")MIX	−0.0155 (0.0047)***
relevel(factor(ethnicity), ref = "WBI")OAS	−0.0006 (0.0054)
relevel(factor(ethnicity), ref = "WBI")OBL	0.0289 (0.0158)
relevel(factor(ethnicity), ref = "WBI")OTH	−0.0150 (0.0105)
relevel(factor(ethnicity), ref = "WBI")PAK	−0.0144 (0.0042)***
relevel(factor(ethnicity), ref = "WBI")WHO	0.0014 (0.0033)
scale(age)	0.0441 (0.0010)***
factor(education_state)1	0.0248 (0.0042)***
factor(education_state)2	0.0128 (0.0018)***
factor(education_state)3	0.0120 (0.0024)***
factor(education_state)5	0.0104 (0.0025)***
factor(education_state)6	0.0204 (0.0022)***
factor(education_state)7	0.0233 (0.0025)***
factor(labour_state)FT Education	0.0629 (0.0037)***
factor(labour_state)FT Employed	0.0178 (0.0033)***
factor(labour_state)Job Seeking	−0.0467 (0.0038)***
factor(labour_state)Not Working	−0.0221 (0.0033)***
factor(labour_state)PT Employed	0.0196 (0.0035)***
relevel(factor(NSSEC), ref = 1)1	0.0182 (0.0043)***
relevel(factor(NSSEC), ref = 1)2	0.0176 (0.0036)***
relevel(factor(NSSEC), ref = 1)3	0.0153 (0.0026)***
relevel(factor(NSSEC), ref = 1)4	0.0203 (0.0029)***
relevel(factor(NSSEC), ref = 1)5	0.0276 (0.0032)***
relevel(factor(NSSEC), ref = 1)6	0.0313 (0.0035)***
relevel(factor(NSSEC), ref = 1)7	0.0202 (0.0026)***
relevel(factor(NSSEC), ref = 1)8	0.0290 (0.0030)***
relevel(factor(region), ref = "London")East Midlands	0.0125 (0.0030)***
relevel(factor(region), ref = "London")East of England	0.0130 (0.0029)***
relevel(factor(region), ref = "London")North East	0.0006 (0.0037)
relevel(factor(region), ref = "London")North West	0.0035 (0.0028)
relevel(factor(region), ref = "London")Northern Ireland	0.0245 (0.0037)***
relevel(factor(region), ref = "London")Scotland	0.0202 (0.0031)***
relevel(factor(region), ref = "London")South East	0.0094 (0.0027)***
relevel(factor(region), ref = "London")South West	0.0161 (0.0030)***
relevel(factor(region), ref = "London")Wales	0.0135 (0.0033)***
relevel(factor(region), ref = "London")West Midlands	0.0008 (0.0029)
relevel(factor(region), ref = "London")Yorkshire and The Humber	0.0053 (0.0030)
scale(hh_income)	0.0094 (0.0006)***
R^2	0.0561
Adj. R^2	0.0556

	Baseline SF_12 OLS
Num. obs.	94908

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.2: Coefficients for the OLS baseline model estimating SF-12 MCS score.

3.4 Presentation of the OLS Model

The baseline OLS model for estimating SF-12 MCS presented in Table 3.2 is a poor fit. This is a white box model, but interpreting the poor fit and formulating next steps for improvement is a challenge. This section demonstrates how changing the presentation of the OLS model can improve interpretability and aid in future development to improve model fit an robustness.

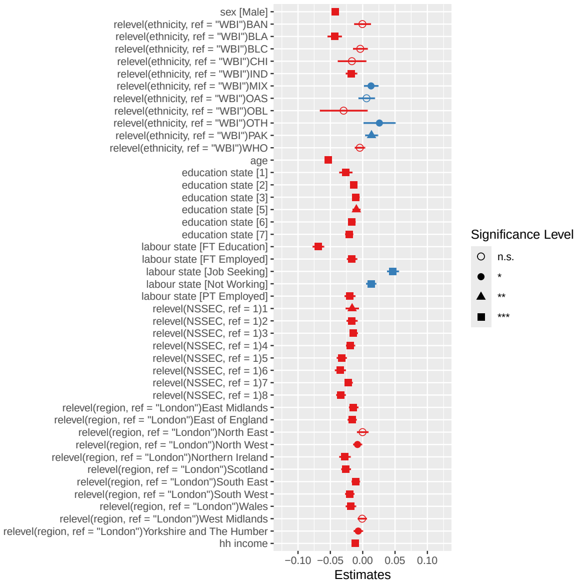
One way to improve interpretability is through better visualisation of the OLS model. Box plots (or forest plots) are increasingly used (Lüdecke and Lüdecke, 2015; Jann, 2014) to display regression coefficients and confidence intervals more clearly. These values are displayed in Figure 3.3a displaying results discussed in Section 3.3 from Table 3.2.

Further goodness of fit testing can be performed using diagnostic plots and tests. For the OLS model it is assumed that errors are homoscedastic (identically and independently distributed) Normal distributed. There are several common diagnostic plots used to test this assumption. Quantile-Quantile (QQ) (Harrell Jr, 2015) plots can be used to determine if a set of observations follows a probability distribution. This plot generates a straight guide line and a plot of all observed points. Ideally the plot of observed points sits exactly on the guideline indicating a perfect fit. The QQ-plot for the baseline model with a Normal guide line is given in Figure 3.3c. It is clear observations closely follow the line at mean 0 but quickly fall away with both tails falling below the guide line. This is a commonly observed pattern that indicates residuals are highly left skewed (Harrell Jr, 2015). Observations are not Normal distributed and more likely follow a Gamma distribution (Harrell Jr, 2015). It also suggests the baseline model is underdispersed and underestimates overall variance in SF-12 MCS values. We can conclude that more information is needed to better estimate variance. This is further supported by a comparison between the predicted and true distribution densities for SF-12 MCS provided in Figure 3.4.

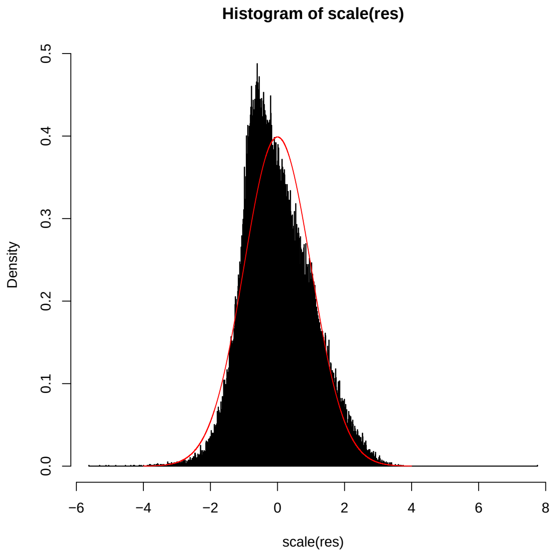
To test homoscedasticity a scale-location plot (Harrell Jr, 2015) is used. This is a scatter plot comparing the square root of residual error terms against fitted values of SF-12 MCS. A red line is also provided indicating the mean residual error term value as fitted values increase. Ideally this red line is horizontal indicating the variance of error terms is constant. Any linear trends suggest the variance changes and homoscedasticity fails (Harrell Jr, 2015). Figure 3.3d shows the scale-location plot for the baseline model. There is a clear negative linear trend (Harrell Jr, 2015) suggesting higher SF-12 MCS values have lower error variance. There is clear heterogeneity in the baseline model discussed further in Section 3.7. The scale-location plot is also useful for identifying outlier observations in the baseline model. There are several points on the right of the graph with very high fitted values. These observations were identified as having extremely high individual household income of over £$500,000$ per month. The OLS model assumes household income linearly increases SF-12 MCS score, producing high predictions (> 60). Handling these outliers can be done in a number of ways. One way is simply to remove or truncate household income values above a certain amount. Another is to assume the baseline model is misspecified. A higher order household income term or piece-wise model can reduced the effect of income on well-being for very large values.

Another approach is to visualise the application of this transition probability model in a larger microsimulation. This is done using internal validation and nowcasting (O'Donoghue and Loughrey, 2014; Archer et al., 2021). The microsimulation is run from the past until the present allowing for comparison with real data. In this case the microsimulation has been run from 2013 when this the baseline OLS model was calibrated until 2021 where the latest United Kingdom Household Longitudinal Survey data is available. The distribution of predicted and real SF-12 MCS values can then be compared to determine goodness of fit over time which is rarely needed outside of microsimulation. Distributions are compared using both boxplots and ridgeline plots (Thrun et al., 2020) in Figure 3.4. It is clear that the OLS model displays severe under dispersion such that the predicted variance in SF-12 MCS is much less than the true value. Fixed effects regression models such as OLS will by definition underestimate the variance of SF-12 MCS (Harrell Jr, 2015) and this problem is compounded by a small choice of variates. In practice this transition probability model used in a microsimulation would estimate mean SF-12 MCS well but increasingly underestimates variance over time. This 'regression to the mean' will substantially effect output for any counterfactual scenario exploring what happens to individuals with extreme SF-12 MCS values and emphasises the need for model selection that preserves statistical moments over time.

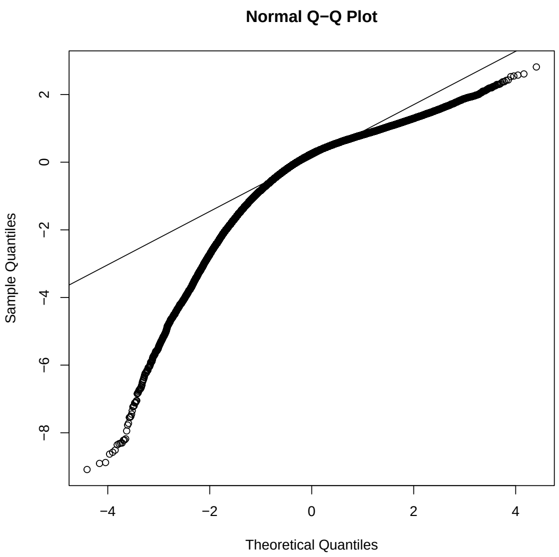
This further visualisation and diagnostics of the baseline model suggests three areas for improvement. First, the data are clearly not Normal distributed. Data transforms such as box-cox and log transforms could adjust the distribution to be approximately Normal. Alternatively use of non-Normal models such as Gamma models may be more appropriate. Second, there is insufficient information to estimate the variance of SF-12 MCS values. Most SF-12 MCS predicted values are severely underestimated by over 10 points. Adding more variables would be a simple next step but is not done here because it is desirable to use as small a set of variables as possible, and any variable included will also need a transition model, exponentially increasing overall microsimulation complexity. Third, the baseline model is not homoscedastic. Relaxing this assumption for the OLS model can be done in a number of ways explored further in Section 3.7.



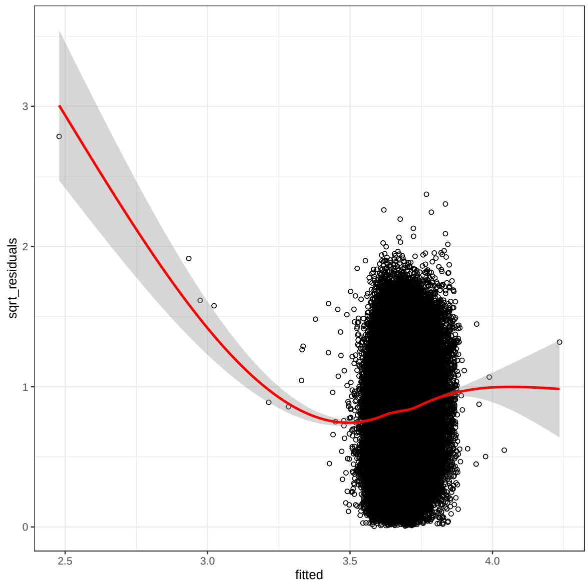
(a) A forest plot of coefficients for the baseline OLS model. These coefficients correspond with the OLS column in Table 3.2.



(b) The distribution of residuals for the baseline OLS model compared against a normal distribution in red.

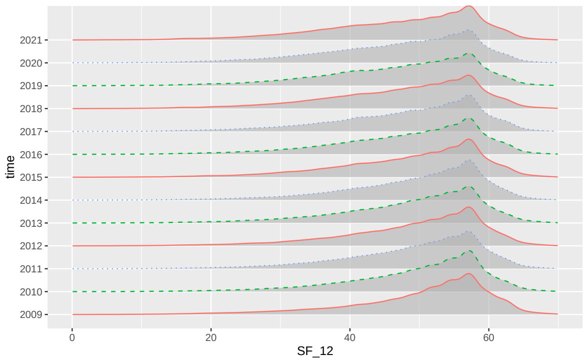


(c) A Quantile-Quantile (QQ) plot for the OLS baseline model.

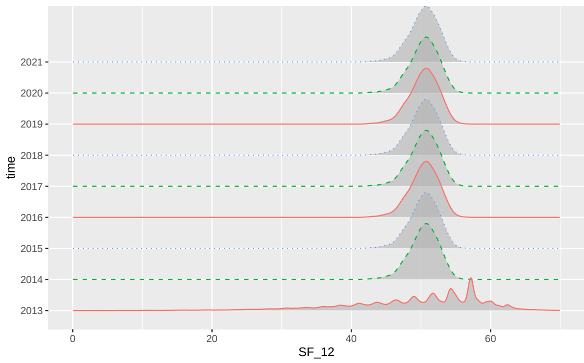


(d) A Scale-Location plot using for the baseline OLS model plotting fitted values against the square root of absolute residual errors. The red line indicates the mean residual value over fitted values of SF-12 MCS.

Figure 3.3: Diagnostic plots for baseline OLS Model.



(a) Ridgeline plots showing the distribution of real SF-12 MCS scores from 2013 − 2021 for UKHLS data.



(b) Ridgeline plots for the predicted density of SF-12 MCS scores from 2013 − 2021 using the baseline OLS model. Data for the year 2013 are real data and all subsequent years are predictions.

Figure 3.4: Ridgeline plots for internal validation of the baseline OLS model against 2013 − 2021 real data.

3.5 Multiple Imputation Chained Equations for Missing UKHLS Data

There is a substantial amount of missing data in the Understanding Society datasets. This initial OLS model uses the complete case analysis assumption such that all observations with missing values are simply omitted. This results in over 60% of observations being removed and variables, for example loneliness cannot be used as they are only available from 2018 onwards. In order to include more data in transition probability models the full UKHLS dataset $2009-2021$ is utilised to perform multiple imputation to include the loneliness variable as well as using more observations to reduce bias.

The missing data structure for UKHLS data is provided in Figure 3.5a. The left part of the plot indicates the proportion of each variable that is missing. For example, there are approximately 16% of individuals missing education state and SF-12 MCS values to be imputed. The right hand portion of the plot indicates the frequency of combinations of missing values. Individuals missing nutrition, loneliness, neighbourhood safety, and ncigs are the most common due to questions not being asked in certain waves. Almost all other missing entries contain some combination of missing SF-12 MCS, education state and labour state. Household income and NSSEC missing values are small ($< 1\%$) each. This missingness structure highlights large missingness in some variables including loneliness, number of cigarettes (ncigs), and neighbourhood safety as these variables are only asked in some waves of UKHLS data. In waves where these data are present they have very limited missing data ($< 5\%$) in which they are suitable for multiple imputation (Jakobsen et al., 2017) techniques.

The Multiple Imputation Chained Equations (MICE) algorithm is implemented following a well established procedure (Jakobsen et al., 2017; Sterne et al., 2009; Davis and Lay-Yee, 2019) in two stages. There are a number of mathematical assumptions when implementing the MICE algorithm that must be considered (Jakobsen et al., 2017). The first is to define some set of auxiliary imputation variables. This is typically a larger dataset than the one used in the transition probability model. These variables are chosen because of high correlation with SF-12 MCS as well as having a large number of observations complete for both the auxiliary value and SF-12 MCS. Without these conditions prediction using further variables in MICE is redundant (Sterne et al., 2009; Jakobsen et al., 2017; Davis and Lay-Yee, 2019). One example is the ghq_depression variable. It has a strong negative correlation with SF-12 MCS of -0.68 and over 86% of observations missing SF-12 MCS have ghq_depression present. Identification of these variables is performed using visualisation techniques from R's 'ggmice' (Josse et al., 2024) and 'VIM' (Prantner, 2011) packages. Correlation and spine plots are used to identify variables that strongly correlate with missing values with spine plots and logistic regression testing to determine if missing values influences distributions providing evidence for inclusion in MICE imputation (Nakagawa and Schielzeth, 2013). All additional variables used in MICE imputation are presented in Table 3.4.

Two further parameters are required to run the MICE algorithm. The number of runs of the MICE algorithm defines how many complete imputed populations are generated. It is recommended to generate a large number $N_{mice} = 30$ of imputed populations due to some variables having a large percentage of missingness (Sterne et al., 2009; Jakobsen et al., 2017) to be able to provide a sufficient sample size. For each imputed population the MICE algorithm is run for some number of iterations. On each iteration a series of linear regressions are fitted estimating missing data. The minimum possible number of iterations should be used ensuring that the mean and variance of imputed variables across all 30 populations suitably converges but preventing overfitting and underestimation of sample variance (Sterne et al., 2009). The R. package 'mice' has an early stopping functionality built in preventing unnecessary further iterations with at least 5 iterations used for each imputed population. MICE imputation is performed using all available UKHLS data from $2011-2020$ giving $n = 500780$ total observations with and without missing data.

Variable Name (UKHLS name)	Description	Units
fridge_freezer (cduse5)	Has a fridge freezer	Binary (0/1)
washing_machine (cduse6)	Has a washing machine	Binary (0/1)
tumble_dryer (cduse7)	Has a tumble dryer	Binary (0/1)
heating (hheat)	Has comfortable heating	Binary (0/1)
dishwasher (cduse8)	Has a dishwasher	Binary (0/1)
microwave (cduse9)	Has a microwave	Binary (0/1)
car_crime (crcar)	Car crime nearby	Ordinal $(1-4)$
drunks (crdrnk)	Drunken disorder nearby	Ordinal $(1-4)$
muggings (crmugg)	Muggings nearby	Ordinal $(1-4)$
racial_abuse (crrace)	Racial Abuse nearby	Ordinal $(1-4)$
teenagers (crteen)	Teenage disorder nearby	Ordinal $(1-4)$
vandalism (crvand)	Vandalism nearby	Ordinal $(1-4)$
fruit_day (wkfruit)	Days per week fruit is eaten	Counts (¿=0)
veg_day (wkvege)	Days per week vegetables are eaten	Counts (¿=0)
fruit_per_day (fruitamt)	Fruit eaten per day	Counts (¿=0)
veg_per_day (vegeamt)	Vegetables eaten per day	Counts (¿=0)

Table 3.3: Variables used to derive the housing quality, nutrition quality, and neighbourhood safety composite variables.

Trajectories for the SF-12 MCS mean and variance for all 30 algorithms are shown in Figure 3.5b. There is convergence of all trajectories to an approximate mean of 49.2 and standard deviation of 10.3 within 10 MICE iterations. These values are close to the true values of 50 and 10 suggesting no further imputation is needed. These 30 imputed populations are each fitted with the baseline model and pooled together into one OLS model (Rubin, 2004) with coefficients in Table 3.5 and Figure 3.5c. Compared with the baseline model model coefficients are similar with smaller confidence intervals due to larger sample size. Variables including certain ethnic groups and education states are significant in the MICE population but not in the baseline model. New variables for nutrition, cigarettes, neighbourhood safety and especially loneliness are all significant and influence prediction. The updated formula is then

Variable Name	Description	Units
behind_on_bills (xphsdba)	Are you behind on bills?	Binary Yes/No
financial_situation (finnow)	How would you describe your current financial situation?	1 – 4 Ordinal Scale
future_financial_situation (finfut)	What do you expect your future financial situation will be like?	Binary 1 – 4 Ordinal Scale
likely_move (lkmove)	Are you likely to move house?	Binary Yes/No
ghq_depression (scghqi)	Do you feel depressed?	1 – 4 Ordinal Scale
ghq_happiness (scghql)	Do you feel happy overall?	1 – 4 Ordinal Scale
clinical_depression (hcondn17)	Have you been diagnosed with clinical depression ?	Binary Yes/No
general_health (scsf1)	How would you describe your overall health?	1 – 5 Ordinal Scale
phealth_limits_work (scsf3a)	Does your physical health limit your ability to work?	1 – 4 Ordinal Scale
mhealth_limits_work (scsf4a)	Does your mental health limit your ability to work?	1 – 4 Ordinal Scale

Table 3.4: Auxiliary variables included in MICE imputation of SF-12 MCS.

$$
\begin{aligned}
\log(\text{SF-12\_MCS}) = \beta_0 &+ (\text{sex} \times \beta_1) + (\text{age} \times \beta_2) + (\text{ethnicity} \times \beta_3) \\
&+ (\text{education\_state} \times \beta_4) + (\text{labour\_state} \times \beta_5) + (\text{region} \times \beta_6) \\
&+ (\text{net\_household\_income} \times \beta_7) + (\text{NSSEC} \times \beta_8) + (\text{loneliness} \times \beta_9) \\
&+ (\text{ncigs} \times \beta_{10}) + (\text{nutrition quality} \times \beta_{11})
\end{aligned}
$$

There are a number of sign changes for MICE imputed coefficients. Most notable the effect of higher quality employment becomes more important. There are far more employed individuals in this dataset, giving a better estimate of their well-being. More data unsurprisingly improves the model fit. Inclusion of further derived variables, particularly loneliness, has also improved model fit with large significant effects for loneliness, housing quality, and smoking on SF-12 MCS. An adjusted r-square score of 0.296 shows substantial improvement vs the baseline model suggesting improved goodness of fit but it is misleading to directly compare different datasets (Harrell Jr, 2015). A Quantile-Quantile plot for the MICE pooled OLS model in Figure 3.5d suggests residuals for the MICE model more closely fit to a Normal distribution than the baseline model with a long left tail suggesting the OLS model overestimates SF-12 MCS for low valued individuals. This is unsurprising as MICE is known to force variables to follow a Normal distribution and application of further augmentations of the MICE algorithm for non-Normal continuous variables are needed (Huque et al., 2018). This also provides further evidence that SF-12 MCS is non-Normal distributed that the OLS model cannot sufficiently estimate.

Five-fold cross validation is also performed on each imputed MICE population independently. Each of the 30 imputed populations is split into 5 folds fitting the OLS model 5 times and producing 5 RMSE scores. This produces an overall $30 \times 5 = 150$ RMSE scores giving mean 0.164 ± 0.008. This is a clear improvement from the baseline model but it is difficult to perform direct comparison due to different datasets. Note that each randomly generated fold groups repeated observations from each individual within each fold. That is all observations with the same pidp identifier are in the same fold. Longitudinal modelling later in this paper requires repeat observations are not in both training and test sets.

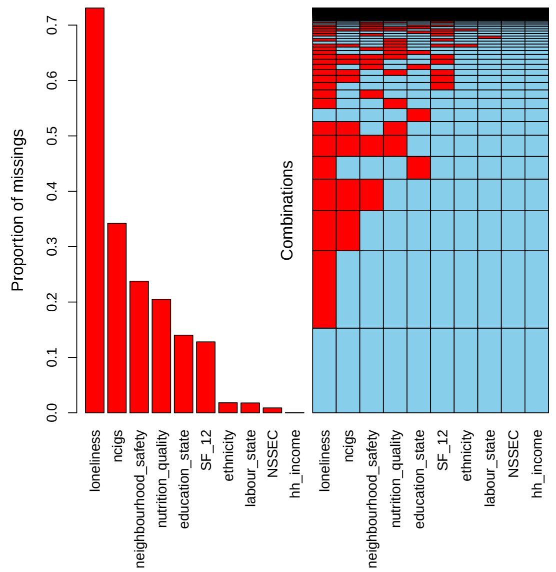
Validation boxplots and ridgeline plots for the MICE OLS model are also provided in Figure 3.6. It is clear the the MICE OLS model better estimates variance in SF-12 MCS compared to the baseline model. It is clear that the loneliness variable has a strong effect on prediction of SF-12 MCS resulting in three distinct peaks in the distribution. This can result in large prediction errors when a person is lonely but does not necessarily see a large drop in well-being. As sample variance is also still underestimated, inclusion of individual SF-12 MCS histories in future sections aims to reduce over-sensitivity to loneliness and better predict low SF-12 MCS values.

	MICE SF_12 OLS Pool
(Intercept)	4.0948 (0.0039)***
factor(sex)Male	0.0234 (0.0011)***
relevel(factor(ethnicity), ref = "WBI")BAN	−0.0034 (0.0047)
relevel(factor(ethnicity), ref = "WBI")BLA	0.0319 (0.0045)***
relevel(factor(ethnicity), ref = "WBI")BLC	0.0140 (0.0045)**
relevel(factor(ethnicity), ref = "WBI")CHI	0.0098 (0.0086)
relevel(factor(ethnicity), ref = "WBI")IND	0.0108 (0.0034)**
relevel(factor(ethnicity), ref = "WBI")MIX	−0.0049 (0.0044)
relevel(factor(ethnicity), ref = "WBI")OAS	0.0092 (0.0049)
relevel(factor(ethnicity), ref = "WBI")OBL	0.0286 (0.0148)
relevel(factor(ethnicity), ref = "WBI")OTH	−0.0073 (0.0092)
relevel(factor(ethnicity), ref = "WBI")PAK	−0.0112 (0.0038)**
relevel(factor(ethnicity), ref = "WBI")WHO	0.0033 (0.0031)
scale(age)	0.0270 (0.0008)***
factor(education_state)1	0.0091 (0.0037)*
factor(education_state)2	0.0035 (0.0016)*
factor(education_state)3	0.0001 (0.0021)
factor(education_state)5	−0.0027 (0.0022)
factor(education_state)6	−0.0002 (0.0020)
factor(education_state)7	0.0008 (0.0023)
factor(labour_state)FT Education	0.0416 (0.0032)***
factor(labour_state)FT Employed	0.0073 (0.0025)**
factor(labour_state)Job Seeking	−0.0180 (0.0036)***
factor(labour_state)Not Working	−0.0121 (0.0027)***
factor(labour_state)PT Employed	0.0110 (0.0027)***
relevel(factor(NSSEC), ref = 1)1	0.0043 (0.0034)
relevel(factor(NSSEC), ref = 1)2	0.0042 (0.0028)
relevel(factor(NSSEC), ref = 1)3	0.0048 (0.0021)*
relevel(factor(NSSEC), ref = 1)4	0.0084 (0.0023)***
relevel(factor(NSSEC), ref = 1)5	0.0153 (0.0025)***
relevel(factor(NSSEC), ref = 1)6	0.0191 (0.0029)***
relevel(factor(NSSEC), ref = 1)7	0.0138 (0.0020)***
relevel(factor(NSSEC), ref = 1)8	0.0204 (0.0026)***
relevel(factor(region), ref = "London")East Midlands	0.0061 (0.0028)*
relevel(factor(region), ref = "London")East of England	0.0071 (0.0025)**
relevel(factor(region), ref = "London")North East	−0.0022 (0.0034)
relevel(factor(region), ref = "London")North West	−0.0032 (0.0025)
relevel(factor(region), ref = "London")Northern Ireland	0.0067 (0.0027)*
relevel(factor(region), ref = "London")Scotland	0.0063 (0.0024)**
relevel(factor(region), ref = "London")South East	0.0039 (0.0023)
relevel(factor(region), ref = "London")South West	0.0068 (0.0027)*
relevel(factor(region), ref = "London")Wales	−0.0028 (0.0027)
relevel(factor(region), ref = "London")West Midlands	−0.0039 (0.0024)
relevel(factor(region), ref = "London")Yorkshire and The Humber	0.0005 (0.0025)

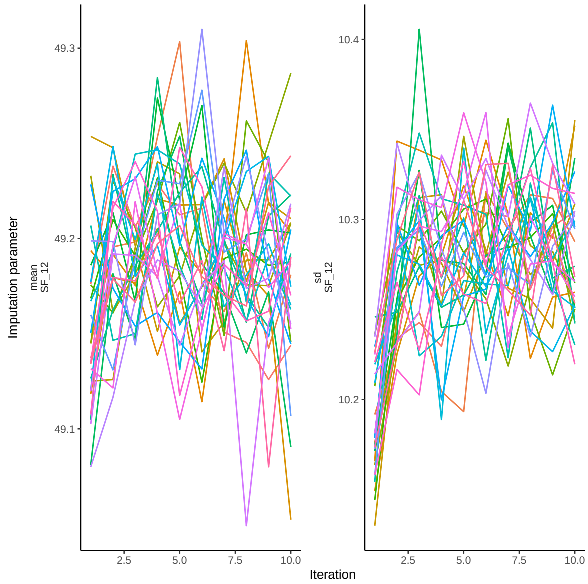
	MICE SF_12 OLS Pool
scale(hh_income)	0.0061 (0.0007)***
scale(hh_income^2)	−0.0043 (0.0006)***
factor(housing_quality)Low	−0.0205 (0.0018)***
factor(housing_quality)Medium	−0.0045 (0.0011)***
factor(neighbourhood_safety)2	0.0153 (0.0012)***
factor(neighbourhood_safety)3	0.0188 (0.0016)***
factor(loneliness)2	−0.1292 (0.0015)***
factor(loneliness)3	−0.3173 (0.0027)***
ncigs	−0.0016 (0.0001)***
nutrition_quality	0.0003 (0.0001)***
nimp	30
nobs	139825
R^2	0.2916
Adj. R^2	0.2914

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

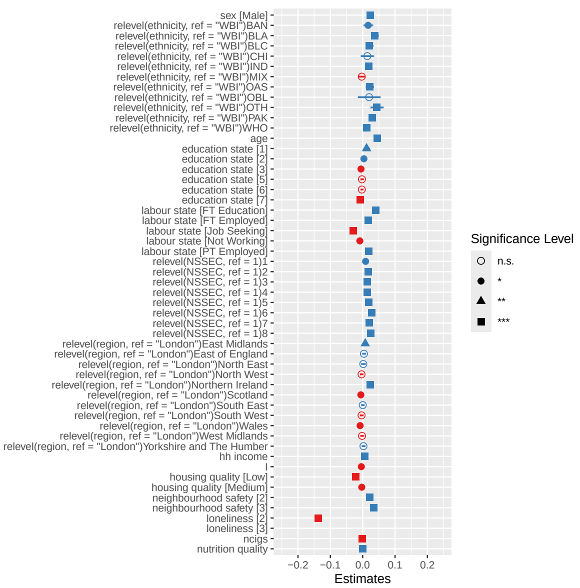
Table 3.5: OLS baseline model coefficients estimating SF-12 MCS score with MICE imputated UKHLS data.



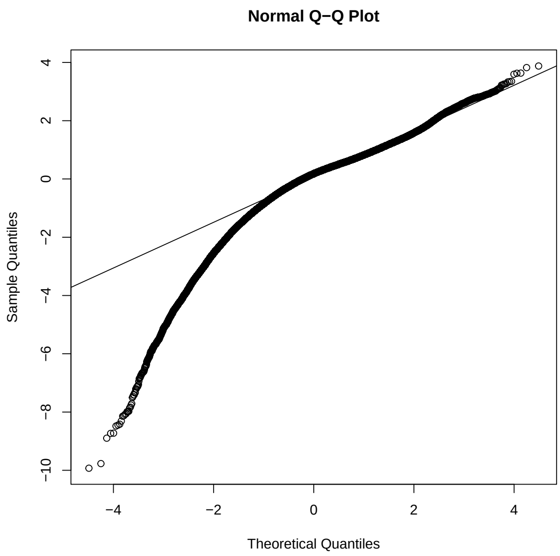
(a) Total missingness structure for UKHLS data between 2011 and 2020.



(b) Trajectories of SF-12 MCS mean and variance for 30 imputed population each across 10 MICE iterations.



(c) A forest plot of OLS coefficients using MICE imputed data. Corresponding with the MICE OLS column in Table 3.5.



(d) A Quantile-Quantile plot for the pooled OLS model using MICE imputed data.

Figure 3.5: Diagnostic Plots for OLS model using MICE imputed data.

3.6 Validation and Uncertainty

Simple validation has been provided for the baseline OLS model using handover plots estimating how the OLS model predicts behaviours going forwards in time from 2013 to 2021. Further validation of this model then requires assessment of out of sample performance and quantification of uncertainty with respect to input data and model structure selection using sensitivity analysis and variable selection methods (Petrik et al., 2020; O'Hagan et al., 2007; Burgard and Schmaus, 2019).

3.6.1 Uncertainty due to Calibration Data and Out of Sample Performance

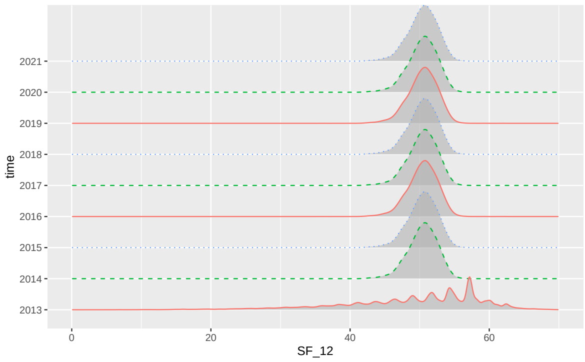
Unweighted sample data used in calibration are not representative of the full UK population in either the observed 2011 − 2013 data (Benzeval et al., 2020) or at future time points. Further validation is required to assess if the sample population is representative of transition dynamics in the full UK population as well as testing out of sample performance using cross validation.

Ensembles of Models

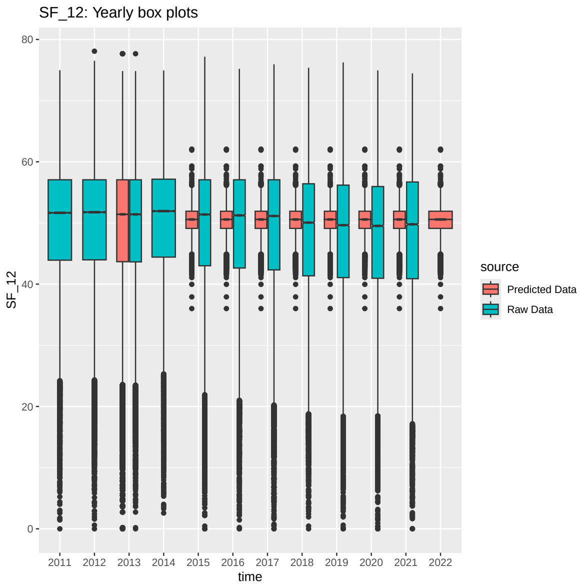
In Section 3.5 application of the MICE algorithm resulted in 30 imputed UKHLS populations each fitted with an OLS model to predict SF-12 MCS and pooled together. To further test uncertainty in UKHLS data due to outliers and out of sample performance for the baseline model, five fold cross validation is used (James et al., 2013) on each imputed population. R's 'caret' package (Kuhn, 2008) is used to split UKHLS data into five pieces. Cross validation is run 5 times each time reserving one piece of data for testing, and fitting the OLS model to all 4 remaining pieces. Each piece of data is omitted sequentially determining if it contains outliers that perturb model coefficients (James et al., 2013). Cross validated model coefficients are then pooled together using Rubin's rules as seen in MICE procedure (Rubin, 2004). Model coefficients and ridgeline validation plots are provided for the cross validated OLS model in Table 3.6 and Figure 3.6. Overall there is little difference shown between the MICE OLS model and cross validated models. This suggests any outliers in the UKHLS population do not effect model coefficients. This is unsurprising here given the UKHLS sample size is very large, but the procedure is informative would be extremely useful for transition models built on smaller datasets.

	5-Fold Cross Validated SF_12 OLS
(Intercept)	$4.0924\ (0.0029)^{***}$
'factor(sex)Male'	$0.0233\ (0.0010)^{***}$
'relevel(factor(ethnicity), ref = "WBI")BAN'	$-0.0025\ (0.0035)$
'relevel(factor(ethnicity), ref = "WBI")BLA'	$0.0364\ (0.0033)^{***}$
'relevel(factor(ethnicity), ref = "WBI")BLC'	$0.0142\ (0.0035)^{***}$
'relevel(factor(ethnicity), ref = "WBI")CHI'	$0.0148\ (0.0070)^{*}$
'relevel(factor(ethnicity), ref = "WBI")IND'	$0.0121\ (0.0026)^{***}$

	5-Fold Cross Validated SF_12 OLS
‘relevel(factor(ethnicity), ref = ”WBI”)MIX‘	−0.0068 (0.0036)
‘relevel(factor(ethnicity), ref = ”WBI”)OAS‘	0.0146 (0.0039)***
‘relevel(factor(ethnicity), ref = ”WBI”)OBL‘	0.0254 (0.0105)*
‘relevel(factor(ethnicity), ref = ”WBI”)OTH‘	0.0007 (0.0069)
‘relevel(factor(ethnicity), ref = ”WBI”)PAK‘	−0.0080 (0.0029)**
‘relevel(factor(ethnicity), ref = ”WBI”)WHO‘	0.0032 (0.0022)
‘scale(age)‘	0.0268 (0.0007)***
‘factor(education_state)1‘	0.0076 (0.0029)**
‘factor(education_state)2‘	0.0044 (0.0013)***
‘factor(education_state)3‘	0.0001 (0.0017)
‘factor(education_state)5‘	−0.0025 (0.0018)
‘factor(education_state)6‘	−0.0001 (0.0016)
‘factor(education_state)7‘	0.0000 (0.0018)
‘factor(labour_state)FT Education‘	0.0409 (0.0027)***
‘factor(labour_state)FT Employed‘	0.0070 (0.0023)**
‘factor(labour_state)Job Seeking‘	−0.0182 (0.0028)***
‘factor(labour_state)Not Working‘	−0.0114 (0.0023)***
‘factor(labour_state)PT Employed‘	0.0101 (0.0025)***
‘relevel(factor(NSSEC), ref = 1)1‘	0.0055 (0.0030)
‘relevel(factor(NSSEC), ref = 1)2‘	0.0058 (0.0025)*
‘relevel(factor(NSSEC), ref = 1)3‘	0.0043 (0.0018)*
‘relevel(factor(NSSEC), ref = 1)4‘	0.0070 (0.0020)***
‘relevel(factor(NSSEC), ref = 1)5‘	0.0152 (0.0021)***
‘relevel(factor(NSSEC), ref = 1)6‘	0.0193 (0.0024)***
‘relevel(factor(NSSEC), ref = 1)7‘	0.0126 (0.0018)***
‘relevel(factor(NSSEC), ref = 1)8‘	0.0217 (0.0020)***
‘relevel(factor(region), ref = ”London”)East Midlands‘	0.0083 (0.0022)***
‘relevel(factor(region), ref = ”London”)East of England‘	0.0079 (0.0021)***
‘relevel(factor(region), ref = ”London”)North East‘	−0.0030 (0.0028)
‘relevel(factor(region), ref = ”London”)North West‘	−0.0026 (0.0021)
‘relevel(factor(region), ref = ”London”)Northern Ireland‘	0.0082 (0.0024)***
‘relevel(factor(region), ref = ”London”)Scotland‘	0.0076 (0.0022)***
‘relevel(factor(region), ref = ”London”)South East‘	0.0045 (0.0020)*
‘relevel(factor(region), ref = ”London”)South West‘	0.0079 (0.0022)***
‘relevel(factor(region), ref = ”London”)Wales‘	−0.0006 (0.0022)
‘relevel(factor(region), ref = ”London”)West Midlands‘	−0.0031 (0.0021)
‘relevel(factor(region), ref = ”London”)Yorkshire and The Humber‘	0.0016 (0.0022)
‘scale(hh_income)‘	0.0064 (0.0006)***
‘scale(hh_income^2)‘	−0.0045 (0.0006)***
‘factor(housing_quality)Low‘	−0.0200 (0.0014)***
‘factor(housing_quality)Medium‘	−0.0045 (0.0010)***
‘factor(neighbourhood_safety)2‘	0.0145 (0.0010)***
‘factor(neighbourhood_safety)3‘	0.0189 (0.0013)***
‘factor(loneliness)2‘	−0.1265 (0.0010)***
‘factor(loneliness)3‘	−0.3183 (0.0017)***



(a) A ridgeline plot demonstrating predicted SF-12 MCS scores by the five-fold cross-validated OLS baseline model from 2013 to 2021.



(b) A boxplot demonstrating the distribution of SF-12 MCS predicted data using five-fold cross validation from 2013 to 2021.

Figure 3.6: Internal five-fold cross-validation nowcasting plots.

	5-Fold Cross Validated SF_12 OLS
ncigs	-0.0016 $(0.0001)^{***}$
nutrition_quality	0.0004 $(0.0001)^{***}$
R^2	0.2892
Adj. R^2	0.2889
Num. obs.	139825
RMSE	0.1642

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.6: OLS coefficients for estimation of SF-12 MCS using five-fold cross validation.

Weighted Regression

Substantial work has been done to ensure that the United Kingdom Household Longitudinal Survey data is representative of the overall UK population (Benzeval et al., 2020). To do this, certain areas of the population that are hard to reach, such as young adults, ethnic minorities, and households in London, are over-sampled (Benzeval et al., 2020) to ensure sufficient sample size for subgroups of interest. The raw sample is then biased when used to estimate transition probabilities for the whole UK population and requires correction using sample weights (Benzeval et al., 2020). These inverse probability weights scale the importance of individual observations according to their frequency in the national population (Benzeval et al., 2020). An individual with a larger weight indicates higher prevalence in the full UK population.

A weighted OLS model (Harrell Jr, 2015) can be implemented using R's 'lm' package (R.

Core Team, 2021). Model coefficients are found in Figure 3.7. Model coefficients are overall very similar in the baseline and weighted OLS models. Some regression coefficients, such as certain ethnic groups including Indian and Bangladeshi, have larger confidence intervals. These aforementioned over-sampled subgroups are correctly down-weighted in the OLS model.

Sample weights can also be treated as another variable to be transitioned forwards in time in dynamic microsimulation (Dekkers and Cumpston, 2012). The intention is to update sample weights and keep the sample population representative into the future. In this case we use simple updating of sample weights according to change in age, sex, and ethnic demographic according to projections from the NEWETHPOP macro-model (Rees et al., 2017), which provides estimates for future population demographic composition by size, age, sex and ethnic group at local authority scale. Sample weight is updated according the cell based reweighting methodology (Tysinger, 2021), whereby individual weights are updated according to change in their demographic group prevalence over time using the formula $w_{t+1} = \frac{w_t p_t}{p_{t-1}}$ given w_t and w_{t+1} are the current and future weights and p_t and $pt + 1$ are the current and future proportion of the population an individual's demographic group represents. If the number of White British males aged $25 - 30$ increases by 2% over 5 years the sampling weight is inflated by 1.02 accordingly.

Weights of the $2011 - 2013$ UKHLS data are projected 5 years forwards in time to $2016 - 2018$. Two further weighted OLS models are fitted (1) using $2016 - 2018$ weights directly and (2) using $2011 - 2013$ data with $2016 - 2018$ projected weights. This allows testing of the assumption that the weighted OLS model is still able to predict the next state when extrapolating five years into the future. Coefficients for both models are provided in Table 3.7. Overall models show similar coefficients and goodness of fit with adjusted R-Squared values of 0.2781 and 0.3025. There are 30 shared variables with statistically significant coefficients between both models. There are some discrepancies particularly for the ethnicity and region variables potentially due to lower sample size in $2011 - 2013$ data.

	SF_12 2018 With 2013 Extrapolated Weights	SF_12 2018 With Weights
(Intercept)	4.0944 (0.0041)***	4.0616 (0.0048)***
factor(sex)Male	0.0272 (0.0012)***	0.0246 (0.0013)***
relevel(factor(ethnicity), ref = "WBI")BAN	−0.0055 (0.0069)	0.0159 (0.0087)
relevel(factor(ethnicity), ref = "WBI")BLA	0.0296 (0.0045)***	0.0342 (0.0062)***
relevel(factor(ethnicity), ref = "WBI")BLC	0.0186 (0.0058)**	0.0174 (0.0071)*
relevel(factor(ethnicity), ref = "WBI")CHI	0.0242 (0.0076)**	0.0017 (0.0117)
relevel(factor(ethnicity), ref = "WBI")IND	0.0138 (0.0036)***	0.0261 (0.0045)***
relevel(factor(ethnicity), ref = "WBI")MIX	−0.0039 (0.0048)	−0.0062 (0.0058)
relevel(factor(ethnicity), ref = "WBI")OAS	0.0142 (0.0054)**	0.0304 (0.0066)***
relevel(factor(ethnicity), ref = "WBI")OBL	0.0367 (0.0147)*	0.0188 (0.0227)
relevel(factor(ethnicity), ref = "WBI")OTH	−0.0109 (0.0082)	0.0273 (0.0111)*
relevel(factor(ethnicity), ref = "WBI")PAK	−0.0179 (0.0046)***	0.0272 (0.0057)***
relevel(factor(ethnicity), ref = "WBI")WHO	0.0108 (0.0029)***	0.0154 (0.0033)***
scale(age)	0.0250 (0.0009)***	0.0399 (0.0010)***
factor(education_state)1	0.0063 (0.0041)	0.0125 (0.0047)**
factor(education_state)2	0.0043 (0.0017)*	0.0040 (0.0019)*
factor(education_state)3	−0.0030 (0.0024)	−0.0006 (0.0026)
factor(education_state)5	−0.0013 (0.0024)	−0.0024 (0.0026)
factor(education_state)6	−0.0004 (0.0021)	0.0037 (0.0022)
factor(education_state)7	0.0003 (0.0025)	−0.0026 (0.0026)
factor(labour_state)FT Education	0.0375 (0.0037)***	0.0499 (0.0045)***
factor(labour_state)FT Employed	0.0061 (0.0031)	0.0177 (0.0036)***
factor(labour_state)Job Seeking	−0.0141 (0.0037)***	−0.0213 (0.0048)***
factor(labour_state)Not Working	−0.0132 (0.0032)***	−0.0111 (0.0038)**
factor(labour_state)PT Employed	0.0102 (0.0034)**	0.0178 (0.0039)***

	SF_12 2018 With 2013 Extrapolated Weights	SF_12 2018 With Weights
relevel(factor(NSSEC), ref = 1)1	0.0045 (0.0038)	0.0092 (0.0039)*
relevel(factor(NSSEC), ref = 1)2	0.0062 (0.0031)	0.0133 (0.0031)***
relevel(factor(NSSEC), ref = 1)3	0.0049 (0.0024)*	0.0115 (0.0023)***
relevel(factor(NSSEC), ref = 1)4	0.0086 (0.0026)**	0.0156 (0.0026)***
relevel(factor(NSSEC), ref = 1)5	0.0165 (0.0029)***	0.0204 (0.0028)***
relevel(factor(NSSEC), ref = 1)6	0.0219 (0.0031)***	0.0203 (0.0033)***
relevel(factor(NSSEC), ref = 1)7	0.0146 (0.0023)***	0.0193 (0.0024)***
relevel(factor(NSSEC), ref = 1)8	0.0234 (0.0027)***	0.0237 (0.0028)***
relevel(factor(region), ref = "London")East Midlands	0.0062 (0.0032)	0.0028 (0.0031)
relevel(factor(region), ref = "London")East of England	0.0070 (0.0026)**	0.0053 (0.0029)
relevel(factor(region), ref = "London")North East	0.0002 (0.0035)	0.0007 (0.0037)
relevel(factor(region), ref = "London")North West	−0.0031 (0.0027)	−0.0044 (0.0027)
relevel(factor(region), ref = "London")Northern Ireland	0.0026 (0.0036)	0.0098 (0.0041)*
relevel(factor(region), ref = "London")Scotland	0.0056 (0.0028)*	−0.0067 (0.0030)*
relevel(factor(region), ref = "London")South East	0.0033 (0.0024)	0.0005 (0.0026)
relevel(factor(region), ref = "London")South West	0.0067 (0.0028)*	−0.0036 (0.0029)
relevel(factor(region), ref = "London")Wales	−0.0022 (0.0033)	−0.0112 (0.0034)**
relevel(factor(region), ref = "London")West Midlands	−0.0042 (0.0028)	−0.0060 (0.0029)*
relevel(factor(region), ref = "London")Yorkshire and The Humber	0.0021 (0.0029)	−0.0007 (0.0028)
scale(hh_income)	0.0058 (0.0008)***	0.0036 (0.0006)***
scale(hh_income^2)	−0.0041 (0.0006)***	0.0004 (0.0006)
factor(housing_quality)Low	−0.0263 (0.0019)***	−0.0435 (0.0023)***
factor(housing_quality)Medium	−0.0053 (0.0013)***	−0.0050 (0.0013)***
factor(neighbourhood_safety)2	0.0145 (0.0013)***	0.0211 (0.0015)***
factor(neighbourhood_safety)3	0.0188 (0.0017)***	0.0265 (0.0018)***
factor(loneliness)2	−0.1277 (0.0015)***	−0.1340 (0.0016)***
factor(loneliness)3	−0.3130 (0.0043)***	−0.3279 (0.0031)***
ncigs	−0.0016 (0.0001)***	−0.0014 (0.0001)***
nutrition_quality	0.0003 (0.0001)***	0.0005 (0.0001)***
nimp	30	30
nobs	115852	98605
R^2	0.2905	0.3018
Adj. R^2	0.2902	0.3014

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.7: OLS SF-12 MCS model coefficients using 2018 data comparing differences between real and projected weights.

3.6.2 Uncertainty due to Model Coefficients

OLS model coefficients themselves are random variables whose mean values are used in prediction of next state. Transition models can be highly sensitive to input parameters such that different samples of model coefficients can produce substantially different outputs (Petrik et al., 2020; O'Hagan et al., 2007). This section explores how varying model coefficients changes prediction.

In the above ensembles of OLS models all regression coefficients are assumed to take their mean values (Rubin, 2004). For example, in Table 3.2 the pooled OLS model fitted to MICE data suggests being male (vs female) improves SF-12 MCS by a score of +1.1941. However this coefficient is a random multivariate Normal variable (Harrell Jr, 2015). When OLS models are used in Monte Carlo simulation studies it is common to treat regression coefficients as input parameters drawing random values for each run (O'Hagan et al., 2007; Petrik et al., 2020; Creedy et al., 2007). Sensitivity analysis typically varies one parameter at a time but this can result in a very large number of model runs. A faster approach can be used (O'Hagan et al., 2007) sampling from the p-dimensional multivariate Normal distribution (Harrell Jr, 2015) with parameters derived from the pooled MICE OLS model. A sample of 100 OLS models is created from the pooled MICE model with randomly generated coefficients. These 100 models are each used to perform internal validation testing model fit from 2013 − 2021 randomly chooses one

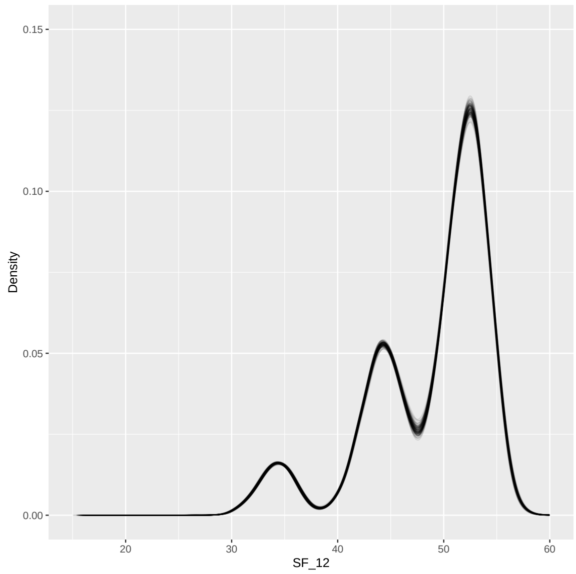


Figure 3.7: Density plots for predicted SF-12 MCS state in 2021 using estimated using 2013 data. A total 100 OLS models are fitted with random coefficients to determine the sensitivity of prediction to coefficient values.

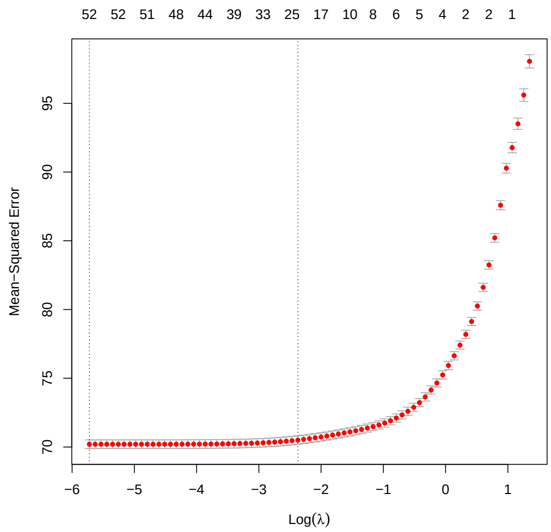


Figure 3.8: Results for the 10-fold cross validated LASSO algorithm applied to the baseline OLS model. The left and right vertical dashed lines indicate the choice of coefficients for the minimum overall error and first standard error rule subsets.

Figure 3.9: Figures for validation testing the sensitivity of the baseline model to model structure and coefficient values.

of the available 30 imputed populations and draws new random model coefficients. The SF-12 MCS densities for all 100 model runs are plotted in Figure 3.7. There is limited difference in the distribution of SF-12 MCS score across all 100 runs suggesting low uncertainty in OLS parameter values.

3.6.3 Uncertainty due to Model Structure Using the LASSO

The set of predictor variables used in the baseline model may be misspecified. Inclusion of statistically insignificant variables will result in a model that is not parsimonious. Omitting these variables would then produce a simpler and more interpretable model. Conversely, if significant variables are not included omitted variable bias would then produce a poorer fitting model. Variable selection (Adhikari et al., 2019) methods provide one approach to screen from a pool of available predictor variables that can be used to predict SF-12 MCS ensuring a parsimonious model is selected.

To validate the choice of coefficients in the OLS model the LASSO variable selection algorithm is used (James et al., 2013) to choose the most parsimonious subset of predictors from a pool of all available variables from Table 3.1. The LASSO is implemented using the ‘glmnet’ package in R. with recommended 10-fold cross-validation (James et al., 2013). The two ‘best’ subsets of coefficients are provided giving model choices with the minimum overall root mean squared error and the most parsimonious subset according to the first standard error (1se) rule (James et al., 2013).

Results are shown in Figure 3.8. Two vertical dashed lines indicate the chosen subset of variables for the minimum error and 1se subsets. The minimum error subset contains all 50 available total variables. This suggests the model with all available variables provides the best performing OLS model and that all variables should be kept. However, according to the 'one standard error (1se)' rule (James et al., 2013) the most parsimonious model only includes 24 out of 50 variables suggesting 30 variables described in Table 3.8 do not significantly improve model fit and should be omitted. Most factor levels for ethnicity, region, and NSSEC codes are suggested for removal. Use of the LASSO suggests all included variables improve prediction of SF-12 MCS and further variable selection with more variables and longitudinal techniques such as the fused LASSO (Adhikari et al., 2019) can assist in validation of model structure and parsimony particularly in scenarios with limited data resources. Further sensitivity analysis of model structure can be used comparing multiple different modelling techniques (McLay et al., 2015) further ensuring optimal model selection.

	Minimum Error	First Standard Error
(Intercept) (1)	50.61	51.43
(Intercept) (2)	0.00	0.00
factor(sex)Male	1.49	1.38
relevel(factor(ethnicity), ref = "WBI")BAN	−0.13	0.00
relevel(factor(ethnicity), ref = "WBI")BLA	1.89	0.72
relevel(factor(ethnicity), ref = "WBI")BLC	0.60	0.00
relevel(factor(ethnicity), ref = "WBI")CHI	0.95	0.00
relevel(factor(ethnicity), ref = "WBI")IND	0.65	0.00
relevel(factor(ethnicity), ref = "WBI")MIX	−0.03	0.00
relevel(factor(ethnicity), ref = "WBI")OAS	0.67	0.00
relevel(factor(ethnicity), ref = "WBI")OBL	2.20	0.00
relevel(factor(ethnicity), ref = "WBI")OTH	0.17	0.00
relevel(factor(ethnicity), ref = "WBI")PAK	−0.63	−0.27
relevel(factor(ethnicity), ref = "WBI")WHO	0.53	0.00
scale(age)	1.34	1.01
factor(education_state)1	0.44	0.00
factor(education_state)2	0.30	0.07
factor(education_state)3	−0.09	0.00
factor(education_state)5	−0.01	0.00
factor(education_state)6	−0.06	0.00
factor(education_state)7	−0.08	0.00
factor(labour_state)FT Education	1.94	0.86
factor(labour_state)FT Employed	0.26	0.00
factor(labour_state)Job Seeking	−0.55	−0.81
factor(labour_state)Not Working	−0.36	−0.20
factor(labour_state)PT Employed	0.52	0.17
relevel(factor(NSSEC), ref = 1)1	0.10	0.00
relevel(factor(NSSEC), ref = 1)2	0.20	0.00
relevel(factor(NSSEC), ref = 1)3	0.11	0.00
relevel(factor(NSSEC), ref = 1)4	0.24	0.00
relevel(factor(NSSEC), ref = 1)5	0.84	0.38
relevel(factor(NSSEC), ref = 1)6	1.11	0.48
relevel(factor(NSSEC), ref = 1)7	0.64	0.13
relevel(factor(NSSEC), ref = 1)8	1.18	0.59
relevel(factor(region), ref = "London")East Midlands	0.35	0.00
relevel(factor(region), ref = "London")East of England	0.40	0.00
relevel(factor(region), ref = "London")North East	−0.11	0.00
relevel(factor(region), ref = "London")North West	−0.02	−0.00
relevel(factor(region), ref = "London")Northern Ireland	0.20	0.00
relevel(factor(region), ref = "London")Scotland	0.45	0.00
relevel(factor(region), ref = "London")South East	0.22	0.00
relevel(factor(region), ref = "London")South West	0.38	0.00
relevel(factor(region), ref = "London")Wales	0.00	0.00
relevel(factor(region), ref = "London")West Midlands	−0.08	0.00
relevel(factor(region), ref = "London")Yorkshire and The Humber	0.20	0.00
scale(hh_income)	0.33	0.13
scale(hh_income^2)	−0.22	−0.03
factor(housing_quality)Low	−1.30	−0.99

	Minimum Error	First Standard Error
factor(housing_quality)Medium	-0.28	-0.05
factor(neighbourhood_safety)2	0.74	0.52
factor(neighbourhood_safety)3	1.05	0.76
factor(loneliness)2	-6.74	-6.58
factor(loneliness)3	-15.49	-15.34
ncigs	-0.09	-0.08
nutrition_quality	0.02	0.01

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.8: Coefficients for the SF-12 MCS OLS baseline model selected using the LASSO first standard error and minimum error rules.

3.7 Application of GLMMs to Baseline SF-12 MCS Model.

Diagnostics for the baseline model in Section 3.4 suggests heterogeneity, that is residual errors are not identically and independently Normal distributed, for UKHLS data. There are three suggested reasons for this. First, observations are not Normal distributed, there is a clear right skew in SF-12 MCS scores suggesting that residuals are not Normal distributed. Second, observations are not independent of each other. There is correlation between repeat observations within individuals and households over time. Third, there is higher volatility in individuals with higher fitted SF-12 MCS values, as denoted by a positive slope in the scale-location plot in Figure 3.3. These three reasons suggests that the homoscedastisity assumption is violated and further modelling techniques are needed to relax this assumption. This section introduces generalised linear mixed models to address these three sources of heterogeneity combined with missing data, presentation, and validation techniques from previous sections in order to produce a final transition model for estimating SF-12 MCS state to be used in future dynamic microsimulation models.

Generalised linear mixed modelling (GLMM) is a technique used to include both mixed fixed and random effects as well as non-Normal response variables (Bates, 2014). Reflecting SF-12 MCS and its maximum value produces a variable that is non-negative and has a right skew that can be approximated by a Gamma distribution (Harrell Jr, 2015). Since individual identifiers are provided in UKHLS (”pidp”) that assign observations to each individual random intercept terms can also be used (Bates, 2014). Intercept values in the OLS model are assumed to be different for each individual with these values drawn from a Normal distribution to be estimated. This results in the Gamma-Normal Generalised linear mixed model defined by

$$log(SF\_12\_MCS + 10) = X\beta + Zb + \varepsilon$$

where $X\beta$ is contains all fixed effect terms used in the MICE OLS model (Equation 3.5) and Zb contains all random intercept terms conditional on pidp. The GLMM is applied to all 30 MICE datasets using the ‘lme4’ and MICE R. packages (Bates, 2014; Schreuder et al., 2021) pooled together into an ensemble final estimate of model coefficients. All variables in the MICE OLS model are used in the GLMM model with an additional lagged dependent variable (LDV) term (McLay et al., 2015; Richiardi et al., 2014). These terms are commonly added to microsimulation transition probability models in order to simulate reaction to variable change due to interventions (Richiardi et al., 2014; McLay et al., 2015).

Model coefficients and diagnostics are given in Table 3.9 and Figure 3.10. The GLMM model provides marginal and conditional R^2 scores (Nakagawa and Schielzeth, 2013) of 0.422 and 0.556 vs 0.290 and AIC score (Harrell Jr, 2015) of 891231.1 vs. 999226.9 suggest an improvement in model fit over the MICE OLS model. Again five-fold cross validation also performed on each imputed MICE population giving $RMSE = 0.152 \pm 6 \times 10^{-3}$ indicating smaller and less uncertain error vs the OLS and MICE models. Model coefficients are given in Table 3.9. Due to the nature of the Gamma-Normal model coefficients are interpreted differently. A coefficient of -0.1 indicates an $1 - e^{-0.1} = 0.0905 = 9.05\%$ decrease in SF-12 MCS score. Note as data is reflected coefficients also have opposite signs. For example, higher levels of loneliness (3 vs. 1) shows a positive model coefficient indicating a decrease in SF-12 MCS. We see overall similarly with the MICE OLS model in Table 3.5. The random terms gives standard deviation explained by pidp personal identifier as 0.153 suggests random intercepts accounts for 53% of variance in residual errors not explained by fixed effects. A QQ plot in Figure 3.5d demonstrate that the Gamma-Normal GLMM is able to better estimate skew in SF-12 MCS score but still underestimated sample variance and a scale-location plot in Figure 3.10d does not exhibit a strong linear trend as in the baseline case suggesting again the the GLMM model is an improvement and accounts for some heterogeneity. However, increasing error for low SF-12 MCS scores (high fitted values) suggests this model does not fully explain heterogeneity and further predictor variables are required. Out of sample cross validation is given in Figure3.11 suggests improved estimation of overall SF-12 MCS mean and variance over time compared with the baseline model. Population variance is still underestimated but may be corrected using alignment methods (Li and O'Donoghue, 2016). The LASSO is not applied to the GLMM model due to limited available software that is left to future work.

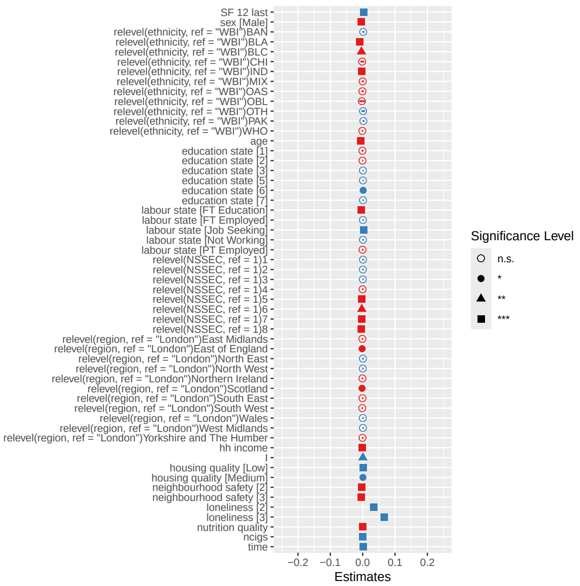
The GLMM model shows substantial improvement in model fit including out of sample RMSE performance and adherence to underlying mathematical assumptions. Further improvements to the GLMM model can be accquired by including further predictor variables, extending validation forwards in time and under counterfactual scenarios, and meta analysis comparison of several models.

	SF_12 GLMM
(Intercept)	−1.1295 (0.3720)**
SF_12_last	0.0027 (0.0001)***
factor(sex)Male	−0.0050 (0.0004)***
relevel(factor(ethnicity), ref = "WBI")BAN	0.0016 (0.0018)
relevel(factor(ethnicity), ref = "WBI")BLA	−0.0085 (0.0017)***
relevel(factor(ethnicity), ref = "WBI")BLC	−0.0039 (0.0016)*
relevel(factor(ethnicity), ref = "WBI")CHI	−0.0015 (0.0032)
relevel(factor(ethnicity), ref = "WBI")IND	−0.0033 (0.0013)*
relevel(factor(ethnicity), ref = "WBI")MIX	−0.0000 (0.0015)
relevel(factor(ethnicity), ref = "WBI")OAS	−0.0010 (0.0018)
relevel(factor(ethnicity), ref = "WBI")OBL	−0.0026 (0.0051)
relevel(factor(ethnicity), ref = "WBI")OTH	0.0011 (0.0033)
relevel(factor(ethnicity), ref = "WBI")PAK	0.0020 (0.0013)
relevel(factor(ethnicity), ref = "WBI")WHO	−0.0012 (0.0011)
scale(age)	−0.0053 (0.0003)***
factor(education_state)1	−0.0010 (0.0013)
factor(education_state)2	−0.0006 (0.0005)
factor(education_state)3	0.0006 (0.0007)
factor(education_state)5	0.0007 (0.0008)
factor(education_state)6	0.0014 (0.0007)*
factor(education_state)7	0.0010 (0.0008)
factor(labour_state)FT Education	−0.0052 (0.0010)***
factor(labour_state)FT Employed	0.0007 (0.0009)
factor(labour_state)Job Seeking	0.0033 (0.0011)**
factor(labour_state)Not Working	0.0007 (0.0009)
factor(labour_state)PT Employed	−0.0006 (0.0009)
relevel(factor(NSSEC), ref = 1)1	0.0010 (0.0012)
relevel(factor(NSSEC), ref = 1)2	0.0005 (0.0010)
relevel(factor(NSSEC), ref = 1)3	0.0003 (0.0008)
relevel(factor(NSSEC), ref = 1)4	−0.0003 (0.0008)
relevel(factor(NSSEC), ref = 1)5	−0.0031 (0.0009)***
relevel(factor(NSSEC), ref = 1)6	−0.0028 (0.0011)**
relevel(factor(NSSEC), ref = 1)7	−0.0024 (0.0007)***
relevel(factor(NSSEC), ref = 1)8	−0.0040 (0.0008)***
relevel(factor(region), ref = "London")East Midlands	−0.0010 (0.0010)
relevel(factor(region), ref = "London")East of England	−0.0017 (0.0009)

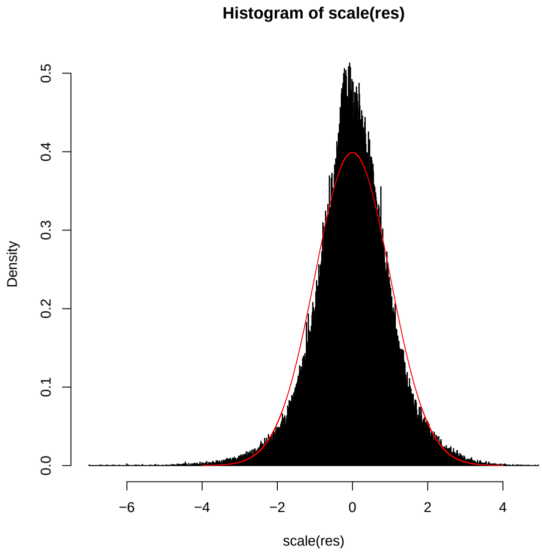
	SF_12 GLMM
relevel(factor(region), ref = "London")North East	0.0007 (0.0012)
relevel(factor(region), ref = "London")North West	0.0002 (0.0009)
relevel(factor(region), ref = "London")Northern Ireland	-0.0009 (0.0010)
relevel(factor(region), ref = "London")Scotland	-0.0019 (0.0009)*
relevel(factor(region), ref = "London")South East	-0.0007 (0.0009)
relevel(factor(region), ref = "London")South West	-0.0014 (0.0010)
relevel(factor(region), ref = "London")Wales	0.0002 (0.0010)
relevel(factor(region), ref = "London")West Midlands	0.0008 (0.0009)
relevel(factor(region), ref = "London")Yorkshire and The Humber	-0.0006 (0.0009)
scale(hh_income)	-0.0010 (0.0002)***
scale(hh_income^2)	0.0006 (0.0002)**
factor(housing_quality)Low	0.0019 (0.0006)**
factor(housing_quality)Medium	0.0007 (0.0004)
factor(neighbourhood_safety)2	-0.0026 (0.0004)***
factor(neighbourhood_safety)3	-0.0042 (0.0006)***
factor(loneliness)2	0.0340 (0.0007)***
factor(loneliness)3	0.0665 (0.0012)***
nutrition_quality	-0.0001 (0.0000)***
ncigs	0.0003 (0.0000)***
time	0.0012 (0.0002)***
nimp	30
nobs	128851

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

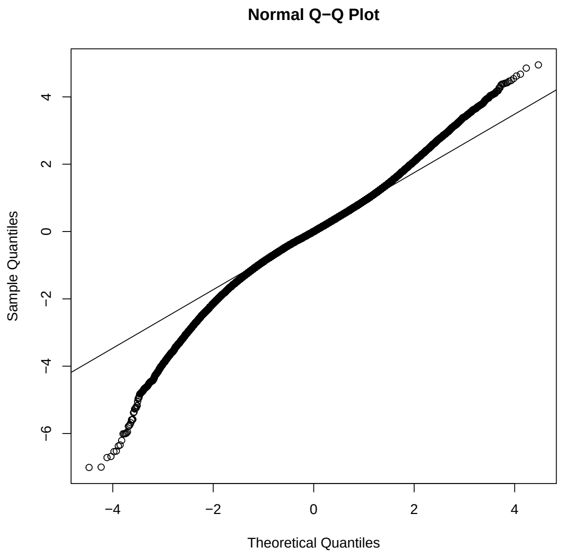
Table 3.9: Gamma-Normal GLMM Coefficients estimating SF-12 MCS score using random intercepts.



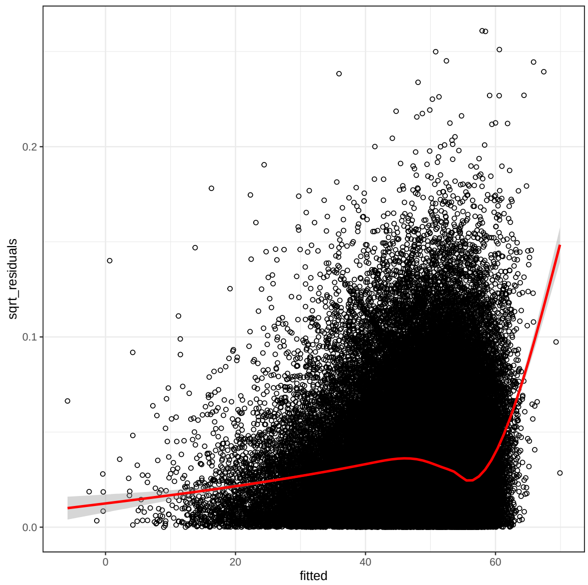
(a) Forest plot of SF-12 MCS coefficients for the SF-12 MCS GLMM model.



(b) Comparison of residual density between actual and predicted observations of SF-12 MCS GLMM model.

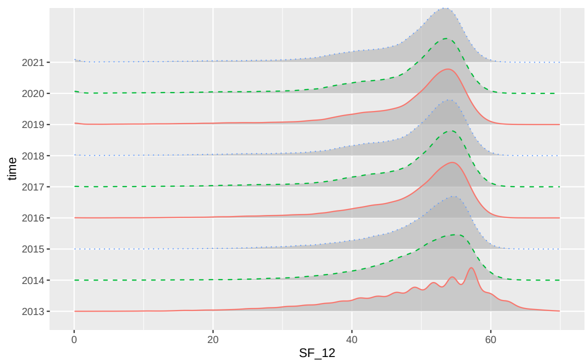


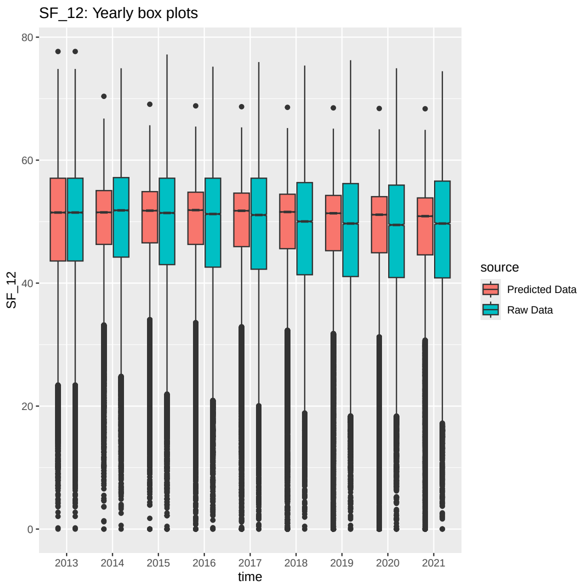
(c) A QQ-plot for the SF-12 MCS GLMM model.



(d) A Scale-Location plot using for the SF-12 MCS GLMM model. The red line indicates the mean residual value over fitted values of SF-12 MCS score.

Figure 3.10: Diagnostic Plots for the SF-12 MCS GLMM Model.





(a) Ridgeline plots for the predicted density of SF-12 MCS scores from $2013-2021$ using the GLMM model. Data for the year 2013 are real data with predictions for all future years derived from this data. Real data are available in Figure 3.4.

(b) Boxplots for the predicted density of SF-12 MCS from $2013-2021$ using the GLMM model and cross-validated nowcasting. Sample variance is still underestimated but provides improvement over the baseline model.

Figure 3.11: Internal five-fold cross-validation nowcasting plots for GLMM Model.

3.8 Discussion

Well calibrated transition dynamics are essential in dynamic microsimulation to ensure accurate, reproducible prediction of future state for individual longitudinal trajectories and aggregate population distributions in complex, non-linear systems. Recent review into the state of the art of dynamic microsimulation (O'Donoghue and Dekkers, 2018) has highlighted a need to codify common dynamic microsimulation methods including transition probabilities standardising best practice to increase accessibility, address common pitfalls and mistakes, and provide a platform for implementation of further novelty. The case study presented in this paper complements an external literature review into transition probability model development [authors work under review in IJM. CITE IF ACCEPTED.] highlighting three key areas of planning, validation, and interpretation that must be addressed in development. An example transition probability model has been developed estimating SF-12 MCS score as part of a wider mental well-being dynamic microsimulation.

Designing an SF-12 MCS model required consideration of existing health microsimulation and policy literature. Extensive background research iteratively refined an existing causal loop diagram conditional on available open source data, statistical software, and required simple model structure. Using an ordinary least square model that is universally available allows for easy testing of model assumptions, simple interpretable model structure that can be easily critically reviewed and applied elsewhere, and can be extended to more complex generalised linear mixed GLMM models relaxing the homogeneity assumption if OLS model assumptions are violated.

Testing baseline model goodness of fit using internal validation showed poor overall fit including a large root mean squared error (RMSE) value, biased estimation of mean and variance over time, and violation of model assumptions including homoscedastisity due to repeated observations and skewed response data. Sensitivity analyses were also used quantifying uncertainty in the baseline model with respect to model structure, model parameter values, and input source data. Overall, the OLS model was not sensitive to randomisation of model coefficients but application of the LASSO variable selection algorithm suggests more than half of the included predictor variables suggested by background literature are not necessary to produce parsimonious prediction of SF-12 state. Application of missing data correction using the MICE algorithm demonstrated prediction is highly dependent on variables such as loneliness that are only recorded in recent survey data suggesting strong model sensitivity to source data. Validation suggests use of longitudinal models that can account for repeat individual observations was required prompting use of GLMM methods. The above analysis was repeated for the GLMM model showing improved model fit with a reduction in RMSE score over 5-fold cross validation from 0.223 to 0.152, better estimation of statistical moments over time using nowcasting, and improvement in diagnostics for underlying modelling assumptions.

Validation and improvement of SF-12 MCS state has been supported with extensive visualisation including common model diagnostics diagrams and internal validation methods including handover plots. Application of standard visualisation methods further increases accessibility, allowing for easy comparison with future external simulated and real data. Full coefficients and model code are provided for all models fitted throughout this article intending to provide a highly interpretable model that can be examined by peers and utilised in wider mental well-being research by non-technical policy partners to construct narratives on how policies enact health change across the whole population and key subgroups.

While this case study provides a useful general application of transition probability development methods there are a number of limitations. This is a single transition model for a specific case and a more generic development plan is required in line with similar work (Burgard et al., 2020). Similarly testing should be expanded to a full-scale microsimulation model providing a more complex testing environment exploring how interaction between multiple transition probability models (Harding et al., 2010a), consistency of repeated application of transition models over time (Petrik et al., 2020; Adhikari et al., 2019) as model coefficients and population demographics evolve, change in prediction under counterfactual scenarios and for specific subgroups of a population, and computational scalability to very large populations can also be integrated in development literature. Application to a well established microsimulation such as EUROMOD (Richiardi et al., 2021) and accessible test scenarios would provide a more standardised example.

Future work aims to address these limitations constructing a full dynamic microsimulation model in order to test further novelty in transition probability development. A full microsimulation allows for consideration of interactions between multiple transition models discussing planning and standardisation of transition models in line with wider microsimulation design (Burgard et al., 2019) and validation considering multiple modules together including the order in which they are run (Harding et al., 2010a; Scott et al., 2003). Counterfactual scenarios can also be applied to this model providing further test scenarios for external validation and goodness of fit testing for transition models under large scale population change. This microsimulation would also serve as a useful test case for replacing individual transition models one at a time with novel transition probability methods. Direct comparison between multiple transition models estimating the same process could build a library of available methods with clear demonstration of benefits and potential trade-offs.

Finally a full microsimulation can be utilised to develop further visualisation methods for mental well-being microsimulation in line with other social geography literature (Ballas, Rossiter, Thomas, Clarke and Dorling, 2005). Demonstrating how individual life paths change across interventions can then further facilitate construction of non-technical narratives to further promote uptake of microsimulation in government decision making.

3.9 Connection between Chapter 3 and Chapter 4.

This chapter has applied methods discussed within the literature review for improving the robustness and goodness of fit for transition probability models in dynamic microsimulation. A case study has demonstrated how the Understanding Society dataset for the UK population can be processed and utilised to produce a transition probability model estimating future Short Form 12 Mental Component Score (SF-12 MCS) state using housing and socioeconomic predictor variables.

An initial ordinary least squares regression was fitted to estimate SF-12 MCS showing poor performance with an root mean square error (RMSE) score of 0.223 under five fold cross-validation, and poor prediction of SF-12 MCS variance over time. Further presentation of the OLS model demonstrated how model assumptions were violated suggesting potential improvement using models that relax the assumption of homoscedasticity. Application of Generalised Linear Mixed Modelling and lagged dependent variables significantly improved prediction giving a RMSE of 0.152. Methods to test robustness of the GLMM model were also applied including imputation of missing values in UKHLS data. Application of deterministic missing data imputation including the Last Observation Carried Forwards (LOCF) algorithm corrected substantial missing data that was a result of survey design. Multiple imputation chained equations algorithms (Krijkamp et al., 2018) were then employed to impute the remaining missing data. Further sensitivity analysis quantified the uncertainty of the imputed input data using an ensemble of MICE imputed populations. Uncertainty due to model coefficients and structure is tested using randomised model coefficients and the LASSO algorithm quantifying how much predicted output changed and influential variables. Application of Monte Carlo noise also ensured consistent prediction of behaviour under random chance due to initial perturbations in population state. Overall, this testing demonstrated limited change in the overall predicted output suggesting consistent prediction across multiple data sources and model structures.

This case study has demonstrated how best practice can improve prediction quality for a transition probability model in discrete-time dynamic microsimulation. Previous chapters have combined literature review and case study examples contribute to the larger required tutorial literature and outline potential future work available including application to a system of multiple transition probability models as continuous time models. This thesis aims to utilise this knowledge in the development of a full dynamic microsimulation model for estimating the health effect of policies that improve household income. The following chapter describes the initial construction of this microsimulation model 'MINOS' and application to several key income policies including the Scottish Child Payment, the Real Living Wage, and the Energy Price Cap Guarantee.

Chapter 4: Estimating the Effects of Income Support Policies on the Mental Well-Being of the UK Population Using a Microsimulation Model

4.1 Abstract

In response to the recent cost-of living crisis, the United Kingdom (UK) Government implemented a number of policies intended to provide income support to vulnerable households. For these policies to be cost effective they must be evidence-based and tailored to benefit individual households and priority subgroups that need the most support. Computational modelling can support the appraisal and optimisation of policy targeting. This paper introduces the dynamic microsimulation framework MINOS, which addresses existing evidence gaps, and provides a tool for policy makers to interrogate and estimate the effect of such policies. MINOS was used to estimate the mental well-being effects of different income policy scenarios for a synthetic population of UK individuals and households over different time horizons. We provide estimated policy effects for (1) child benefit uplifts for existing welfare recipients, (2) payment of the national living wage to all workers, and (3) energy price cap interventions. Our modelling results suggest that the child benefit and living wage uplift policies have the capacity to improve mental well-being for the UK population with the living wage intervention providing well-being gains affecting a larger proportion of the population with lower uncertainty. The Energy Price Cap guarantee does lessen some of the ongoing energy crisis' detrimental effects on mental well-being. However, the policy as modelled is expensive and on its own, insufficient to return well-being to pre-crisis levels.

4.2 Introduction

The United Kingdom (UK) saw real household disposable income fall by 4.3% in the 2022 – 2023 fiscal year - the largest annual decrease since records began in 1956 (Office for Budget Responsibility, 2023; Office for National Statistics, 2022). A combination of socioeconomic and geopolitical factors has sharply increased the cost of fuel, food, and other household expenses resulting in a cost of living crisis that risks having a profound impact on health (Ballesteros-Arjona et al., 2022; Broadbent et al., 2023; UK Fuel Poverty Monitor, 2022). In 2022, more than half of UK households were estimated to be unable to afford winter heating bills (Bradshaw and Keung, 2022), which is likely to have exacerbated respiratory disease due to damp and mould (Ballesteros-Arjona et al., 2022; Senedd Cymru and Welsh Parliament, 2023; UK Fuel Poverty Monitor, 2022). Increasing food costs are reducing food security, driving households to food banks, low quality diets and skipping meals altogether, resulting in both malnutrition and obesity (Lonnie and Johnstone, 2023; Senedd Cymru and Welsh Parliament, 2023). Incidence of mental illnesses including depression, anxiety, and suicidal ideation have increased (Senedd Cymru and Welsh Parliament, 2023; Wetherall et al., 2023; UK Fuel Poverty Monitor, 2022; Ballesteros-Arjona et al., 2022), with rising debt and financial strain as likely contributing factors.

The national and devolved governments of the UK have a range of policy levers available to protect disposable income (Inequalities in London, 2023; Wetherall et al., 2023; Prowse and Fells, 2016), which include making direct payments to eligible households and enacting legislation to ensure employers pay above a certain level. In an attempt to mediate some of the impact of rising fuel costs, the energy price cap guarantee support scheme was implemented to cap energy prices for households in England, Scotland and Wales (Norman and Corfe, 2022). In September 2022, energy prices were capped such that the average annual household energy bill would not exceed £2500 per year, which increased to £3000 by April 2023, with any difference in expenditure deferred to future taxation (UK Fuel Poverty Monitor, 2022; Norman and Corfe, 2022). In Scotland only, child uplift payments provide households in receipt of certain welfare benefits with extra income support (Scottish Government, 2017). In its commitment to reduce child poverty and in response to rising inflation and energy costs, the Scottish Government doubled the Scottish Child Payment from £10 to £20 in April 2022, and the continued cost of living crisis was cited as a reason for a further increase to £25 in November 2022 (Scottish Government, 2022). As of February 2024, there is ongoing discussion in Scottish Parliament to again increase the child payment to £30 in line with inflation (Pybus, 2023) despite evidence submitted by the Save the Children (Statham et al., 2022) and Child Poverty Action Group (Pybus, 2023) child poverty charities strongly recommending an increase to £40 is required meet child poverty targets with expansion to further vulnerable subgroups of the population including single parents. Simulated research suggests payments of up to £85 per week while expensive could be cost effective (Congreve et al., 2022) for decreasing child poverty. An example of legislation is the UK Government's statutory living wage, which as of April 2024 requires employers to pay their employees aged 21 years old and above £11.44 per hour (Ogden and Thomas, 2024). However, this falls short of the Living Wage Foundation's recommendation for a level of income that would be sufficient to sustain employees and their families (Prowse and Fells, 2016; Cominetti and Murphy, 2022), of £12.00 per hour outside of London and £13.15 within London for 2024. The real living wage is voluntarily adopted by employers, for example as part of good employment charters.

Understanding the efficacy of these policy levers is essential, however multiple independent reviews into policy making decisions across all sectors of UK Government have highlighted a lack of empirical 'evidence-based policy', (Cairney, 2019, 2021; Norman and Corfe, 2022; Yang, 2020; Brownson et al., 2009) instead there is reliance on historical or out of context evidence and expert opinion. Pragmatically there is no time to perform national-level evidence gathering for best policy in crises that require urgent intervention (Yang, 2020; Lenihan, 2015). Moreover, such policies risk being implemented by siloed departments with a lack of wider consultation that would consider consequences over longer periods of time and in other sectors such as health (Sasse and Thomas, 2022; Meier et al., 2019; Höhn et al., 2023). Increasingly a 'health in all policies' approach, whereby policy decisions across multiple sectors are understood as drivers of well-being (De Leeuw, 2017; Meier et al., 2019) is being adopted by local government in England, who have responsibility for public health (Greszczuk, 2019) and advocated by public health professionals from across the UK (Committee of the Faculty of Public Health in Scotland, 2017).

Further concerns can arise around a lack of precision in the targeting of such policies.For example, during the 2022 energy crisis, the Energy Price Cap Guarantee and Energy Bill Support Scheme were very expensive fiscal interventions, exceeding £100bn, which were widely criticised (UK Fuel Poverty Monitor, 2022; for Social Justice, 2023; Hodgkin and Sasse, 2022) as unsuitable for protecting the most vulnerable low-income households . Alternate measures have since been proposed including extensive upgrading of insulation for social housing or the banning of prepayment meters, which typically increase the per-unit energy costs for poorer households, that may have been more precisely targeted, created environmental co-benefits and cost less (UK Fuel Poverty Monitor, 2022; Norman and Corfe, 2022; for Social Justice, 2023). Likewise, income support policy during the coronavirus pandemic has been described as 'scatter gun' (Milne, 2020), for example providing blanket stimulus payments and furloughing to those who may have not needed assistance, while ignoring other industrial sectors including aviation and recreation (Cairney, 2021). Again, researchers have since proposed more efficient schemes to protect both household and commercial finances (Milne, 2020; Cairney, 2021; Bhattacharjee et al., 2023). With respect to voluntary approaches such as the real living wage, evidence on benefits and costs can encourage their adoption in legislation (Brownson et al., 2009; Lenihan, 2015). Given the complexity involved in developing effective interventions that protect public health in a crisis, there is an urgent need for methods that quickly and robustly provide evidence that informs policy.

One such mechanism for generating policy evidence is through the implementation of dynamic microsimulation models (O'Donoghue and Dekkers, 2018). A population of individual units, such as people or households, is propagated forwards in time under transition probability mechanics and counterfactual policy scenarios (Burgard et al., 2021). Such microsimulation models have a history of informing economic, but not health, policy in the UK, for example the Policy Simulation Model 'PENSIM2' used by the UK Department for Work and Pensions to assess taxation and benefits policy (Li and O'Donoghue, 2013). As data, methods, and computing power become more readily available, dynamic microsimulation is seeing application to new areas of research including household disposable income change in response to geopolitical shocks (Mosley et al., 2022; Bhattacharjee et al., 2023; Katikireddi et al., 2022; Broadbent et al., 2023) and public health scenarios concerning disease incidence and resource allocation (Rutter et al., 2011; Archer et al., 2021; Smith et al., 2021; Spooner et al., 2021; Craig et al., 2022; Skarda et al., 2021). A number of research organisations are proposing implementation of dynamic microsimulation for assessing the relationship between household disposable income and mental well-being (Meier et al., 2019; Skarda et al., 2021; Katikireddi et al., 2022; Craig et al., 2022). Income support policy can then be parameterised to estimate change in household disposable income under intervention exploring how this propagates through facets of an individual's life into tangible health outcomes. Such a microsimulation provides results for an attribute rich dataset of individuals that can be used to assess change in outcomes at national level as well as interrogation outlining subgroups of a population that require further targeted policy (e.g. vulnerable individuals) (Burgard et al., 2021). Evidence for a wide spectrum of policies can be quickly generated over both long and short term time horizons. Ultimately an income-health dynamic microsimulation provides a useful tool for policy makers to synthesize evidence for income support policy that is cost-effective and provides help to those who need it the most.

In this paper we use the Microsimulation for Interrogation of Social and Health Systems (MINOS) model to estimate the effect that change in household disposable income has on mental well-being for the adult population of the United Kingdom (UK). We implement three scenarios: the first models the effect that the energy bill support scheme had at time of implementation and extends the policy in to the future; the second models the effect that the current (£25) and proposed (£50) Scottish Government child payment would have if implemented for the whole of the UK; and the third assesses the impact of raising the statutory living wage to the higher national living wage figure. MINOS has been developed by the Systems Science in Public Health and Health Economics Research (SIPHER) Consortium (Meier et al., 2019) to model identified links between policy intervention and health well-being outcomes (Meier et al., 2019; Hill O'Connor et al., 2023; Aki Tsuchiya, Guoqiang Wu, 2021; Höhn et al., 2023). In this paper MINOS is used to provide an associative simulation of income to health pathways that can be interrogated to investigate income support policy effect. The remainder of this paper is structured as follows: Section 4.3 provides an overview of methods, Section 4.4, provides results of the scenario interventions, Section 4.6 provides discussion outlining limitations and future work and Appendix 8.1 provides appendices of further methodological detail and data tables.

4.3 Methods

This section describes the data, methods, and structure of the MINOS framework and its application to scenarios that impact on household disposable income.

4.3.1 Complex Systems Modelling and Causal Loop diagrams.

Estimating the effect candidate policy will have health outcomes for the UK population is performed using the complex systems modelling methodology (CSM) (Haraldsson, 2004; Meier et al., 2019) performed in two stages. The first stage gathers evidence from literature and experts from academia and communities into causal loop diagrams (CLDs) (SIPHER Consortium et al., 2023; Campbell et al., 2023) describing how individual, household, neighbourhood, and national characteristics are connected. Characteristics are represented as individual nodes with any arrows between them signifying causal relationships either positive or negative. A node between income and nutrition quality suggests evidence that as income changes nutrition quality also changes. In this case the effect is positive as more money implies greater access to higher quality food including fruits and vegetables. This research utilises an existing causal loop diagrams (SIPHER Consortium et al., 2023; Campbell et al., 2023; Consortium, 2023) describing pathways between household disposable income and health whose derivation is described elsewhere (SIPHER Consortium et al., 2023; Campbell et al., 2023).

4.3.2 Microsimulation for Interrogation of Social Science Systems (MINOS)

The second stage of CSM then emulates these pathways using individual-level modelling techniques such as microsimulation (Meier et al., 2019). The intention is to operationalise a given causal loop map estimating the effect of hypothetical policy on the UK population. The Microsimulation for Simulation of Social Science Systems (MINOS) (Clay et al., 2023; Lomax et al., 2023) is a dynamic microsimulation used in SIPHER approximating the income-health systems map provided online (SIPHER Consortium et al., 2023). MINOS is designed to be completely open source written in the R. and python languages using publicly available datasets. The full income-health system is estimated in MINOS using a series of discrete-time modules predicting next state using current population information and associative modelling described in Figure 4.1.

Each module contains a series of transition probability models each estimating some subset of the future population state. For example, the tobacco module estimates whether individuals smoke at all as well as monthly cigarette consumption. Development of these transition models was done collaboratively within the SIPHER group. Model code was compiled into R. notebooks containing model formulae, model structure based on dependent variable, independent variable selection literature, model assumptions and validation methods, sources of uncertainty, and model diagnostics allowing for iterative discussion and refinement. Data processing was also provided including dataset description and variable summaries, derivation of required variables including household disposable income and housing quality composites, and handling of missing data. Individual transition models were then arranged into the order provided in Figure 4.1 including policy intervention. These notebooks were used to extensively plan and design the initial microsimulation structure according to planning questions discussed in previous literature review chapters. Output from these notebooks is provided later in this chapter and in supplementary material. Model design was easily interpretible highlighting potential flaws and how each module in the microsimulation system may respond to the status quo over time or to policy intervention.

All transition probability methods used apply associative prediction method to estimate future state in yearly time increments. Causal predictive methods are not used due to preference of high interpretability and data constraints due to annual data.

This module structure then allows the MINOS microsimulation model to project the UK population forwards in time under various hypothetical policy strategies. The larger CSM framework can then provide policy makers an overall view of how a potential policy may effect a system based on causal evidence both qualitatively using causal diagrams as well as quantitatively using synthetic microsimulation evidence.

4.3.3 Data

The Understanding Society Household Longitudinal Study (UKHLS) (Fumagalli et al., 2017) is the primary data source used in MINOS. UKHLS is a household panel survey with 13 waves of data currently available (2009 - 2021). UKHLS data contains hundreds of individual and household attributes including demographics, employment status, income, and health. The subset of UKHLS variables used in MINOS for the income to mental health model are shown in Table 4.1. The key variable that is intervened on is household disposable income while Short Form 12 Mental Component Score (SF-12 MCS) is the outcome. Household disposable income is defined as income available to a household after taxes, national insurance and housing costs to spend on other needs (Aki Tsuchiya, Guoqiang Wu, 2021). This provides an exposure variable for MINOS with all policy scenarios parameterised as change in household disposable income. SF-12 MCS is a measure developed (Lawrence and Fleishman, 2004) to provide a continuous score with higher values indicating higher mental well-being. SF-12 MCS is used as an outcome variable to determine any change in mental well-being resulting from the policy intervention. Both income and SF-12 MCS variables are continuous highly skewed variables with medians of £1350 and 50 respectively.

There are several stages of pre-processing applied to UKHLS data before use in MINOS. Initially many variables are relabelled from integer codes to strings to improve readability in model coefficients and outputs. Missing data handling is then applied. In instances where values are not missing, such as unemployed individuals having no salary, are replaced to improve readibility and differentiate between missing values and non-applicable values. Some variables are only recorded when change occurs. for example as an individual enters the dataset in 2011 their ethnicity is recorded but then registered as missing in subsequent waves because it has not changed. These variables are forward filled using Last Observation Carried Forwards imputation. Some individual age values are missing and are corrected using linear interpolation. Derived composite variables are then generated. Some values such as household disposable income must be derived from net income, household outgoings such as rent, and adjusted for inflation. Finally complete case analysis is used to remove all observations with truly missing values. After correction there are $n = 14737$ observations in 2020 which serves as the jump-off point for our simulations. This population is the same for every model run. Fertility and mortality rates by age, sex, and ethnicity (Rees et al., 2017) from the NEWETHPOP project are used to evolve the population forward in time within MINOS. A summary of key variables used in the MINOS model is provided in Table 4.1 with more detail on data pre-processing provided in Section 4.3.5.

Name	Category	Description
pidp/hidp	Demographics	Individual/household unique identifiers
age	Demographics	Individual Age.
nkids	Demographics	Children under 16 years old living in the household.
sex	Demographics	Biological sex
ethnicity	Demographics	Ethnicity*
region	Demographics	Government region*
education_state	Demographics	Highest Education State*
housing_quality	Housing	Household quality+
neighbourhood_safety	Housing	?+
marstat	Housing	Marital Status*
hh_comp	Housing	Number of people in the household
household_disposable_income	Employment	Household Income after taxes, national insurance, and housing costs.
job_sector	Employment	What job sector (jobsec07) are you employed in?*
NSSEC	Employment	What is the NSSEC socioeconomic code for your employment?*
nutrition_quality	Health	How many fruits and vegetables do you consume per week?+
SF12_MCS	Health	What is your SF-12 MCS score?
loneliness	Health	How often are you lonely?

Table 4.1: Key Understanding Society panel dataset variables used in MINOS estimating the effect of income support policy on mental well-being. Further information on discrete variables (marked *) and derived variables (market +) are provided in the following section.

4.3.4 Model Structure

MINOS follows typical dynamic microsimulation design (Spielauer et al., 2020; Burgard et al., 2020) as a discrete time Markov model that projects the population forwards in yearly increments using only current time information. Transition probabilities for MINOS are divided into a series of modules. Each module handles an intermediary pathway variable that is influenced by change in disposable income and goes on to influence mental well-being. These pathways have been identified by synthesising published evidence and engaging with policy makers (Hill O'Connor et al., 2023; Aki Tsuchiya, Guoqiang Wu, 2021) and summarised in Figure 4.1. For example, the nutrition quality module estimates the number of fruit and vegetables an individual consumes in a week. If household disposable income decreases there is less expenditure on good quality food which will then have a negative impact on individual mental well-being.

The overall model structure for MINOS is given in Figure 4.2. A starting population for MINOS is synthesised by directly importing processed Understanding Society data. MINOS then implements a series of transition probability modules. The order in which these modules are run is important and occurs in four sets. The first set updates household disposable income. All policy interventions are parameterised as change in household disposable income; placing this module first allows change to propagate through the rest of the system immediately. The second set contains auxiliary 'life path' birth and death modules that estimate the life trajectories of individuals. The third set of modules represents the intermediary pathways between income and mental well-being. The final (fourth) set contains a single mental well-being module that estimates the change in SF-12 MCS, given change in all other modules. An overview of all transition probability models used in MINOS is provided in Table 4.2. The choice of model type depends primarily on the outcome variable data type. More detailed description of formulae, coefficients and software implementation are provided in Section 4.3.6.

After running each module set, the third stage of MINOS is population replenishment. Here, a new cohort of individuals is added to the model in each year to maintain population size and representativeness. Individuals $n = 306$ in UKHLS data aged $15 - 16$ years old are added to the population with sample weights (Fumagalli et al., 2017) updated in order to match projected change in age, sex, and ethnicity distributions (Rees et al., 2017). For example, if the number of White British 16 year old males increases in the UK population, then corresponding individuals in the replenishing data see larger sample weights. The replenishment and transition probability modules are run repeatedly in 15 annual waves until 2035. MINOS's final output is cross-sectional yearly projections of individuals who's state in the income, pathway and SF-12 MCS modules has been transitioned by the model.

MINOS is built on the existing Vivarium dynamic microsimulation framework written in Python (Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA, 2023). The R. programming language is used to calculate all transition probability models. All modelling was undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. All results were run on a standard ARC4 node with 40 cores and 768GB of RAM. MINOS iterations were run in batches each requiring 1 core and $5GB$ of RAM requiring 15 ± 2 minutes of runtime. It took approximately 2 hours depending on HPC availability to generate all 600 iterations used in the results presented in this paper. This large computational cost is primarily due to use of the R. and python languages, repeated application of large non-linear regression models and ensuring consistent alignment of household characteristics over time including household disposable income and housing quality.

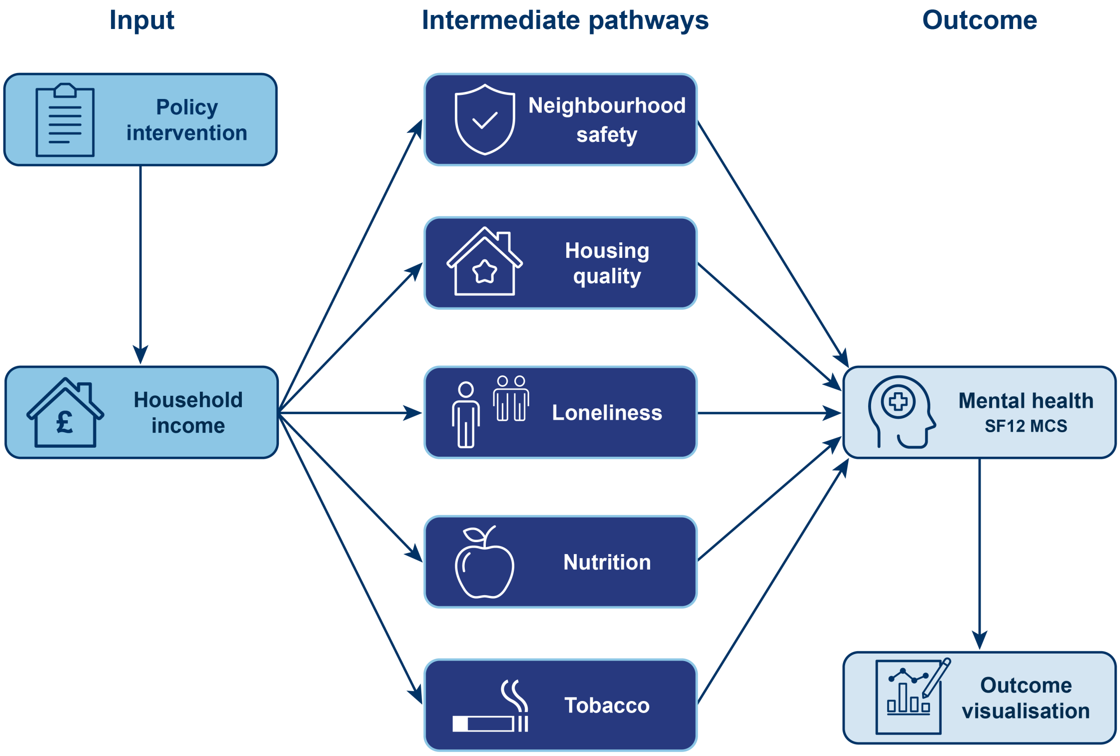


Figure 4.1: MINOS high level pathways. Each box indicates a pathway that is influenced by income and goes on to cause change in mental well-being. These boxes are represented as a series of modules each containing some set of related transition probability models.

4.3.5 Data Preprocessing

Several stages of data preprocessing are applied to Understanding Society data in order to impute missing values and improve the readability of model data. This is performed in several stages described below beginning with raw UKHLS panel data for reproducibility (Lovelace and Dumont, 2017). Final tables describing key continuous and discrete variables are provided in Section 8.1.1.

Raw Data Processing

Understanding Society data are downloaded directly from the UK data service (Fumagalli et al., 2017) as yearly interview cohort data divided into several sections for individual and household response, children only data, and other datasets not used in MINOS. Initial data processing extracts required variables for MINOS from individual and household response datasets (Fumagalli et al., 2017) from UKHLS that are then merged together to create rows for each individual person with complete personal and household information. Certain categorical variables are renamed from integer codes to string values for clarity. It is much easier to read labour state 'employed' rather than code 3. This also allows for much easier fitting of transition probability models in R. that automatically recognises strings as factors.

Module Name (Module Set)	Outcome Variable	Variable Type	Model Used
Household disposable income (1)	Monthly household disposable income (£s)	Continuous	Generalised Linear Mixed Models (GLMMs)
Mortality (2)	Is the subject alive? (yes/no)	Binary	Rate Tables
Fertility (2)	Has subject given birth in last year? (yes/no)	Binary	Rate Tables
Housing quality (3)	Household quality composite (1-3 Likert)	Ordinal	Cumulative Link Model
Neighbourhood safety (3)	Neighbourhood composite (1-3 Likert)	Ordinal	Cumulative Link Model
Loneliness (3)	Is subject lonely? (1-3 Likert)	Ordinal	Cumulative Link Model
Nutrition (3)	How many fruit and vegetables per week? 0+	Continous	Generalised Linear Mixed Models (GLMMs)
Tobacco (3)	Cigarettes smoked per week (ncigs) 0+	Continuous	Zero Inflated Poisson (ZIP)
Mental Well-Being (SF12) (4)	SF-12 MCS Mental Component Summary score (0-100)	Continuous	Generalised Linear Mixed Models (GLMMs)

Table 4.2: Table of transition probability models used in MINOS. Note these transition models operate independently in a single time step with no interaction effects. Full model formulae, implementation and methods are available in Section 4.3.6.

Data Cleaning/ Correction

Within Understanding Society there are also a number of variables that are deterministically missing by non-applicability (Fumagalli et al., 2017) that must be cleaned (Lovelace and Dumont, 2017). For example, if an individual is unemployed they will not be provided with NSSEC employment socioeconomic code as they do not have a job that can be graded. We assign separate encodings to these variables to differentiate them from missing data.

Similarly there are a number of observations that are missing due to not being carried forwards (Fumagalli et al., 2017). Many values in Understanding Society are not carried forwards such that they are recorded only when the individual enters the dataset or change state, otherwise they are registered as missing (Fumagalli et al., 2017). For example, an individual could enter the dataset recording their ethnicity once but this value is then missing for all remaining years. Some variables such as ethnicity are entered into the model once and never change. These values are filled in using previous information providing substantially more complete observations. Correction is performed using Last Observation Carried Forwards Imputation (LOCF) (Zhang, 2016). Missing values are replaced with the last previous complete value. If an individual has their initial ethnicity recorded in 2009 their ethnicity state is forward filled for 2010 data and beyond. Care is taken to ensure truly missing individual values are not replaced. Only forward LOCF imputation is used as no assumption can be made about years of data before an observation was made. Only for immutable variables such as ethnicity forward and backward filling are used. Only the missing data code −8 for non-applicability is replaced and not other codes for refusal, missing by proxy and so on. The variable for age cannot be forward filled as age increases over time. In this case their age is filled in with linear interpolation (Zhang, 2016). After deterministic correction only truly missing data values are left that require imputation.

Composites

Some variables required for estimating MINOS pathways do not have corresponding variables in Understanding Society and must be derived as composites based on available Understanding Society data to ensure a complete input population (Wei et al., 2022). This section describes the derivation several composite variables used in the MINOS model estimating income policy.

Household Disposable Income - the key exposure in MINOS - is a derived variable according to its OECD formula (Eurostat-Glossary, n.d.; Anyaegbu, 2010). Using the Understanding Society variables listed in Table 4.4, household income is calculated by taking the household net income, subtracting monthly household rent, council tax, and mortgage costs , and adjusting it by the OECD equivalence factor (to account for household composition and number of children). This disposable income is then finally adjusted for inflation using consumer price index data (Johnson, 2015) to provide disposable income in terms of 2020 £s. The household income is the same for each member of the household. The final disposable income formula is then

$$\begin{aligned}\text{disposable income} &= (\text{Net Household Income} - \text{Housing Costs}) \\ &\times \left(\frac{\text{CPI}}{\text{OECD Equivalence}}\right).\end{aligned}$$

There is no direct variable for material housing quality in Understanding Society. SIPHER research has suggested (Aki Tsuchiya, Guoqiang Wu, 2021) that a suitable proxy can be generated according to the presence of certain household appliances and the ability to heat the home adequately. Individuals are assigned into three tiers based on possession of five household appliances and adequate heating, given in Table 4.4. These appliances are divided into core items (fridge freezer, washing machine) and heating that are essential to comfortable living standards, and a second set of luxury items (microwave, tumble dryer, dishwasher) which are non-essential but represent an improved housing quality. Tier 1 is the lowest tier of housing quality for individuals missing at least one core item, tier 2 is for individuals with all core items but some missing the second set of luxury items, and tier 3 is for individuals with all core and luxury items.

Neighbourhood Safety is also derived. Research undertaken by the SIPHER consortium has developed a proxy (Aki Tsuchiya, Guoqiang Wu, 2021) for safety based on the frequency of several crimes occurring in a neighbourhood given in Table 4.4. Individuals are assigned a state from 1, 2, or 3 based on the frequency of these crimes in their neighbourhood. A value of 1 is assigned for the worst state, where any of the listed crimes are 'Very common' or 'Fairly common' in the neighbourhood. A value of 2 is assigned where any crime is 'not very common', and 3 is assigned for the best case scenario where all crimes are 'not at all common'.

Nutrition Quality is estimated based on the frequency and number of servings of both fruit and vegetable consumed by a household in an average week (French et al., 2019). This is a known measure of nutrition quality known to be influenced by disposable income and improve physical and mental health (French et al., 2019). The variables from Understanding Society used to create this proxy are listed in Table 4.4 based on the frequency and quantity of fruits and vegetables eaten. The number of days eating fruit or vegetables is multiplied by the number of portions on an average day. These two values are then added together producing total fruits and vegetables consumed as a measure of nutrition quality.

$$\begin{aligned}\text{Nutrition Quality} = &(\text{Days eating fruit} \times \text{Portions per day})\\ &+ (\text{Days eating veg} \times \text{Portions per day})\end{aligned}$$

Hourly wage is another derived variable, based on US variables in Table 4.4. If the variable *basrate* (Basic pay hourly rate) is present, the hourly wage is set to this value. If not (I.E. for salaried employees), the hourly wage is calculated based on the usual gross pay per month and the number of hours normally worked in a week (multiplied by 4.2 for average number of weeks in a month), according to the following formula

$$\text{Hourly Wage} = \frac{\text{Gross pay per month}}{(\text{Hours worked per week} \times 4.2)}.$$

Complete Case Correction

Final data correction involves complete case analysis (Jakobsen et al., 2017; Gilbert and Troitzsch, 2005) of several critical variables that are required for MINOS to run. Any observations missing these key variables provided in Table 4.3 are removed from the dataset. The structure of remaining missing data for these variables are summarised in Figure 4.3. All variables excluding loneliness and education state have less than 5% missing data suggesting they are suitable for complete case analysis (Jakobsen et al., 2017). For the remaining two variables testing of the missing at random (MAR) assumption (Sterne et al., 2009) is required analysing whether missing data is conditional on other available data. While these is no definitive formal test for the MAR assumption analyses of potential bias in data is required (Sterne et al., 2009; Jakobsen et al., 2017). This requires assessment of variables used in MINOS for variables that are correlated with loneliness and education state as well as prevalence of missing values in these states (Sterne et al., 2009). A pairwise correlation plot in Figure 4.4 providing correlations between all MINOS variables. Loneliness shows correlation $\rho > 0.2$ with age, SF-12 MCS, and marital status. Education state show correlation with only age. Spineplots (Prantner, 2011) can visually assess the differences in distributions for these correlation variables given missing or not missing loneliness and education state variables. Overall these plots show limited distributional differences between SF-12 MCS, age, and marital status distributions when loneliness is missing. Similarly there is no difference in age for missing education state. Additional statistical testing for the MAR assumption are performed using logistic regression (Nakagawa, 2015). A response variable for loneliness and education state (0 is data is missing, 1 if it is missing) is regressed upon using all available MINOS variables. This logistic model for loneliness suggested only a widowed marital status ($p = 0.0436$) was a statistically significant indicator of missing data. For education state only age was statistically significant ($p = 2.09e - 4$). All significant coefficients have small values < 0.001 suggesting small influence on the probability of missing values. There is overall limited evidence that missing values of loneliness and education state follow an MAR missingness pattern suggesting complete case is viable.

The following chapter intends to correct these missing values using multiple imputation techniques particularly the MICE algorithm. Data matching methods may also be necessary in order to include variables not in Understanding Society. Summary statistics for key discrete and continuous variables in appendix Tables 8.1 - 8.2 below.

Replenishing Population

In order to maintain population size and adjust the population to match demographic change projected in to the future, a replenishing cohort is added to MINOS model in each year. This replicates methods used in similar studies (Archer et al., 2021; Birkin, Wu and Rees, 2017) and avoids a reduction in sample size and a population that becomes unrepresentative due to the processes of ageing and mortality. This population consists of all $15 - 16$ year old individuals within the Understanding Society sample in the year 2020. Individual sample weights for this population are updated every year according to age, sex, and ethnicity projections from the NEWETHPOP project (Rees et al., 2017) and Office for National Statistics midyear population estimates (Nash, 2019). Population weights are re-sampled according to the prevalence of their age, sex, and ethnicity within the current and projected future population. For example, if their demographic is twice as large in 2040 their sample weight is doubled.

Variable Name (Understanding Society Code)	Description	Deterministic Correction?	LOCF Imputation?	Composite?	Complete Case?
pidp	Personal Identifier	N	N	N	N
hidp	Household Identifier	N	N	N	N
sex	Biological sex	N	FB	N	N
ethnicity (racel_dv)	Ethnicity	N	FB	N	N
age (dvage)	Age	N	L	N	Y

region (gor_dv)	Government Region	N	N	N	Y
education_state (qfhigh_dv)	Highest Qualification	N	F	N	Y
nkids (nkids_dv)	Number of children aged $0-16$	N	N	N	N
fridge_freezer (cduse5)	Has a fridge freezer	N	N	housing_quality	Y
washing_machine (cduse6)	Has a washing machine N	N	N	housing_quality	Y
tumble_dryer (cduse7)	Has a tumble dryer	N	N	housing_quality	Y
heating (hheat)	Has comfortable heating	N	N	housing_quality	Y
dishwasher (cduse8)	Has a dishwasher	N	N	housing_quality	Y
microwave (cduse9)	Has a microwave	N	N	housing_quality	Y
hh_composition (hhtype_dv)	Type of house (detached etc.)	N	N	N	N
marital_status (marstat_dv)	Marriage status	N	N	N	Y
house_tenure (tenure_dv)	Housing tenure (rent/mortgage etc.)	N	N	hh_income	N
hhsize	People in household	N	N	N	N
car_crime (crcar)	Car crime nearby	N	N	neighbourhood_safety	N
drunks (crdrnk)	Drunken disorder nearby	N	N	neighbourhood_safety	N
muggings (crmugg)	Muggings nearby	N	N	neighbourhood_safety	N
racial_abuse (crrace)	Racial Abuse nearby	N	N	neighbourhood_safety	N
teenagers (crteen)	Teenage disorder nearby	N	N	neighbourhood_safety	N
vandalism (crvand)	Vandalism nearby	N	N	neighbourhood_safety	N
council tax band (ctband_dv)	Council tax band	N	N	hh_income	N

hh_netinc (fihhmnnet1_dv)	Net household income	N	N	hh_income	N
oecd_equiv (iegmoecd_dv)	OECD Equivalence Scaling	N	N	hh_income	N
hh_rent (rentgrs_dv)	Household rent	N	F	hh_income	N
hh_mortgage (xpmg_dv)	Household mortgage payment	N	N	hh_income	N
yearly_ electricty (xpelecy)	Electricity expenditure	N	F	yearly_energy	Y
yearly_gas (xpgasy)	Gas Expenditure	N	F	yearly_energy	Y
yearly_gas_electric (xpduely)	Gas and electricity combined expenditure	N	F	yearly_energy	Y
yearly_oil (xpoily)	Fuel oil expenditure	N	F	yearly_energy	Y
yearly_other_fuel (xpsgly)	Other fuel expenditure	N	F	yearly_energy	N
hourly_rate (basrate)	Basic hourly pay	N	N	hourly_wage	N
gross_paym (paygu_dv)	Gross monthly pay	N	N	hourly_wage	N
gross_pay_se (jspayg)	Gross self employed monthly pay	N	N	hourly_wage	N
job_hours (jbhrs)	Weekly job hours	N	N	hourly_wage	N
job_hours_se (jshrs)	Self employed weekly job hours	N	N	hourly_wage	N
job_inc (jspayu)	Income from job/business	N	N	hourly_wage	N
jb_inc_per (jspayw)	Job pay period	N	N	hourly_wage	N
job_sector (jbsect)	Public or private sector	Yes. Set unemployed values from −8 to 0.	N	hourly_wage	N
emp_type (jbft_dv)	Full or part time employed	N	N	N	N

job_sec (jb-nssec8_dv)	NSSEC Employment code	Yes. Set unemployed values from −8 to 0.	N	N	Y
ncigs	Weekly cigarettes consumed	Yes set non-smokers from −8 to 0.	N	N	N
loneliness (sclonely)	Loneliness frequency	N	N	N	N
SF12_MCS (sf12mcs_dv)	Short Form 12 Mental Component Score	N	N	N	N
fruit_day (wk-fruit)	Days per week fruit is eaten	N	N	nutrition_quality	N
veg_day (wkvege)	Days per week vegetables are eaten	N	N	nutrition_quality	N
fruit_per_day (fruitamt)	Fruit eaten per day	N	N	nutrition_quality	N
veg_per_day (vegeamt)	Vegetables eaten per day	N	N	nutrition_quality	N

Table 4.3: Imputation Strategies for each variable used in MINOS.

Variable Name	Description
Disposable Income	
fihhmnnet1_dv	Household net income
rentgrs_dv	Gross monthly rent.
xpmg_dv	Gross monthly mortgage.
ctband_dv	Council tax band.
ieqmoecd_dv	Modified OECD equivalence scale.
Housing Quality	
cduse5	Do you have a fridge freezer in your household?
cduse6	Do you have a washing machine in your household?
cduse7	Do you have a tumble dryer in your household?
cduse8	Do you have a dishwasher in your household?
cduse9	Do you have a microwave in your household?
hheat	Do you have sufficient heating in your household?
Neighbourhood Safety	
crburg	Are there burglaries in your neighbourhood?
crcar	Is there car crime in your neighbourhood?
crdrnk	Is there drunken disorder in your neighbourhood?
crmugg	Are there muggings in your neighbourhood?
crrace	Are there racial attacks in your neighbourhood?

crteen	Are there issues with teenagers in your neighbourhood?
crvand	Is there vandalism in your neighbourhood?
Nutrition Quality	
wkfruit	How many days per week do you eat fruit?
fruitamt	How many portions of fruit do you eat on a day when you eat fruit?
wkveg	How many days per week do you eat vegetables?
vegeamt	How many portions of vegetables do you eat on a day when you eat veg?
Hourly Wage	
basrate	Basic bay hourly rate
paygu_dv	Usual gross pay per month: current job
jbhrs	Number of hours normally worked in a week

Table 4.4: Understanding Society variables used for generation of derived composites in MINOS.

4.3.6 Transition Probability Models

This section contains detail on the implementation of all transition probability models (Rutter et al., 2009) used in MINOS that are specified in Table 4.2 and Figure 4.1 as part of a standard series of modules (Tikka et al., 2021). For each transition probability model the model type is defined dependent on outcome variables and user requirements. Full model formulae are providing along with a short literature review on choice of independent variables and brief model diagnostics. Coefficient tables for each transition model are also provided in Section 8.1.2 in the appendix.

GLMMs

Continuous variables are estimated using Generalised Linear Mixed Models (GLMMs) (Breslow and Clayton, 1993). GLMMs build on ordinary linear regression allowing for application of fixed and random effects as well as non-normal distributed outcome data (Breslow and Clayton, 1993; Bates, 2014; Gelman, 2005). This allows for inclusion of multiple years of observations from repeated individuals when calibrating transition probability models using a much larger dataset and accounting for individual longitudinal heterogeneity in behaviour (Burgard et al., 2019). For all GLMM models here only random intercept effects are used for simplicity providing each individual random intercept scores drawn from a Normal distribution. All models are used creating the 'lme4' package (Bates et al., 2014).

Household disposable income data are highly skewed requiring use of a Gamma-Normal GLMM model (Hanum et al., 2018). This again assumes Normal distributed random intercepts but assumes outcome data are drawn from a Gamma distribution with strictly positive right skewed values. The Gamma-Normal GLMM has general formula

$$E[y|X, Z] = exp(X\beta + Zb)$$

with response variable y, fixed effects X and random effects Z design matrices, and regression coefficients for fixed β and random b effects. Note coefficients for Gamma-Normal models are interpreted differently to the standard OLS model. A coefficient of 0.1 indicates that the coefficient multiplicatively increases the outcome by $e^{0.1} = 10.5\%$. Estimation of household income is performed using a Gamma-Normal GLMM using the formula

$$\begin{aligned} hh\_income = exp((X\beta) + (Zb)) = exp((&hh\_income\_last + hh\_income\_diff + age + age^2 + age^3 \\ &+ sex + ethnicity + region + education\_state + NSSEC) + (intercept|pidp)). \end{aligned}$$

This formula utilises common socio-demographic predictors used to estimate household disposable income (Ali et al., 2019; Mallet and Weale, 2018) including age, sex, ethnicity, region, education and socio-economic code. This model also utilises random intercepts by individual identifiers (pidp). A lagged dependent variable (McLay et al., 2015) is also used estimating current income using previous income states. Model coefficients are provided in Table 8.3.

To estimate SF-12 MCS score a Gamma-Normal GLMM is also used. SF-12 MCS data are reflected converting them from left-skewed to right-skewed and translated to be strictly positive as required for Gamma distributed data. The SF-12 MCS model has formula

$$\begin{aligned} log(SF12\_MCS) = exp((&log(previous\_SF\_12\_MCS) + SF\_12\_MCS^2) \\ &+ sex + ethnicity + age + age^2 + housing\_quality \\ &+ hh\_income + neighbourhood\_safety + nutrition\_quality \\ &+ smoker) + (intercept|pidp)). \end{aligned}$$

Independent variables used to predict SF-12 MCS are sourced from existing studies on UKHLS data (Parra-Mujica et al., 2023). Independent variables include typical age, sex, and ethnicity. Additional variables have also been included as part of SIPHER analysis on predictors of mental well-being including housing quality, neighbourhood safety, nutrition, and smoking (Aki Tsuchiya, Guoqiang Wu, 2021; Consortium, 2023). Lagged dependent variables and random intercepts are also used (Parra-Mujica et al., 2023). Model coefficients are provided in Table 8.5. Additionally, as SF-12 MCS values are reflected in order to be right skewed effect signs are inverted. A coefficient of $e^{0.1} = 10.5\%$ indicates a *decrease* in SF-12 MCS score.

Nutrition quality data is approximately Normal distributed hence a Normal-Normal GLMM (Hanum et al., 2018) is used with formula

$$y = X\beta + Zb + \varepsilon.$$

This Normal-Normal GLMM assumes that responses y and random effects b are both Normal distributed with mean and variance parameters to be estimated. The formula for estimating nutrition quality is given as

$$\begin{aligned} nutrition\_quality = & \ previous\_nutrition\_quality \\ & + age + sex \\ & + education\_state + ethnicity + hh\_income \\ & + ncigs + SF12\_MCS + (1|pidp). \end{aligned}$$

Nutrition quality independent variables are derived from social determinants of nutrition and obesity literature (Javed et al., 2022; French et al., 2019; Hughes and Kumari, 2017). Variables for common demographics including age, sex, education, and ethnicity are provided. Household income, previous health conditions (SF-12 MCS), smoking, and lagged nutrition quality are also used as recommended by the literature. This model utilised random intercepts due to multiple longitudinal waves of data present and was found to improve model fit. Model coefficients are provided in Table 8.4.

Zero Inflated Poisson (ZIP) Models

Some count variables such as the number of cigarettes smoked per month are zero inflated (Salonen et al., 2019). The majority of the population do not smoke and have value 0 skewing outcome data. In order to correctly estimate cigarette consumption a zero-inflated Poisson regression (Jackman et al., 2015) is typically used in microsimulation (Salonen et al., 2019). This is a two stage model that estimates the probability a person smokes at all and the number of cigarettes a month they would smoke. Random sampling is used to determine whether a person smokes or not depending on their probability. If they do not smoke they are assigned value 0. If they do smoke they are assigned their predicted counts value. ZIP models are implemented using the pscl package in R that utilised the EM algorithm (Jackman et al., 2015). The ZIP model requires two sets of predictors estimating whether a person smokes at all (zero model) and given they do smoke how many cigarettes do they smoke (counts model) respectively. These are referred to as zero and counts models with the following formulae and coefficients.

Two formulae are provided estimating cigarette (ncigs) consumption for the zero model

$$\begin{aligned} ncigs\_zero = & \ ncigs\_zero\_last + previous\_smoker + ethnicity + age + sex + NSSEC \\ & + hh\_income + labour\_state + SF12\_MCS \end{aligned}$$

and counts model

$$
\begin{aligned}
ncigs\_count = {} & previous\_ncigs + age + sex + education\_state \\
& + SF12\_MCS + NSSEC + labour\_state + hh\_income + ethnicity.
\end{aligned}
$$

For both models prediction is based on age, sex, and ethnicity demographics all found to significantly affect consumption (Chen et al., 2019; Chandola et al., 2004). Income, employment, socio-economic state (NSSEC) and previous smoking behaviour are all also included when predicting both states. Additinoal random effects modelling is desirable (Chandola et al., 2004) but not possible with the pscl R. package used to fit these models. Model coefficients are provided in Table 8.6.

Cumulative Link Modelling for Ordinal Data

There are three ordinal likert scales estimated in MINOS for housing quality, neighbourhood safety and loneliness. The simplest approach to estimate next state for ordinal data is using cumulative link models (CLMs) (Christensen, 2018). This approach is very similar to logistic regression but extends prediction to some generic number of states. It is very useful in microsimulation for its speed and output prediction of probability matrices that can be immediately used to sample next state. Extensions to cumulative link mixed modelling is desirable but was found to be unstable for large datasets. CLMs implemented using the 'ordinal' package in R (Christensen, 2018). Coefficients and formulae for the three ordinal outcome variables are provided below.

Estimating the housing quality composite as a proxy for material deprivation is performed using a CLM model with formula

$$
\begin{aligned}
housing\_quality = {} & housing\_quality\_last + age + age^2 + sex \\
& + SF12\_MCS + ethnicity + hh\_income + hh\_income^2 \\
& + NSSEC + labour\_state + housing\_tenure.
\end{aligned}
$$

Estimation of household material deprivation uses common demographic indicators including age, sex, and ethnicity (Whelan and Maitre, 2012). Income related terms including net household income, housing tenure, socio-economic class (NSSEC), previous housing tenure, and labour state. Mental-wellbeing is also used due to significance of similar variables including global happiness index (GNDH) (Whelan and Maitre, 2012). Model coefficients are provided in Table 8.7.

Loneliness is also estimated using a CLM model with formula

$$\begin{aligned} loneliness = {} & previous\_loneliness + age + sex + SF12\_MCS \\ & + education\_state + labour\_state + hh\_income \\ & + hh\_comp + marital\_status + ethnicity. \end{aligned}$$

Demographic predictors including age, sex, ethnicity, marital status (living alone), and education have shown to significantly predict loneliness before and after the coronavirus pandemic (Bu et al., 2020). Economic variables for household income and labour state are also recommended. Previous loneliness and SF-12 MCS health scores have also been included. Model coefficients are provided in Table 8.9.

Finally, a CLM model has been fitted estimating neighbourhood safety using formula

$$\begin{aligned} neighbourhood\_safety = {} & previous\_neighbourhood\_safety + age + sex \\ & + NSSEC + ethnicity \\ & + hh\_income + housing\_quality + region + loneliness + housing\_tenure. \end{aligned}$$

Prediction of neighbourhood safety requires estimating the frequency anti-social events (Warr et al., 2009) a household experiences. This is difficult to do without specific spatial data but there are studies estimating the social determinants of these events (Livingston et al., 2014; Baum et al., 2009). Deprivation is the strongest predictor of crime rates including material deprivation (housing_quality) as well as socio-economic state (NSSEC) and household income. Age and sex variables are often included in studies but are often not statistically significant. Other demographic variables including ethnicity and housing tenure, particularly prevalence of short-term social renting, are also strong indicators of crime rates. Loneliness is used as an indicator of mental well-being and social inclusion. Again previous neighbourhood safety is also included. Model coefficients are provided in Table 8.8

NNET Single Layer Neural Networks Estimating Categorical Data

Discrete categorical variables such as labour and education states are difficult to predict. The simplest solution is to use a multinomial single layer neural network provided by R's multinom package (Templ et al., 2017). Similar to CLM models this will also provide a vector of probabilities for belonging to each possible state that is sampled from to generate next state. implemented using the nnet (Ripley and Venables, 2016) package in R. A variety of similar models are used to estimate discrete state movement in microsimulation such as multi-state markov models (Krijkamp et al., 2018) but the nnet package allowed for speed particularly fitting higher dimensional models with a large number of independent variables.

An nnet model is used to estimate UK labour state with formula

$$\begin{aligned} labour\_state = {} & previous\_labour\_state + financial\_situation + age + sex + ethnicity + region \\ & + education\_state + marital\_status + housing\_tenure + nkids + SF\_12MCS. \end{aligned}$$

Estimating labour state requires predicting employment participation as well as reasons for not working including being a carer, in education, retired, or otherwise economically inactive. Demographic variables used to predict labour state include age, sex, ethnicity, region, education state, marital status, housing tenure, and number of children (Quarina, 2017). Previous health state SF-12 MCS and previous labour state are also used (Quarina, 2017).

Mortality and Fertility Rate Tables.

Rate tables are also typically used to transition discrete states forwards in time (Rutter et al., 2011) but have substantially larger data requirements compared with nnet models. For transitioning fertility and mortality states we have pre-existing lookup rate tables available from the NEWETHPOP project (Rees et al., 2017). These tables provide a rate of birth or death for 2020 based on age, sex (for mortality only), ethnicity, parity (for fertility only) and region. Parity is simply defined as a mothers current number of children. For each individual the rate of birth and death is converted into a probability for the given yearly interval. A random draw is then used to determine whether an individual changes state. For mortality individuals are assigned a 'dead' state and removed from the model. For fertility the number of children in the household assigned to the mother (nkids) is incremented by 1. These rate tables are preferable to predictive methods used in most of MINOS as they exhaustively cover all possible states. However, they have large data requirements and cannot be derived for most variables in Understanding Society. When children reach 16 years of age within a household nkids is then decremented by 1 reducing the number of children in the household.

4.3.7 Uncertainty in Transition Probability models

Dynamic microsimulation modelling is subject to a number of sources of uncertainty due to input parameters (Sharif et al., 2012; Petrik et al., 2020; Kopec et al., 2010). Testing how much these uncertainties influence microsimulation prediction is essential for unbiased estimation of variance in policy outcomes and making potential modelling caveats clear to model users where predicted results may fail due to underlying model assumptions. This section briefly reviews sources of uncertainty in microsimulation and methods used in this thesis to address them.

Uncertainty due to Input Data

Input data used to calibrate transition models can be a large source of uncertainty (Sharif et al., 2012; Petrik et al., 2020). Bias in sample data as well as bias due to missing data correction methods can cause poor prediction of national-scale behaviour particularly forwards in time and under counterfactual scenarios.

There are two primary sources of uncertainty due to input data for the MINOS model. Uncertainty due to input UKHLS source data and population outliers is tested using five-fold cross validation below in Section 4.3.8. Uncertainty due to missing data is limited due to application of deterministic missing data correction techniques including complete case analysis.

Uncertainty due to Model Coefficients

Transition probability models used in microsimulations are often fitted with model coefficients that are themselves random variables (O'Hagan et al., 2007; Petrik et al., 2020). Different values of these coefficients may exhibit large differences in projected behaviour across a full microsimulation model run that must be included in sensitivity analyses.

Sensitivity analysis is used varying each transition model coefficient systematically observing how microsimulation output varies (Burgard et al., 2019). This can be done systematically adjusting one coefficient at a time but this often leads to an exponential increase in the number of microsimulation model runs required (O'Hagan et al., 2007). Instead sensitivity analysis may draw model coefficients from a probability distribution such as a multivariate normal distribution identifying and potential outliers (Petrik et al., 2020; Richiardi and Richardson, 2017). This random sampling of model coefficients is also performed during microsimulation model runs to determine any change in confidence intervals for model output. This thesis uses random sampling of coefficients for regression style transition probability models during microsimulation model runs. Further noise may be added neural network models used to estimate discrete data but this is left to future work.

For uncertainty in model coefficients all regression models used have randomised coefficients. For each of the 100 MINOS runs models draw coefficients according to a multi-variate Normal distribution and use these coefficients throughout the model run. Further analyses determining the effect of changing each individual coefficient one at a time are needed but are computationally intensive (Petrik et al., 2020).

Uncertainty due to Model Structure

Uncertainty quantification due to model structure assesses the choice of independent variables used in regression models. Iterative refinement of selected variables is updated combining discussion developing causal loop diagrams into transition models described in Section 4.3.2 combined with application of variables selection methods such as the LASSO in Chapter 3 showed limited evidence to reject any independent variables defined by literature review.

Uncertainty due to Monte Carlo Error

Many transition models produce point estimates of future state. For individuals with the same characteristics they will be predicted identical future states. In order to prevent this Monte Carlo random noise is typically added to predicted states to induce stochastic behaviour (Sharif et al., 2012; Petrik et al., 2020; Wolf, 2001). This also typically helps prevent regression to the mean in microsimulation prediction preventing underestimation of population variance over time. Noise is typically drawn from simple distributions such as Normal random variables (Wolf, 2001) but more complex MCMC methods are also used (Sharif et al., 2012).

In this thesis Normal distributed random noise is added to prediction of continuous variables. All discrete CLM models calculate probability tables for belonging to each discrete state that are drawn from. Continuous variables have added Normal noise (Sharif et al., 2012). Probability of birth and deaths calculated from rate tables in vivarium (Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA, 2023) are treated as Bernoulli variables such that an individual dies with some probability p and lives otherwise.

Small amounts of Normal noise are added to predicted GLMM output in order to prevent overfitting and simulate Monte Carlo error in model coefficients (Koehler et al., 2009; Sharif et al., 2012).

Variance Reduction.

Often accounting for all sources of uncertainty can result in large confidence intervals for microsimulation output (Zaidi and Rake, 2001). In an effort to reduce large computation time microsimulation models use variance reduction techniques to 'seed' part of the model run and reduce confidence intervals (Zaidi and Rake, 2001; Andreassen et al., 2020; O'Hagan et al., 2007; Petrik et al., 2020).

One common example is use of common random numbers (Murphy et al., 2013) which partially seeds microsimulation behaviour making individuals by seeding some transition probability models. If a mortality module is seeded such that they always die at the same time in each microsimulation model run this can greatly reduce uncertainty albeit producing further model assumptions that must be clarified in results (Wolf, 2001). This thesis does not use CRN seeding in any microsimulation models due to acceptable confidence intervals but it is implemented as part of the vivarium framework (Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA, 2023) and is applied in application elsewhere.

4.3.8 Validation

Validation of the MINOS input population that ensures key variables match Local Authority level aggregates is well described elsewhere (Höhn et al., 2024). When projecting the MINOS population forwards in time and under intervention it is difficult to perform any kind of validation without direct comparison of the same policies and spatio-temporal context in other simulation literature (Rutter et al., 2011).

Cross-Validated Nowcasting

Nowcasting in microsimulation requires predicting from the past to the present time allowing for comparison of predicted output against real data (O'Donoghue and Loughrey, 2014). Nowcasting is then performed over a six year interval from $2014 - 2020$ where all required UKHLS variables are available to run the MINOS model.

In order to perform nowcasting five-fold cross validation is used (McLay et al., 2015). UKHLS data is split into five equally sized parts each containing repeated observations from the same individuals. One fifth of UKHLS data is set aside for testing, fitting transition models to the remaining training data. This test data is then projected from 2014 to 2020 using transition models fitted on training data. This process is repeated five times excluding each portion of the initial UKHLS data sequentially. This produces a set of five projected UKHLS populations in 2020 that are compared against real data to assess MINOS prediction quality. Example cross-validated nowcasting for key MINOS variables are provided in Section 4.4.

Handover Plots

Many microsimulation models often use 'eye test' (Archer et al., 2021) visualisation observing how the population behaves when projecting into the future ensuring results look sensible and reliable. Example visualisation for key yearly energy and SF-12 MCS variables are provided in Figures 4.5a - 4.5b showing reasonable preservation of statistical moments and linear trends over time. Handovers for key MINOS variables including income and SF-12 MCS are also provided in Section 4.4. Further visualisation for validation of other variables used in MINOS is provided in online repositories.

4.3.9 Policy Intervention Scenarios

Three interventions are presented: the Child Payment Uplift, Living Wage, and the Energy Price Cap Guarantee. Intervention results are compared to a baseline model run.

Baseline

The baseline intervention assumes no change in circumstances for the UK population. Household disposable income remains at 2020 levels indefinitely. This is useful as a benchmark scenario that emulates status-quo and intentionally does not make any dynamic changes (e.g. inflation adjustment) that would make it more difficult to compare to. All other modelled interventions are compared against this baseline to determine the net change in mental well-being that results from the specific policy under investigation.

Child Payment Uplift

As part of the 2017 Poverty Act (Scottish Government, 2017) the Scottish government launched the Child Payment Scheme providing households in receipt of income benefits schemes with children under 6 years old £10 per week (Scottish Government, 2017). This payment expanded in November 2022 to £25 per week covering eligible children up to 16 years of age (Scottish Government, 2022) on universal credit and related legacy benefits. This policy is implemented in MINOS for the whole of the UK by selecting all households where a member is in receipt of universal credit and where there are children present. In these households, £25 per week per child under 16 years old is added to the total disposable income. Model results are produced for every adult within that household. By implementing the policy in this way, we can model the impact on SF-12 MCS for a specific sub-population (given our model jump off of 2020) and also assess the improvements that might be seen if the policy were extended to the rest of the UK. As there are proposals to increase the Scottish Child payment to £30, £40, and beyond (Pybus, 2023; Congreve et al., 2022) we also implement the Scottish Child Payment in increasing increments of £5 from £25 to £50 per child per week. This work intends to determine if these more expensive interventions are cost effective and show significant improvement in SF-12 MCS vs the standard £25 child payment.

Living Wage

The Living Wage Foundation (Prowse and Fells, 2016; Cominetti and Murphy, 2022) suggests that the current minimum wage of £10.45 for those over 23 years old in the UK is not sufficient to maintain a good quality of life. A voluntary minimum hourly wage of £12.00 (£13.15 in London) has been proposed by the Living Wage Foundation (Cominetti and Murphy, 2022) as necessary to sustain employees and their families. This intervention is operationalised by identifying individuals who are employees and paid an hourly wage which is below the living wage. Their income is boosted to exactly the threshold. This income is then translated to household income and becomes available to all members of the household in which the individual resides, which is the population upon which the intervention is applied.

Energy Price Cap Guarantee

The Energy Price Cap Guarantee (EPCG) (Broadbent et al., 2023) intervention is designed to emulate the role that government support has in protecting mental health during an ongoing energy cost crisis (UK Fuel Poverty Monitor, 2022). Between our model jump-off in 2020 and 2022, energy prices rose sharply and were capped by the UK Government so that the average household would pay around £2,500. This intervention applies three scenarios where (1) a baseline with no energy price change provides a status-quo where household energy costs did not increase; (2) the sharp increase in energy cost is passed on to households with no government support; and (3) both the price increase and EPCG support is applied. The main utility of the results under this scenario is in comparing the protection offered by government support vs the impact of a price increase with no support. Given that prices remain high (the current cap is around £3,500 for the average household) the scenario maintains support until 2035 to assess the potential for long term mitigation of the detriment to mental health.

4.3.10 Outcomes and Visualisation

In order to describe and measure the effect of policy scenarios on the UK population several visualisation and statistical testing methods are used. This section outlines testing microsimulation model outcomes and visualisation tools to demonstrate change in microsimulation data due to counterfactual intervention of single and multiple model runs.

Individual Populations

Demonstrating the effect of a policy for a single microsimulation model run can be useful for providing more detailed descriptions of change in variable distributions and individual trajectories that are more demonstrate to demonstrate for multiple model runs.

For continuous variables including disposable income and SF-12 MCS, traditional boxplots are used to measure change in distribution over time (Archer et al., 2021). Measuring median and interquartile range is a simple, effective measure to determine if population median and variance follow consistent trends over time. Standard lineplots (Zaidi and Rake, 2001) are also provided demonstrating change in just median or mean over time for simple results that can be applied to multiple model runs. Alternatively Kernel Density Plots and Ridgeline Plots (Thrun et al., 2020) are used estimating the entire distribution of continuous variables particularly with strong skewing seen in household income.

For discrete variables stacked bar-plots (Zhou et al., 2022; Salonen et al., 2019) are used showing the absolute and percentage values of individuals belonging to each state over time. Confidence bars can be added to these bar-plots when displaying change for multiple microsimulation runs.

Visualising individual dynamism is also desirable for microsimulation output (Salonen et al., 2019; McLay et al., 2015). Spaghetti plots display individuals trajectories such as income over time. These plots are not formal diagnostics but rather an eye test to again ensure trends in mean and variance over time. Combination with metrics such as PRESS statistics (McLay et al., 2015) can be used as more formal validation to ensure sufficient oscillation and response to intervention. Similar techniques are used in representing spatial migration including chloropleth maps (Zhou et al., 2022). Many other visualisation methods are available but are utilised depending on variable type, paper, and research field requirements (Zinn et al., 2014).

This thesis relies on all of the above methods throughout to display change in health, income, and other covariate behaviour over time. No formal testing is used to compare output from single model runs due to low sample size.

Multiple Populations

Change due to counterfactual scenarios in all microsimulation models is made using a sample of many model iterations (O'Donoghue and Dekkers, 2018). Visualisation methods such as mean line-plots used above are also used for multiple model runs with confidence intervals attached (Archer et al., 2021). For a continuous variable, the mean over time is taken for each of some n model runs. This produces a set of n means that are then used to construct final confidence intervals in line-plots. Similar aggregation can be used for discrete data measuring the percentage of individuals belonging to each state (Archer et al., 2021).

This aggregation is also performed for certain sub-populations of the microsimulation population (Archer et al., 2021). The desired population subset is selected taking aggregates for each model run as normal. These sub-populations may be individuals living in certain geographical areas (Burgard et al., 2021). Mean change by geographic area can then be aggregated further into a grand mean (I.E. a scalar mean of means) and plotted in a chloropleth map useful for differentiating spatial change due to policy (Lomax and Smith, 2017a).

Other subgroups include key priority individuals of a population of interest to policy partners and model users (Meier et al., 2019). One example would be a subset of single mother households when investigating child poverty policy (Scottish Government, 2022). Similarly a treatment on the treated sub-population is also commonly used (Archer et al., 2021). In order to correctly estimate treatment effect on a population treatment on the treatment methods are used. (Archer et al., 2021)In this case only the subset of the population that receives treatment in the counterfactual scenarios are compared directly with those who would have received treatment under baseline. This allows for accurate estimate of policy effect by directly comparing the same subgroup of the population that would have received intervention in a counterfactual but did not receive it in the baseline.

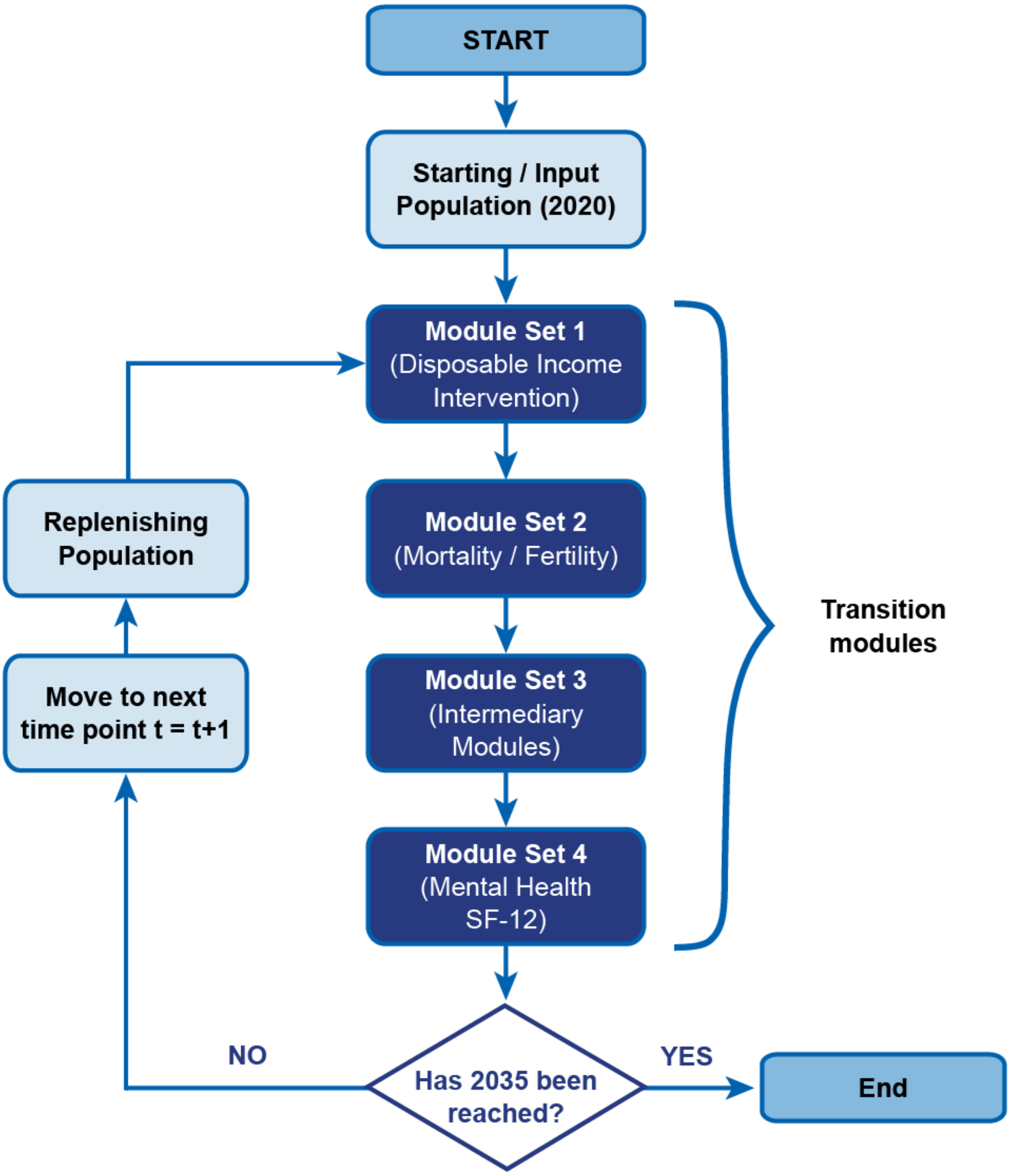


Figure 4.2: An operational flow chart for the MINOS model.

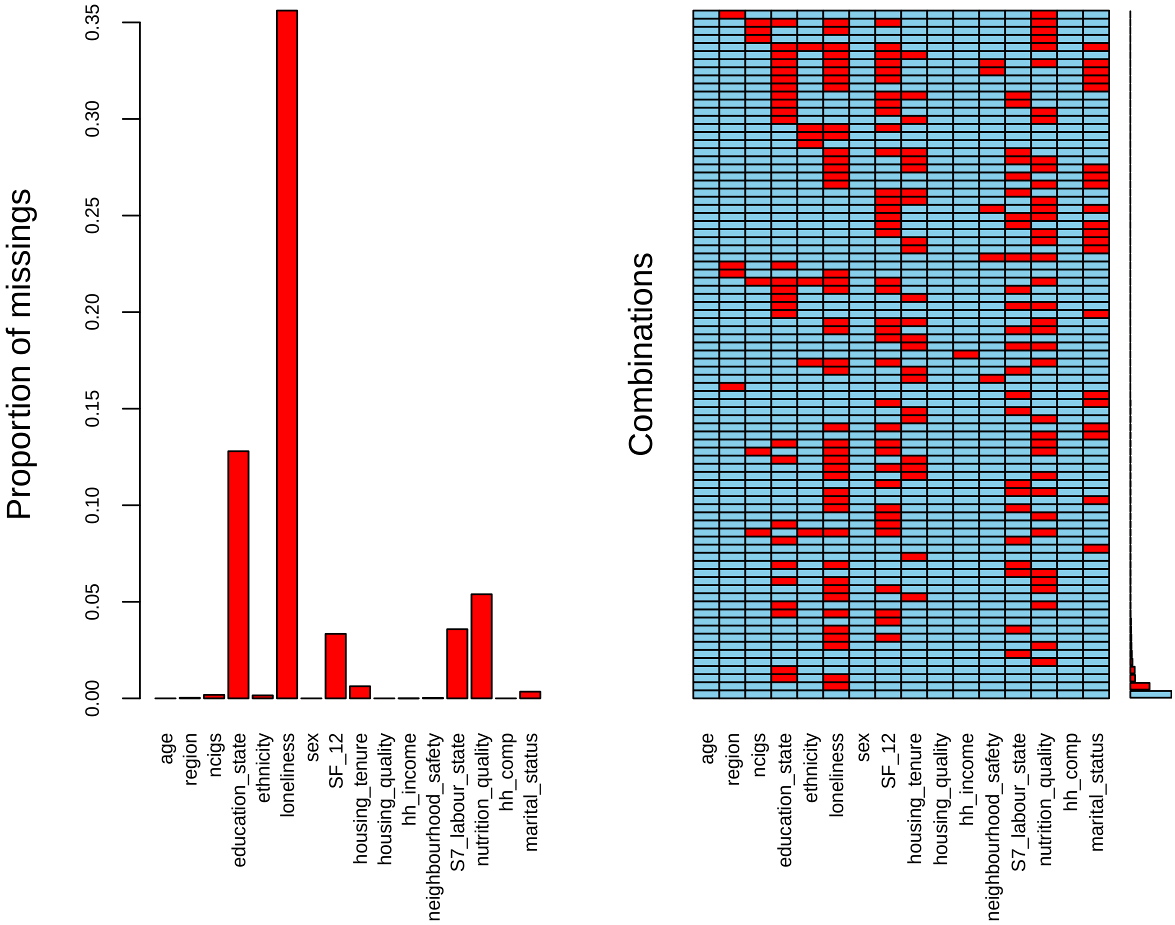


Figure 4.3: A missingness structure (aggr) plot for key MINOS variables after preprocessing. All variables excluding loneliness and education state show less that 5% of missing data making them suitable for complete case analysis.

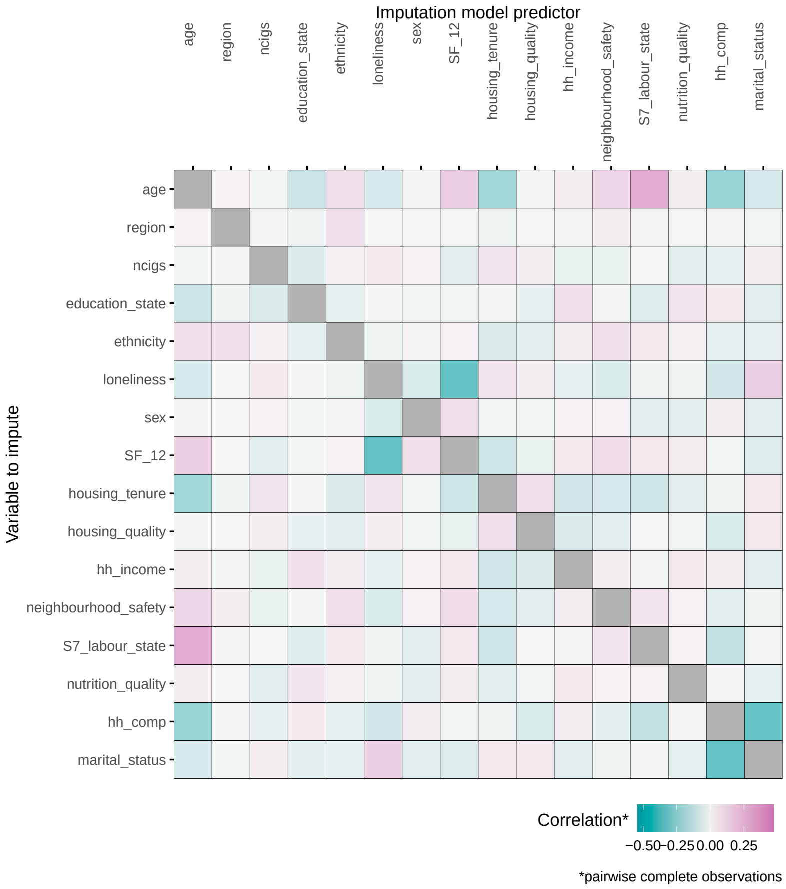
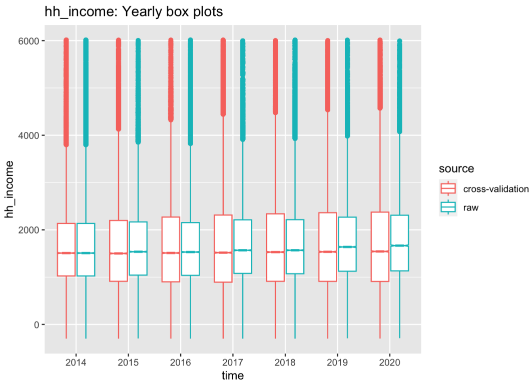
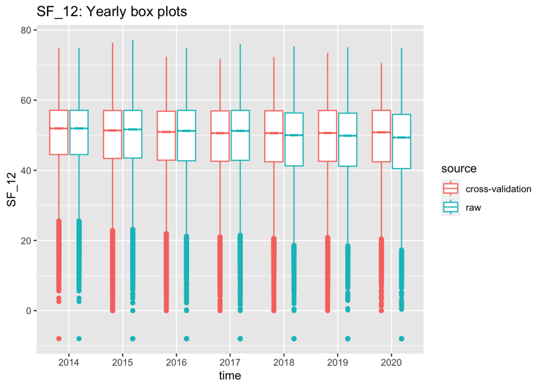


Figure 4.4: Pairwise correlation between key MINOS variables from Table 4.1 used to inform choice of complete case analysis.



(a) Boxplots for Nowcasting Cross Validation for Household Income.



(b) Boxplots for Nowcasting Cross Validation for SF-12 MCS

Figure 4.5: Five-fold cross-validated nowcasting boxplots for key household income and SF-12 MCS variables.

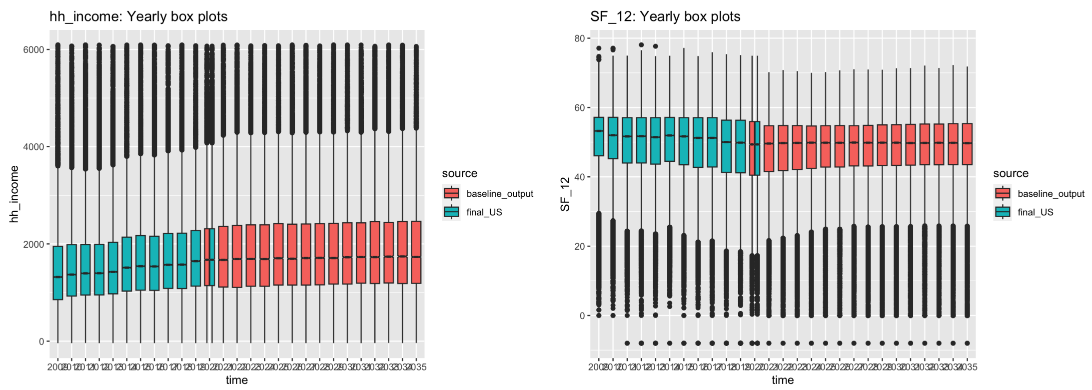
4.4 Validation

Given that we present the results from a simulation in this paper, this section provides validation for two key variables in the MINOS model: the variable which is intervened on (household disposable income) and our key outcome variable (SF-12 MCS) under the baseline model run. Robust validation of this model run provides assurance that variations from the baseline in results for the policy models are a result of the explicit changes made to represent each scenario.

4.4.1 Cross Validation

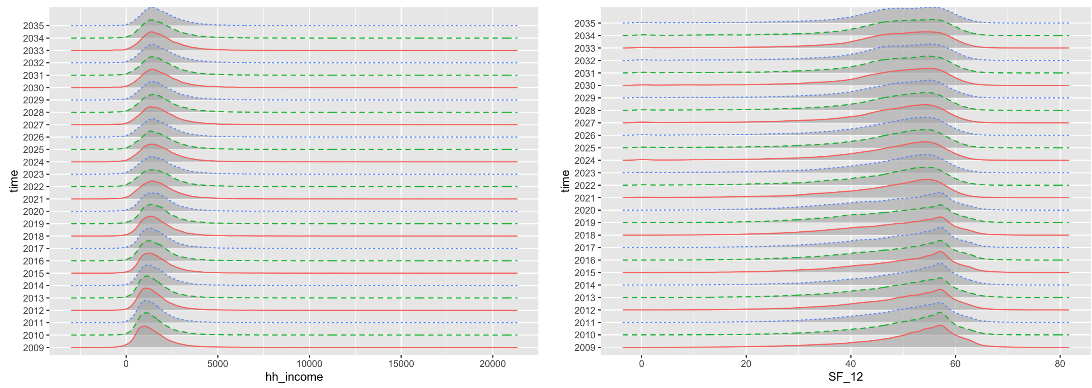
As is undertaken in other microsimulation studies (Archer et al., 2021) we undertake nowcasting that projects data from 2014 to 2020, comparing results with observed data to determine goodness of fit. The year 2014 reflects data availability constraints, with many key variables needed for to estimate MINOS pathways not available in earlier years.

Five fold cross-validation (McLay et al., 2015) is used to test sensitivity of the MINOS model to input data and transition model coefficients. UKHLS data is first split into five equal sets, four of these sets are used to train transition probability models and the fifth test set used as the input population for the simulation. This is repeated five times with different observations in the test data each time, producing five projections that are pooled together for final comparison with real data. Each of these cross validated models is run from 2014-2020 where observed UKHLS data are available, allowing for direct comparison of modelled against observed data. Figure 4.5 shows results of the household disposable income and SF-12 MCS distributions which represent the range of outcomes for all individuals in the model. Median and variance for both variables are well preserved suggesting MINOS sufficiently replicates the distribution of income and is able to predict SF-12 MCS in the baseline model.



(a) Handover Boxplots for household disposable income

(b) Handover Boxplots for SF-12 MCS



(c) Ridgeline Handover plots for household disposable income.

(d) Ridgeline Handover plots for SF-12 MCS

Figure 4.6: Handover plots for MINOS for baseline household disposable income and SF-12 MCS variables.

Handover Plots

Handover plots are used to assess if projected trends are in line with historic data and are useful to assess if the model is behaving as expected (Archer et al., 2021). Income and SF-12 MCS are estimated for the baseline projection between 2020 to 2035 and visualised using boxplots and ridgeline plots (Thrun et al., 2020) in Figure 4.6 to assess the distribution and skew of results. Results shows stable and plausible projection into the future maintaining median, variance and skew for the distribution of both variables. Taken together, the cross-validation and handover plots suggest that for baseline runs (i.e. where no intervention is introduced), MINOS is able to replicate observed data and produce projections to a 15 year time horizon that are stable and plausible. This provides confidence that counterfactual scenarios assessed in results can be compared to the baseline status-quo model runs to provide insight in to the effect of different policy interventions.

4.5 Results

Change in SF-12 MCS is estimated for each intervention by running 100 model iterations from 2020 to 2035 for the baseline, child uplift, living wage, and EPCG scenarios. Mean SF-12 MCS for each population sub-group of interest is calculated in every year of the scenario run providing a sample of 100 population means used to construct confidence intervals for overall change in SF-12 MCS for that sub-group. These absolute changes in well-being can be very small so relative change in SF-12 MCS against baseline values is used. The choice of baseline population varies for different policy interventions, to reflect the group to whom the intervention applies. Relative change alone does not convey the absolute size of the intervened on population, or the intervention cost. In Table 4.5 we report the cost per capita of each intervention. We also report the percentage of the population who have an SF-12 MCS score of 45.6 or less. This has been identified as the optimal cutoff when screening for 30-day depressive disorders (Vilagut et al., 2013). It is estimated that 2% of this population suffer from depression and anxiety disorders and provides a tangible estimate of the positive health effects for change in SF-12 MCS score.

Figure 4.7a provides estimates for change in SF-12 MCS for individuals within households where someone is in receipt of Universal Credit, where the £25 to £50 per child per week payments are applied in £5 increments. The £25 and £50 interventions have positive effects on mental well-being of $1.5 \pm 0.25\%$ and $2.75 \pm 0.25\%$ improvement by 2035 respectively. Initial SF-12 MCS improvement is small but over time this increases and is sustained. Both interventions consistently support 7.5% of the population costing an average of £44 and £89 per capita respectively. It is clear the the £25 intervention does not significantly improvement SF-12 MCS and increasing the payment to £40 shows significant improvement in well-being vs the baseline and £25 uplift as early as 2024. This results suggests evidence towards the recommendation to improve the Scottish child payment to £40. Improvement is not completely linear with the £50 uplift providing 83% more SF-12 MCS improvement suggesting diminishing returns outline elsewhere (Congreve et al., 2022). These interventions also show 1.64% and 3.96% reductions in individuals with SF-12 MCS score below 45.6 suggesting approximately 1300 and 3150 fewer cases of anxiety and depression, assuming 2% of individuals below 45.6 SF-12 MCS score contract these conditions (Vilagut et al., 2013), in the UK population of $4.4m$ households on universal credit as of July 2023.

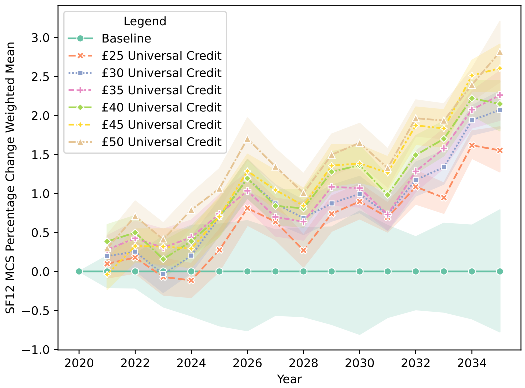
Figure 4.7b shows overall change in SF-12 MCS score for all individuals in households where there is an increase in disposable income under the living wage scenario. SF-12 MCS score improves by $2.0 \pm 0.05\%$ by 2035 showing a consistent linear trend. The living wage intervention is more expensive than both child uplift payments on a per capita basis, costing an average of £125 per person but affects a larger proportion of people (33%) and providing a larger and less uncertain gain in SF-12 MCS. The intervention results in a 2.68% reduction in individual with SF-12 MCS below 45.6, which would equate to approximately 2600 fewer potential incidences of mental illness nationally in the $5.1m$ individuals paid below the living wage.

Figure 4.7c shows change in well-being under the energy crisis scenario with no government support and under the Energy Price Cap Guarantee against a baseline of households with non-zero energy expenditure. The energy crisis with no government support shows a clear decrease in mental well-being reaching $-1.2 \pm 0.05\%$. The EPCG clearly alleviates some effect of detriment, with mental well-being decline reduced to $-0.5 \pm 0.05\%$ but this is not enough

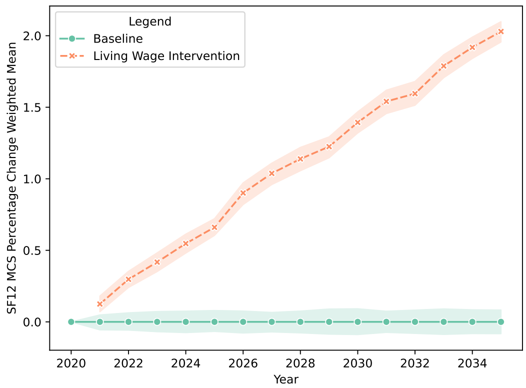
Year	Intervened Population Percentage	Cost Per Capita (£s)	Baseline Percentage Below 45.6.	Intervention Percentage Below 45.6.	Difference.
£25 Universal Credit					
2021	7.48	52.36	53.23	52.71	-0.52
2025	7.38	46.74	48.19	47.22	-0.97
2030	7.60	41.57	48.48	47.55	-0.93
2035	7.30	36.93	49.93	48.29	-1.64
£50 Universal Credit					
2021	7.48	104.73	53.23	52.56	-0.67
2025	7.37	93.54	48.19	46.59	-1.60
2030	7.59	83.23	48.48	46.16	-2.32
2035	7.27	73.95	49.93	45.97	-3.96
Living Wage Intervention					
2021	33.45	138.76	37.51	37.18	-0.33
2025	33.26	133.93	34.70	33.56	-1.14
2030	33.21	126.58	35.15	33.15	-2.00
2035	33.21	118.91	35.41	32.73	-2.68
Energy Price Cap Guarantee					
2021	99.18	74.60	34.49	34.34	-0.15
2025	96.07	71.02	32.93	32.19	-0.75
2030	92.16	65.75	33.99	33.07	-0.92
2035	88.27	59.81	35.18	34.12	-1.07

Table 4.5: Statistics for Policy Cost and the Number of People Who Receive Support for the Scottish Child Payment, Living Wage, and Energy Price Cap Guarantee policies.

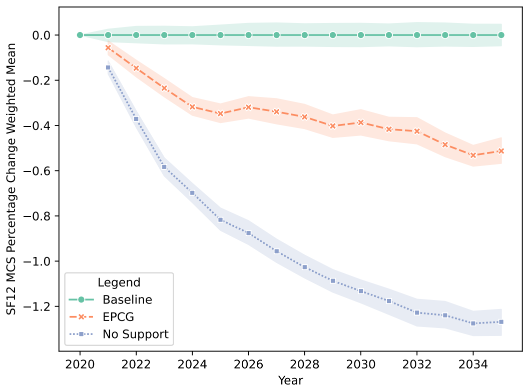
to return average SF-12 MCS to pre-energy crisis levels. While this intervention is not the most expensive per capita (£68) it has the largest overall cost (£13m) per year for the MINOS sample population as almost all households ($> 90\%$) are intervened upon. A 1.07% reduction in individuals with SF-12 MCS below 45.6 suggests approxiamtely 13000 fewer incidences of mental illness assuming 95% of $67.3m$ UK citizens have energy expenditure. Extrapolation using sample weights (Fumagalli et al., 2017) suggests this intervention will cost £$78 \pm 8bn$ per year applied to the full scale United Kingdom population which aligns with similar, albeit highly uncertain, estimates (UK Fuel Poverty Monitor, 2022). The EPCG is beneficial versus doing nothing but shows the smallest well-being gain of all scenarios.



(a) £25 to £50 child payment interventions in £5 increments for households in receipt of Universal Credit.



(b) Living Wage Intervention vs Baseline for households containing individuals below the living wage.



(c) EPCG and No Government Support vs Baseline for Households that Purchase Energy.

Figure 4.7: Aggregated line plots denoting relative percentage difference in SF-12 MCS score for the child uplift, living wage, and EPCG interventions applied to MINOS.

4.6 Discussion

Our modelled results suggest that the per-capita relative mental well-being improvements following the same income supplementation would be approximately $1.43 - 2.44$ times larger when providing a £50 vs £25 Scottish child payment. Implementation of the living wage intervention shows an increase in mental well-being 2.05 times larger than application of the £50 Scottish child payment. This intervention is more expensive per-capita (£89 vs. £125) but affects a much larger proportion of households (7.5% vs 33%) with lower uncertainty. The Energy Price Cap Guarantee reduces the decline in SF-12 MCS score by 60% versus no Government support and while not the most expensive intervention per-capita (£68) has high overall cost due to intervention on almost all of the population. The key innovation of these findings is that the MINOS microsimulation model can be used to test a diverse range of counterfactual income support scenarios and outcome metrics. MINOS can provide evidence advising Government policy across sectors promoting health in all policies. There is also potential to produce efficient, targeted policy interventions providing support to vulnerable households that most need support while balancing resource constraints.

The income to mental health model applied using MINOS in this paper is specified by translating causal loop diagrams into actionable, associative modules, constructed from real data, that represent five key pathways (neighbourhood safety, housing quality, loneliness, nutrition and tobacco), alongside demographic modules within the microsimulation. The advantage of taking this approach, rather than looking at the direct relationship between income and mental health, is that it provides a way of explicitly operationalising the mechanisms that exist in a wide body of literature, translating a health in all policy approach (Meier et al., 2019; De Leeuw, 2017) in to actionable insight, and allow users to assess the relative importance of each pathway. These pathways are transitioned independently within the discrete time framework but changes in individual state delivered by other pathways are taken in to account for each yearly run. It also means that MINOS is scaleable and adaptable - further modules can be added, or these modules can be used in other formulations of the MINOS microsimulation. For example a current focus of development is adaptation to a model for physical health (where SF-12 Physical Component Score is the outcome). Where many models are created within policy silos, we intend for MINOS to be adaptable for a wide range of applications, capitalising on considerable investment in producing an open source framework.

A key limitation for MINOS is data availability. Longitudinal UKHLS data used calibrated MINOS only offers yearly interviews for individual household and does not provide sufficient granularity to estimate more acute mental well-being change that can occur in times of crisis. There are also undoubtedly other pathways between a change in income and mental health that are not measured in this application of MINOS - the model in necessarily a tractable representation of the system but further work could investigate the effects of other (as yet unidentified) pathways.

This paper has demonstrated the utility of MINOS framework for providing insights in to the policy impacts that a change in household disposable income might have on mental health. But the potential for the model to be used as a sandbox for further policy research is substantial. These include combinations of multiple policies, policies that vary over time and space, policies that optimise resource allocation, and application to policy pathway systems beyond Figure 4.1. These are all developments we intend to explore using the MINOS framework. Extension to further health outcomes including physical well-being and quality life years is also on our agenda, which will provide further outcome metrics for policy makers. Inclusion of other outcome metrics not related to health such as child poverty rates allow for comparison of diverse policy options and their effects across different sectors in order to promote health in all policies. Most dynamic microsimulation models are large, complex and continuously evolving, and MINOS is no exception. Application of further methodologies including alignment and more nuanced life path processes are on the development agenda for MINOS. Full documentation and user tutorials are regularly updated [1] to ensure accessibility for critical review and provide a starting point for modellers who may wish to utilise this framework elsewhere. Ultimately MINOS provides a tool for assessing a range of interventions, providing evidence needed for government to make informed decisions that protect public health in all policies.

[1] Online documentation will be unblinded after review.

4.7 Connection between Chapter 4 and Chapter 5.

This chapter has described creation of the Microsimulation for Interrogation of Social Science Systems (MINOS) dynamic microsimulation to operationalise income to mental health causal loop diagrams developed by the SIPHER consortium. Understanding Society data from previous chapters was utilised to create an input population for the MINOS model including additional creation of composite variables contained in the loop diagram but not present in the original data (e.g. household disposable and housing material deprivation proxy). A series of transition probabilities were also constructed with this data including models for estimating future income state, several intermediary variables including housing quality and alcohol consumption, SF-12 MCS mental health outcomes from chapter 3, and demographic processes to maintain population representativeness including births and deaths. After validation ensuring sensible prediction of individual and population level behaviour three income support policies have been applied to the MINOS model. The Scottish Child payment provided between £25 and £50 per child per week to household on universal credit with children. The real living wave policy increased the minimum wage to the real living wage £12.00/13.15 outside and inside London respectively. Finally the Energy Price Cap Guarantee was implemented fixing yearly household energy bills assuming a large increase in fossil fuel costs. Each of these policies demonstrated a positive change in mental well-being. Implementing a larger Scottish Child payment of £50 over the original £25 demonstrated a significant improved in SF-12 MCS that began to show diminishing returns. The real living wage showed consistent, low uncertainty improvement although it affected a small proportion of the population and was expensive per capita. Finally, the energy price cap guarantee did alleviate the effect of high energy bills but was not sufficient to preserve mental well-being shown before fossil fuel prices increased.

Initial application of the MINOS model has demonstrated a successful ability to evidence the mental well-being effect of income support policy. There are several limitations to this model. First, it uses the aspatial weighted Understanding Society sample as its input population. While this sample is designed to be representative of the full UK population using a weighted sample with each individual representing a portion of the UK, it ignores potential individual behaviour over time. Including a full population of individuals that have associated spatial information allows for testing of the scalability of the MINOS model to larger populations, exploration of spatially orientated policy such as housing policy, and more detailed exploration of small subsections of the population with a much larger sample size. Second, the MINOS model only includes SF-12 MCS score to estimate change in public health. This variable does not consider physical health and it is difficult to interpret what a single unit of 'SF-12 MCS' is without specific contexts such as post-surgery treatment. Application of the corresponding Short Form 12 Physical Component Score (SF-12 PCS) allows for consideration of physical health and can then be used to estimates change in Quality Adjusted Life Years (QALYs) and Incremental Cost-Effectiveness Ratios (ICERs) that are much more tangible measures of cost-effectiveness for policy makers. Finally, the MINOS model has so far only intervened upon household disposable income. Implementing policy that intervenes on other variables including subjective household heating can provide more realistic policy scenarios. Directly comparing two competing policies also allows for direct discussion on future UK government policy on household energy bills and provides a framework to interpret simulated evidence from the MINOS model for future policy combinations.

Chapter 5: Evaluating the Impact of Household Energy Policy on Adult Mental and Physical Health in the United Kingdom

5.1 Abstract

Energy policy is set to dominate the global political landscape indefinitely as countries transition towards carbon neutral targets and renewable technologies. In the United Kingdom, subsidies for insulation and electric heating system retrofitting have seen limited uptake relative to neighbouring countries due to lack of evidence identifying vulnerable households and skepticism over cost-effectiveness. The UK is now falling behind with surging rates of energy poverty adversely affecting physical and mental health.

The new UK Labour Government, as well as devolved Governments at local authority level, are expanding these retrofitting policies but require micro-scale evidence to allocate personalised policy according to individual needs as well as national large scale infrastructure planning and justify potential health benefits to encourage investment of private capital.

This paper presents methodology to generate synthetic evidence for the effect of energy policy at a household level using the dynamic microsimulation model MINOS. Three energy policies including a Good Heating 'morning after' intervention, the Energy Price Cap Guarantee (EPCG) which is already implemented, and the Great British Insulation Scheme (GBIS/ECO4) are assessed using Short Form 12 Health Scores, Quality Adjusted Life Years, and final Incremental Cost Effectiveness Ratios. These three intervention scenarios are applied to a synthetic 10% scale sample of the Greater Manchester Combined Authority population within the United Kingdom from 2020 to 2035. Results indicate that while the EPCG was beneficial protecting public mental well-being during high energy pricing it was not cost-efficient and proactive insulation scheme would have provided high ICER scoring as early as 2024. These results applying microsimulation to energy policy can be greatly expanded to include further combinations of policy, areas of the United Kingdom, and application to other areas of policy entirely including child poverty.

5.2 Introduction

The UK has seen one of the largest increases in household energy bills in Europe and the highest disparity in percentage income spent on energy across income quintiles (Evans, 2022; Department of Energy Security and Net Zero., 2024; Ari et al., 2022). Since the 1973 oil crisis, which saw dramatic fossil fuel price increases, the UK has done little to improve its energy infrastructure (Hodgkin and Sasse, 2022) and over-reliance on natural gas. This is in comparison to neighbouring countries e.g. Germany, Finland, and France (Hodgkin and Sasse, 2022; Ward et al., 2024; Ari et al., 2022), where upgrades including use of electrical heating systems such as heat pumps and large scale insulation retrofitting subsidies (Hodgkin and Sasse, 2022) have significantly reduced excess mortality due to extreme cold compared with the UK (Ballesteros-Arjona et al., 2022). Researchers (Hill O'Connor et al., 2023; Broadbent et al., 2023; Booth, 2023; UK Fuel Poverty Monitor, 2022; Ballesteros-Arjona et al., 2022) and UK Healthcare representatives (National Health Service (NHS) Confederation, 2022) are describing high energy bills as an emerging public health crisis, with inadequate heating driving an increase in the incidence of physical respiratory illness due to damp and mould, and financial stress caused by high energy prices resulting in an increase in incidences of mental illness. Recent geopolitical events, notably the ongoing Russo-Ukrainian war, demonstrate the UK is still vulnerable to increasingly volatile fossil fuel pricing and must make vital infrastructure upgrades in order to protect public health and develop energy resilience (Hodgkin and Sasse, 2022; UK Fuel Poverty Monitor, 2022).

A number of policies have been proposed by the UK government in order to reduce household energy bills. During the winter of 2022 the Energy Price Cap Guarantee (EPCG) directly intervened on household energy bills by deferring excess cost to future taxation (UK Fuel Poverty Monitor, 2022). The EPCG is by far the most expensive income support policy ever implemented in the UK with costs continuing to increase above £100bn, but it has been widely criticised as an ineffective and unsustainable stop-gap measure (UK Fuel Poverty Monitor, 2022) that does not provide help to those most in need. Initially driven by Net Zero carbon emission targets, the UK government is now subsidising installation of retrofitted insulation and electrical heating systems to reduce household energy bills (United Kingdom Government: Department for Energy Security and Net Zero., 2023; Li et al., 2022; Panakaduwa et al., n.d.). The Great British Insulation Scheme (GBIS) (Regan, n.d.), also known as the Energy Company Obligation (ECO4), is one such policy providing Local Authorities with £1bn to retrofit household insulation for low energy efficiency households within the UK (Regan, n.d.). The GBIS has seen limited uptake and is proving difficult to implement due to a lack of evidence and low political will (Regan, n.d.; Snell et al., 2018). The general public is largely skeptical of any income and health benefits provided by insulation retrofitting and are reluctant to provide private capital (Regan, n.d.) fearing being driven further into poverty. At the same time, Local Authorities are having trouble identifying households (Snell et al., 2018; UK Fuel Poverty Monitor, 2022) based on existing poverty metrics and are reluctant to spend money as their budgets are reduced by disinvestment. As of August 2024, the new Labour Government has made another highly criticised U-turn (*ECIU Comment of Labours Revised Green Prospeity Plan*, n.d.) reducing budget to schemes like the GBIS further prompting the need for precise policy.

Both issues may potentially be addressed by providing evidence for the positive health effect of policies such as the GBIS at a household level. The general public are substantially more willing to invest in retrofitting if health and economic benefits can be clearly proven (Regan, n.d.). Likewise, UK Governments could optimise cost-effectiveness of the GBIS by identifying vulnerable households most in need of support (UK Fuel Poverty Monitor, 2022; Regan, n.d.) while providing population-level health benefits that are highly desirable as part of a 'health in all policies' approach (Meier et al., 2019). Generation of this household-level health evidence has the potential to increase uptake of the GBIS at a lower cost, ideally creating a positive feedback that insulates the UK and prevents future energy and public health crises.

Dynamic microsimulation is an individual-level modelling technique that can be used to interrogate the potential impacts of energy policies (Burgard et al., 2021). A population of individual units is generated, either synthetically or from real panel data, and evolved forwards in time under transitions dynamics and hypothetical policy scenarios (Burgard et al., 2021; Spielauer et al., 2020). Dynamic microsimulation allows for simultaneous analysis where both national-level targets and assistance for any vulnerable small scale subgroups and even individual households are met(Clay et al., 2023). Inclusion of other attributes including long-term individual histories and spatial components (O'Donoghue et al., 2013) make microsimulation a useful tool for estimating how the UK population public health is effected under energy policy providing policy makers a useful tool to evidence and optimise new candidate policy in an artificial but realistic environment without extensive clinical trials.

This paper aims to expand on an existing dynamic microsimulation framework MINOS (Clay et al., 2023; Lomax et al., 2023); developed to estimate the effect of change in household disposable income on mental well-being for the United Kingdom population. MINOS is calibrated using the Systems Science in Public Health and Health Economics Research (SIPHER) consortium synthetic spatial population dataset (Lomax et al., 2024; Hoehn et al., 2024) constructed from United Kingdom Longitudinal Household Survey (UKHLS) data (Fumagalli et al., 2017) and simulated annealing methodology to create a spatially representative population of the Greater Manchester Combined Authority (GMCA) region with respect to key aggregates (Hoehn et al., 2023). This input population is then projected forward in time under a number of policy interventions. First, a scenario assuming energy prices do not increase with all households are assumed to have comfortable heating instantly in a 'morning after' (Kopasker et al., 2024) scenario. This is an idealised scenario exploring how change in key energy poverty variables effects health and serves as a useful benchmark for comparison with real policy. Second, emulation of two actual government policies, the Energy Price Cap Guarantee and the Great British Insulation Scheme, is implemented to improve health by improving heating quality and energy bills. These interventions are assessed by change in Quality Adjusted Life Years (QALYs) (Lawrence and Fleishman, 2004) and Incremental Cost Effectiveness Ratios (ICERs) (Gafni and Birch, 2006). These provide established measures for recorded change in health and cost-effectiveness. This remainder of this paper is structured as follows: Section 5.3 provides an overview of methods, Section 5.6, provides results of the scenario interventions, with Section 5.7 providing discussion outlining limitations and future work and Appendix 8.2 provides appendices of further methodological detail and data tables.

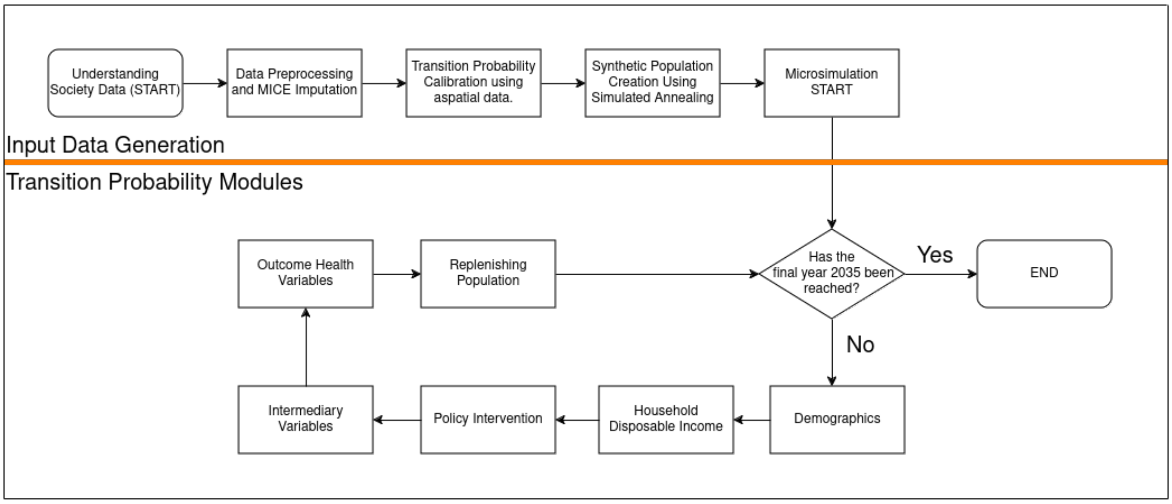
5.3 Methods

5.3.1 The MINOS Dynamic Microsimulation

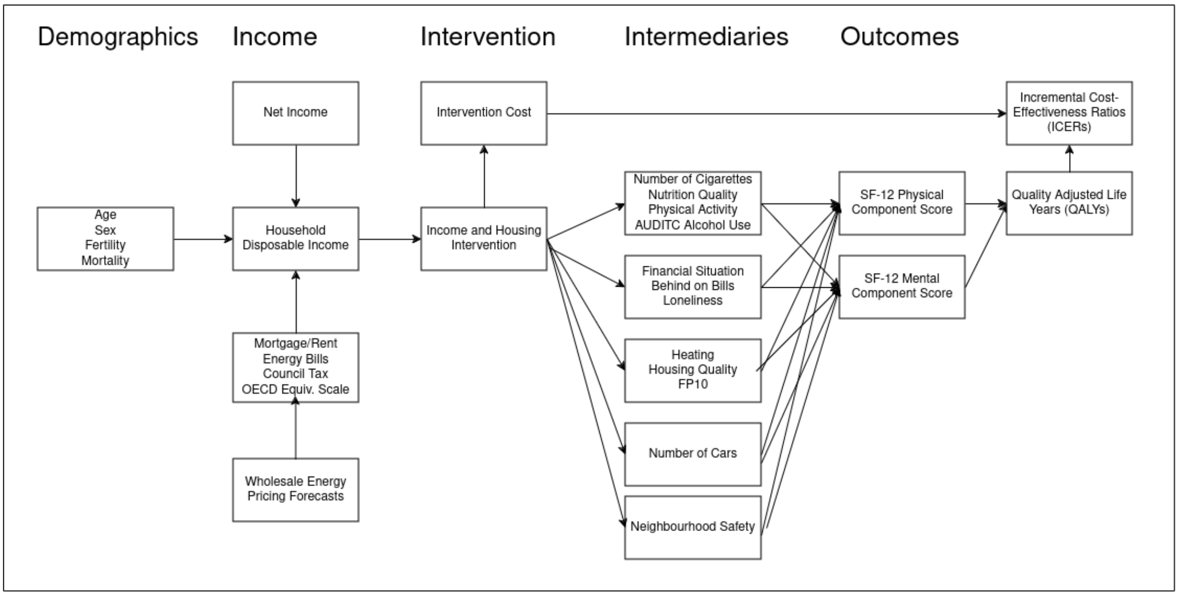
The 'theory of change' framework (Haraldsson, 2004) is a methodology used to identify and evidence potential candidate policy to improve human quality of life. In order to evidence the effect of potential policy, causal loop Modelling (CSM) (Meier et al., 2019) is used to identify how a policy would alter a population and how it propagates through complex causal pathways to key outcomes such as health measures. Systematic review of literature evidence is combined with opinions from experts and community members into causal loop diagrams (SIPHER Consortium et al., 2023; Campbell et al., 2023) describing how individual, household, and neighbourhood characteristics are connected. If household income and household quality are directly connected it suggests a causal relationship. If household income increases, the household has more money to spend on maintenance and housing quality would also be expected to increase. This research utilises an existing causal loop diagrams (SIPHER Consortium et al., 2023; Campbell et al., 2023; Consortium, 2023) describing pathways between household disposable income, energy bills, and health whose derivation is described elsewhere (SIPHER Consortium et al., 2023; Campbell et al., 2023).

The second stage of CSM then emulates these pathways using individual-level modelling techniques such as microsimulation (Meier et al., 2019). The intention is to operationalise a given causal loop map estimating the effect of hypothetical policy on the UK population. The Microsimulation for Simulation of Social Science Systems (MINOS) (Clay et al., 2023; Lomax et al., 2023) is a dynamic microsimulation used in SIPHER approximating the income-health systems map provided online (SIPHER Consortium et al., 2023). The full income-health system is estimated in MINOS using a series of discrete-time modules predicting next state using current population information and associative modelling described in Figure 5.1a. The MINOS model aims to build on several existing spatial dynamic microsimulation models built at Leeds including SimBritain (Ballas, Rossiter, Thomas, Clarke and Dorling, 2005) and static time spatial microsimulation projects (Lomax and Smith, 2017b) in several ways. MINOS is using individual household level data, where previously coarser spatial areas have been used, and is additionally projecting these small spatial area individuals forwards in time providing rich micro-level data requires to implement personalised policy. MINOS is designed to be completely open source written in the R. and python languages using publicly available datasets. While previous iterations of the MINOS model (Clay et al., 2023) focused solely on policy that affects change in household disposable income, this paper extends MINOS to policies that alter housing characteristics including energy bills and insulation quality. This allows for application of the health in all policies ideology (Meier et al., 2019) exploring the health outcomes of non-health policy to exhaustively asses all potential consequences when implementing housing and energy policy. Additionally multiple policies can be combined together exploring interaction effects and any areas where households incorrectly receive no or multiple interventions (Waddams, 2023).

The development of this microsimulation again required creation of a series of transition probability models defined in Figures 5.1 - 5.1b considerate of requirements specified in chapter 2. These models were constructed using a series of R. notebooks containing information on justification of model design including independent variable literature review, data processing, and validation methods. These notebooks we're designed to be interpretable as possible allowing critical review and refinement within the SIPHER consortium into this final series of modules and the order in which they are run. Further information from these notebooks is provided in the following input population data processing and transition model in Sections 5.3.3 - 5.3.5.



(a) The overall MINOS microsimulation structure including input population generation and transition dynamics. Derivation of all components and data sources starting from the Understanding Society dataset (Fumagalli et al., 2017) are explained below.



(b) The complete list of modules used in the MINOS model. Modules are divided into four income, intervention, intermediate, and health outcomes groups. Each box represents a series of transition probability models such as for OECD Equivalence Scale (Anyaegbu, 2010) are explained in detail below.

Figure 5.1: The overall MINOS model structure and pathways used to estimate housing policy effect on well-being.

5.3.2 Input Populations

MINOS uses a standard dynamic microsimulation structure (Spielauer et al., 2020) consisting of an input population and transition dynamics components described in Figure 5.1a. The input population is a dataset (details below) containing information for individual households and individuals that are representative of the GMCA area containing information required to estimate change in health due to energy policy. There a number of auxiliary datasets also used in MINOS to inform behaviour including mortality rates and energy pricing. This section describes the datasets used in the MINOS model including details on preprocessing and the generation of a synthetic population with spatial information.

The MINOS input population is constructed from Understanding Society (UKHLS) (Fumagalli et al., 2017) administrative data. This data currently consists of 12 waves of annual interviews for individual UK households ($2009 - 2021$) containing approximately 30000 observations per year. Hundreds of individual variables are recorded in the UKHLS including information on income, health, and housing characteristics. It is the largest public dataset of its type for the United Kingdom and one of the best available candidates for constructing a microsimulation input population. All variables from UKHLS data used in MINOS are specified in the Table 5.1.

5.3.3 Data Preprocessing

Several stages of data preprocessing are applied to Understanding Society data in order to impute missing values and improve the readability of model data. The following section describes how this spatially representative input population for the GMCA area was created providing final descriptive data tables in appendix Section 8.2.1.

Raw Data Processing

Understanding Society data are downloaded directly from the UK data service (Fumagalli et al., 2017) as yearly interview cohort data divided into several sections for individual and household survey responses. As in Chapter 4 only individual and household response data sections are used. Initial data processing extracts required variables for MINOS from both datasets that are then merged together to create rows for each individual person with complete personal and household information. Certain categorical variables are renamed from integer codes to string values. It is much easier to read labour state 'employed' rather than code 3. This also allows for much easier fitting of transition probability models in R. that automatically recognises strings as factors. Additionally discrete variables elements with very small group sizes are binned to prevent overparameterisation. For example, in the category education state, individuals in 'Government Training' who make up less than 5 observations per year are merged into 'Full Time Education' instead.

Name	Category	Description
pidp/hidp	Demographics	Individual/household unique identifiers
age	Demographics	Individual Age.
nkids	Demographics	Children under 16 years old living in the household.
sex	Demographics	Biological sex
ethnicity	Demographics	Ethnicity*
region	Demographics	Government region*
education_state	Demographics	Highest Education State*
housing_quality	Housing	Household quality+
neighbourhood_safety	Housing	?+
housing_tenure	Housing	What is your housing tenure status? (Rented, Owned Outright, etc.)
marstat	Housing	Marital Status*
hh_comp	Housing	Number of people in the household
ncars	Housing	How many cars are used by the household?
number_of_rooms	housing	How many rooms are in the household?
net_income	Income	What is your total household income after tax and national insurance?
rent	Income	How much does the household pay towards rent?
mortgage	Income	How much does the household pay towards mortgages?
council tax	Income	How much does the household pay in council tax?
OECD Equivalence Score	Income	What is the household OECD Equivalence Score?
job_sector	Income	What job sector (jobsec07) are you employed in?*
NSSEC	Income	What is the NSSEC socioeconomic code for your employment?*
yearly_energy	Income	What is your yearly energy expenditure?
behind_on_bills	Income	Are you behind on payment for any household bills?
nutrition_quality	Health	How many fruits and vegetables do you consume per week?+
SF12_MCS	Health	What is your SF-12 MCS score?
SF12_PCS	Health	What is your SF-12 PCS score?
loneliness	Health	How often are you lonely?
auditc	Health	What is your auditc score?
active	Health	Do you mean the government guidelines for 'phyiscally active' with a minimum amount of exercise per week?

Table 5.1: Understanding Society Data Variables used in MINOS. Variables marked * required further preprocessing or derivation described in detail in supplementary online material at https://github.com/Leeds-MRG/Minos.

Data Cleaning/ Correction

A large proportion of UKHLS missing data can again be corrected deterministically. As in Chapter 4 deterministic correction is applied to non-applicable data such as income for unemployed individuals. Last observation carried forwards imputation (Zhang, 2016) and linear interpolation are used to correct for observations that are not carried forwards in UKHLS data. After deterministic correction only remaining missing data values are left that require imputation.

Composites

Again several variables are not provided in UKHLS data and must be derived from other available information. All variables used to calculate these composites are provided in Table 5.4. Household income is again calculated using the standard OECD formula (Eurostat-Glossary, n.d.; Anyaegbu, 2010) by taking the household net income, subtracting monthly housing outgoings and scaling by OECD equivalence factor accounting for household size and number of children. Finally income is scaled by consumer price index into 2020 £s giving final formula

$$\text{household disposable income} = \frac{\text{(net income) - ((rent) + (mortgage) + (council tax) + (yearly energy)}}{.}$$

There is no direct variable for material housing quality in Understanding Society. SIPHER research has suggested (Aki Tsuchiya, Guoqiang Wu, 2021) that a suitable proxy can be generated according to the presence of certain household appliances and the ability to heat the home adequately. Individuals are assigned into three tiers based on possession of five household appliances and adequate heating, given in Table 5.4. These appliances are divided into core items (fridge freezer, washing machine) and heating that are essential to comfortable living standards, and a second set of luxury items (microwave, tumble dryer, dishwasher) which are non-essential but represent an improved housing quality. Tier 1 is the lowest tier of housing quality for individuals missing at least one core item, tier 2 is for individuals with all core items but some missing luxury items, and tier 3 is for individuals with all items.

Neighbourhood Safety is also derived. Research undertaken by the SIPHER consortium has developed a proxy (Aki Tsuchiya, Guoqiang Wu, 2021) for safety based on the frequency of several crimes occurring in a neighbourhood given in Table 5.4. Individuals are assigned a state from $1-3$ based on the frequency of these crimes in their neighbourhood. A value of 1 is assigned for the worst state, where any of the listed crimes are 'Very common' or 'Fairly common' in the neighbourhood. A value of 2 is assigned where any crime is 'not very common', and 1 is assigned for the best case scenario where all crimes are 'not at all common'.

Nutrition Quality is another derived variable. This is estimated based on the frequency and number of servings of both fruit and vegetable consumed in an average week. The variables from Understanding Society used to create this proxy are listed in Table 5.4. The number of days eating fruit or veg is multiplied by the number of portions on an average day, then added together to generate a continuous score of nutrition quality.

$$\begin{aligned}\text{Nutrition Quality} &= (\text{Days eating fruit} \times \text{Portions per day}) \\ &\quad + (\text{Days eating veg} \times \text{Portions per day})\end{aligned}$$

Hourly wage is another derived variable, based on US variables in Table 5.4. If the variable *basrate* (Basic pay hourly rate) is present, the hourly wage is set to this value. If not (i.e. for salaried employees), the hourly wage is calculated based on the usual gross pay per month and the number of hours normally worked in a week (multiplied by 4.2 for average number of weeks in a month), according to the following formula.

$$\text{Hourly Wage} = \frac{\text{Gross pay per month}}{(\text{Hours worked per week} \times 4.2)}$$

AUDITC score (Bush et al., 1998) is a commonly used measure to identify individuals at risk of contracting non-communicable disease due to alcohol consumption who are not alcohol dependent. A series of 3 likert scale questions taking values $0-4$ are asked about alcohol consumption related to frequency of consumption and dependent. These Scales are then added together producing a final AUDITC score between 0 and 12. These scores are then binned into 4 categories of no risk $(0-4)$, low risk $(5-7)$, medium risk $(8-10)$, and high risk or possible dependence $(11-12)$. Note these categories are adjusted slightly by sex with women's categories all reduced by one point. For example low risk is $(4-6)$ rather than $(5-7)$.

Physical activity is defined by government guidelines as individuals who perform at least 150 minutes of moderate activity or 75 minutes of physical activity per week. Variables for the number of minutes for both moderate and vigorous physical activity are both included in the UKHLS. A binary variable is produced given value 1 if the individual meets these government guidelines defining them as physically activate and 0 otherwise defining them as inactive.

To estimate yearly energy consumption UKHLS provides four variables for yearly expenditure on consumption of electricity, gas, oil, and other solid fuels. Additionally variables for payment methods for gas and electricity are provided determining energy payment methods including variable or fixed rate tariffs and prepayment meters. Using these variables approximate kWh consumption can be estimated by subtracting standing charges from yearly energy bills. These kWh consumption are then adjusted according to changes in wholesale commodity prices. The final yearly energy consumption is then the sum of expenditure for all four fuel types.

Quality Adjusted Life Years (QALYs) are a measure used to described the efficacy of an intervention scenario on public health (Lawrence and Fleishman, 2004). One QALY is defined as a population gaining one year of quality of life under an intervention versus some counterfactual baseline. QALY score can be approximated by first calculating a utility score (Lawrence and Fleishman, 2004) from SF-12 MCS and SF-12 PCS scores using a formula

$$\begin{aligned} u &= -1.6984 + (0.07927 * PCS) + (0.02859 * MCS) \\ &+ -0.000126(PCS * MCS) + -0.00141 * PCS^2 \\ &+ -0.00014 * MCS^2 + 0.0000107 * PCS^3. \end{aligned}$$

This utility score is filtered according to dead or alive individuals. If an individual is deceased their utility score is 0. These utilities scores are then added up across a whole population providing a final QALY score for a population for a given time point and intervention scenario.

Final data correction involves complete case analysis of several critical variables that are required for MINOS to run. Any observations missing these key variables provided in Table 4.3 are removed from the dataset. Future work intends to correct these missing values using multiple imputation techniques particularly the MICE algorithm. Data matching methods may also be necessary in order to include variables not in Understanding Society. Summary statistics for key discrete and continuous variables in Table 5.1 are provided in Tables 8.10 - 8.11 below.

MICE

All remaining missing data are then 'truly' missing and must be imputed using the Multiple Imputation Chained Equations (MICE) algorithm. MICE has been implemented according to procedure (Jakobsen et al., 2017) ensuring the right missingness structure for imputed variables, inclusion of other auxiliary variables with high cross coverage and correlation with missing values, and multiple imputations and testing for model convergence. Finally, a number of composite variables that are not defined in UKHLS are derived. This includes values such as disposable household income derived from net income and outgoings as well as variables such as housing quality and neighbourhood safety proxy measures for housing characteristics derived from previous SIPHER research (Aki Tsuchiya, Guoqiang Wu, 2021).

Remaining missing data for all variables are summarised in Figure 5.2. All variables, apart from nutrition quality which is more complete in later UKHLS waves so is included, show less than 10% missing data present suggesting they are suitable for MICE imputation (Jakobsen et al., 2017). Procedure applying the MICE algorithm requires estimating missing values for each variable using a series of linear regression models. Independent variables chosen for MICE, known as auxiliary variables (Sterne et al., 2009; Jakobsen et al., 2017), are used in MICE imputation but are then removed and are not part of the MINOS input population.

One example application of the MICE algorithm is provided exploring imputation of the $SF-12MCS$ variable and eligible auxiliary variables. A pairwise correlation plot in Figure 5.3 demonstrates correlations between $SF-12MCS$ and variables available in the UKHLS dataset(Fumagalli et al., 2017). Variables including loneliness, ghq depression and overall happiness, and health limiting social activity all have strong correlations $|\rho| > 0.2$ suggesting they should be included in imputation. Auxiliary variables used for SF-12 MCS MICE imputation are given in Table 5.2.

Variable (UKHLS name)	Definition
future financial situation (finfut)	Expected future financial situation
Likely to move house (lkmove)	Prefer to move house.
GHQ Depression (scghqi)	General Health Question for unhappy or depressed.
GHQ Happiness (scghql)	General Health Question for general happiness.
SCSF1 (scsf1)	Short Form 1 Score for overall well-being. Part of SF-36.
Health Limits Social Activity (scsf7)	Does health limit social activity.

Table 5.2: Additional auxiliary variables used for example SF-12 MCS MICE imputation but not present in the final MINOS input population.

Visualisation including spineplots and histograms (Prantner, 2011) using R.'s VIM package are also used to explore differences in distributions for auxiliary variables when an imputed variable is missing and when it is not. These differences provide further evidence of correlation and inclusion in MICE as well as evidence for desired cross-coverage. Example split histograms are provided in Figures 5.4,5.5 showing clear differences in distribution for loneliness and *ghq_depression* when SF-12 MCS is missing. Both variables have higher prevalence of depression and loneliness suggesting inclusion would be beneficial. Both variables have very strong cross-coverage only missing 34/29381 and 96/29381 values when SF-12 MCS is also missing providing further strong evidence as auxiliary variables in MICE. Further MICE implementation is provided in supplementary technical documentation including overall model fit, statistics, and differences in SF-12 MCS distribution before and after imputation.

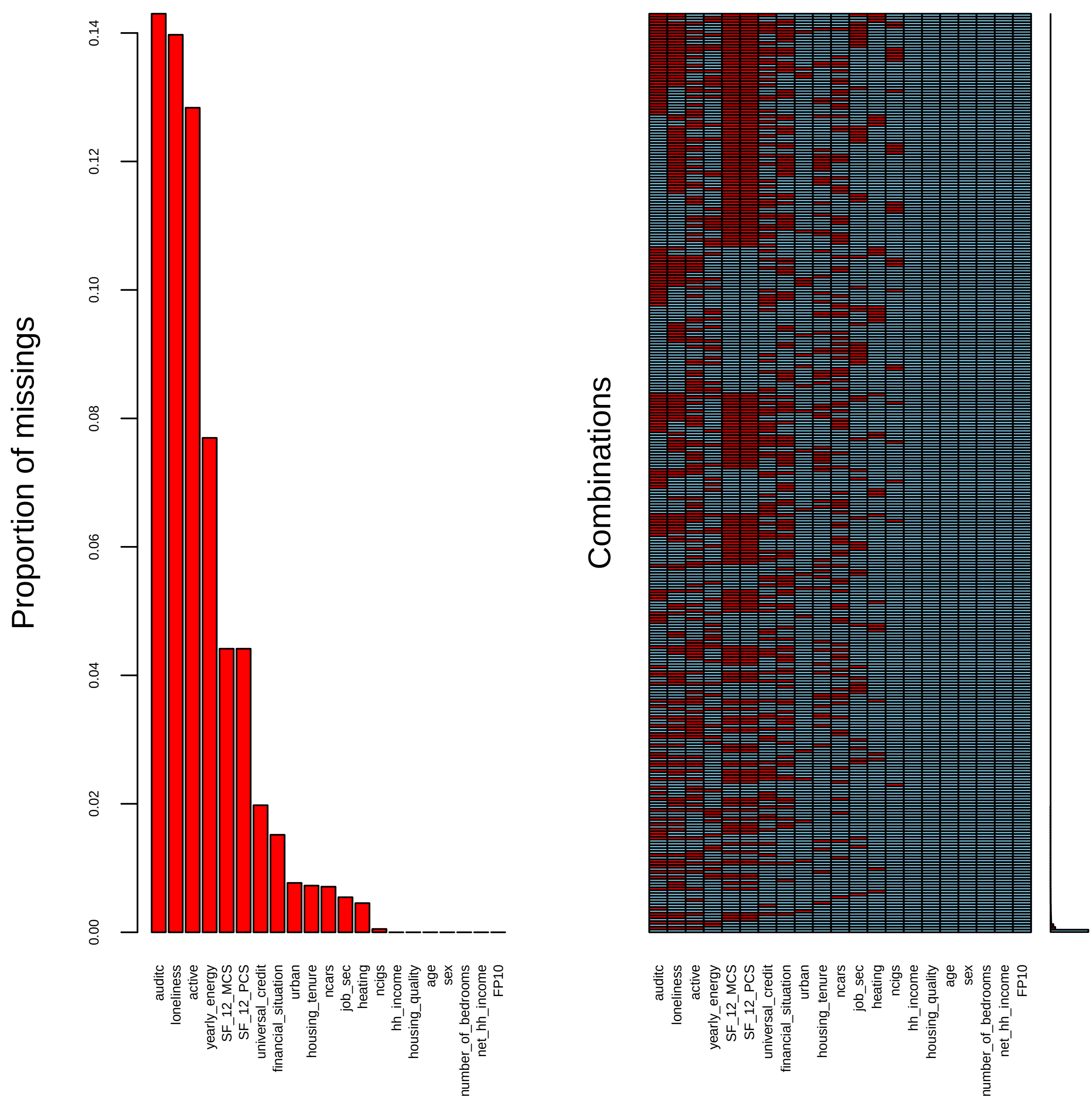


Figure 5.2: A missingness structure (aggr) plot for key MINOS variables after preprocessing.

Synthetic Data

Imputed UKHLS data is not representative of the GMCA area. It is a weighted sample (Fumagalli et al., 2017) designed to represent the nationwide UK population. Many individuals that are hard to access, such as ethnic minorities and those without internet, are intentionally over-sampled to ensure sufficient sample size. Additionally UKHLS data provides no spatial information such that individuals from the GMCA area cannot be identified. An output of the SIPHER consortium has been the creation of a synthetic population using UKHLS data (Hoehn et al., 2024; Lomax et al., 2024) for the full UK population with Lower Super Output Area (LSOA) spatial information. This population is generated using the simulated annealing algorithm (Lidbe et al., 2017) validated extensively against national and local area aggregates. Data from ONS and NOMIS sources (Lomax et al., 2024) containing attributes on household income as well as general demographics including age, sex, ethnicity, education, and household

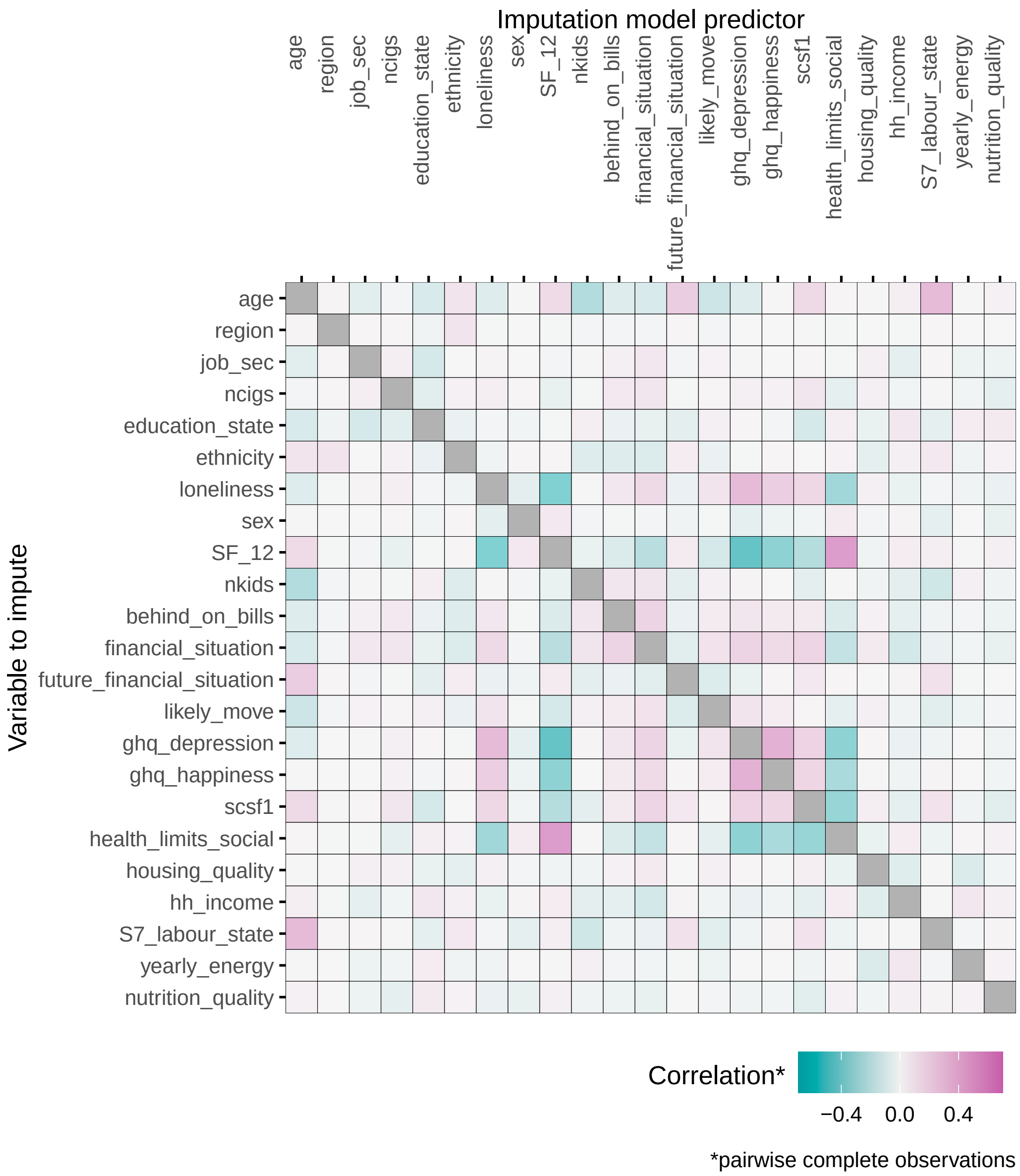


Figure 5.3: Pairwise correlation between SF-12 MCS and variables considered for auxiliary use in MICE imputation.

structure at LSOA level are used to constrain the simulated annealing algorithm that populates each LSOA with individual units from UKHLS until constraints are satisfied (Höhn et al., 2024; Lomax et al., 2024). Further details on the application of the simulated annealing algorithm, which is not a part of this thesis, are now publicly available on the UK data service (Lomax et al., 2024).

This data consists of two columns mapping individual household ID codes (hidp), that correspond to hidps in UKHLS, to corresponding LSOA areas. MINOS then merges this synthetic data with imputed UKHLS data producing the full UK population taking a subset for all households whose LSOAs are in the GMCA area. Post-processing was required to align spatial attributes ensuring individuals in rural areas possess the correct rural/urban identification and correct assignment of Government region to the 'North West' region of England (Taylor and Wren, 1997). Additional variables are also added dependent on LSOA including index of multiple deprivation and income quintiles for later disaggregation in results.

The full GMCA population then consists of $n = 1.70m$ individual adults and is used directly as the

MINOS input population. Due to computational constraints 10% samples of households are taken with replacement for the GMCA population resulting in a final samples of $N = 53869 \pm 100$ households with $n = 100244 \pm 2500$ adults over 16 years of age [1]. Descriptive statistics of key variables are provided in Table 5.1. This data serves as the final input population for the MINOS model. Note all weights from UKHLS data (Fumagalli et al., 2017) are reassigned to 1 as the synthetic population (Hoehn et al., 2023) assumes each individual observation is representative of exactly one individual household. Summary tables for continuous and discrete variables respectively are provided in Tables 8.10 - 8.11. This synthetic population ensures that it is representative for the UK and GMCA areas for the year 2020 only. Summary statistics for this synthetic data describe key discrete and continuous variables in Table 5.1 are provided in Tables 8.10 - 8.11 below.

Ideally this population would be validated at smaller spatial areas such as Output Areas (OAs) to better estimate neighbourhood behaviour but this is not possible due to a lack of available constraint data at this level. Additionally observing change in health outcomes at OA level can be highly uncertain due to low sample size require a large number of microsimulation runs to produce consistent results. Estimating spatial interactions at this level would also be highly computationally expensive in order to account for interactions between these areas such as migration. Using LSOAs provides lower precision but has substantially more available data, is much more efficient, and the units have approximately the same population with a large enough sample size limiting uncertainty./

Replenishing Population

Finally, in order to maintain population size over time replenishing data is created from the final input populations. A sample of households with 16 year old young adults is added to the population every year. The sampling of these replenishing households is performed using weighted bootstrapping. These weights are updated over time using cell based reweighting (Tysinger, 2021) according to mid year estimates of future age, sex, and ethnicity distributions for the GMCA area derived from the NEWETHPOP project (Rees et al., 2017) and NOMIS data (Nash, 2019).

Variable Name (Understanding Society Code)	Description	Deterministic Correction?	LOCF Imputation?	Composite?	MICE Imputed?
pidp	Personal Identifier	N	N	N	N
hidp	Household Identifier	N	N	N	N
sex	Biological sex	N	Y	N	Y
ethnicity (racel_dv)	Ethnicity	N	Y	N	Y
age (dvage)	Age	N	L	N	N
region (gor_dv)	Government Region	N	N	N	Y
education_state (qfhigh_dv)	Highest Qualification	N	F	N	Y

[1] Individuals under 16 years are not included in UKHLS data explicitly but values such as the number of children per household still influence variables such as household income and OECD Equivalence scale.

nkids (nkids_dv)	Number of children aged $0-16$	N	N	N	N
fridge_freezer (cduse5)	Has a fridge freezer	N	N	housing_quality	Y
washing_machine (cduse6)	Has a washing machine N	N	N	housing_quality	Y
tumble_dryer (cduse7)	Has a tumble dryer	N	N	housing_quality	Y
heating (hheat)	Has comfortable heating	N	N	housing_quality	Y
dishwasher (cduse8)	Has a dishwasher	N	N	housing_quality	Y
microwave (cduse9)	Has a microwave	N	N	housing_quality	Y
hh_composition (hhtype_dv)	Type of house (detached etc.)	N	N	N	Y
marital_status (marstat_dv)	Marriage status	N	N	N	Y
house_tenure (tenure_dv)	Housing tenure (rent/mortgage etc.)	N	N	hh_income	Y
hhsize	People in household	N	N	N	N
car_crime (crcar)	Car crime nearby	N	N	neighbourhood _safety	Y
drunks (crdrnk)	Drunken disorder nearby	N	N	neighbourhood _safety	Y
muggings (crmugg)	Muggings nearby	N	N	neighbourhood _safety	Y
racial_abuse (crrace)	Racial Abuse nearby	N	N	neighbourhood _safety	Y
teenagers (crteen)	Teenage disorder nearby	N	N	neighbourhood _safety	Y
vandalism (crvand)	Vandalism nearby	N	N	neighbourhood _safety	Y
council tax band (ctband_dv)	Council tax band	N	N	hh_income	N
hh_netinc (fihhmnnet1_dv)	Net household income	N	N	hh_income	N
oecd_equiv (iegmoecd_dv)	OECD Equivalence Scaling	N	N	hh_income	N

hh_rent (rentgrs_dv)	Household rent	N	F	hh_income	N
hh_mortgage (xpmg_dv)	Household mortgage payment	N	N	hh_income	N
yearly_ electricty (xpelecy)	Electricity expenditure	N	F	yearly_energy	Y
yearly_gas (xpgasy)	Gas Expenditure	N	F	yearly_energy	Y
yearly_gas_electric (xpduely)	Gas and electricity combined expenditure	N	F	yearly_energy	Y
yearly_oil (xpoily)	Fuel oil expenditure	N	F	yearly_energy	Y
yearly_other_fuel (xpsgly)	Other fuel expenditure	N	F	yearly_energy	N
hourly_rate (basrate)	Basic hourly pay	N	N	hourly_wage	N
gross_paym (paygu_dv)	Gross monthly pay	N	N	hourly_wage	N
gross_pay_se (jspayg)	Gross self employed monthly pay	N	N	hourly_wage	N
job_hours (jbhrs)	Weekly job hours	N	N	hourly_wage	N
job_hours_se (jshrs)	Self employed weekly job hours	N	N	hourly_wage	N
job_inc (jspayu)	Income from job/business	N	N	hourly_wage	N
jb_inc_per (jspayw)	Job pay period	N	N	hourly_wage	N
job_sector (jbsect)	Public or private sector	Yes. Set unemployed values from -8 to 0.	N	hourly_wage	Y
emp_type (jbft_dv)	Full or part time employed	N	N	N	Y
job_sec (jbnssec8_dv)	NSSEC Employment code	Yes. Set unemployed values from -8 to 0.	N	N	Y
ncigs	Weekly cigarettes consumed	Yes set nonsmokers from -8 to 0.	N	N	Y

loneliness (sclonely)	Loneliness frequency	N	N	N	Y
SF12_MCS (sf12mcs_dv)	Short Form 12 Mental Component Score	N	N	N	Y
SF_12_PCS (sf12pcs_dv)	Short Form 12 Physical Component Score	N	N	N	Y
fruit_day (wkfruit)	Days per week fruit is eaten	N	N	nutrition_quality	Y
veg_day (wkvege)	Days per week vegetables are eaten	N	N	nutrition_quality	Y
fruit_per_day (fruitamt)	Fruit eaten per day	N	N	nutrition_quality	Y
veg_per_day (vegeamt)	Vegetables eaten per day	N	N	nutrition_quality	Y
energy_payment_methods (gaspay,elecpay,duelpay)	Methods of gas, electric, or combined energy bill payment	N	Y	Y	Y
behind_on_bills (xphsdba)	Subjectively behind on bills?	N	N	N	Y
active	moderate and physical activity minutes (mwmin,vwmin)	N	N	Y	Y
auditc	alcohol dependent risk score	N	N	Y	Y
housing_tenure (tenure_dv)	housing tenure (owned/rented/etc.)	N	N	N	Y
ncars	Number of cars	N	Y	N	Y
number of rooms (hsrooms)	Number of bedrooms	N	Y	N	Y

Table 5.3: Imputation Strategies for each variable used in MINOS to estimate housing policy.

Variable Name	Description
Disposable Income fihhmnnet1_dv	 Household net income

rentgrs_dv	Gross monthly rent.
xpmg_dv	Gross monthly mortgage.
ctband_dv	Council tax band.
ieqmoecd_dv	Modified OECD equivalence scale.
Housing Quality	
cduse5	Do you have a fridge freezer in your household?
cduse6	Do you have a washing machine in your household?
cduse7	Do you have a tumble dryer in your household?
cduse8	Do you have a dishwasher in your household?
cduse9	Do you have a microwave in your household?
hheat	Do you have sufficient heating in your household?
Neighbourhood Safety	
crburg	Are there burglaries in your neighbourhood?
crcar	Is there car crime in your neighbourhood?
crdrnk	Is there drunken disorder in your neighbourhood?
crmugg	Are there muggings in your neighbourhood?
crrace	Are there racial attacks in your neighbourhood?
crteen	Are there issues with teenagers in your neighbourhood?
crvand	Is there vandalism in your neighbourhood?
Nutrition Quality	
wkfruit	How many days per week do you eat fruit?
fruitamt	How many portions of fruit do you eat on a day when you eat fruit?
wkveg	How many days per week do you eat vegetables?
vegeamt	How many portions of vegetables do you eat on a day when you eat veg?
Hourly Wage	
basrate	Basic bay hourly rate
paygu_dv	Usual gross pay per month: current job
jbhrs	Number of hours normally worked in a week
auditc1 to audict5	AUDITC questions on alcohol consumption (Bush et al., 1998).
moderate_activity_hours	How many hours of moderate physical activity do you perform per week?
vigorous_activity_hours	How many hours of vigorous physical activity do you perform per week?
yearly_electric	What is your yearly electricity expenditure?
yearly_gas	What is your yearly gas expenditure?
yearly_gas_electric	What is your yearly combined gas and electricity expenditure?
yearly_oil	What is your yearly oil expenditure?
yearly_other	What is your yearly solid and other fuel expenditure?

energy_payment_methods	Methods of payment for each energy type. (Variable Tariff, prepayment meter, etc.)

Table 5.4: Understanding Society variables for generation of derived composite variables in MINOS to estimate housing policy.

5.3.4 Transition Dynamics

Estimating policy effect requires projecting the input population forwards in time using transition dynamics. Individual states are updated using Markov transition probability models estimating future state using current time information as well as longitudinal histories and within household interaction information. Transition dynamics are represented as a series of modules in five specific groups for demographics, income, intervention, intermediaries and outcomes show in Figure 5.1a and Figure 5.1b.

Demographics Stage

The first set of MINOS modules update individual demographics. These are variables not contained within the income-health causal pathways (SIPHER Consortium et al., 2023; Consortium, 2023) (see Figure 5.1b) but are required for preservation of household structure. Individual ages are updated deterministically adding one year to their age on every time step.

A mortality module is used to determine if an individual dies within a given year utilising stratified lookup table data from the NEWETHPOP project (Rees et al., 2017) derived from NOMIS midyear estimates (Nash, 2019). This provides probability of death within a given year stratified by age, sex, and ethnicity. This probability is then used to create a Bernoulli random variable (Andreassen et al., 2020) with a binary outcome determining if an individual dies or stays alive for a given year. Once individuals or their entire household dies, they are no longer transitioned within the model.

Similar rate tables are provided for fertility stratified using age, sex, and ethnicity (Rees et al., 2017) but includes child parity such that the number of children a woman already has adjusts their probability of giving birth.

Income

The most complex component of the MINOS model requires updating of household income subject to policy intervention. The goal is to estimate the final household disposable income at future state $t+1$ as a function of net household income and overheads predicted using information from current t time. The formula for disposable income is given as follows:

$$disposableincome = \frac{(nethouseholdincome) - (rent) + (mortgage) + (counciltax) - (yearlyenergy}{OECDequivalencescale}. \tag{5.1}$$

All these quantities, excluding yearly energy bills, are calculated using generalised linear mixed models (GLMMs) from R's lme4 package calculating next state using long term individual income histories and current information. The OECD equivalence scale is deterministic dependent on number of adults and children under 16 within a household (Anyaegbu, 2010).

Yearly energy consumption is not included in household disposable income according to the definition of household disposable income used in MINOS. To calculate yearly energy consumption yearly wholesale commodity price is estimate for gas, electricity, and other fuels (coal, oil, etc.) using ARIMA modelling (Gabrielli et al., 2022) on historical quarterly energy prices (Department of Energy Security and Net Zero., 2024). These prices as well as household kilo-watt hour usage are used to update energy bill costs for each year nationally across the UK. Details for these models and projected forecasts of wholesale pricing are provided in Appendix 5.3.5. Using yearly energy estimates, the FP10 score, a binary indicator that is true if a household spends more than 10% of their net income on energy bills, is used as an indicator of energy poverty. More modern measures for energy poverty such as low-income low energy efficiency (LILEE) (Qaseem et al., 2024) are preferable but not available due to no data on household energy efficiency ratings. Finally, two subjective measures for financial well-being are used providing ordinal indicators for a person's subjective financial situation as well as their perceived ability to pay bills.

Other 'employment quality' variables are estimated and used to determine household including education, labour state, and the National Statistics Socio-Economic Classification (NSSEC) (Chandola and Jenkinson, 2000) employment socio-economic score. These variables are all categorical and updated using single layer neural network models from R's nnet package, described further in Appendix 5.3.5.

Intervention Scenarios

After household disposable income is calculated any policy intervention is then applied to the population. Interventions used in this paper intervene solely upon yearly energy consumption, and hence household disposable income, as well as probability of subjective thermal comfort measured by a 'heating' variable. The cost of a policy intervention is also calculated by household and used in final cost-effectiveness calculation. Further specification for each policy intervention applied as well as adjustment in household income, heating, and intervention cost are provided in Section 5.4

Intermediary Modules

Changes in income and other variables are propagated through other individual characteristics and can result in changes in behaviours that are ultimately connected to health. To emulate these changes a set of intermediary modules are used estimating variables that are influenced by change in income and heating. For example, if a household gains more disposable income they may be able to spend money on more nutritional food, which can influence both physical and mental well-being. Definitions of response variables and further detail of implementation for these modules is found in Table 5.5 and Appendix 5.3.5.

These modules are dependent on response variables. Health intermediaries include lifestyle habits that go on to influence health. These include tobacco and alcohol consumption (number of cigarettes smoked and alcohol use disorder test (auditc) (Bush et al., 1998) score), general nutrition quality, and loneliness.

Housing quality variables are directly related to the condition of the physical properties of the households. These include composite measures of material deprivation 'housing_quality' (Aki Tsuchiya, Guoqiang Wu, 2021) and neighbourhood safety, subjective heating ability, housing tenure state (renting, mortgaged, etc.), and number of cars.

Finally a series of subjective financial situation variables are provided determining the household opinion on finances. Two households on the same income can have drastically different opinions on their fiscal health which strongly influences mental well-being. Two ordinal variables rated from 1 to 5 are provided for whether a household feels like they are behind on household bills as well as their overall financial situation. For example, a score of 1 for being behind on bills implies a household is struggling heavily to pay household bills.

Health Outcome Modules

The final stage of transition dynamics involves estimation of health outcomes. This stage is not Markovian but rather estimates latent health outcomes for the current year. There are four health variables of interest. Short-Form 12 Scores (Lawrence and Fleishman, 2004) are derived from a short questionnaire providing approximately Normal distributed $N(50, 10^2)$ measures of individual mental (SF-12 MCS) and physical (SF-12 PCS) well-being. These are well-established publicly available health measures that are present across multiple years of survey data (Johnson et al., 2023) and commonly used in microsimulation (Katikireddi et al., 2022).

The main drawback of SF-12 scores is they are difficult to interpret; a one point change in SF-12 MCS score can be highly contextual. Alternatively, Quality Adjusted Life Years (QALYs) (Lawrence and Fleishman, 2004) are a more interpretable measure used by the UK government to determine change in quality of life. There exists a quadratic formula (Lawrence and Fleishman, 2004) between SF-12 scores and QALY used to derive final QALY scores in MINOS. Additionally Incremental Cost Effectiveness Ratios (ICERs) (Gafni and Birch, 2006) are utilised dividing total policy cost by total QALY gain providing a final measure of cost-effectiveness. This measure is also widely used in the UK Government according to Green Book guidelines (Government Finance Function, HM Treasury, United Kingdom Government, 2024) whereby an ICER ratio of £70000 per QALY is considered to be a cost-effective policy.

Replenishment

Finally replenishing the input population is implemented at the end of the microsimulation year. This entire process is then repeated for every year until the time horizon is met.

Module	Description	Model Type
Demographics		
Births	How many children are born per woman each year.	Rate Tables
Deaths	Which individuals die each year.	Rate Tables
Ageing	Ageing each individual and their children one year for each time step (years).	Deterministic
Education State	Calculating a new highest education state (The UK Government provides levels 1 to 9 (UK Government, 2019)).	Categorical nnet
Labour State	Calculating a new labour state. (Retired/Employed/etc.)	Categorical nnet
NSSEC	Calculating a new NSSEC employment score (1-8)	Categorical nnet
Income		
Net Income	Calculating new net household income (£s)	Gamma-Normal GLMM
FP10	Calculate if more than 10% of income is spent on energy bills.	Deterministic

Rent	Calculating new net household rent (£s)	Normal-Normal GLMM
Mortgage	Calculating new net household mortgages (£s)	Normal-Normal GLMM
Council Tax	Calculating new yearly energy bills subject to change in wholesale pricing (£s).	Normal-Normal GLMM
Energy Wholesale Pricing	Forecast wholesale pricing of gas, electricity, and other fuel pricing.	ARIMA time series models
Intervention		
Good Heating Dummy	Adjusts all households to have comfortable heating.	Deterministic
Energy Price Cap Guarantee	Adjusts household yearly energy bills according to a £3000 energy cap	Deterministic
Great British Insulation Scheme	Reduce household energy bills and increase change of comfortable heating subject to housing characteristics.	Deterministic
Intervention Cost	Calculate the population-level cost of an intervention.	Deterministic
Intermediaries		
Nutrition Quality	Calculate weekly fruit and vegetable consumption.	Normal-Normal GLMM
Neighbourhood Safety	Calculate neighbourhood composite safety (1-3)	Ordinal CLM
Financial Situation	Calculate subjective financial situation (1-5)	Ordinal CLM
Behind On Bills	Calculate whether an individual is subjectively behind on bills (1-5).	Ordinal CLM
AUDITC Alcohol Consumption	Calculate alcohol consumption risk score (1-3)	Ordinal CLM
Physical Activity	Calculate whether an individual qualifies for adequate physical activity (0/1).	Logistic Regression
Housing Quality	Estimate new housing material deprivation composite (1-3)	Ordinal CLM
Loneliness	Calculate sujective loneliness scale (1-3).	Ordinal CLM
Heating	Calculate whether a household has comfortable subjective heating (0/1).	Logistic Regression
ncigs	Calculate weekly cigarette consumption.	Mixed ZI-Hurdle model

ncars	Calculate the number of cars used by a household.	Mixed ZI-Hurdle model
Health Outcomes		
SF-12 MCS Score	Calculate the Short Form 12 Mental Component Score	Gamma-Normal GLMM
SF-12 PCS Score	Calculate the Short Form 12 Physical Component Score	Gamma-Normal GLMM
QALY Score	Calculate Quality Adjusted Life Years from	Deterministic
ICER Score	Calculate the incremental cost-effectiveness ratio using QALYs and intervention cost.	Deterministic

Table 5.5: All modules used in the MINOS dynamic microsimulation model to estimate housing policy effect on health.

5.3.5 Transition Probability Models

This section describes all transition models used to generate the MINOS model. Descriptions are provided on modelling methods [2] used dependent on outcome variables, literature justification for each independent variable, and some goodness of fit diagnostics. This information is repeated here for simplicity and easier splitting this thesis up into separate articles. Again model coefficient tables are provided in thesis appendices in Section 8.2.2.

Continuous Variables

This section contains detail on the implementation of all transition probability models used in MINOS that are specified in Table 5.5. All transition models here generate randomised fixed effects coefficients at the start of each MINOS run assuming parameters are multi-variate normal distributed.

GLMMs GLMM models described in Section 4.3.6 are fitted in MINOS estimating household net income, rent and mortgage spending, council tax, yearly energy spending, nutrition quality, and SF-12 MCS/PCS scores. Again Gamma-Normal and Normal-Normal GLMMs are used dependent on skew in outcome variables. The first series of GLMM models are used to estimate household disposable income. Initially household net income is estimated using the formula

$$\begin{aligned} net\_hh\_income = exp(&previous\_net\_hh\_income + (previous\_net\_hh\_income)^2 + age + age^2 \\ &+ age^3 + sex + ethnicity + region + education\_state \\ &+ NSSEC + SF\_12\_MCS + (intercept|pidp)). \end{aligned}$$

This formula utilises common socio-demographic predictors used to estimate household net income (Ali et al., 2019; Mallet and Weale, 2018) including age, sex, ethnicity, region, education and socio-economic code. This model also utilises random intercepts by individual identifiers (pidp). A lagged dependent

[2]There is substantial overlap with transition probability methods used in 4. These methods are not repeated but references to previous sections are provided.

variable (McLay et al., 2015) is also used estimating current income using previous income states. Model coefficients are provided in Table 8.12.

For estimating household rent a similar formula is used

$$\begin{aligned}hh\_rent = & net_h h_i ncome + age + age^2 + sex + ethnicity + region \\ & + education\_state + NSSEC + labour\_state + time + SF\_12 \\ & + marital\_status + bedrooms + housing\_quality \\ & + (intercept|pidp)).\end{aligned}$$

This model uses the same age, education, region, marital status, and ethnicity (Samarin et al., 2024) are used to estimate household income. Economic variables including net income, socioeconomic code (NSSEC), labour state, and housing tenure are also significant in literature (Samarin et al., 2024). Some household characteristics including housing quality and number of bedrooms (Amenyah and Fletcher, 2013) are used but further spatial variables including commute time and amenities access would ideally be included but are not present in UKHLS data. Lagged dependent variables for previous rent costs are also used. Note all households not in the rental housing tenure state are automatically assigned 0 rent costs. Model coefficients are provided in Table 8.13.

Estimating monthly household mortgage is performed using a similar model with formula

$$\begin{aligned}hh\_mortgage = & exp(previous\_hh\_mortgage + age + age^2 + age^3 + region + education\_state \\ & + NSSEC + time + bedrooms + housing\_quality + marital\_status \\ & + (intercept|pidp)).\end{aligned}$$

Age, education state, and government region, and martial status demographic variables are used (Koblyakova et al., 2014). Household size (bedrooms), and material deprivation housing quality and used to approximate household value and are used to predict monthly mortgage spending (Koblyakova et al., 2014). Income variables include previous mortgage and net income expenditure, employment status, and socio-economic codes (NSSEC). Model coefficients are provided in Table 8.14.

Estimation of monthly council tax expenditure uses the formula

$$\begin{aligned}council\_tax = & exp(hh\_income + age + age^2 + age^3 + sex + ethnicity + region \\ & + education\_state + time + NSSEC + housing\_tenure + (intercept|pidp)).\end{aligned}$$

Council tax expenditure is conditional on estimated house price (Giles and Ridge, 1993). Assuming households do not change in the MINOS model future council tax is estimated solely using current council tax spending as well as demographics including age, sex, ethnicity, and government region. Model coefficients are provided in Table 8.15.

Finally household yearly energy expenditure is estimated using a GLMM model with formula

$$\begin{aligned} yearly\_energy = {} & previous\_yearly\_energy + financial\_situation \\ & + housing\_tenure + age + net\_hh\_income + heating \\ & + housing\_quality + region + (intercept|pidp)). \end{aligned}$$

Demographic variable added include ethnicity (Graff et al., 2021), age, labour state particularly pension age and unemployed adults and number of household occupants including young children (Jones and Lomas, 2015; Taneja and Mandys, 2022). Income variables include net income, previous gas and electricity prices, and subjective financial state (Graff et al., 2021; Taneja and Mandys, 2022; Jones and Lomas, 2015). Household variables including urban or rural location, number of bedrooms, heating ability and general material deprivation (housing quality) (Jones and Lomas, 2015; Taneja and Mandys, 2022). Further household variables including floor space and number of shared walls, and water heating equipment (Jones and Lomas, 2015; Taneja and Mandys, 2022) would ideally be included but are not present in UKHLS data. Model coefficients are provided in Table 8.16.

OECD Equivalence scale (Anyaegbu, 2010) is a measure used to adjust household income according to the number of inhabitants within a household. It is broadly used as a method to account for 'economy of scale' within a household such that a household with more individuals is generally cheaper to run per person. OECD scale it calculated as a simple score adding 1 for the initial household adult, 0.5 for every additional adult greater than or equal to 16 years old, and 0.25 for every child under 16 years old. For example, a household with two adults and two child would be given score $1 + 0.5 + 0.25 + 0.25 = 2$. These models are then used to construct household disposable income described in composite generation above.

There are a series of other GLMM models used to predict health intermediate and outcome variables. Nutrition quality is estimated using a Normal-Normal GLMM as response data is not skewed using formula

$$\begin{aligned} nutrition\_quality = {} & previous\_nutrition\_quality \\ & + age + sex \\ & + education\_state + region + ethnicity + hh\_income \\ & + ncigs + SF12\_MCS + (1|pidp) \end{aligned}$$

Nutrition quality independent variables are derived from social determinants of nutrition and obesity literature (Javed et al., 2022; French et al., 2019; Hughes and Kumari, 2017). Variables for common demographics including age, sex, education, and ethnicity are provided. Household income, previous health conditions (SF-12 MCS), smoking, and lagged nutrition quality are also used as recommended by the literature. This model utilised random intercepts due to multiple longitudinal waves of data present and was found to improve model fit. Model coefficients are provided in Table 8.4

To estimate SF-12 MCS score a Normal-Normal GLMM is also used. SF-12 MCS data are reflected converting them from left-skewed to right-skewed then taking logarithms to produce an approximately Normal distributed outcome variable often used to estimate highly skewed data (Manning, 2012). Independent variables used to predict SF-12 MCS are sourced from existing studies on UKHLS data

(Parra-Mujica et al., 2023). Indepdent variables include typical age, sex, and ethnicity. Additional variables have also been included as part of SIPHER analysis on predictors of mental well-being including housing quality, neighbourhood safety, nutrition, and smoking (Aki Tsuchiya, Guoqiang Wu, 2021; Consortium, 2023). Lagged dependent variables and random intercepts are also used (Parra-Mujica et al., 2023). Model coefficients are provided in Table 8.5.

$$
\begin{aligned}
SF12\_MCS = exp((&previous\_SF\_12\_MCS + previous\_SF\_12\_MCS^2 \\
&+ sex + ethnicity + age + housing\_quality \\
&+ behind\_on\_bills + financial\_situation + yearly\_energy \\
&+ net\_hh\_income + heating + region + neighbourhood\_safety \\
&+ housing\_quality + loneliness + nutrition\_quality \\
&+ ncigs + (intercept|pidp))
\end{aligned}
$$

A GLMM model is also provided estimating SF-12 PCS score using formula

$$
\begin{aligned}
SF12\_PCS = exp((&previous\_SF\_12\_PCS + (previous\_SF\_12\_PCS)^2 \\
&+ sex + ethnicity + financial\_situation + yearly\_energy \\
&+ heating + sex \\
&+ age + age^2 + education\_state \\
&+ region + education\_state + housing\_quality + net\_hh\_income \\
&+ nutrition\_quality + NSSEC + ncigs + loneliness + active \\
&+ auditc + (intercept|pidp)).
\end{aligned}
$$

Demographic variables used to estimate SF-12 PCS include age, education, housing tenure, sex, and ethnicity (Mason et al., 2016; Callan et al., 2015). Income variables include household income and subjective financial well-being (Downward et al., 2020), energy bills (Grey et al., 2017), labour state, car usage, and socioeconomic (NSSEC) codes (Callan et al., 2015; Mason et al., 2016). Lifestyle variables including alcohol and cigarette consumption (Dissing et al., 2013), physical activity, nutrition quality, loneliness (Grey et al., 2017), and previous physical health state (Mason et al., 2016) particularly chronic illnesses. Finally housing variables include subjective thermal comfort and overall housing material deprivation (Dissing et al., 2013; Grey et al., 2017). Model coefficients are provided in Table 8.19.

Counts Variables Some count variables such as the number of cigarettes smoked per month are zero inflated. The majority of the population do not smoke and have value 0 skewing outcome data. Non-zero counts data are also highly overdispersed with some extreme values of cigarettes consumption exceeding 100 units per week. Standard zero inflated Poisson modelling (Jackman et al., 2015) was insufficient to correctly estimate cigarette consumption and can only be utilised for a single yearly wave of data due to repeat observations. In order to correctly estimate cigarette consumption a mixed effect hurdle model is used (Altinisik, 2023) is used. This is a two stage model that estimates the probability a person smokes at all and the number of cigarettes a month they would smoke. Random sampling is used to determine whether a person smokes or not depending on their probability. If they do not smoke they are assigned value 0. If they do smoke they are assigned their predicted counts value. The hurdle model is implemented in R. using the GLMMAdaptive package (Rizopoulos, 2022) that requires two sets of predictors for smoking at all and how many cigarettes smoked respectively. These are referred to as zero and counts models with the following formulae and coefficients. Formulae and coefficients for the cigarettes consumption and number of cars are provided below. Note both models implicitly utilised random intercept effects for each individual identifier (pidp) when estimating both consumption and non-zero status.

Two formulae are provided estimating cigarette (ncigs) consumption for the zero model

$$ncigs\_zero = previous\_ncigs + age + sex + ethnicity + (1|pidp)$$

and the counts model

$$\begin{aligned} ncigs\_count = {} & previous\_ncigs + age + age^2 + nutrition\_quality + sex + ethnicity + time \\ & + (1|pidp). \end{aligned}$$

For both models prediction is based on age, sex, and ethnicity demographics all found to significantly affect consumption (Chen et al., 2019; Chandola et al., 2004). Income, employment, socio-economic state (NSSEC) and previous smoking behaviour are all also included when predicting both states. Additinoal random effects modelling is desirable (Chandola et al., 2004) but not possible with the pscl R. package used to fit these models. Model coefficients are provided in Table 8.20.

The hurdle model is also utilised to estimate car ownership by household which still has a large number of non-zero values that cannot be explained by non-inflated counts models. Formulae for the ncars zero and counts models are given a s

$$\begin{aligned} ncars\_zero = {} & previous\_ncars + net\_hh\_income + region + nkids + urban + age + time \\ & + (1|pidp) \end{aligned}$$

and

$$\begin{aligned} ncars\_count = {} & previous\_ncars + age + age^2 + region + nkids + net\_hh\_income \\ & + labour\_state + time + (1|pidp). \end{aligned}$$

Demographic variables include region, urban or rural classifier, housing tenure, education, age, household size using number of children, and sex (Thomas et al., 2016; Goodman et al., 2012). Previous physical well-being household income, socio-economic code (NSSEC), and a linear time trend are also included (Goodman et al., 2012; Thomas et al., 2016). Model coefficients are provided in Table 8.21.

Energy Pricing ARIMA models To estimate yearly whole energy pricing UK BEIS data for yearly electricity, gas, solid fuel, and other fuel price data are used (Department of Energy Security and Net Zero., 2024). These provide time series data from $1990 - 2023$ used to estimate yearly energy pricing. ARIMA models (Jan et al., 2022) are fitted to each wholesale pricing time series in R. (Dhamo and Puka, 2010) choosing the best model structure based on AIC score. These models are demonstrated below in Figure 5.6. In the MINOS run these values are used to inform energy pricing and update household income accordingly. Energy pricing is randomised each time drawing additional noise from normal distributions $N(0, 50^2)$.

Ordinal Variables

Several models are utilised to estimate ordinal variables in MINOS.

Cumulative Link Modelling Cumulative link modelling in Section 4.3.6 is again used to estimate change in ordinal variables. Estimating the housing quality composite as a proxy for material deprivation is performed using a CLM model with formula

$$\begin{aligned} housing\_quality = {} & housing\_quality\_last + age + age^2 + sex + region + education\_state \\ & + SF12\_MCS + ethnicity + hh\_income + hh\_income^2 \\ & + housing\_tenure + loneliness + neighbourhood\_safety. \end{aligned}$$

Estimation of household material deprivation uses common demographic indicators including age, sex, and ethnicity (Whelan and Maitre, 2012). Income related terms including net household income, housing tenure, socio-economic class (NSSEC), previous housing tenure, and labour state. Mental-wellbeing is also used due to significance of similar variables including global happiness index (GNDH) (Whelan and Maitre, 2012). Model coefficients are provided in Table 8.7

Loneliness is also estimated using a CLM model with formula

$$\begin{aligned} loneliness = {} & previous\_loneliness + age + sex + ethnicity + region + education\_state \\ & + housing\_quality + neighbourhood\_safety + nutrition\_quality + ncigs + NSSEC \\ & + net\_hh\_income + marital\_status \end{aligned}$$

Demographic predictors including age, sex, ethnicity, marital status (living alone), and education have shown to significantly predict loneliness before and after the coronavirus pandemic (Bu et al., 2020). Economic variables for household income and labour state are also recommended. Previous loneliness and SF-12 MCS health scores have also been included. Model coefficients are provided in Table 8.9

Finally, a CLM model has been fitted estimating neighbourhood safety using formula

$$\begin{aligned} neighbourhood\_safety = {} & previous\_neighbourhood\_safety + age + sex + region + NSSEC \\ & + ethnicity + hh\_income + housing\_quality + loneliness \\ & + nutrition\_quality + ncigs \end{aligned}$$

Prediction of neighbourhood safety requires estimating the frequency anti-social events (Warr et al., 2009) a household experiences. This is difficult to do without specific spatial data but there are studies estimating the social determinants of these events (Livingston et al., 2014; Baum et al., 2009). Deprivation is the strongest predictor of crime rates including material deprivation (housing_quality) as well as socio-economic state (NSSEC) and household income. Age and sex variables are often included in studies but are often not statistically significant. Other demographic variables including ethnicity and housing tenure, particularly prevalence of short-term social renting, are also strong indicators of crime rates. Loneliness is used as an indicator of mental well-being and social inclusion. Again previous neighbourhood safety is also included. Model coefficients are provided in Table 8.8

A CLM model is also used to estimate household subjective financial well-being.

$$\begin{aligned} financial\_situation = {} & previous\_financial\_situation + age + sex + NSSEC + ethnicity \\ & + hh\_income + region + yearly\_energy + nkids + SF\_12\_PCS \\ & + SF\_12\_MCS \end{aligned}$$

Demographic variables include age, sex, ethnicity, marital status, number of children, and education (Fan and Babiarz, 2019; Gray, 2013).Income variables include previous financial situation, NSSEC, and labour state (Fan and Babiarz, 2019; Gray, 2013). Previous physical health is also used (Gray, 2013). Model coefficients are provided in Table 8.25.

A CLM model is similarly used to estimate households being behind on bills

$$\begin{aligned} behind\_on\_bills = {} & previous\_behind_o n_b ills + yearly_e nergy + ethnicity + age \\ & + SF\_12\_MCS + net\_hh\_income + net\_hh\_income^2 + financial_s ituation \end{aligned}$$

Demographic variables age, sex, education, number of children, and marital status are included (Oksanen et al., 2015). Income variables include previous behind on bills status, household income, and energy bills (Oksanen et al., 2015; Stone et al., 2023). Previous mental well-being state is also used (Gray, 2013). Model coefficients are provided in Table 8.26.

NNET Again NNET single layer neural networks from Section 4.3.6 are used to estimate change in categorial variables. An nnet model is used to estimate audit-c score

$$\begin{aligned} auditc = {} & previous\_auditc + age + sex + education\_state + financial\_situation \\ & + yearly\_energy + SF\_12\_MCS + heating + SF\_12\_PCS + hh\_income \\ & + active + ethnicity + housing\_quality + neighbourhood\_safety + region \end{aligned}$$

Variables used to estimate audit-c are cited from wider high alcohol consumption determinants literature. Demographic variables include age, sex, education, marital status, ethnicity, and number of children (Dias et al., 2011; Wilson, 2020; Pollack et al., 2005). Health and lifestyle variables include physical activity, tobacco use, nutrition quality, and SF-12 PCS (Dias et al., 2011; Wilson, 2020).Income variable for labour state and household income variables are also used (Wilson, 2020; Dias et al., 2011; Pollack et al., 2005).

Finally variables for household material deprivation and neighbourhood safety are also provided (Pollack et al., 2005).

Another nnet model is also used to estimate UK labour state with formula

$$\begin{aligned} labour\_state = {} & previous\_labour\_state + financial\_situation + age + sex + ethnicity + region \\ & + education\_state + marital\_status + housing\_tenure + nkids + SF\_12MCS. \end{aligned}$$

Estimating labour state requires predicting employment participation as well as reasons for not working including being a carer, in education, retired, or otherwise economically inactive. Demographic variables used to predict labour state include age, sex, ethnicity, region, education state, marital status, housing tenure, and number of children (Quarina, 2017). Previous health state SF-12 MCS and previous labour state are also used (Quarina, 2017).

Binary Variables

The standard logistic regression is implemented to predict binary outcomes such as heating. This is implemented using the R.'s glm (Venables and Smith, 2003) function which is part of the built in 'stats' package.

Subjective thermal comfort (heating) is estimated using a logit model with formula

$$\begin{aligned} heating = {} & previous\_heating + yearly\_energy + financial\_situation + behind\_on\_bills \\ & + net\_hh\_income + nkids + housing\_tenure + marital\_status. \end{aligned}$$

Demographic variables include age, sex, ethnicity, region, housing tenure, marital status, and number of children (Ruse et al., 2019; Sawyer et al., 2022). Income variables include household income and labour state (Ruse et al., 2019; Sawyer et al., 2022). Housing variables include number of bedrooms, material deprivation, and previous heating. (Ruse et al., 2019; Sawyer et al., 2022). Model coefficients are provided in Table 8.27.

Finally, a logistic model is used to estimate individual weekly physical activity

$$\begin{aligned} active = {} & age + sex + education\_state + financial\_situation + behind\_on\_bills \\ & + yearly\_energy + SF\_12\_MCS + heating + SF\_12\_PCS + net\_hh\_income \\ & + education\_state + marital\_status + ethnicity \end{aligned}$$

Demographic variables include age, sex, ethnicity, education, and marital status (Choi et al., 2017). Income variables include purely household income and yearly energy spending (Choi et al., 2017). Several health and lifestyle variables are used including alcohol and tobacco use, SF-12 MCS and SF-12 PCS, loneliness, and previous physical activity (Choi et al., 2017). Model coefficients are provided in Table 8.28.

Rate Tables.

Again rate tables are used to estimate fertility and mortality updated as described in Section 4.3.6.

Deterministic Modules.

Ageing Individual ages are incremented yearly by 1 every time step. The ages of their children are also incremented by 1 such that when children reach 16 they are considered to be 'adults' and removed from the count of children 'nkids' within the household.

5.3.6 Validation

Validation methods are identical to those in chapters 3 - 4 using cross-validated nowcasting (Mallet and Weale, 2018) and handover plots (Archer et al., 2021) to ensure predicted behaviour matches real data and is 'reasonable and credible' preserving linear trends in mean and variance when predicting into the future. Five-fold cross validated nowcasting plots are provided for key financial situation, heating, cigarette consumption, and SF-12 MCS variables in Figures 5.7a - 5.7d. Further validation can be found in online supplementary material.

Handover plots for key yearly energy and SF-12 MCS variables are provided in Figures 5.8a - 5.8c showing reasonable preservation of statistical moments and linear trends over time. Further visualization for validation of other variables used in MINOS is provided in online repositories.

5.3.7 Uncertainty

Again from chapters 3 - 4 there are four key uncertainty sources in MINOS to address due to input data sources, model coefficients, model structure, and Monte Carlo noise.

Methods to address model coefficients, model structure, and Monte Carlo noise are the same as in Chapter 4 using randomisation of model coefficients, variable selection methods and iterative assessment of independent variables, and addition of random nrmal noise to transition models where possible.

There are now three source of uncertainty due to input data. Uncertainty due to data source again is considered using cross-validation methods. Uncertainty due to missing data is now influenced by (MICE) algorithm. A recommended set of 30 (Sterne et al., 2009) MICE imputed populations are provided and used in the MINOS input population. Assessment of variance within these imputed population is performed as described in Chapter 3.5 testing stable converge towards population summary statistics for variables such as SF-12 MCS. Additionally for each model run of MINOS a random choice of one population from this sample is taken and used as MINOS input data. Due to computational constraints the stochastic simulated annealing algorithm used to generate this population is only run once. In order to try and account for uncertainty, bootstrapping is combined with a random MICE imputed population producing a random 10% sample of the full synthetic population for each MINOS model run. Future work may further vary the size of this sample population, choice of validation constraints for the simulated annealing algorithm, and inclusion of further variance reduction techniques. Additionally this work may utilise all 30 MICE populations at once using methods such as probabilistic microsimulation (Manoukian et al., 2022) or data assimilation (Birkin, 2021).

5.4 Interventions

Baseline

In the baseline scenario we assume that the energy prices increase in line with data from from $2020-2023$ and remain high indefinitely according to ARIMA forecasted (Gabrielli et al., 2022) wholesale electricity and gas prices with forecasted values provided in Section 5.3.5. It is assumed that the UK government does nothing to intervene upon household energy prices. This is the worst case scenario with the lowest expected overall public health. It serves as a benchmark to compare how the impact of the following scenarios on health. In this scenario the pricing of wholesale electricity, gas, oil, and other solid fuels are all estimated using ARIMA modelling described in more detail in Appendix 5.3.5. Estimates are calibrated on the Digest of United Kingdom Energy Statistics (DUKES) quarterly energy pricing available from 1970 to present. Wholesale pricing estimates are then used to estimate final energy bills from kWh consumption of each fuel. The change in yearly energy bills is provided in results Figure 5.9a.

Idealistic Energy Scenarios - No Price Increase and the Good Heating Dummy

The first two interventions run are idealistic scenarios that are useful for comparison against real policies. These are 'morning after' scenarios that at the start year 2020 - these interventions are immediately applied as if 'waving a magic wand' and instantly occur. First, we assume that energy pricing does not increase from 2020 onwards and follows previous approximate linear trends. The intent is to determine what effect rising energy costs has on the health of the GMCA population.

Similarly, the good heating dummy intervention *does* assume increases in energy pricing, but also sets every household to have comfortable subjective heating (variable 'heating=1'). This intervention determines how the health of the GMCA population would change if energy prices increase, but everyone remained comfortable. This intervention serves to test how the UK population health responds to perfect thermal comfort and provides a useful benchmark to compare against real policies.

Energy Price Cap Guarentee (EPCG)

The Energy Price Cap Guarantee (EPCG) is a policy implemented by the UK Government from 2022 onwards in order to limit high energy bills. A price cap is set by the UK Government body OFGEM such that limits mean yearly energy bills for households on variable rate tariffs. Any excess is then paid to energy suppliers by the UK government who intend to recoup these costs through future taxation. The price cap is variable dependent on energy pricing varying from £$2000-3500$ between 2021 and 2024 (Bolton, 2023). For simplicity, we assume a constant fixed rate price cap of £3000 for the full life cycle of the MINOS model.

Eligible households on variable gas and electricity energy tariffs have their yearly energy consumption multiplicatively scaled until the mean reaches the £3000 target. Note post processing is applied such that energy bills do not go below the minimum 'floor price' of $34p$ and 10.3 per kWh for gas and electricity respectively (Bolton, 2023). Ineligible members of the UK population include households who are on fixed rate tariffs, whose energy consumption is included in rent, and households relying on other fuels such as coal and oil. These assumptions are highly subject to change in response to market forces, e.g. fixed rate tariffs were no longer offered as energy prices began to rise, as well as behavioural change and housing upgrades away from fossil fuel consumption.

The Great British Insulation Scheme (GBIS)

The UK Government is increasingly implementing policies to try and reduce energy consumption within residential properties (Regan, n.d.; Rafique and Williams, 2021). The Energy Company Obligation (ECO4) is an early example utilised since 2010 subsidising retrofitting of insulation for low income households (Regan, n.d.). This policy saw limited uptake due to low eligibility and poor marketing (Regan, n.d.). Its successor, the Great British Insulation Scheme (GBIS) (Regan, n.d.) has over£$1bn$ invested from UK government to retrofit households that are either in lower tax bands $A-E$ (only $A-D$ in England) or have low energy efficiency ratings of D to G. Council tax bands are categories denoted by letters A to I indicating how much income an individual pays towards council tax dependent on the household value and household structure including number of children (UK Fuel Poverty Monitor, 2022). Eligible households are subsidised to implement a wide variety of possible insulation on a case-by-case basis (Regan, n.d.).

The GBIS scheme is emulated in MINOS by intervening on all households on low council tax bands and households with a positive FP10 energy poverty measure, that is their energy bills exceed 10% of their net income, earning below the median household income. Ideally, household energy efficiency ratings (Regan, n.d.) and modern fuel poverty measures such as low-income low energy efficiency (LILEE) (Qaseem et al., 2024) would also be utilised to identify households in fuel poverty but are not yet present in UKHLS data. Each household is given a recommended £7000 (Emden and Murphy, 2023) contributing towards household retrofitting costs for various insulation types dependent on household characteristics such as stone walled households that can only use solid insulation. Households that receive this payment are assumed to retrofit their household saving £600 in yearly energy bills and increasing the probability of comfortable subjective heating.

5.5 Experiment Technical Specifications

For each intervention the MINOS model is run 100 times producing 500 overall model runs. Each model run uses a random input population with different a MICE imputation population and a different random sample of 10% of households. All models are run in batches on the ARC4 High Performance Computing (HPC) system at the University of Leeds. An individual MINOS model run takes 11 ± 0.45 hours with the full experiment taking approximately 2.5 days. The population is run until the year 2035 for computational to limit overall runtime while still providing a strong medium to long term time horizon to estimate policy cost-effectiveness. Further work may extend this time horizon ideally to 2050 based on Government net-zero targets.

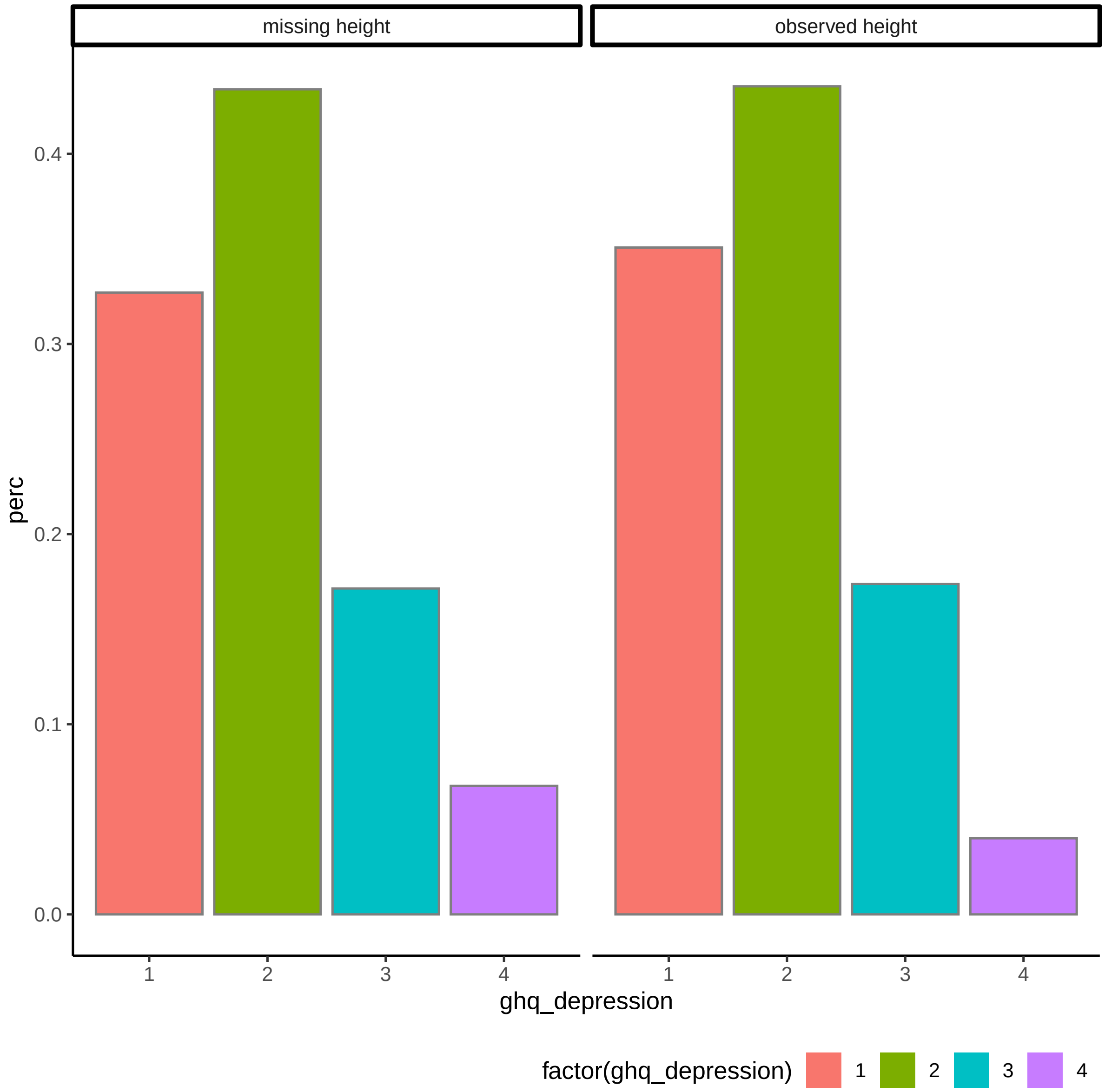


Figure 5.4: Split histogram for ghq depression distributions for observations with missing SF-12 MCS values and those with complete SF-12 MCS values. Overall depression is higher in the population missing SF-12 MCS values providing evidence for inclusion as an auxiliary variable.

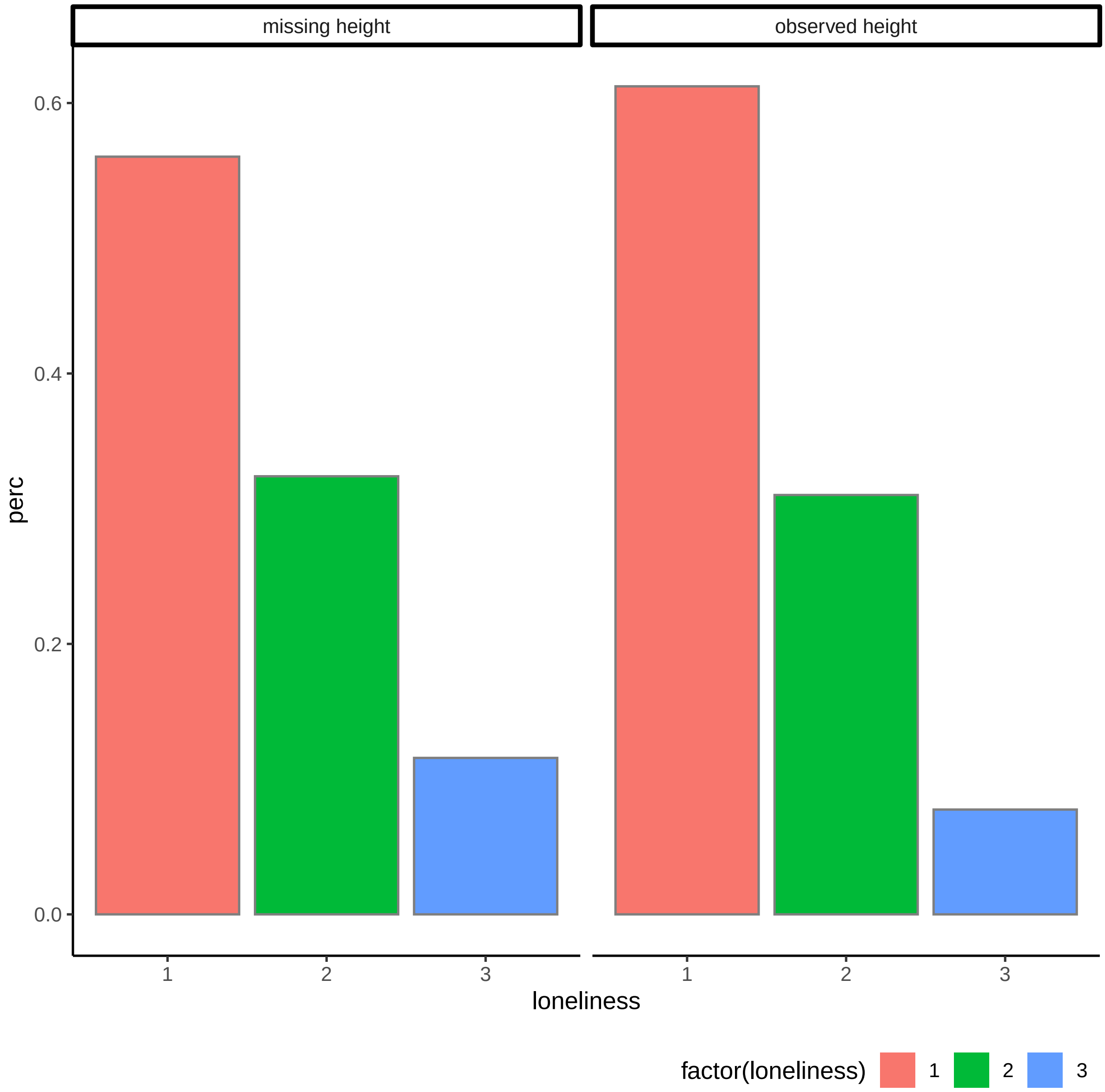
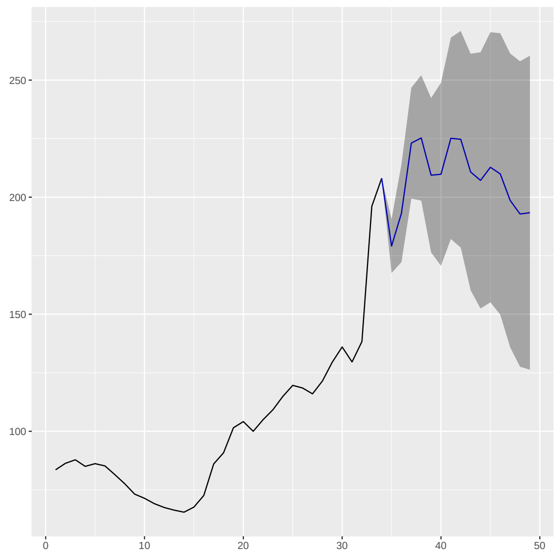
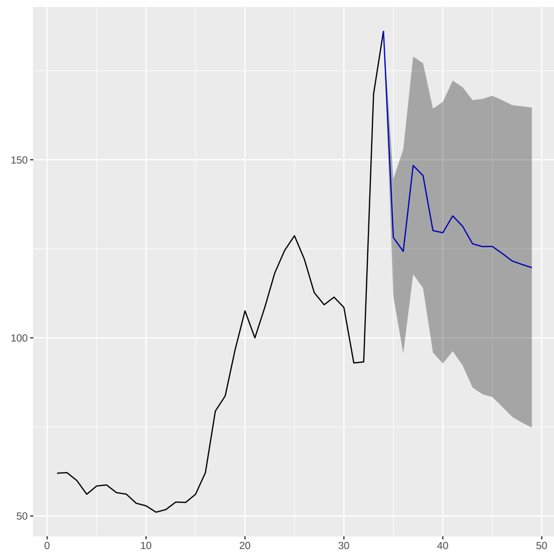


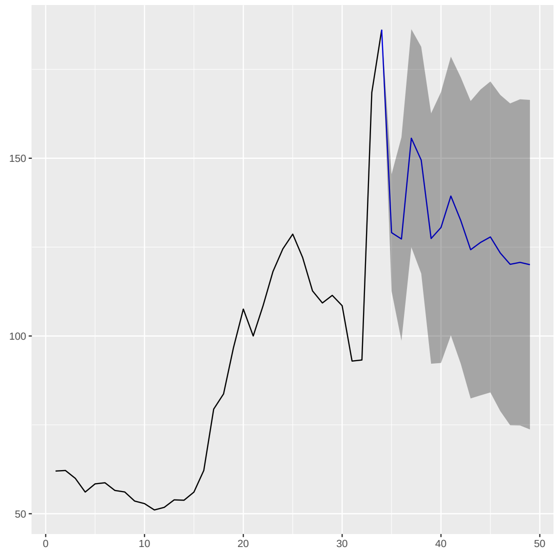
Figure 5.5: Split histogram for loneliness distributions for observations with missing SF-12 MCS values and those with complete SF-12 MCS values. Again loneliness is higher when missing SF-12 MCS suggesting inclusion in MICE imputation is beneficial.



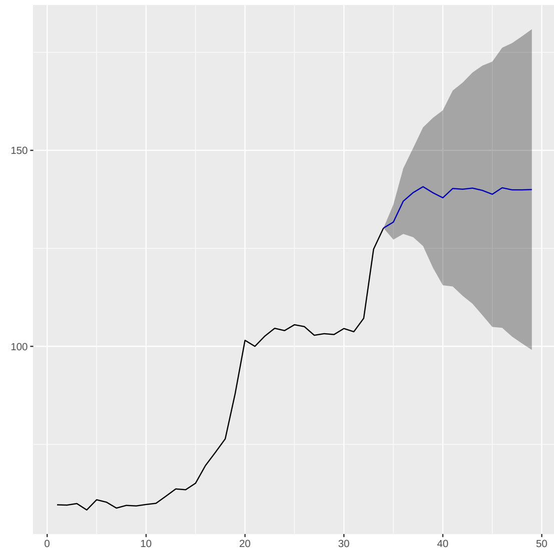
(a) Electricity Pricing ARIMA Forecast.



(b) Gas Pricing ARIMA Forecast.

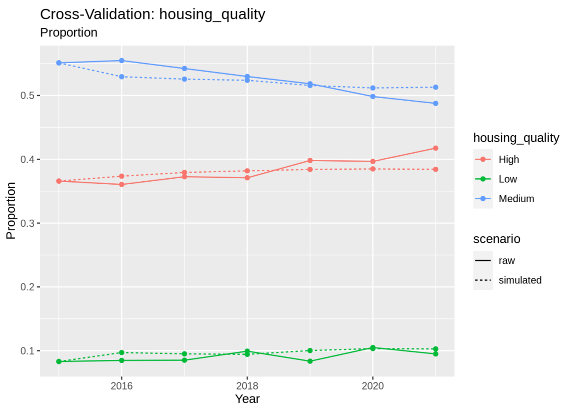


(c) Oil Pricing ARIMA Forecast.

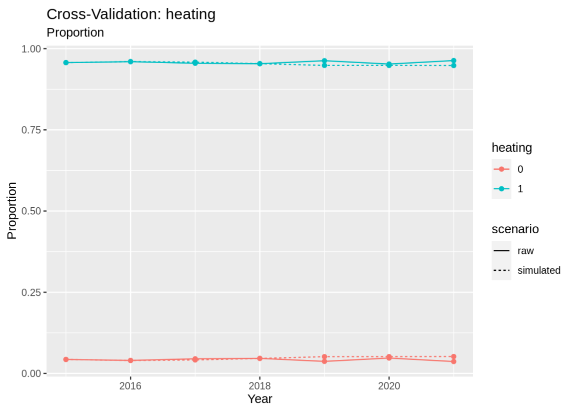


(d) Other Solid Fuel Pricing ARIMA Forecast.

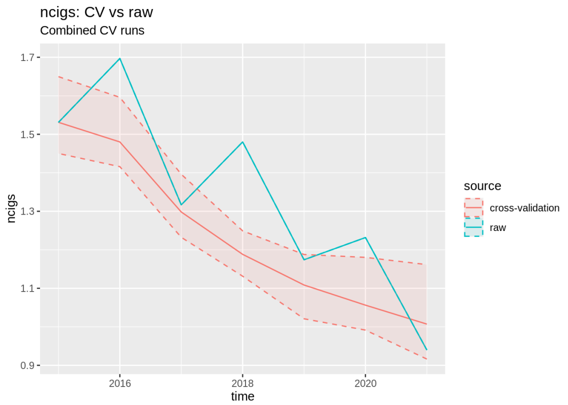
Figure 5.6: Forecasts for wholesale energy pricing using ARIMA models.



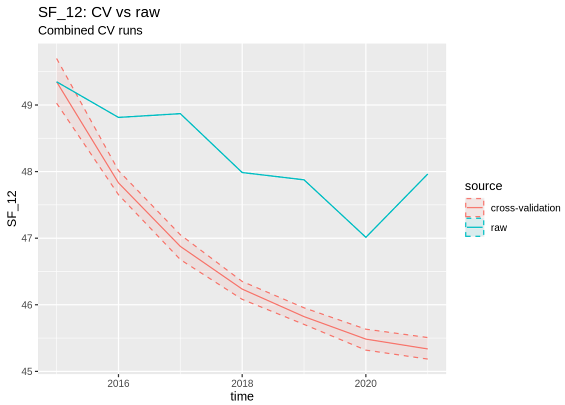
(a) Five-fold cross validated nowcasting estimating housing quality state from 2014 to 2020.



(b) Five-fold cross validated nowcasting estimating subjective thermal comfort (heating) state from 2014 to 2020.

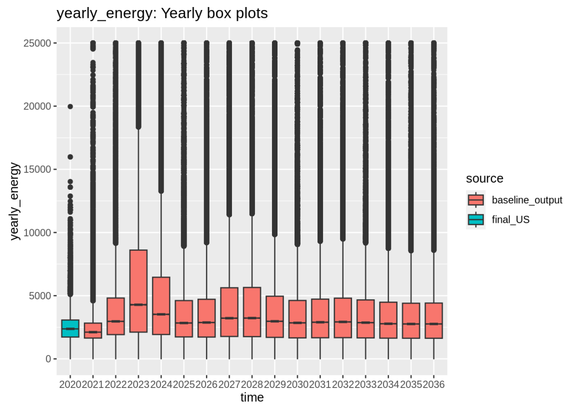


(c) Five-fold cross validated nowcasting estimating cigarette consumption (ncigs) from 2014 to 2020.

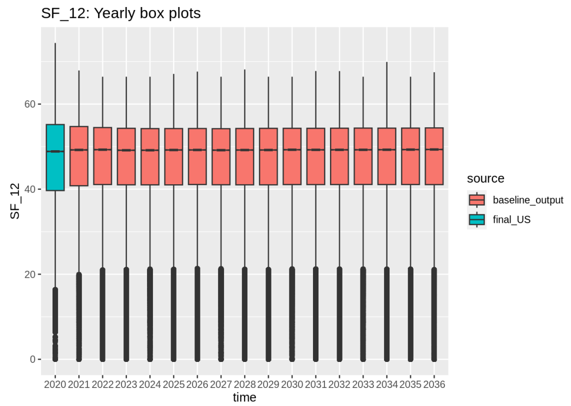


(d) Five-fold cross validated nowcasting estimating SF-12 MCS mean state from 2014 to 2020.

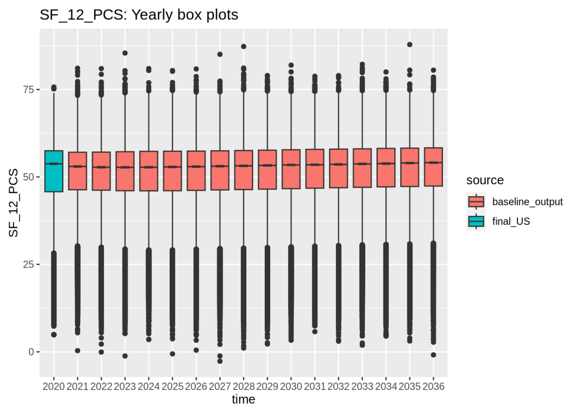
Figure 5.7: Five-fold cross-validation nowcasting plots for four key MINOS variables.



(a) Handover boxplots for yearly household energy bills over time under the baseline intervention. There is a sharp increase in energy bills from 2021 onwards that peaks at 2023 and gradually declines until 2035. There is also a substantial increase in energy variance dependent on initial energy consumption.



(b) Handovers boxplots for SF-12-MCS score under the baseline intervention. The distribution of SF-12-MCS score displays sensible trends over time preserving variance but showing a decrease in median score due to increasing energy pricing.



(c) Handovers boxplots for SF-12-PCS score under the baseline intervention. The distribution of SF-12-PCS score displays sensible trends over time preserving variance but showing a decrease in median score due to increasing energy pricing.

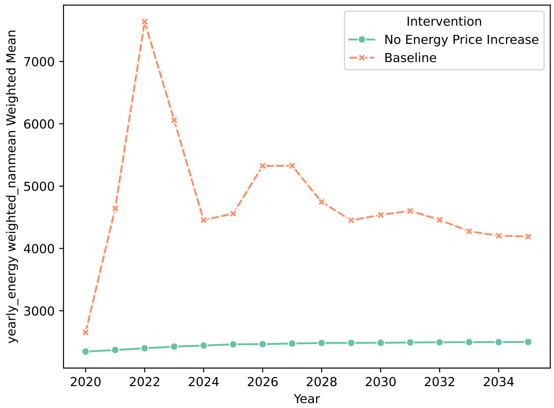
Figure 5.8: Handover boxplots for internal validation of yearly energy, SF-12 MCS, and SF-12 PCS distributions.

5.6 Results

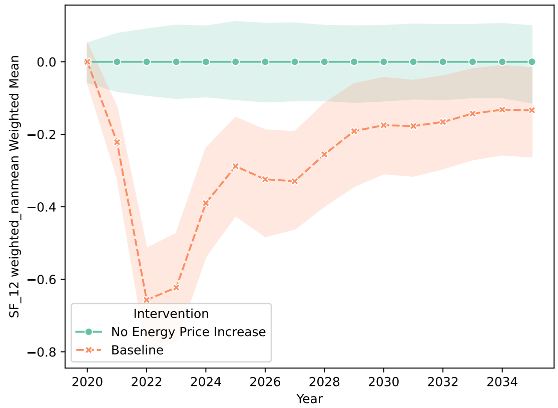
For each of the 100 model runs the mean SF_12_MCS, SF_12_PCS, yearly energy consumption and energy cost is taken over time. This provides 100 mean trajectories for each variable that can be plotted into confidence intervals in Figure 5.9a - 5.9f. Note these are relative change plots indicating the percentage change in each variable that are more readable particularly for small effect sizes. A score of 2.50 indicates a 2.50% improvement in SF-12 MCS or PCS score vs the baseline of no government intervention. The results of comparing the baseline against a scenario where energy prices do not increase is presented in Figure 5.9a - 5.9b. It can be seen that the price of energy doubles by 2023 and gradually decreases over time, but does not return to non-crisis energy baseline. Surprisingly while the SF-12 MCS score does deteriorate with higher energy pricing well-being levels gradually reduce back to baseline over time. The heating dummy, EPCG, and GBIS interventions all show positive movement in the SF-12 MCS score. The GBIS and EPCG schemes both perform well initially but improvement in the EPCG declines as energy pricing falls back down below the market cap. The good heating dummy does not show statistically significant increase from the baseline. For SF-12 PCS only the GBIS intervention shows significant positive change in the PCS score suggesting combined changes in yearly energy bills and comfortable heating are both required.

Change in SF-12 scores can be converted into QALY and ICER ratios. Using a set formula (Lawrence and Fleishman, 2004) an individuals PCS and MCS scores are converted into a scalar 'Quality Years' score. This score is then adjusted to remove deceased individuals from the population. Summing individual scores results in a final population QALY score for a given year and policy intervention. QALY scores are then added cumulatively over time giving 100 QALY trajectories for each intervention. For ICER scores, QALY values are divided by intervention cost each year providing change in the QALY score per unit pound. Confidence intervals are constructed for each intervention giving a final estimate of change in QALY score. QALY and ICER plots for the EPCG and GBIS interventions are presented in Figures 5.9e - 5.9f. It is clear that both interventions have a positive effect on health for the GMCA population by protecting household income. While the EPCG is initially more cost effective it is overtaken by the GBIS intervention as early as 2024 suggesting the proactive measure, while requiring high capital, is more beneficial long term for public health. While the final cost effectiveness of the GBIS does not reach the Green Book guidelines (Government Finance Function, HM Treasury, United Kingdom Government, 2024) recommended £70,000 ICER ratio for a cost effective policy this is subject to change as energy and insulation pricing changes further over time.

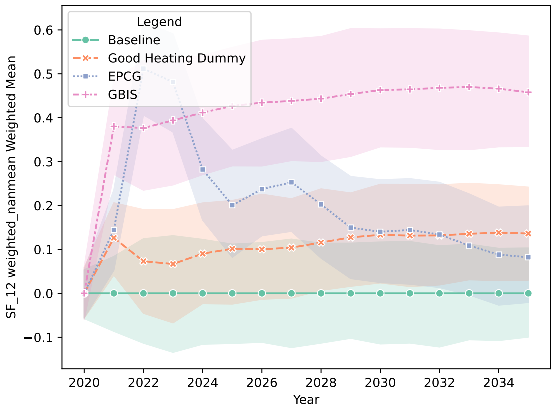
As the MINOS population is individual-based policy can be examined over sub-populations and individual units. Several examples are provided here demonstrating who benefits most from the EPCG and GBIS policies respectively. First, the effect on SF-12 scores by income quintile is provided in Figures 5.10a,5.10b. The GMCA population is stratified by income quintiles with the first quintile having the lowest 20% of household incomes. It is clear in the EPCG case that all income quintiles benefit from the intervention but there is no significant difference across quintiles. On the other hand the GBIS intervention also sees improvement for the entire population but the poorest income quintile sees a larger significant improvement in SF-12 MCS. The EPCG intervention applies multiplicative scaling such that the more income spent on energy, the larger the savings applied. On the other hand, the GBIS intervention applies more uniform additive scaling dependent on household size and lower council tax bands (lower income).



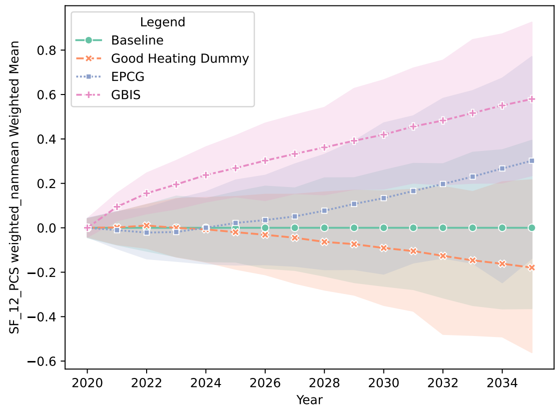
(a) Yearly energy expenditure over time with and without the UK energy crisis causing wholesale commodity pricing to increase.



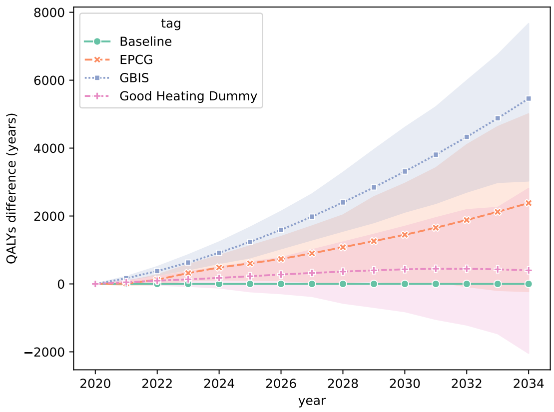
(b) SF-12 MCS over time with and without an increase in yearly energy prices according to the UK energy crisis.



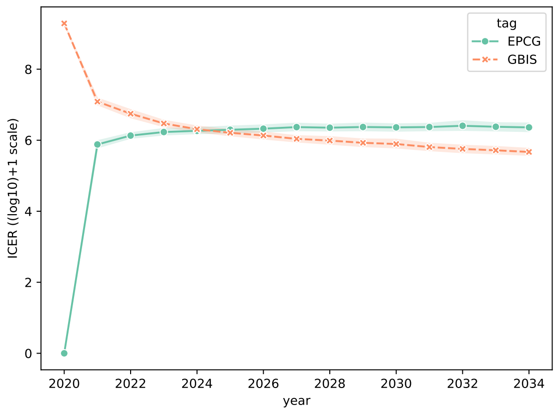
(c) SF-12 MCS score over time for each of the four energy policy interventions.



(d) SF-12 PCS score over time for each of the four energy policy interventions.

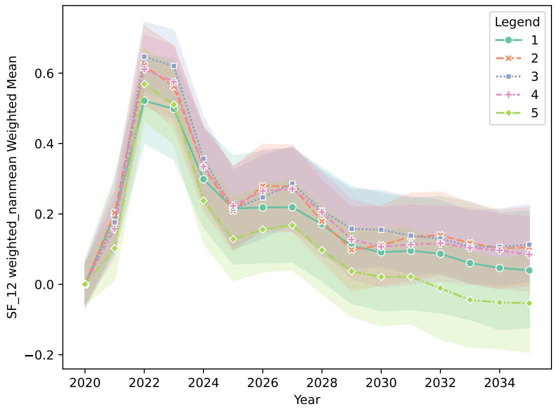


(e) Cumulative differences in QALY scores for each intervention against the baseline.

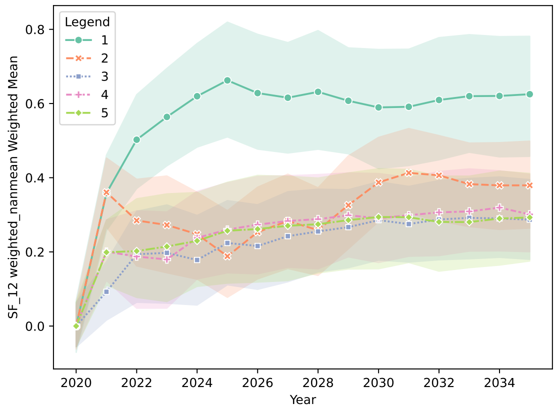


(f) ICER ratios for the GBIS and EPCG interventions comparing the cost of each intervention against the cumulative gain in QALYs against the baseline.

Figure 5.9: Lineplots for change in SF-12 MCS, SF-12 PCS, QALYs, and ICERs under energy poverty policy interventions.



(a) Gain in SF-12 MCS score under the EPCG intervention stratified by income quintile.



(b) Gain in SF-12 MCS score under the GBIS intervention stratified by income quintile.

Figure 5.10: Disaggregation of EPCG and GBIS SF-12 MCS benefit by income quintiles. Only the GBIS intervention sees significantly higher SF-12 MCS improvement for the poorest 20% of households.

5.7 Discussion

Household energy policy is set to dominate the UK political landscape for the foreseeable future, the challenge for policymakers is to strike a fine balance between carbon emissions targets and public opinion. The transition to renewable energy requires a reduction in household energy consumption via infrastructure overhaul that is able to sufficiently insulate households and provides electrical heating systems. The UK public has been reluctant to pay for updated heating systems due to skepticism on the cost-effectiveness of potential health and income benefits. The recent surges in energy pricing and net-zero initiatives are beginning to impact public opinion on these subsidies. However, it is proving difficult to deploy these schemes on a national level and identify vulnerable households that need urgent or additional assistance. The newly incumbent UK government is keen to expand existing household heating subsidies, however it still requires evidence of their cost-effectiveness to enable large scale deployment.

This work has demonstrated application of the MINOS dynamic microsimulation model to provide evidence of the effects of housing energy policy on health. This model utilises a synthetic input population for the GMCA population combined with a series of transition dynamics modules to estimate the effects of three policy interventions. A 'good heating dummy' morning after (Kopasker et al., 2024) intervention is applied exploring if subjective heating was provided for all households overnight as well as emulations of the real Energy Price Cap Guarantee and Great British Insulation Scheme (ECO4) policies implemented during the $2022 - 2023$ energy crisis and beyond. Results explore how these three interventions cause changes in Quality Adjusted Life Years (QALYs), Increment Cost Effectiveness Ratios (ICERs), as well as other key variables such as heating and energy costs.

Results show that all policies alleviate health effects of energy crisis. EPCG is expensive and performs well initially but is not as cost effective as the GBIS over time. While the GBIS is still not an effective policy according to guidelines this is highly subject to change given updated energy pricing, lower retrofitting costs and combination with other policy over time. Additional disaggregation is presented by income quintile and Local Authority spatial areas demonstrate how the MINOS model can be used in further research exploring policy effects by specific subgroups in order to identify and optimise targeted policy.

The primary limitation of the MINOS model is available data and computation power. UKHLS data has limited variables pertaining to household characteristics such as energy efficiency rating used to estimate energy poverty in more modern measures, insulation type and household size. This limits the amount of detail that can be implemented in terms of implementing more specific insulation policy such as those focusing solely in rural areas with solid wall insulation. Additionally this data will allow for more accurate estimate of change in household energy bills due to more specific insulation types. This can potentially be addressed in future work as the UKHLS dataset will be linked with UK Department for Energy Strategy and Net Zero's National Energy Efficiency Data-Framework (NEED) (Department for Business, Energy and Industrial Strategy (BEIS) London, 2013) containing further individual household insulation quality and energy usage variables. However, this linkage is likely to have significant missing data and production of further synthetic data may be required. Additionally, estimations of certain processes including household income must make large assumptions on individual household income and household behaviour. It is assumed that household incomes are reasonably consistent over time and not subject to factors such as employment. Additionally household energy spending habits remain consistent, which is unlikely as households use less and do not want to go into debt, and further inclusion of social mechanisms could provide more robust results. Finally it is assumed there are no further geopolitical events beyond the energy $2022 - 2023$ energy price increases. Exploration of sensitivity analyses introducing further socio-economic shocks, e.g. future price surges would facilitate use in policy analysis.

Future application of the MINOS model has three primary objectives; (1) implement further social and economic dynamics to better estimate change in income state under energy pricing scenarios, (2) apply further more realistic and heterogeneous income and housing support policy dependent on further data, and (3) more complex combinations of policy including several policies together, further time horizons, implementation of long-term renewable energy profiles, socio-economic shocks, and observation of other outcomes beyond health are all viable conditional on updates to government strategy. The MINOS housing prototype model has yielded sensible, tangible results for exploring the health consequences of future energy policy in the UK, and could provide part of a toolkit to balance transition to renewable energy and public opinion protecting public health and quality of life against potential future crises.

5.8 Connection between Chapter 5 and Chapter 6.

The MINOS dynamic microsimulation model has been expanded to estimate the mental and physical health effect of support policy during the ongoing UK energy crisis. Incorporation of new causal loop diagrams including energy poverty pathways, new health outcomes for physical health and Quality Adjusted Life Years, and use of a full synthetic population for the Greater Manchester area have all been incorporated to better quantify the effect of energy bill policy. Two further policies have been implemented exploring both a 'good heating dummy' morning after solution assuming if all households in the UK had comfortable subjective heating to examine how idealistic good heating would affect health as well as a new Great British Insulation Scheme (GBIS) policy emulating real existing government strategy to retrofit household insulation. Unsurprisingly, the good heating shows that comfortable heating results in an overall increase in physical health score but if energy bills remain high, mental health will still be negatively impacted. Similarly the energy price cap guarantee appears to protect mental health by reducing energy bills but does little for physical health. The GBIS scheme that improves both thermal comfort and household energy bills provides the largest gain in QALYs overall. Moreover, the GBIS scheme appears to be less cost effective initially due to high initial capital required, but overtakes the EPCG as early as 2024 which is faster suggested by other available literature but includes a much higher and faster uptake rate. The GBIS also appears to show larger improvement in the lowest income decile compared to the EPCG which equally improves all quintiles. The EPCG scheme saves money according to how much initial energy expenditure a household has, whereas the GBIS saves a flat amount of income leading to lower income households saving a much larger percentage of their income. Overall, this policy aligns with the proactive improvement of housing seen in countries such as Germany and Finland, is cost effective in the long term.

This chapter has begun to highlight how policy can vary across sub-populations of the UK but has not yet utilised the spatial component of the MINOS input population. The following final chapter provides a brief spatial analysis of the EPCG and GBIS policies producing plain narratives describing how behaviour changes at smaller spatial resolutions to try and identify any further groups left behind by these interventions.

Chapter 6: Analysis of MINOS Dynamic Microsimulation Output Data Exploring Energy Poverty Efficacy Heterogeneity Over Space and Vulnerable Sub-Populations

6.1 Introduction

A 'Health In All Policies' approach is increasingly popular within UK Government bodies such that all policies implemented across all sectors should be considerate of potential collateral health impact (Meier et al., 2019; Greszczuk, 2019). Previous literature highlights the broad range of determinants that affect public health including socioeconomic, environmental, and behavioural factors (Greszczuk, 2019) that are utilised to produce personalised, dynamic policy according to each individual's needs. One notable exemption in policy making is the inclusion of spatial heterogeneity (Augustin et al., 2023). Many public health crises have been shown to vary across UK geographies including the coronavirus pandemic causing increased unemployment and mental distress in city centres and coastal towns (Collard et al., 2021; Hancock and Tyler, 2024) as well as high energy prices affecting cost of living in rural areas, among the elderly, and particularly in Scotland and Northern Ireland (UK Fuel Poverty Monitor, 2022). Many UK policies are then criticised as 'leaving people behind' and are not cost-effective providing unnecessary funding to more affluent regions of the country (Bridgen and Robinson, 2023). Implementing just and equitable policy must then factor in how public health varies over space into policy making to reduce government spending and improve public health (Chatterton et al., 2016; Bouzarovski and Simcock, 2017; Snell et al., 2018; Bridgen and Robinson, 2023).

While it is clear spatially heterogeneous policy would benefit the UK (Bouzarovski and Simcock, 2017; Augustin et al., 2023), government bodies are reluctant to implement such 'postcode lottery' style policies (Snell et al., 2018; Weinberg et al., 2024). Implementation of spatial policy in conjunction with nationwide policy targets has proven awkward due to limited communication between Government Bodies particularly at Local Authority (LA) level (Best et al., 2023). Spatial policy has a public perception as being unfair (Snell et al., 2018; Weinberg et al., 2024) and confusing near geographical boundaries (Cushion et al., 2020; Bridgen and Robinson, 2023). There is usually limited evidence available on how policy effect varied across the UK at LA level to justify specific implementation without negative political ramifications (Bridgen and Robinson, 2023). Case study experiments are needed to codify methodology for optimising spatial policy and encouraging further uptake within UK Government.

High cost of living and net zero carbon policy have facilitated generation of spatial energy policy evidence at the LA-level (Bridgen and Robinson, 2023). This evidence highlights existing discrepancies in the Energy Company Obligation (ECO4/GBIS) scheme for retrofitting household insulation demonstrating that existing allocation of funding proportional to population size and limited consideration of demographic differences between regions is highly inefficient (Bridgen and Robinson, 2023; UK Fuel Poverty Monitor, 2022). This provides an ideal case study leveraging differences in energy poverty characteristics across Local Authorities exploring how GBIS intervention effect health changes over space. Identifying vulnerable subgroups of the population allows for better allocation of public funding and increased assurance that a high minimum standard of public health can be met.

This chapter contributes a short analysis building on output from the MINOS (Clay et al., 2023) spatial dynamic microsimulation model discussed in Chapter 5 exploring the Energy Price Cap Guarantee (EPCG) and Great British Insulation Scheme (GBIS) policies (UK Fuel Poverty Monitor, 2022). Spatial information from the MINOS synthetic population (Hoehn et al., 2024) is used to map how public health changes for the Greater Manchester (GMCA) area over time at LSOA, Ward, and Local Authority levels identifying features in areas that benefit more from each policy intervention. Brief discussion explores policy implications for difference in groups affected by these policies and identifies which subgroups of the population are still being overlooked as well as variance in health improvement across Local Authorities that may be useful to personalise existing strategy. Future work identifies potential further policy that could address these gaps in just energy poverty policy and begin to implement decision support tools optimising policy cost-effectiveness and leveraging Local Authority energy poverty evidence for improved resource allocation.

6.2 Methods

This section provides a review of methods and definitions used in spatial analysis of MINOS spatial population output. While there are overall limited additional methods definition of specific energy poverty variables and common visualisation tools are provided.

There are several additional data requirements attaching spatial information to individual observations in MINOS data. These variables are not included in the original Understanding Society dataset (Fumagalli et al., 2017) and must be attached using the simulated annealing algorithm that is well described elsewhere(Hoehn et al., 2024; Lomax et al., 2024). ONS lookup tables and geometry tiles (Office for National Statistics, 2024) are provided allowing for mapping distribution of health change in chloropleth maps and translation to coarser spatial resolutions including Ward and Local Authority levels. This chapter does not provide information on how to generate synthetic spatial data but uses this data to produce spatially representative microsimulation output for the GMCA area that can be mapped and used to observed spatial trends.

These lookup tables also contain several variables characterising each LSOA based on how rural they are, and their level of material deprivation. The Rural-Urban Classification (Bibby and Brindley, 2013) identifies how rural each LSOA is using four categories for Rural, Rural Towns, Urban Towns, and Major Conurbations. As the GBIS policy is expected to benefit rural areas more, examining health improvement over these classifiers will attempt to verify this claim. Output area classification variables (Payne and Abel, 2012) are also used to understand socio-economic differences for UK LSOAs. There are 24 total 'mid-level' groups used in this chapter to further differentiate differences in LSOAs that may contribute to heterogeneity in policy effect. Finally, a set of inclusive economy indicators (Lupton et al., 2023) have at LSOA level have been used to further differentiate the policy effects. A total of 7 out of 13 available indicators have been chosen providing characteristics on poverty, employment opportunity, and general inclusion for each LSOA identified from the synthetic UKHLS data used in MINOS with methods that are well described elsewhere (Lupton et al., 2023). A legend of these indicators is provided for each GMCA Local Authority in Tables 6.7,6.8.

The Index of multiple deprivation (IMD) is a measure of deprivation for a given LSOA area provided by the UK government based on a number of factors including economic opportunity and accessibility of other amenities (Payne and Abel, 2012). In this case IMD is divided into global deciles such that decile 1 contains households within the 10% most deprived LSOAs nationally. Use of local deciles whereby decile 1 is the most deprived 10% in Manchester may yield different results.

Fuel Poverty 10% (FP10) (Davillas et al., 2022) is one method to identify energy poor households who spend more than 10% of their net household income on energy bills. This measure is used to determine energy poor households within the GMCA population. While this measure has received criticism due to considering high income, high energy expenditure households to be fuel poor more modern definitions such as Low-Income Low Energy Efficiency (LILEE) (Davillas et al., 2022) measures cannot be used in this research due to there being no available energy efficiency rating data in Understanding Society (Fumagalli et al., 2017).

Several methods are used to aggregate MINOS individual-level data over LSOA and Ward spatial resolutions. Estimation of non-cumulative variables including SF-12 MCS and SF-12 PCS score are aggregated by subsetting households contained within each LSOA and taking mean values. For each of 100 MINOS model runs this provides 100 mean values for each LSOA area that are then further aggregated into 1673 grand mean values for each LSOA. For cumulative variables such as QALYs cumulative summation is taken before taking means giving the mean cumulative QALY gain for each LSOA. These values are used to identify how a policy changes the health of a given LSOA. Mapping at an LSOA level is difficult to visualise due to very small spatial tiles and vulnerability to low sample size. Instead means are often further aggregated up to the ward level giving 215 mean values for change in health allowing for clearer interpretation of consistent spatial trends.

Incremental Cost-Effectiveness Ratios (ICERs) are also calculated for each LSOA area. This is performed using a simple formula $ICER = \frac{\Delta C}{\Delta Q}$ such that the ICER ratio is the cost of an intervention divided by change in QALY score (Lawrence and Fleishman, 2004). These ICER ratios are plotted in simple scatter plots with the intervention cost on the x-axis and QALY gain on the y-axis (Pirc et al., 2018). This allows for visualisation of ICER ratios in a geometric plane that is beneficial vs traditional ICER values that struggle with divide by zero errors for near 0 QALY values (Pirc et al., 2018). This also allows for easy identification of cost-effective policies in the top left region of the graph with low cost and high gain. Lines passing through the origin on this graph can then be used to identify policies of fixed ICER ratios. According to UK Government Green Book guidelines a policy is cost-effective if it costs less than £$70,000$ per QALY (Government Finance Function, HM Treasury, United Kingdom Government, 2024). This provides a minimum cost-effectiveness threshold line with equation $y = x/70000$. This line is drawn on all ICER plane graphs providing easy comparison such that if an LSOA point is above this line it is cost-effective (Pirc et al., 2018). Analysis of LSOAs that are above this cost-effectiveness threshold as well as below the threshold with high cost are explored further in the following section. All scatter plot values are also coloured according to their IMD decile to determine quickly if a policy affects more deprived deciles disproportionately.

Determining if a policy is more effective in certain GMCA LAs is performed using simple linear regression. QALY gain is compared against a number of variables including IMD, Local Authority area, intervention cost, to identify areas that benefit more or less from each policy. Other variables are also assessed using simple statistics including mean and percentages to explore differences between interventions.

Brief validation for fossil fuel consumption for the GMCA region is also provided. Mapping of gas, electricity, fuel oils, and other fuel consumption is provided with brief comparison to available external LSOA-level simulated consumption data. This mapping is provided at both the starting time 2020 as well as at 2023 when energy prices were at maximum.

6.3 Results

6.3.1 Energy Validation

Validation for small scale energy consumption for the UK population is provided using LSOA level maps and histograms. Median yearly energy consumption for the five fuel types are calculated for each LSOA over the GMCA area and plotted below in Figures 6.1-6.3. Reference maps can be found describing the GMCA Local Authorities ((Hincks, 2017), Figure 1.) including IMD and energy poverty (Catapult Energy Systems, 2024) distribution. For electricity, gas, and combined gas and electricity consumption the majority of LSOAs have approximately £1200 combined median gas and electricity costs. Consumption is overall lower in urban regions of Manchester including the city centre and outlying towns which aligns with external research (Chatterton et al., 2016; Catapult Energy Systems, 2024) suggesting lower costs due to higher population density. There are some outlier consumption values particularly in the Stockport and Oldham town centres, potentially due to different demographic structure or missing data.

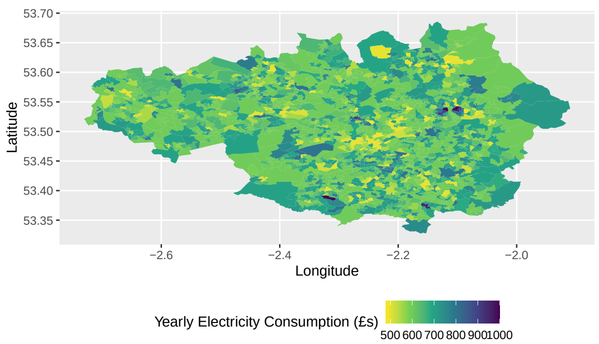
Gas and electricity consumption can be compared directly with Department of Energy Security and Net Zero (DESNZ) median kWh consumption data by LSOA (Department of Energy Security and Net Zero, United Kingdom Government, 2024b). There is a strong linear relationship between MINOS population yearly gas and electricity bills with these yearly median kWh consumption data. Simple linear regressions is fitted to gas and electricity giving formulae $gas\_bill = 0.0464 \times gas\_kwh$ and $electric\_bill = 0.210 \times electric\_kwh$ each with adjusted R-squared scores greater than 0.95. Given prices of $17p$ and $3p$ per kWh for electricity and gas in 2020 (Department of Energy Security and Net Zero, United Kingdom Government, 2024b) plus standing charges suggests energy expenditure in the MINOS population matches real DESNZ data at the LSOA level. Additionally, comparison can be made with other simulated energy consumption at LSOA level for the GMCA region (Chatterton et al., 2016; Catapult Energy Systems, 2024) and shows similar patterns of low consumption in urban areas.

Fuel oil consumption disparity between rural and urban areas is much more pronounced with very high rural consumption. This also aligns with external literature given increased farmland and decreased connectivity to mains supply (Chatterton et al., 2016). Consumption of other fuels is overall low and shows no clear pattern between rural and urban areas. No validation data is available for direct comparison to the authors knowledge for fuel oil and other fuels but this may change in the future if UKHLS data is aligned with other DESNZ datasets (Department of Energy Security and Net Zero, United Kingdom Government, 2024a).

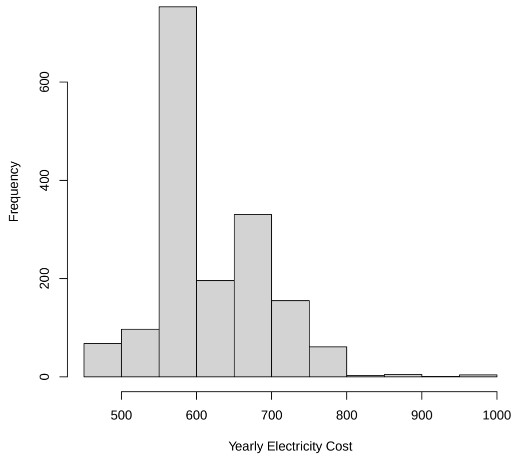
6.3.2 QALY Ward-Level Mapping

Change in SF-12 MCS, SF-12 PCS, and QALYs scores are displayed at the ward spatial level in Figure 6.4 in the year 2025. For the EPCG intervention it is clear that the intervention shows near universal improvement in mental well-being but limited improvement in physical well-being in rural areas. This equates overall to mixed QALY gains with strongest improvement of approximately 20% in areas that show combined MCS and PCS improvement.

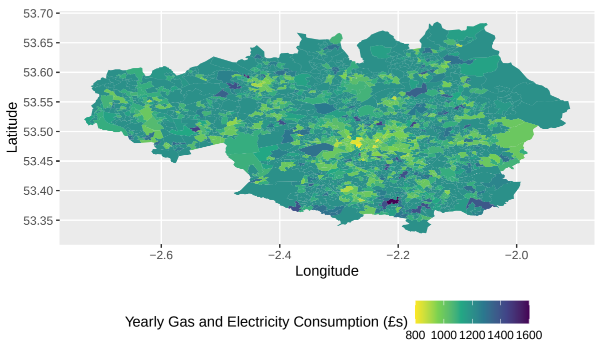
On the other hand the EPCG intervention shows virtually no improvement in SF-12 MCS but large



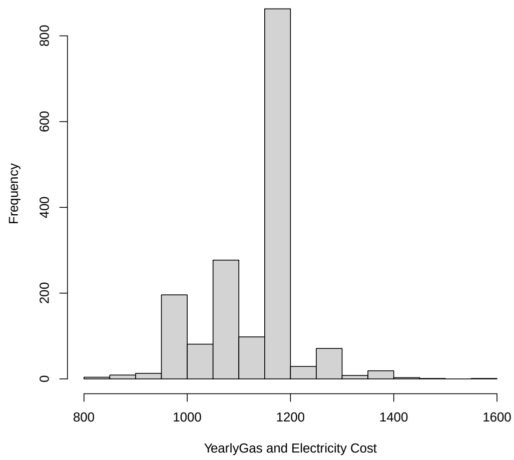
(a) Electricity expenditure chloropleth map of the GMCA by LSOA.



(b) Electricity Bills.

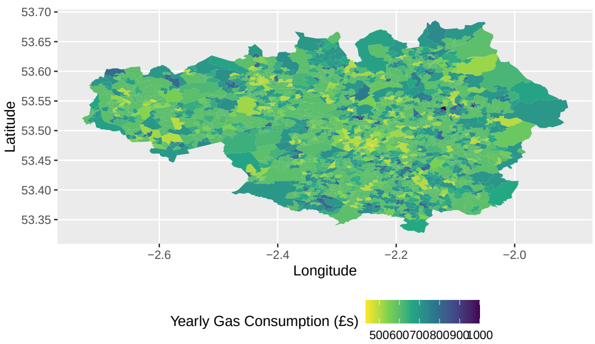


(c) Gas expenditure chloropleth map of the GMCA by LSOA.

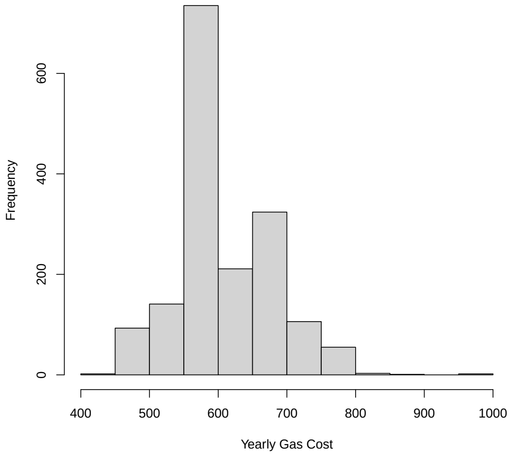


(d) Gas Bills.

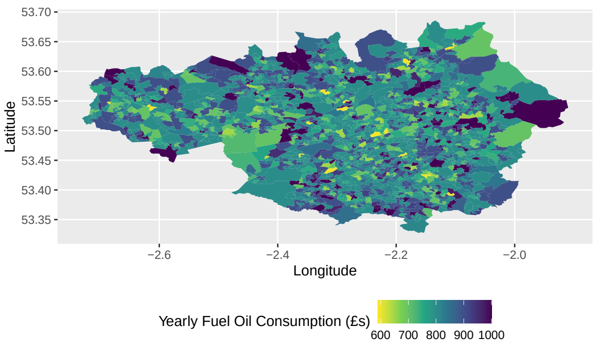
Figure 6.1: Descriptive statistics for gas and electricity expenditure for the GMCA synthetic input population.



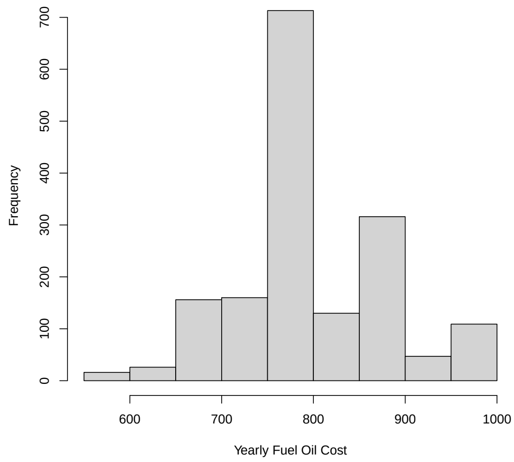
(a) Gas and Electricity Combined expenditure chloropleth map of the GMCA by LSOA.



(b) Gas and Electricity Combined.

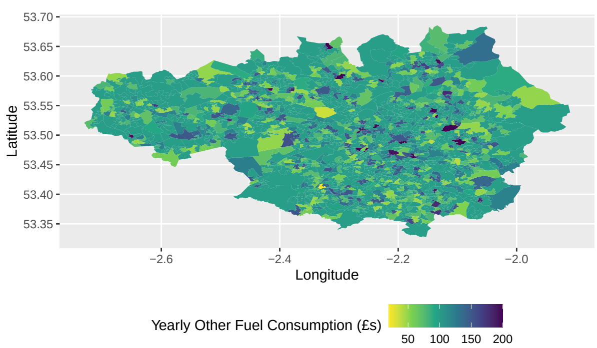


(c) Fuel Oil expenditure chloropleth map of the GMCA by LSOA.

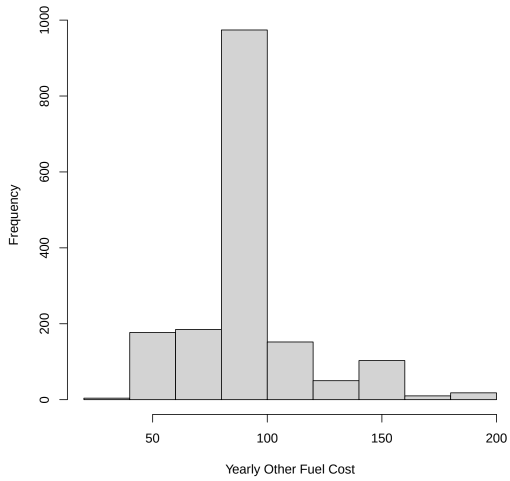


(d) Fuel Oil Bills.

Figure 6.2: Descriptive statistics for combined gas electricity and fuel oils expenditure for the GMCA synthetic input population.



(a) Other Fuels expenditure chloropleth map of the GMCA by LSOA.



(b) Other Fuels.

Figure 6.3: Descriptive statistics for other fuels expenditure for the GMCA synthetic input population.

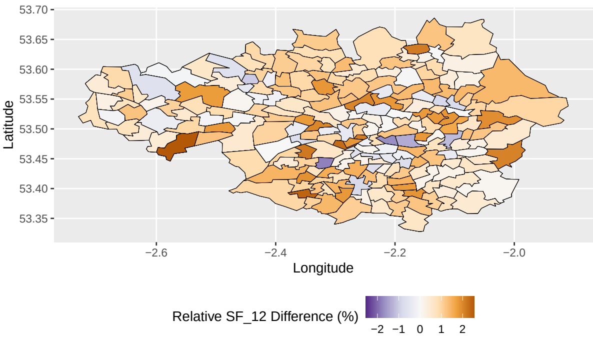
improvement in SF-12 PCS. This equates to near universal positive QALY improvement for the whole GMCA area. This demonstrates both the importance of PCS in QALY calculation (Lawrence and Fleishman, 2004) and good thermal comfort as well as the need for further reduction in household energy bills in order to support mental well-being improvement as well. Some combination of the two policies may yield higher QALY gain than either policy individually.

While these plots cannot be used for formal testing they provide an interpretable glance at what a policy is doing and allow for easier identification of overlooked areas in order to further tailor policy.

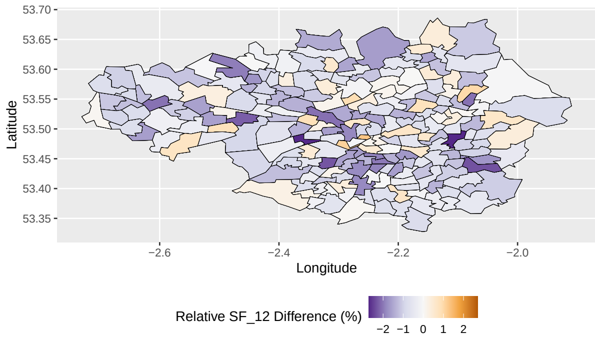
6.3.3 QALY Change Linear Regression

Initial examination of how these two policies vary over space is also performed using linear regression. For the GBIS and EPCG interventions we first calculate the cumulative gain in QALYs for each household against the baseline intervention where no energy policy is applied. These QALY values are then aggregated by LSOA providing 1673×100 mean QALY gains for each LSOA and each model run. These are then aggregated further into grand means giving a set of 1673 scalar QALY gain values for each LSOA. We now have two sets of QALY gains by LSOA for the EPCG and GBIS interventions. Final relative differences in QALY gain for the GBIS intervention are calculated for each LSOA using a simple formula $QALY\_DIFF = \frac{(GBIS\_QALY - EPCG\_QALY)*100}{EPCG\_QALY}$ giving a final set of 1673 scalar QALY relative difference values. For example, if an LSOA has a score of 20 it implies the GBIS intervention sees a 20% improvement in QALY score versus the EPCG intervention. These relative difference scores are then used as the response variable in linear regression to determine if spatial features in the GMCA area are predictive of higher QALY gain under the GBIS or EPCG interventions. These QALY differences are compared against several features including rural urban classifiers, output area classifiers, inclusive economy indicators, Local Authority, and by financial cost exploring where the QALY intervention sees better performance. Coefficients for this model are provided in Table 6.1.

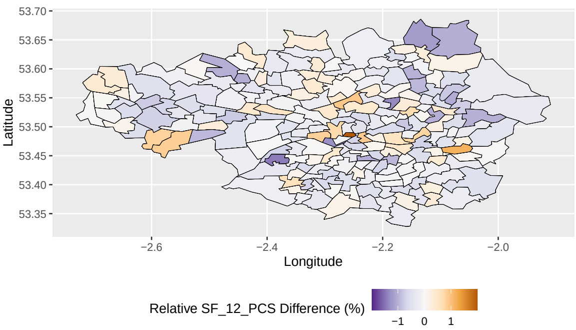
The GBIS intervention sees a positive gain in QALY score over the EPCG intervention in both rural area classifications. Both urban areas see negative overall gain in QALYs suggesting the EPCG intervention



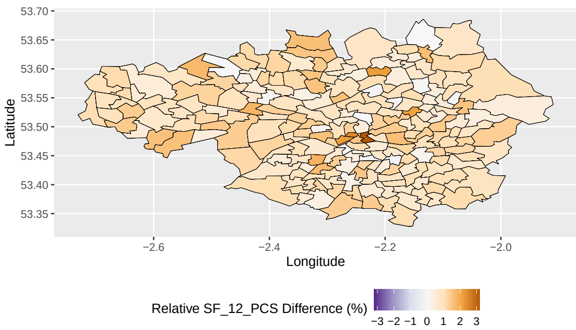
(a) SF-12 MCS relative percentage change at ward level for the EPCG intervention intervention.



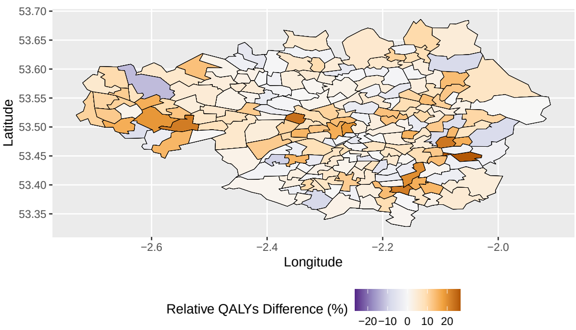
(b) SF-12 MCS relative percentage change at ward level for the GBIS intervention.



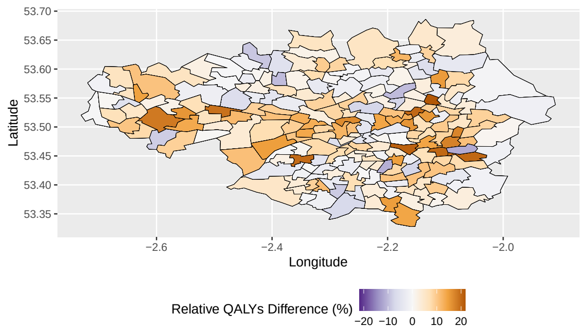
(c) SF-12 PCS relative percentage change at ward level for the EPCG intervention.



(d) SF-12 PCS relative percentage change at ward level for the GBIS intervention.



(e) % QALY gain for the EPCG intervention.



(f) % QALY gain for the GBIS intervention.

Figure 6.4: Health Change Maps for the EPCG and GBIS Interventions at ward-level for the GMCA.

targets rural areas as desired. Intervention cost demonstrates a negative coefficient suggesting that the EPCG sees a larger gain in QALYs in LSOAs where more money is allocated. This aligns with previous discussion where support given by the EPCG is proportional to initial energy consumption and improves mental well-being. Areas with low energy consumption see low application of the EPCG, and have several other issues discussed later, such that the GBIS intervention on physical health becomes more beneficial.

For output area classifiers compared against the baseline of 'aspiring neighbourhoods' only ageing rural neighbourhoods and primary sector workers (I.E. farmers) see further QALY gain both of which are strongly rural areas. Many other OAC classifiers have negative coefficients but still demonstrate positive QALY gain under the GBIS intervention including Asian Traits, Ageing Suburbanites, Challenged White Communities, Cosmopolitan Student Neighbourhoods, Young Ethnic Communities that have either low-income, high age, or high prevalence of ethnic minorities that again are prioritised by the GBIS intervention. Only Propspering Countryside Life, Highly Qualified Professionals, and Hard-Pressed Flat dwellers show no evidence the GBIS increases QALY gain. These areas are either high earners and have higher energy consumption that are favoured by the EPCG intervention, or show high deprivation and multiple issues again discussed further in Section 6.3.5.

Inclusive economy indicators also demonstrate strong trends in behaviour. Skills and qualifications, Digital Connectivity, Physical Connectivity, Poverty, and Cost of Living indicators all suggests improvement in QALY. Skills and Qualifications and Poverty measures are both strongly related to income suggesting lower income households benefit more from the GBIS intervention. Similarly digital and physical connectivity are both related to rural areas and potentially increased car usage suggesting further evidence the GBIS intervention affects rural areas. Only decent pay has a negative coefficient value but this is a double negative suggesting areas with lower employment above the living wage, I.E. poorer areas, benefit more from the GBIS intervention.

All Local Authorities see improvement in QALY gain versus the reference Bolton level with Manchester, Trafford, Tameside, and Oldham seeing the largest improvement. Comparison with Inclusive Economies indicators by Local Authority given in Tables 6.7-6.8 suggests improvement in the Manchester, Trafford and Oldham LAs is due to existing high energy poverty and child poverty as well as low wages, skills and qualifications, and physical connectivity. These areas have lower incomes and potentially higher income to energy expenditure ratios due to energy poverty and poor public transport accessibility which aligns with other available evidence on the benefits of insulation schemes from previous chapters and literature (UK Fuel Poverty Monitor, 2022). Rochdale and Tameside LA behaviour is not well explained by these indicators. The Tameside LA has strong performance under the GBIS intervention but generally has high incomes and low energy poverty. The Tameside LA is generally older, is over 90% White British, has a strong urban rural split, and has most of the both rural and deprived areas of the GMCA area (Hincks, 2017) which may all contribute to stronger performance under the GBIS intervention. Conversely the Rochdale LA does not perform well under the GBIS intervention despite poor performance on these inclusive economy indicators and further inclusion of additional demographics and higher order terms are required to explain these discrepancies.

Overall the GBIS intervention appears to benefit the desired rural areas, low income households and ethnic minority households. There are some discrepancies in this data including why higher expenditure benefits the EPCG intervention but this analysis suggests the GBIS intervention works as intended.

	Estimate	Std. Error	t value	Pr>|t|)
(Intercept urban rural code: Rural Towns, LA Name: Bolton, SOAC code: Achieving Neighbourhoods)	0.6330	0.1795	3.53	0.0004
urban rural code: Rural village and dispersed	-0.3339	0.2777	-1.20	0.2293
urban rural code: Urban city and town	-1.0037	0.1753	-5.73	0.0000
urban rural code: Urban major conurbation	-0.7657	0.1589	-4.82	0.0000
Local Authority Name: Bury	0.2572	0.0566	4.54	0.0000
Local Authority Name: Manchester	0.4629	0.0802	5.77	0.0000
Local Authority Name: Oldham	0.2931	0.0536	5.47	0.0000
Local Authority Name: Rochdale	0.1135	0.0577	1.97	0.0494
Local Authority Name: Salford	0.3795	0.0598	6.35	0.0000
Local Authority Name: Stockport	0.2015	0.0592	3.40	0.0007
Local Authority Name: Tameside	0.3786	0.0579	6.54	0.0000
Local Authority Name: Trafford	0.3150	0.0690	4.56	0.0000
Local Authority Name: Wigan	0.1616	0.0578	2.79	0.0052
Intervention Cost	-0.1102	0.0195	-5.64	0.0000
Super Output Area Classifiers (SOAC): Affluent communities	-0.1075	0.0868	-1.24	0.2156
Super Output Area Classifiers (SOAC): Ageing rural neighbourhoods	1.1362	0.2963	3.83	0.0001
Super Output Area Classifiers (SOAC): Ageing suburbanites	-0.2000	0.0829	-2.41	0.0158
Super Output Area Classifiers (SOAC): Ageing urban communities	-0.0156	0.1032	-0.15	0.8796
Super Output Area Classifiers (SOAC): Asian traits	-0.0096	0.0942	-0.10	0.9186
Super Output Area Classifiers (SOAC): Aspiring urban households	-0.1512	0.0835	-1.81	0.0701
Super Output Area Classifiers (SOAC): Challenged white communities	-0.2208	0.0855	-2.58	0.0099
Super Output Area Classifiers (SOAC): Comfortable neighbourhoods	-0.1582	0.0886	-1.79	0.0743
Super Output Area Classifiers (SOAC): Comfortable suburbia	-0.2183	0.0816	-2.68	0.0074
Super Output Area Classifiers (SOAC): Constrained renters	-0.2965	0.1043	-2.84	0.0045
Super Output Area Classifiers (SOAC): Cosmopolitan student neighbourhoods	-0.0801	0.1078	-0.74	0.4576
Super Output Area Classifiers (SOAC): Endeavouring social renters	-0.0412	0.0944	-0.44	0.6628
Super Output Area Classifiers (SOAC): Hampered neighbourhoods	-0.1191	0.0795	-1.50	0.1343
Super Output Area Classifiers (SOAC): Hard-pressed flat dwellers	-0.4261	0.1474	-2.89	0.0038
Super Output Area Classifiers (SOAC): Highly qualified professionals	-0.3147	0.1023	-3.08	0.0021
Super Output Area Classifiers (SOAC): Households in terraces and flats	-0.1999	0.0837	-2.39	0.0169
Super Output Area Classifiers (SOAC): Inner city cosmopolitan	0.2192	0.1489	1.47	0.1411
Super Output Area Classifiers (SOAC): Primary sector workers	1.3713	0.4117	3.33	0.0009
Super Output Area Classifiers (SOAC): Prospering countryside life	-0.4586	0.2566	-1.79	0.0739
Super Output Area Classifiers (SOAC): Rural traits	-0.3104	0.2319	-1.34	0.1807
Super Output Area Classifiers (SOAC): Urban cultural mix	-0.0352	0.0863	-0.41	0.6835
Super Output Area Classifiers (SOAC): Young ethnic communities	-0.1231	0.0965	-1.28	0.2021
Participation in Paid Employment	0.0075	0.0455	0.17	0.8682
Skills and Qualifications	0.2137	0.0458	4.67	0.0000
Involuntary Exclusion from the Labour Market	0.0626	0.0402	1.56	0.1197
Digital Connectivity	0.0772	0.0315	2.45	0.0144
Wealth Inequality	0.0183	0.0115	1.59	0.1112
Physical Connectivity	0.0315	0.0136	2.32	0.0204
Earnings Inequality	0.0096	0.0179	0.54	0.5894
Housing Affordability	0.0024	0.0247	0.10	0.9236
Poverty	0.1336	0.0301	4.44	0.0000
Cost of Living	0.1354	0.0468	2.89	0.0038
Decent Pay	-0.0751	0.0294	-2.56	0.0106
Decision Making Inclusion	0.0540	0.0217	2.48	0.0131
Job Security	-0.0028	0.0271	-0.10	0.9177

Table 6.1: Linear Regression Coefficients estimating differences in QALY gain for the GBIS intervention against the EPCG intervention.

6.3.4 ICER Planes and Description of Cost-Effective and Not Cost-Effective LSOAs

Intervention cost-effectiveness is demonstrated using ICER planes. The mean QALY gain and intervention cost is taken for each LSOA in the GMCA region. For Example, a marker at coordinates $(x, y) = (10, 1, 000, 000)$ would provide 10 QALY costing a million pounds with an ICER score of $1, 000, 000/10 = 100, 000$ and is not cost-effective according to the $70, 000$ minimum threshold. These 2-d points can be displayed using a scatter plot directly but there are too many points to make meaningful inference. Instead two dimensional kernel density estimate contours (Chen, 2017) are plotted to estimate the distribution of ICER scores for each LSOA. ICER planes have been plotted for the Rochdale, Tameside, and Manchester Local Authorities and both energy interventions in Figures 6.5-6.6. For each of these plots Green Book £$70, 000$ minimum threshold guideline is also plotted as a diagonal white straight line.

These planes vary by both Local Authority and intervention. For the Rochdale there is overall small gain in QALY scores for both interventions. The GBIS intervention is cheaper but it is clear both interventions have limited effect. For Tameside both interventions see positive QALY gain and similar costs. The GBIS intervention shows larger, more consistent QALY improvement but fewer overall LSOAs above the Green Book line. These plots support above evidence suggesting Tameside performs better under the GBIS intervention and Rochdale sees limited improvement under either intervention.

The Manchester Local Authority planes in Figure 6.6 is extremely expensive for both interventions but shows increased QALY gain under the GBIS intervention with almost no cost-effective LSOAs.

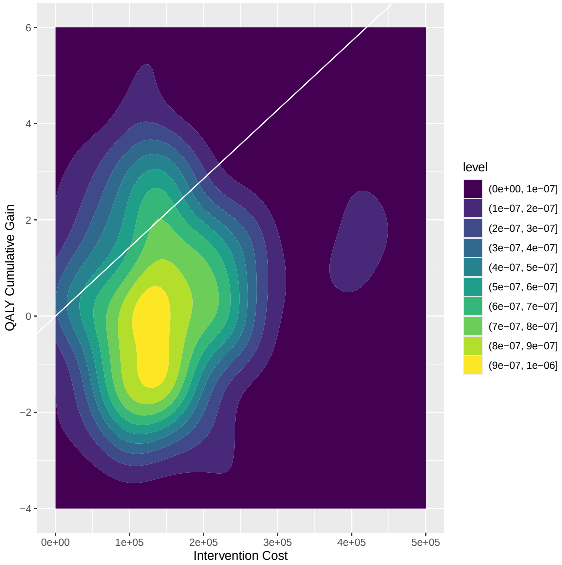
The GBIS intervention sees a positive QALY gain whereas the EPCG intervention does not but it is much more expensive costing over >£$750k$ per LSOA. The Manchester, Oldham, and Bury LSOAs all exhibit similar behaviour containing some LSOAs with very high overall expenditure. The remainder of this chapter explores why certain LSOAs are cost-effective or not using analysis of subgroups and individual trajectories attempting to construct narratives describing performance and potential tweaks to the applied policies. For cost-effective LSOAs, exploration of circumstances tries to demonstrate why each EPCG and GBIS intervention has a positive effect versus the baseline intervention.

Individuals belonging to both cost-effective LSOAs and expensive not cost-effective LSOAs under both interventions are taken from the 2020 synthetic MINOS input population. Key variables within MINOS data are described to determine any differences in LSOAs that may lead to them being cost-effective or expensive and not cost-effective under each intervention.

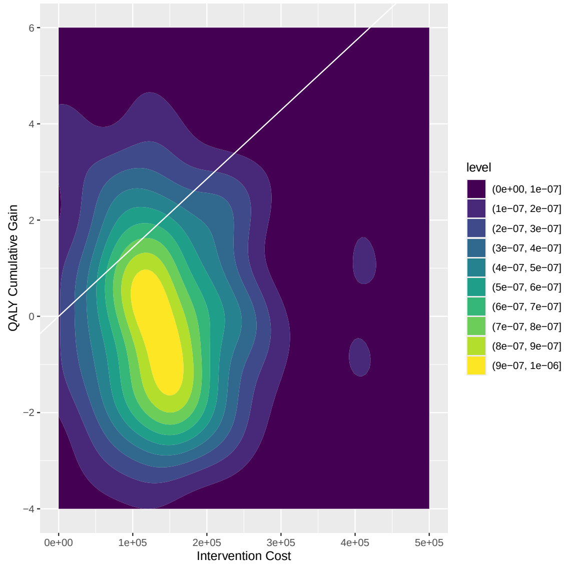
Statistics are provided for several MINOS population variables in Tables 6.2-6.3. Continuous variables mean values are provided for each intervention for cost-effective and expensive LSOAs. For discrete variables percentages are provided for the number of individuals present in each state again under both interventions as well as cost-effective and expensive, non cost-effective LSOA types.

For cost-effective LSOAs there are overall limited differences between continuous variables in both interventions against the total population. The GBIS sees slightly better performance for LSOAs with lower income, lower energy consumption and lower age. While individual yearly electricity and yearly gas expenditure are lower than in the total population combined gas and electricity expenditure is higher. This could be due to error in calculation of combined gas and electricity usage or the influence of different payment methods for separate and combined fuel consumption such as prepayment meters or direct debit and requires further UKHLS variables to investigate further.

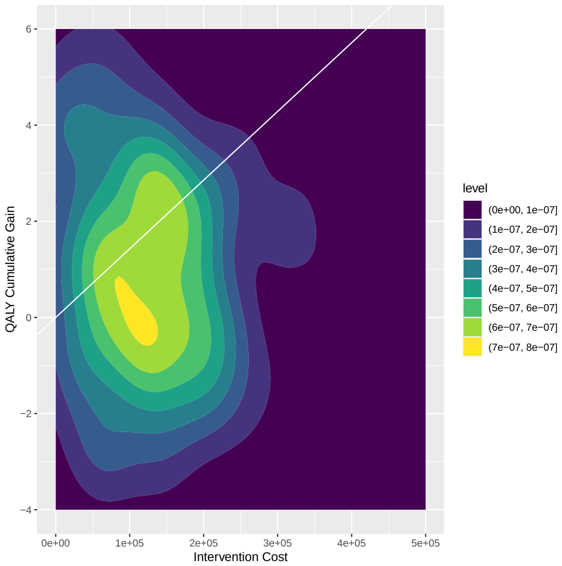
Expensive and not cost-effective LSOAs show substantial differences in continuous variable distributions.



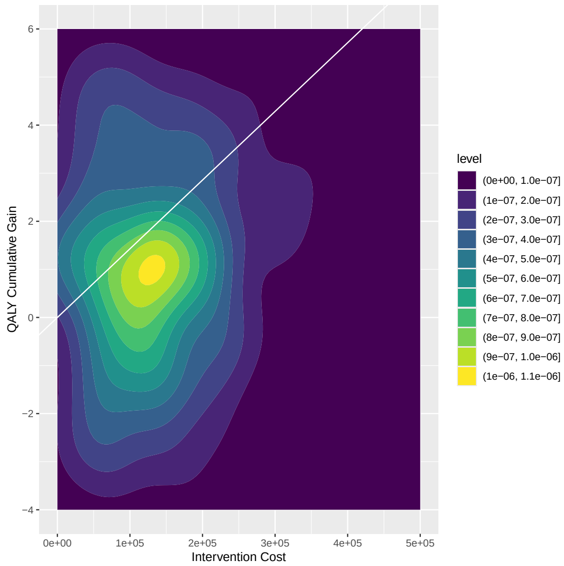
(a) Two dimensional kernel density estimates plotting change in QALY against intervention cost for each LSOA over the EPCG intervention and Rochdale Local Authority.



(b) Two dimensional kernel density estimates plotting change in QALY against intervention cost for each LSOA over the GBIS intervention and Rochdale Local Authority.

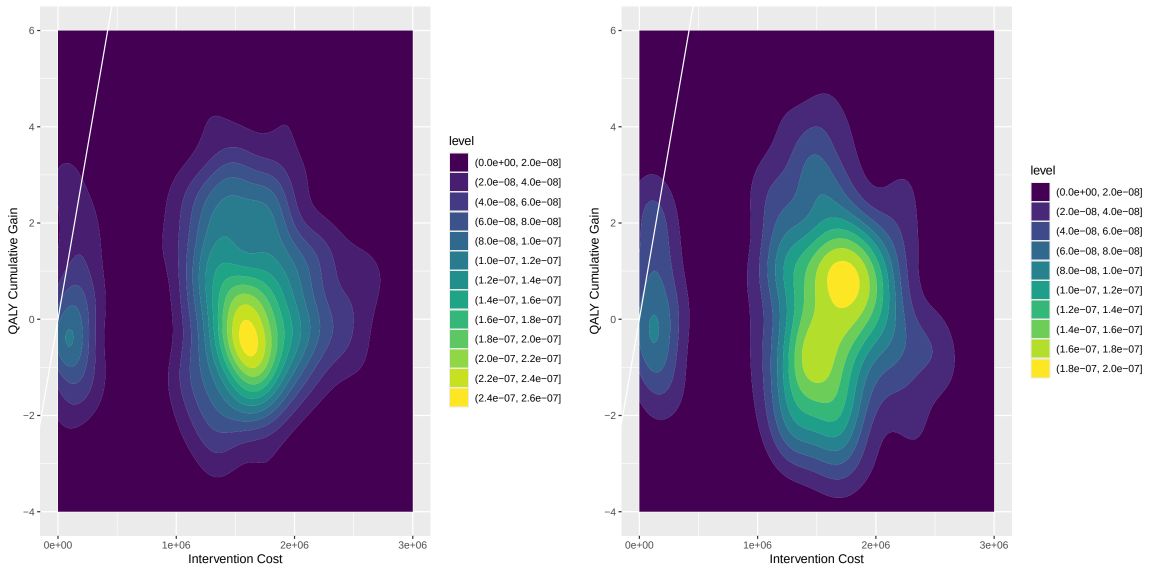


(c) Two dimensional kernel density estimates plotting change in QALY against intervention cost for each LSOA over the EPCG intervention and Tameside Local Authority.



(d) Two dimensional kernel density estimates plotting change in QALY against intervention cost for each LSOA over the GBIS intervention and Tameside Local Authority.

Figure 6.5: ICER planes plotting two dimensional kernel density estimate contours for mean change in QALY and intervention cost by LSOA for the GBIS and EPCG interventions over the Rochdale and Tameside Local Authorities. The White Line Indicates the Green Book £70k per QALY recommended threshold.



(a) Two dimensional kernel density estimates plotting change in QALY against intervention cost for each LSOA over the EPCG intervention and Manchester Local Authority.

(b) Two dimensional kernel density estimates plotting change in QALY against intervention cost for each LSOA over the GBIS intervention for the Manchester Local Authority.

Figure 6.6: ICER planes plotting two dimensional kernel density estimate contours for mean change in QALY and intervention cost by LSOA for the GBIS and EPCG interventions over the Manchester Local Authority. The White Line Indicates the Green Book £70k per QALY recommended threshold.

There are much lower incomes for both interventions. However, there is a much smaller reduction in yearly energy consumption suggesting these LSOAs have a higher income to energy expenditure ratio. Mortgages are lower in this group but rent is much higher suggesting a combination of high cost renters typically in urban areas and larger housing as well as low mortgage costs typically in rural areas and low quality housing stock. Separate electricity and gas consumption are higher than the total population but combined gas and electricity consumption is lower against suggesting these interventions are more beneficial to households who pay both bills together. Expensive LSOAs are also much younger than the population average. These data suggests expensive low cost-effectiveness LSOAs are a low income households that are renting, pay separate gas and electricity bills, and are generally younger. This group having high rent and low mortgages is somewhat contradictory and could be from a number of sources including renters with high costs and low quality housing stock with small mortgages. This data corresponds with previous regression analysis on QALY gain above suggesting high cost renters as well as affluent rural households and very poor households with low mortgages see little gain from either intervention. Further cross tabulation of variables may elucidate more specific subgroups of the population such as low mortgage, low income groups providing more detail.

Several discrete variables are also provided. For housing quality both interventions are more cost-effective for low material quality houses. Medium quality houses are more prevalent for cost-effective and expensive LSOAs and the GBIS intervention benefits these households more in both cases. The EPCG intervention unsurprisingly benefits high quality households more and they are featured less in both cost-effective and expensive groups versus the total population.

For the FP10 measure in cost-effective LSOAs the GBIS intervention benefits these energy poor households more. However, for expensive LSOAs the opposite is true and the EPCG intervention benefits fuel poor households. In both cases energy poverty is more prevalent in these groups than in the total population. This may be due to FP10 generally being a poor measure of energy poverty due to the inclusion of households with both high income and high energy expenditure (Davillas et al., 2022).

Cost-Effective LSOAs demonstrate the GBIS is more effective for every group besides Bangladeshi, Pakistani, White British, and White Other. Additionally every ethnic group besides Bangladeshi, Pakistani and, White British are more prevalent in cost-effective LSOAs than in the total population for both interventions. This initially suggests that these interventions are more beneficial for most ethnic minority groups. For expensive and not cost-effective LSOAs all ethnicities apart from White British and White Other have much higher prevalence than in the total population. The GBIS intervention effect is much more mixed and only benefits some ethnic groups. Again the four Bangladeshi, Pakistani, White British and White Other stand out as outliers. White British ethnicities benefit more from the EPCG intervention and have a very low prevalence in the expensive LSOA group. The White Other ethnicity seems to get the most benefit from the GBIS intervention out of all ethnicities and they are also less prevalent in expensive LSOAs. Pakistani and Bangladeshi ethnicities seem to see the least benefit from both interventions. They both do not feature more in cost-effective LSOAs, whereas every other ethnicity minority does, but they do see a large increase in prevalence in expensive LSOAs. This suggests that these ethnicities are see limited benefit from either intervention and supports why the Rochdale Local Authority, which features high concentrations of both ethnicities (Phillips et al., 2008), sees limited QALY gain.

Finally the distribution of IMD deciles for cost-effective and expensive LSOAs are provided. For cost-effective LSOAs decile $2, 3, 4$ and $9, 10$ are over-represented versus the total population. The GBIS intervention also appears to perform better in these deciles. All other deciles are underrepresented against the total population and benefit more from the EPCG intervention suggesting the GBIS intervention benefits lower-middle incomes in deciles $3-5$ but does not sufficiently protect the most deprived GMCA areas.

For expensive LSOAs the lowest two deciles $(1-2)$ are much more prevalent than in the total population. There are fewer households in this group for the GBIS intervention suggesting while this intervention is expensive it is providing more help than existing strategy. Deciles $3-4$ are also over-represented but the EPCG intervention appears to benefit this group more. This group has a large number of households eligible for the GBIS intervention due to low council tax but has higher overall thermal comfort and lower income to energy bill ratios so appears to see less benefit despite large investment. Other deciles $5-10$ see similar behaviour as in cost-effective LSOAs such that they are underrepresented versus the total population and benefit more from the EPCG intervention due to high energy expenditure. Overall these results suggest that the GBIS intervention does better lower IMD deciles more than the EPCG intervention but still does not provide enough help to the most deprived LSOAs in the GMCA area.

Overall it appears the GBIS intervention is more beneficial to the desired target areas of the UK population including rural areas and ethnic minority groups. However, there are a still a number of groups that do not see improvement under the GBIS intervention including those with high rents, Pakistani and Bangladeshi ethnic groups, and the lowest IMD decile. These gaps provide some explanation as to the performance of the Rochdale and Tameside local authorities described above and suggests clean subgroups of the population that require more support.

Variable	LSOA Type	GBIS Variable Mean	EPCG Variable Mean
Monthly Household Disposable Income (£s)	Cost-Effective LSOAs	1579.51	1583.80
	Expensive and not Cost-Effective LSOAs	1430.08	1432.89
	Total Population	1584.78	1584.78
Yearly Energy Consumption (£s)	Cost-Effective LSOAs	2585.93	2593.60
	Expensive and not Cost-Effective LSOAs	2557.80	2558.07
	Total Population	2598.89	2598.89
Monthly Household Rent (£s)	Cost-Effective LSOAs	222.86	221.28
	Expensive and not Cost-Effective LSOAs	343.55	329.96
	Total Population	219.85	219.85
Monthly Household Mortgage (£s)	Cost-Effective LSOAs	334.97	337.14
	Expensive and not Cost-Effective LSOAs	267.44	271.13
	Total Population	336.36	336.36
Nutrition Quality	Cost-Effective LSOAs	16.53	16.53
	Expensive and not Cost-Effective LSOAs	16.19	16.20
	Total Population	16.54	16.54
Number of Cars	Cost-Effective LSOAs	1.31	1.31
	Expensive and not Cost-Effective LSOAs	1.06	1.07
	Total Population	1.32	1.32
Council Tax (£s)	Cost-Effective LSOAs	159.77	159.93
	Expensive and not Cost-Effective LSOAs	154.20	154.18
	Total Population	160.12	160.12
Yearly Electric Consumption (£s)	Cost-Effective LSOAs	736.34	736.11
	Expensive and not Cost-Effective LSOAs	744.79	744.11
	Total Population	737.64	737.64
Yearly Gas Consumption (£s)	Cost-Effective LSOAs	693.52	694.81
	Expensive and not Cost-Effective LSOAs	702.05	696.01
	Total Population	696.74	696.74
Yearly Fuel Oil Consumption (£s)	Cost-Effective LSOAs	903.65	903.90
	Expensive and not Cost-Effective LSOAs	900.42	899.61
	Total Population	905.84	905.84
Yearly Combined Gas and Electricity Consumption (£s)	Cost-Effective LSOAs	1206.87	1207.65
	Expensive and not Cost-Effective LSOAs	1181.88	1184.56
	Total Population	1209.04	1209.04
Yearly Other Fuel Consumption (£s)	Cost-Effective LSOAs	202.74	201.50
	Expensive and not Cost-Effective LSOAs	205.89	209.98
	Total Population	202.23	202.23
Age (years)	Cost-Effective LSOAs	45.49	45.61
	Expensive and not Cost-Effective LSOAs	43.18	43.20
	Total Population	45.55	45.55

Table 6.2: Continuous variable statistics for cost-effective and expensive not cost-effective LSOAs.

Variable	LSOA Type	Variable Levels	GBIS Percentages	EPCG Percentages
Housing Quality (1-3)	Cost-Effective LSOAs	Low	25.17	25.17
	Cost-Effective LSOAs	Medium	44.24	44.08
	Cost-Effective LSOAs	High	30.58	30.75
	Expensive and not Cost-Effective LSOAs	Low	29.18	29.21
	Expensive and not Cost-Effective LSOAs	Medium	45.96	45.79
	Expensive and not Cost-Effective LSOAs	High	24.87	25.00
	Total Population	Low	24.86	24.86
	Total Population	Medium	44.18	44.18
	Total Population	High	30.96	30.96
FP10 (False/True)	Cost-Effective LSOAs	FALSE	84.28	84.31
	Cost-Effective LSOAs	TRUE	15.72	15.69
	Expensive and not Cost-Effective LSOAs	FALSE	83.83	83.76
	Expensive and not Cost-Effective LSOAs	TRUE	16.17	16.24
	Total Population	FALSE	84.31	84.31
	Total Population	TRUE	15.69	15.69
Ethnicity	Cost-Effective LSOAs	BAN	3.40	3.39
	Cost-Effective LSOAs	BLA	2.92	2.87
	Cost-Effective LSOAs	BLC	2.61	2.50
	Cost-Effective LSOAs	CHI	0.50	0.50
	Cost-Effective LSOAs	IND	4.80	4.73
	Cost-Effective LSOAs	MIX	2.86	2.80
	Cost-Effective LSOAs	OAS	1.30	1.24

	Cost-Effective LSOAs	OBL	0.23	0.23
	Cost-Effective LSOAs	OTH	0.49	0.46
	Cost-Effective LSOAs	PAK	4.50	4.56
	Cost-Effective LSOAs	WBI	71.74	72.06
	Cost-Effective LSOAs	WHO	4.64	4.66
	Expensive and not Cost-Effective LSOAs	BAN	4.88	4.74
	Expensive and not Cost-Effective LSOAs	BLA	4.30	4.11
	Expensive and not Cost-Effective LSOAs	BLC	3.09	3.13
	Expensive and not Cost-Effective LSOAs	CHI	0.47	0.43
	Expensive and not Cost-Effective LSOAs	IND	4.92	4.76
	Expensive and not Cost-Effective LSOAs	MIX	3.47	3.37
	Expensive and not Cost-Effective LSOAs	OAS	1.32	1.48
	Expensive and not Cost-Effective LSOAs	OBL	0.31	0.31
	Expensive and not Cost-Effective LSOAs	OTH	0.66	0.69
	Expensive and not Cost-Effective LSOAs	PAK	6.61	6.59
	Expensive and not Cost-Effective LSOAs	WBI	65.43	65.93
	Expensive and not Cost-Effective LSOAs	WHO	4.53	4.46

	Total Population	BAN	3.40	3.40
	Total Population	BLA	2.93	2.93
	Total Population	BLC	2.55	2.55
	Total Population	CHI	0.48	0.48
	Total Population	IND	4.72	4.72
	Total Population	MIX	2.82	2.82
	Total Population	OAS	1.25	1.25
	Total Population	OBL	0.22	0.22
	Total Population	OTH	0.49	0.49
	Total Population	PAK	4.60	4.60
	Total Population	WBI	71.93	71.93
	Total Population	WHO	4.60	4.60
Index of Multiple Deprivation Decile	Cost-Effective LSOAs	1	22.19	22.39
	Cost-Effective LSOAs	2	14.02	13.35
	Cost-Effective LSOAs	3	11.73	10.46
	Cost-Effective LSOAs	4	9.84	9.47
	Cost-Effective LSOAs	5	6.69	8.58
	Cost-Effective LSOAs	6	6.64	7.56
	Cost-Effective LSOAs	7	7.10	7.15
	Cost-Effective LSOAs	8	7.81	8.55
	Cost-Effective LSOAs	9	7.52	6.21
	Cost-Effective LSOAs	10	6.47	6.27
	Expensive and not Cost-Effective LSOAs	1	37.60	39.68
	Expensive and not Cost-Effective LSOAs	2	16.03	16.79
	Expensive and not Cost-Effective LSOAs	3	15.33	14.58
	Expensive and not Cost-Effective LSOAs	4	10.48	8.31

	Expensive and not Cost-Effective LSOAs	5	3.39	3.54
	Expensive and not Cost-Effective LSOAs	6	4.23	5.36
	Expensive and not Cost-Effective LSOAs	7	3.33	2.53
	Expensive and not Cost-Effective LSOAs	8	5.96	4.97
	Expensive and not Cost-Effective LSOAs	9	2.63	4.23
	Expensive and not Cost-Effective LSOAs	10	1.01	
	Total Population	1	22.17	22.17
	Total Population	2	13.97	13.97
	Total Population	3	11.57	11.57
	Total Population	4	8.99	8.99
	Total Population	5	7.58	7.58
	Total Population	6	6.78	6.78
	Total Population	7	7.38	7.38
	Total Population	8	8.23	8.23
	Total Population	9	7.39	7.39
	Total Population	10	5.94	5.94

Table 6.3: Discrete variable statistics for cost-effective and expensive not cost-effective LSOAs.

6.3.5 Change in Circumstances.

Finally, this work aims to explore changes in circumstances between subgroups of the GMCA population to further explain variation in policy benefits across Local Authority geographies. Using existing microsimulation disaggregation methods (Ballas, Rossiter, Thomas, Clarke and Dorling, 2005) the intention is to explore percentage change in key MINOS variables constructing plain, interpretable narratives for why a policy causes health change and promote potential uptake of this evidence in real policy strategy.

For the subpopulation with positive $FP10$ scores in 2020, that is households that spend more than 10% of their income on energy bills, percentage changes in key variables from the MINOS model described in Figure 5.1 are provided in Tables 6.4-6.6 for the Manchester, Tameside, and Rochdale Local Authorities. These tables include the absolute number of individuals, or households for the first row only, belonging to a number of poor quality household states including poor housing quality, neighbourhood safety, and alcohol consumption risk. Percentages are also provided describing what proportion of the population these groups represent. For the first four rows percentages are compared against the total population in each Local Authority. For all remaining rows the percentage is compared against the number of

FP10 positive households in each Local Authority. For example, the percentage of households in the FP10 subpopulation is calculated as a percentage of the total number of all Local Authority households since calculating the percentage of FP10 households in the FP10 subpopulation will always trivially be 100%. On the other hand for variables such as housing quality we are only interested in the percentage of individuals in the FP10 positive population who see improvement over time to determine change in circumstances.

Change in circumstances for the Manchester Local Authority is provided in Table 6.4. Above evidence suggests the Manchester LA (Figure 6.6) sees positive health gain under both interventions but performs better under the GBIS intervention. Both interventions also have high costs due to the large number of households overall and a large percentage of private rental properties. Neighbourhood safety, financial situation, nutrition quality, and high audit-c scores all improve under both interventions suggesting widespread health improvement. Only the GBIS intervention shows positive improvement in material housing quality which may explain discrepancies in health benefits. Additionally the yearly energy expenditure for the GMCA is small due to lower energy requirements in dense urban areas suggesting the EPCG will provide less funding overall to reduce the mean energy expenditure below the price cap. Both interventions are expensive for the large Manchester Local Authority but the GBIS shows larger health improvement due to consistent improvement in intermediaries.

	EPCG 2025	EPCG 2030	GBIS 2025	GBIS 2030
Households (% Total Households)	2139 (7.895%)	1971 (5.795%)	2121 (7.834%)	1945 (5.725%)
Individuals (% of Total Individuals)	2700 (4.28%)	2502 (2.876%)	2679 (4.242%)	2483 (2.854%)
Children (% of Total Children)	1068 (5.05%)	1267 (4.278%)	1073 (5.067%)	1268 (4.275%)
Pensioners (% of Total Pensioners)	793 (12.84%)	750 (8.8%)	768 (12.415%)	727 (8.545%)
Low Housing Quality (% of Total FP10 Individuals)	1323 (49%)	1226 (49.001%)	1328 (49.571%)	1169 (47.08%)
Low Neighbourhood Safety (% of Total FP10 Individuals)	1236 (45.778%)	1087 (43.445%)	1259 (46.995%)	1083 (43.617%)
Poor Financial Situation (% of Total FP10 Individuals)	684 (25.333%)	617 (24.66%)	684 (25.532%)	620 (24.97%)
Total Private Rental Individuals (% of Total FP10 Individuals)	856 (31.704%)	812 (32.454%)	858 (32.027%)	815 (32.823%)
Total Lonely Individuals (% of Total FP10 Individuals)	1353 (50.111%)	1321 (52.798%)	1327 (49.533%)	1331 (53.605%)
Total high risk auditc score(% of Total FP10 Individuals)	134 (4.963%)	120 (4.796%)	145 (5.412%)	124 (4.994%)
Median Monthly Household Disposable Income	-175.14	34.847	-140.985	34.173
Median Yearly Energy Expenditure	2537.56	2512.576	2497.081	2496.824
Median Monthly Household Rent.	485.73	505.115	487.195	504.283
Mean Nutrition Quality	14.593	15.276	14.702	15.288

Table 6.4: Percentage Change in Key Variables for the FP10 positive population for the Manchester Local Authority.

Statistics for the Tameside Local Authority are provided in Table 6.5. This Local Authority also performs well under both interventions seeing larger but more uncertain QALY gain under the EPCG intervention. Tameside is a much smaller Local Authority than Manchester with fewer households in energy poverty and a much smaller percentage of private rental properties. Under both interventions there is improvement in a number of intermediaries including housing quality and nutrition quality. Neighbourhood safety, financial situation, and auditc change are mixed for both interventions. Yearly energy consumption is larger for the more rural Tameside LA with only the EPCG intervention showing improvement in energy bills. These results suggests the Tameside area was already performing well with respect to intermediate variables but saw a large increase in energy consumption. The EPCG addressing this area's primary problem was more effective than the GBIS.

	EPCG 2025	EPCG 2030	GBIS 2025	GBIS 2030
Households (% Total Households)	782 (6.335%)	709 (4.661%)	782 (6.34%)	706 (4.636%)
Individuals (% of Total Individuals)	1003 (3.714%)	917 (2.549%)	1002 (3.715%)	915 (2.546%)
Children (% of Total Children)	382 (5.19%)	452 (4.635%)	381 (5.18%)	447 (4.585%)
Pensioners (% of Total Pensioners)	332 (8.98%)	319 (6.422%)	330 (9.004%)	314 (6.373%)
Low Housing Quality (% of Total FP10 Individuals)	472 (47.059%)	404 (44.057%)	474 (47.305%)	422 (46.12%)
Low Neighbourhood Safety (% of Total FP10 Individuals)	431 (42.971%)	371 (40.458%)	417 (41.617%)	390 (42.623%)
Poor Financial Situation (% of Total FP10 Individuals)	203 (20.239%)	200 (21.81%)	209 (20.858%)	205 (22.404%)
Total Private Rental Individuals (% of Total FP10 Individuals)	183 (18.245%)	174 (18.975%)	185 (18.463%)	179 (19.563%)
Total Lonely Individuals (% of Total FP10 Individuals)	527 (52.542%)	487 (53.108%)	541 (53.992%)	527 (57.596%)
Total high risk auditc score(% of Total FP10 Individuals)	45 (4.487%)	51 (5.562%)	42 (4.192%)	32 (3.497%)
Median Monthly Household Disposable Income	-293.505	-27.689	-289.807	61.74
Median Yearly Energy Expenditure	2619.65	2655.417	2696.385	2669.094
Median Monthly Household Rent.	452.214	462.457	452.286	461.267
Mean Nutrition Quality	15.097	15.941	15.264	15.873

Table 6.5: Percentage Change in Key Variables for the FP10 positive population for the Tameside Local Authority.

Finally statistics for the Rochdale area are presented in Table 6.6. Figure 6.5a suggests this LA performs relatively poorly under both interventions seeing limited QALY gain. This authority is approximately the same size as Tameside with a similar number of households and percentage of private rental properties. Only neighbourhood safety and nutrition quality show consistent improvement across both interventions with mixed results for all other intermediaries. Additionally there is high prevalence of other health issues including loneliness and audit-c high alcohol consumption compared to other Local Authorities that are not addressed by either intervention. Combined with the highest overall yearly energy consumption it is clear that neither intervention is able to sufficiently address change in income due to energy prices and hence shows insufficient positive change in intermediaries to see any health benefit. Further policy intervention is required here providing further support or addressing other issues such as loneliness.

	EPCG 2025	EPCG 2030	GBIS 2025	GBIS 2030
Households (% Total Households)	814 (6.881%)	760 (5.142%)	819 (6.932%)	766 (5.173%)
Individuals (% of Total Individuals)	1068 (3.902%)	1000 (2.692%)	1080 (3.946%)	1014 (2.727%)
Children (% of Total Children)	367 (4.456%)	432 (3.897%)	368 (4.472%)	430 (3.878%)
Pensioners (% of Total Pensioners)	296 (8.438%)	315 (6.737%)	308 (8.752%)	328 (6.945%)
Low Housing Quality (% of Total FP10 Individuals)	450 (42.135%)	437 (43.7%)	484 (44.815%)	441 (43.491%)
Low Neighbourhood Safety (% of Total FP10 Individuals)	466 (43.633%)	421 (42.1%)	465 (43.056%)	427 (42.11%)
Poor Financial Situation (% of Total FP10 Individuals)	235 (22.004%)	227 (22.7%)	222 (20.556%)	239 (23.57%)
Total Private Rental Individuals (% of Total FP10 Individuals)	197 (18.446%)	192 (19.2%)	198 (18.333%)	190 (18.738%)
Total Lonely Individuals (% of Total FP10 Individuals)	589 (55.15%)	559 (55.9%)	557 (51.574%)	548 (54.043%)
Total high risk auditc score(% of Total FP10 Individuals)	65 (6.086%)	59 (5.9%)	65 (6.019%)	80 (7.89%)
Median Monthly Household Disposable Income	-56.106	130.367	-64.904	175.408
Median Yearly Energy Expenditure	2900.502	2863.873	2890.868	2861.953
Median Monthly Household Rent.	461.142	464.043	460.115	462.702
Mean Nutrition Quality	15.051	15.93	15.223	15.809

Table 6.6: Percentage Change in Key Variables for the FP10 positive population for the Rochdale Local Authority.

Overall the GBIS performs better in areas with poor existing housing stock and low to medium yearly energy expenditure such as Manchester. In this case the GBIS provides households larger improvement in circumstances doing more than just protecting their household income but supporting housing quality and other intermediates such as alcohol consumption as well. The EPCG performs better in areas that are primarily only suffering loss in well-being due to high energy prices such as Tameside. If energy prices are very high and existing housing stock is already good quality neither intervention is effective suggesting these interventions do not reduce energy expenditure enough or some unobserved further prominent issues such as loneliness and neighbourhood safety that are not addressed.

6.4 Discussion

There is increasing justification for the inclusion of spatial heterogeneity in health policy to ensure efficient assignment of resources and a high minimum standard of well-being (Augustin et al., 2023). Recent datasets and surveys commissioned in response to high energy prices and carbon neutrality targets have highlighted existing UK Government insulation retrofitting policy that applies funding uniformly to Local Authorities based on population size is wasteful and does not sufficiently protect all households (Bridgen and Robinson, 2023). Application of asymmetric and dynamic policy according to individual household needs is also difficult with limited evidence available and low popularity with policy makers and the general public. Generating evidence for the effect of energy policy at small spatial scales would facilitate optimisation highlighting sub-populations that require more support while demonstrating cost and health benefits at national scale that are interpretable and can be adapted into existing Government budgets.

This chapter provides a demonstration of how output from a spatial dynamic microsimulation model can be used to provide evidence for the spatial effect of existing energy poverty. Two existing Energy Price Cap Guarantee and Great British Insulation scheme policies are described over LSOA, Ward, and LA spatial resolutions for the GMCA area highlighting where each intervention works well and gaps that can be addressed in future energy policy. Validation and mapping of energy consumption for four key fuel types is provided demonstrating expenditure in the synthetic population strongly aligns with available external real and simulated data. This validation confirms that energy expenditure behaviour performs as expected and gives insight into energy consumption across the GMCA area.

Mapping at the Ward-level is provided for change in SF-12 MCS, SF-12 PCS, and QALY score for the EPCG and GBIS interventions. While the EPCG intervention provides larger improvement in mental well-being due to lower household energy bills the GBIS intervention shows larger improvement in physical health which contributes to larger gain in QALY of up to 20% per LSOA versus the baseline intervention where no energy policy is applied. This mapping suggests each intervention addresses different issue caused by high energy pricing within the UK population and potentially a combination of each policy may provided further health benefits.

Linear regression is then used to compared percentage gain in QALY score under the GBIS and EPCG interventions against baseline to determine if certain spatial features are predictive of health improvement. Overall the GBIS intervention sees positive improvement in QALY versus the EPCG intervention with additional improvement for rural households including farmers and retirees, ethnic minority groups, and low income households including students and ageing urban communities. QALY gain is also compared against SIPHER Inclusive Economies indicators highlighting low incomes, high energy poverty, and poor transport links are key drivers of improved QALY gain and potentially high energy consumption overall. All Local Authorities see improvement in QALY gain versus the baseline Bolton dependent primarily on household income and energy poverty rates with notable exceptions of the Tameside which over-performs and Rochdale which under-performs relative to expectations according to Inclusive Economy Indicators.

Further analysis using Incremental Cost-Effectiveness Ratio (ICER) planes identifies three groups of LSOAs including individuals that are above the Green Book $70k$ ICER threshold, LSOAs that have low intervention cost and low QALY gain, and LSOAs that are expensive and not cost-effective. Further exploration of cost-effective and expensive groups is performed describing statistics and differences in a number of variables in the initial 2020 MINOS synthetic population. Overall it is clear the expensive but not cost-effective LSOAs for the GBIS intervention are a mixture of households with high rent, Pakistani and Bangladeshi ethnicities, and belong to the lowest Index of Multiple Deprivation decile which explain previous discrepancies in Rochdale and Tameside Local Authority performance.

Finally, a short change in circumstances analysis is provided exploring difference in policy response for the Manchester, Rochdale, and Tameside local authorities. This work helps to explain performance in the Manchester and Tameside local authorities with respect to the GBIS and EPCG interventions. Manchester is expensive but sees notable improvement in energy spending and housing stock so sees a large health benefit. Tameside has improved housing stock but larger energy bills so sees larger benefit from the EPCG intervention. The Rochdale local authority is still difficult to explain but appears to have high energy bills and good existing housing stock suggesting some further health issue in this area that neither policy can address.

Overall these results suggest the GBIS intervention shows greater improvement in QALY and ICER scores for rural, low income, and ethnic minority households. However, there are still groups overlooked by the GBIS intervention that could be addressed with more specific policy intervention as well as combination with other energy policies such as minimum efficiency standards, cheaper energy pricing for separate electricity and gas payment, and a 'fairer' market that bans practices such as prepayment meters (UK Fuel Poverty Monitor, 2022). Additionally some households in the UK have multiple issues and energy poverty alone are not sufficient to raise public health. Further combination with other policy sectors such as employment may provide further health gain and improved cost-effectiveness. While these LSOAs are expensive and see limited QALY gain omitting these LSOAs from intervention is also not practical. Basing policy decisions purely on hard ICER thresholds is obviously naive prompting the need for policy maker discretion and potentially more dynamic measures of cost-effectiveness. Assessment purely on QALY and limited subgroups of the GMCA population is just a brief exploration into how spatial analysis could be used to identify specific subgroups omitted from an intervention providing cases for more targeted policy in UK Government.

6.4.1 Limitations

There are a number of limitations to this analysis of MINOS output data. The MINOS model itself does not include a number of demographic processes.Internal and external migration is not included allowing individuals to move to higher quality housing and different areas. There is also limited cohabitation such that individuals cannot leave households and cannot form new households and relationships. Children do not fully transform into adults and eventually move out of the household. Inclusion of these further demographic processes will ensure more accurate estimation of the long term GMCA population. There is also limited inclusion of social behaviour including policy uptake and potential different attitudes in different areas. Further inclusion of LSOA level behavioural data using methods such as propensity matching may give a more realistic estimation of uptake particularly of the GBIS intervention.

The MINOS population is only a 10% sample of the full GMCA population. This sample size is required due to computational power but LSOAs can be highly sensitive to random sampling. Larger samples of model runs and variance reduction techniques are required and may provide clearer changes in population characteristics between interventions.

A number of key characteristics are also not included in the MINOS population notably details on housing insulation types and specific features including being in a detached/semi detached house and existing insulation and energy efficiency rating. Inclusion of these variables may allow for more detailed exploration of how the GBIS policy can be tailor to better effect target LSOAs.

Finally these energy policies are considered in isolation and ignore any interaction effects between other policies and potential geopolitical events that have occurred since 2020 such as the coronavirus pandemic. While this is done intentionally for counterfactual scenarios to provided clearer estimation of policy effect inclusion of other policies and events will improve realism and uptake in Government policy analysis.

6.4.2 Future Work

There are a number of opportunities available for potential future work. Exploration of energy poverty policy that is more spatially detailed as opposed to existing largely uniform blanket policy is required. Policies such as the upcoming reduction in the winter fuel payment only effects the elderly population, rent controls such that landlords must provide minimum energy efficiency standards, and policies targeting fuel oils and may provide further distinct spatial features. Combinations of policies together and variation over time and space produce much more complex policies that can be used in conjunction with this spatial analysis providing more complete case studies for improvement of existing Government strategy. Existing policies can also be applied to other areas besides the GMCA region that are more spatially diverse such as the Scotland or the entirety of the UK subject to computation power.

Further development of the MINOS synthetic population would be useful to incorporate further household-level spatial variables as opposed to LA and ward level data used throughout this chapter. Combination with existing data on insulation type and energy efficiency rating would greatly improve the precision that can be applied. Further development of the population is being undertaken as part of SIPHER consortium research with particular focus on longitudinal consistency of synthetic data.

Incorporation of transition probability models into the MINOS model that incorporate spatial heterogeneity is also desirable. Consideration in mixed effects, hierarchical Bayesian models (Leknes and Løkken, 2021), and interactions including spillover effects and clustering could provide further insight into how energy poverty causes development of health in larger neighbourhoods.

This research can also be extended by introducing further spatial analytics methods. Further cross tabulation and regression analysis can be used to identify more specific groups of the population that benefit more or less from energy interventions. For the GBIS example exploration of multiple groups such as high rent and belonging to specific local authorities may shed further light on better allocation of policy funds. Use of SIPHER decision support tools is also of interest exploring implementation of policy that varies at an LSOA level trying to find optima over very high dimensional state spaces (Stewart et al., 2024) but this has very large computation costs.

This chapter provides brief analysis of individual-level datasets with spatial attributes produced by the MINOS microsimulation. The GBIS intervention clearly benefits the GMCA population more than existing strategy but still has a number of clear shortcomings that can be addressed. Initial spatial analysis methodology provided here has huge potential application in policy analysis beyond energy policy and the UK. Providing a toolkit for policy makers to optimise dynamic microsimulation and policy over space could lead to substantial improvement in resource allocation and application of just, equitable policy.

Local Authority	Bolton	Bury	Manchester	Oldham	Rochdale	Salford	Stockport	Tameside	Trafford	Wigan
Participation in Paid Employment	0.705	0.739	0.651	0.693	0.681	0.697	0.756	0.718	0.755	0.726
Skills and Qualifications	0.857	0.884	0.861	0.842	0.841	0.862	0.900	0.852	0.907	0.864
Involuntary Exclusion from the Labour Market	0.050	0.043	0.062	0.050	0.059	0.060	0.037	0.054	0.036	0.053
Digital Connectivity	0.520	0.273	0.516	0.594	0.565	0.524	0.207	0.593	0.208	0.484
Wealth Inequality	2.063	1.994	1.878	2.008	2.000	2.036	1.979	1.925	1.986	2.018
Physical Connectivity	0.474	0.543	0.168	0.330	0.680	0.379	0.619	0.479	0.293	0.684
Earnings Inequality	2.418	2.451	2.351	2.404	2.405	2.363	2.445	2.395	2.465	2.400
Housing Affordability	6.913	8.768	9.779	7.093	7.106	8.662	11.538	7.719	14.185	7.013
Poverty	0.291	0.216	0.306	0.316	0.274	0.224	0.138	0.232	0.113	0.184
Cost of Living	0.058	0.052	0.083	0.058	0.061	0.066	0.046	0.057	0.049	0.052
Decent Pay	0.709	0.724	0.681	0.703	0.703	0.699	0.740	0.700	0.742	0.713
Decision Making Inclusion	0.362	0.364	0.271	0.343	0.300	0.262	0.358	0.278	0.410	0.278
Job Security	0.948	0.949	0.933	0.948	0.948	0.945	0.951	0.953	0.948	0.954

Table 6.7: Inclusive economy indicators proportion by Local Authority.

Inclusive Economy Indicator	Definition
Participation in Paid Employment (1A)	Percentage of working-age people who are employed.
Skills and Qualifications (1B)	Percentage of adults aged $20 - 49$ with a Level 2 or higher NVQ qualification.
Involuntary Exclusion from the Labour Market (2A)	Share of working-age people who are long-term unemployed or inactive due to ill health or disability.
Digital Connectivity (2B)	Engagement with digital as the LSOA level based on Internet User Classification (IUC).
Wealth Inequality (3A)	Ratio of median house prices in most expensive neighbourhood to least expensive neighbourhood.
Physical Connectivity (3B)	Public transport accessibility measure.
Earnings Inequality (4A)	Ratio of weekly earnings (residents in full-time work) between 80th and 20th percentiles.
Housing Affordability (4B)	Ratio of median house prices to median workplace earnings.
Poverty (5A)	Percentage of children living in low-income households.
Cost of Living (5B)	Percentage of households in poor subjective thermal comfort.
Decent Pay (6A)	Proportion of employee jovs that are pain above the statutory living wage.
Decision Making Inclusion (6B)	Percentage voter turnout in local elections.
Job Security (7A)	Share of employees in permanent work.

Table 6.8: Definitions for parliamentary ward-level SIPHER inclusive economy indicators Lupton et al. (2023) by Local Authority.

Chapter 7: Discussion and Conclusion

Local and national governments in the United Kingdom struggling to adopt a Health In All Policies approach to prevent non-communicable disease due to limited available evidence (Greszczuk, 2019; Meier et al., 2019). Real evidence from small-scale randomised controlled policy trials (Rutter et al., 2011; Skarda et al., 2021) and natural experiments (Marois and Aktas, 2021) will always be preferable but often this data cannot be gathered due to limited time and resources. Policy evidence is often required in response to both urgent crises as well as long term policy targets at the national-level. Many policies are often implemented blindly without any understanding of the health outcomes that can potentially negatively impact public health.

One alternative is to simulate artificial evidence approximating how a candidate policy would effect the health of the UK population. However, predicting future health is highly complex due to multiple determinant factors and strong heterogeneity across households and neighbourhoods. Complex systems modelling aims to capture these relationships between a candidate policy, what characteristics it changes in a population, and how this propagates through to impacting health using theory of change and causal loop diagrams (Meier et al., 2019). These diagrams are then emulated using individual-level modelling that creates a copy of the UK population and projects it forwards in time to observe the effects of a policy. This allows policy makers to test a spectra of potential candidate policies quickly to inform decisions that preserve public health.

The stated primary aim of this thesis was the development of a dynamic microsimulation model emulating causal pathways between income and health to robustly evidence and optimise the effect of income and housing support policy in the United Kingdom at individual and neighbourhood levels. This aim was broken down into three objective research questions to review transition probability methods implementing robust causal loop diagram evidence into a microsimulation model, develop a dynamic microsimulation model to estimate the effect of income support policy on health for the UK population, and implement more complex housing support policy for the UK using this microsimulation to begin to explore cost-effective energy policy poverty over key sub-populations and sub-geographies. These research questions are addressed with several key outputs.

First, a literature review explored best practice for implementation of transition probability models in dynamic microsimulation supplement with applied case studies estimating change in Short Form 12 Mental Well-Being score given household characteristics and longitudinal histories. This review serves to collect transition probability methods in a single tutorial paper that are desirable for increasing the impact of dynamic microsimulation by increasing accessibility to new users and applications. This review also serves to highlight application of ATOMIC principles (Wasserstein et al., 2019) in transition probability development that consider uncertainty, thoughtfulness of other research, open data and methods, and modesty to produce reproducible and interpretable projections fostering trust with policy makers and the general public when using simulated evidence to justify new policy. The second research question of this thesis then seeks to develop an initial dynamic microsimulation prototype model MINOS emulating income to health causal loop diagrams to estimate change in mental well-being and clinical depression outcomes for the UK population for a series of simple income support policies. The final research question of this thesis then intends to expand upon the MINOS model implementing further energy poverty pathways and a synthetic spatial population to compare the cost-effectiveness and equity of energy poverty policy for the Greater Manchester Combined Authority area.

Addressing these research gaps provides a microsimulation model for use beyond this thesis in estimation of further income and housing support policy and their effect on non-communicable disease. Simulated evidence from MINOS can be used to ensure that vulnerable subgroups of the population are protected in the event of immediate crises, to influence the application of current policy ensuring that policy is financially and politically feasible, and ensure long term policy goals such as carbon neutral targets are met. Work from the MINOS model is now being used providing evidence for internal SIPHER policy partners including the GMCA and Scottish governments as well as external bodies such as the Mental Health Foundation. This chapter presents key findings from this thesis, limitations in results, and scope for potential future work.

7.1 Key Findings

Can I contribute to literature codifying state of the art best practice for developing transition probability models in dynamic microsimulation?

Recent reviews into the state of the art for dynamic microsimulation (O'Donoghue and Dekkers, 2018) highlighted a gap in literature for the best practice development of transition probability models for dynamic microsimulation. Addressing this gap would facilitate the standardisation of microsimulation knowledge to increase accessibility to new users and promote uptake in evidence-based policy making. A review of literature evidence is supplemented with tangible case studies focusing on development of a single transition probability model as well as an entire microsimulation in order to provide further tangible examples highlighting any potential difficulties and pitfalls in applying best practice. This research question resulted in three key findings.

The first key finding identifies that methods to improve reproducibility and prediction quality for transition probability models can be divided into five categories for presentation, missing data, validation, uncertainty, and heterogeneity. Improved presentation of transition probability models ensure transparent and interpretable models that are open to critical review. Consideration of missing values ensures inclusion of as many individual observations as possible reducing bias and ensure sample data remains representative. Quantification of uncertainty and validation ensures model predictions are sensible and match available data while also being robust over time and under new hypothetical scenarios. Heterogeneous modelling methods allow for better prediction of individual behaviour by including more of their information in prediction including longitudinal individual histories, relation to other nearby households and family members, and spatial information. While these methods are used throughout statistical modelling this review has highlighted aspects unique to dynamic microsimulation. Presentation and validation must consider how predicted outcomes change under repeated application of a transition probability model. Often transition models will 'regress to the mean' as many models such as linear regression can underestimate population variance over time. Out of sample performance must be considered as a microsimulation model predicts into the future and under counterfactual scenarios. Validation techniques including alignment and variance reduction are used to create reliable simulation data that matches macro-level simulated aggregates ensuring sensible prediction at future time points and under counterfactual interventions.

The second key finding shows that these methods can be applied to a transition probability to improve model fit, robustness, and reproducibility. Using Understanding Society (UKHLS) data an ordinary least squares transition probability model was fitted estimating future Short Form 12 Mental Component Score between 2014 – 2020 for individuals within the UK population. This initial model showed a poor adjusted R-square score (0.05) and Root Mean Square Error score (0.223) and failure to predict long term SF-12 MCS population mean and variance. Presentation of model diagnostics and validation were used to identify where underlying model assumptions were violated identifying how this model could potentially be improved. Accounting for missing data using deterministic methods and multiple imputation increased the available sample size and allowed for use of variable not available in all waves of UKHLS data. Sensitivity analysis testing input data, model coefficients and structure, and Monte Carlo noise were used to test robustness to inputs showing limited overall change in prediction. Finally use of Generalised Linear Mixed Modelling and lagged dependent variables allowed for consideration of heteroscedasticity due to repeat observation showing better fit of model assumptions and ten-fold improvement in R-squared score to 0.55 and improvement in RMSE (0.152).

The final key finding demonstrates how these transition probability model best practices can be applied to development of a full dynamic microsimulation model. In empirical Chapters 4-5 I have demonstrated using these best practice guidelines with respect to planning, validation, and interpretation developing causal loop diagrams between income, housing, and health into a series of transition probability modules.

Consideration of user requirements, existing microsimulation frameworks and available datasets quickly resulted in writing transition models using R. Markdown notebooks. Using the R. programming language allowed for construction of models in a widely used programming language with substantial built-in diagnostics and visualisation tools to increase the speed of development and facilitate iterative critical review within the multidisciplinary SIPHER consortium. These notebooks promoted clear presentation of transition model selection justifying choice model selection choices such as overall modelling techniques based on outcome variables, choice of model structure including independent variables, analysis of goodness of fit using validation and uncertainty methods, and other user requirements such as computational cost.

Strong planning then supported strong validation and interpretation. Standardised presentation of model coefficients and diagnostics as well as common validation techniques including cross-validation and visualisation techniques such as handover plots made these transition models easy to critical review in online documentation as part of wider systems thinking ideology considerate of all policy evidence including other health simulation. Model formulae and coefficients can easily be made available online allowing for users to interpret how a policy would change each given variable in a microsimulation. Common visualisation techniques to measure change in health outcomes are used for the MINOS model allowing for easy comparison with potential future health simulation modelling. These outcomes are represented both at full national population level as well as by key subgroups and small geographies.

It was difficult to validate and interpret the effect of transition models particularly when investigating interactions between transition models and how behaviour may change under policy intervention. These interactions can be highly non-linear and involve tens of transition models making them difficult to validate and highly sensitive to overall microsimulation structure. This corroborates with difficulties highlighted in larger microsimulation validation literature (Harding et al., 2010b) and further case studies exploring these interactions using the MINOS model would be beneficial. Additionally validation of a synthetic spatial population was also difficult due to very limited available data to compare against. Expansion of the simulated annealing algorithm to incorporate more waves of data and multiple populations may facilitate further validation. There is also limited other energy policy simulated literature to compare against but this will likely change very quickly as demand increases. Further visualisation methods may also support higher interpretability of interactions between transition models incorporating more detailed individual information such as long term behaviours. Describing change in circumstances allows for construction of simple plain-language narratives for policy effect and variance over vulnerable subpopulations promoting model uptake in real government planning by describing potential inequity.

How can dynamic microsimulation be used to operationalise causal loop diagrams as part of a complex systems modelling framework to generate synthetic evidence for policy effect on change in health outcomes?

The first finding was that a dynamic microsimulation model can be developed using publicly available datasets. Causal loop diagrams representing pathways between household disposable income and physical and mental well-being from the SIPHER consortium were adapted into microsimulation transition probability modules based on present variables in Understanding Society. These pathways between income and Short Form 12 Mental Component Score (SF-12 MCS) showed sensible projection into the future preserving statistical moments and linear trends in the original UKHLS data. This model could also be used to estimate reduction in risk of clinical depression based on the prevalence of individuals below a 45.6 SF-12 MCS threshold score.

Several income support policies were then applied within the MINOS model estimating how change in household disposable income affects change in mental well-being. The Scottish Child Payment was implemented first providing households £25 per child per week given they are on universal credit and numerous legacy benefits. Increasing this payment, which is currently being debated in Scottish parliament, up to £50 saw a near linear increase in SF-12 MCS from 1.5% to 2.8% by 2035 beginning to show diminishing returns. Likewise a real living wage policy was implemented raising the living wage £12 outside of London and £13.15 inside London exceeding the current UK living wage. Again an increase in SF-12 MCS score of 2% by 2035 shows positive improvement with low uncertainty but this policy is highly targeted and more expensive per capita.

Finally the energy price cap guarantee policy was implemented capping mean household energy bills to £3500 as fossil fuel prices increased from 2022 onwards. If the UK Government simply allowed energy prices to increase, this would result in an overall reduction in mental well-being of -1.2%. Implementation of the EPCG policy reduced this decline to only -0.4% by 2035. While this policy does alleviate reduction in SF-12 MCS score by protecting household disposable income, it is not sufficient to return health to pre-2020 scores. This suggests that the EPCG was not as effective as stated in published literature (UK Fuel Poverty Monitor, 2022; Chai et al., 2021). While this intervention has a relatively low cost per capita it is highly expensive as the entire UK population must be intervened upon.

How can evidence from dynamic microsimulation be used to compare the cost-effectiveness of competing policy exploring how priority subgroups and spatial regions are affected differently?

The first finding demonstrates that the MINOS model could be expanded to further incorporate energy poverty pathways and new health outcome variables. Short Form 12 Physical Component Score, Quality Adjusted Life Years, and Incremental Cost Effectiveness Ratios have all been integrated into the initial MINOS model. Again visualisation suggests that estimation of these health outcomes is stable over time although there is very limited data to validate against.

The second finding demonstrates that the SIPHER consortium synthetic input population can be adapted to provide a 10:1 scale, spatially representative estimation for the Greater Manchester population. The MINOS model combined with High Performance Computing was able to simulate the larger population of approximately 200k individuals. Again validation suggests prediction of this population is sensible over time.

The third finding is that proactive insulation retrofitting policy may be more cost effective than existing UK Government energy policy as early as 2024. The Energy Price Cap Guarantee policy and Great British Insulation Scheme policies were applied to the UK population from 2020 onwards. It is assumed that energy prices increase in line with existing quarterly wholesale fossil fuel prices and beyond using forecasted values. Both of these interventions see improvement in SF-12 MCS 0.5% for both interventions and SF-12 PCS scores of 0.3 and 0.5 respectively for the EPCG and GBIS interventions. This results in the EPCG intervention gaining approximately 2500 QALYs, which is not statistically significant from 0, and the GBIS gaining over 6000 QALYs. While the GBIS intervention has very high initial fiscal capital requred it quickly becomes more cost-effective as early as 2024 with a final ICER ratio of approximately £$300,000$ per QALY which is still not cost-effective by Green Book guidelines. While the EPCG intervention is initially more cost-effective according to ICER ratios, this quickly changes for two main reasons. First, the EPCG requires constant capital to maintain mean household energy bills whereas the GBIS only requires large initial capital. Households only need to be insulated once at least for the foreseeable future. Over time the GBIS is then ultimately cheaper and any gain in QALYs gives higher ICER scores. Second, the GBIS intervention has a larger effect on subjective thermal comfort for households. Energy bill savings under the EPCG intervention are proportional to the initial energy consumption of the household where as the GBIS savings are proportional to household size. The result is lower income households, whom contain many of those individual with poor initial thermal comfort, see the most benefit from the GBIS intervention. The GBIS then sees a large gain in physical well-being, and hence ICER scores. This is further supported exploring how SF-12 PCS gain varies across income quintiles for the Manchester populations. The lowest income decile sees significantly more benefit from the GBIS intervention than all other quintiles. On the other hand, the EPCG supports all households with no significant difference between quintiles. While the GBIS intervention is still not cost-effective according to Green Book guidelines this could change quickly as insulation costs decrease or fossil fuel prices increase further. Overall this finding has strong implications for how the UK government achieves its future net zero targets and reduces vulnerability of its public health to fuel markets.

Finally, data from the spatial population data from the MINOS model was analysed determining why the EPCG and GBIS policies varied spatially and within particular subgroups to determine any individual left behind or if the affects were different by each intervention. A brief review highlighted why spatially heterogeneous policy making may benefit the UK providing methods and specifications for spatial data attached to the MINOS model (including Rural Urban Classifiers and Output Area Classifiers). Initial results demonstrated validation of energy consumption's within the MINOS model over space showing strong alignment with existing Department for Energy Security and Net Zero data. The effect of the EPCG and GBIS interventions were mapped at ward level for SF-12 MCS, SF-12 PCS and QALYs showing results that aligned previous results in Chapter 5 demonstrating the GBIS intervention shows larger improvement in physical health which in turn causes larger improvement in QALY. Overall the EPCG showed strong mental well-being improvement and the GBIS intervention showed stronger physical well-being improvement. As physical well-being contributes more to QALY calculation the GBIS intervention showed stronger overall maximum improvement of 20% in QALY score. Linear regression explored how the GBIS intervention saw improved QALY gain against the EPCG intervention with several key spatial variables are predictors. This regression demonstrated that the GBIS intervention showed larger gains in QALYs for regions that are lower income, more rural, high ethnic minority prevalence, and high existing energy poverty suggesting the GBIS intervention is affecting its desired targets (Regan, n.d.). Finally, analysis of ICER scores for each small LSOA region was performed using ICER planes (Pirc et al., 2018) showing two main groups of LSOAs with low and high intervention cost. A small percentage of the LSOAs in the low intervention cost group were cost-effective while the high investment group saw limited QALY gain despite the required capital.

Further analysis explored distributional differences in LSOAs for the 2020 synthetic population to identify subgroups and sub-geographies that performed abnormally under the GBIS and EPCG interventions. Cost-effective LSOA intervention under the GBIS intervention were identified to be rural households, high income to energy expenditure households, most ethnic minority groups, and IMD deciles 3 to 5. LSOAs that were not cost-effective under the GBIS intervention include households with high rent, Pakistani and Bangladeshi ethnic groups in the Rochdale area, and the lowest 2 IMD deciles. Households in these expensive and not cost-effective groups particularly in the bottom 2 IMD deciles typically have multiple problems that the GBIS or EPCG interventions alone cannot address with additional funding for these poorly affected groups or combination with other policy including renting energy efficiency standard and unrelated policies including employment opportunities. While these interventions are not cost-effective for these LSOAs it is clear that these areas should not be omitted from the GBIS intervention but inclusion of additional policy and potentially further measures of cost-effectiveness rather than that scalar green book threshold.

7.2 Contributions and Impacts

This section outlines wider outputs of this thesis and contribution to gaps in existing literature.

7.2.1 Theoretical Contributions

The primary output of this thesis is a significant contribution to the MINOS dynamic microsimulation model. This framework has seen strong initial success and application is now being expanded to work beyond the scope of this thesis. Application to child poverty interventions has been expanded upon and is now being considered for utilisation within the Scottish Government and the Mental Health Foundation. Within the SIPHER consortium this model is also being used by other modelling teams applying both applying the existing model and expanding to new applications including decision support optimisation.

Another major output of this thesis is the contribution to the knowledge for the application of causal systems maps into individual-level dynamic microsimulation modelling as part of a complex systems methodology. Iterative discussion with policy partners and SIPHER academics to construct causal systems maps and constrain them on available data combined with best practice development for transition probability models has been used to demonstrate application of two causal systems maps for income to health and energy poverty. This practice will ideally be used going forwards within the SIPHER consortium further extending application of the MINOS model to new scenarios discussed in future work.

7.2.2 Methodological Contributions

Substantial literature is available for the development of synthetic spatial individual-level populations. However these populations are typically static such that they are only representative for a single fixed time point. This thesis demonstrates the potential application of dynamic microsimulation as a means of updating these synthetic populations forwards in time to estimate future states subject to validation and alignment to maintain representativeness.

Simple spatial analysis is also provided here identifying how the GBIS policy varies over space including areas that did not benefit from this intervention, notably the Rochdale Local Authority in the GMCA area, in order to improve the existing intervention. Further analysis of spatially heterogeneous policy presents a large area of potential future research particularly for energy poverty interventions.

7.2.3 Application Contributions

Application of microsimulation to health economics outcomes such as QALYs for income and housing support policy is a novel contribution that is being implemented by multiple research groups (Broadbent et al., 2023; Katikireddi et al., 2022; Kopasker et al., 2024; Skarda et al., 2021). Measures such as QALYs have been used for over 20 years to assess policy and medical treatment efficacy but dynamic microsimulation provides the ability to assess these measures over both national scale and small subgroups and sub-geographies simultaneously. Government bodies can then use models such as MINOS as part of a wider evidence toolkit optimising policy over small areas justifying the application of precise, dynamic policy to both policy makers and the general public improving average and worst case health outcomes including incidence of non-communicable disease.

As part of the first research question of this thesis literature review and tutorial case study application for best practice development of dynamic microsimulation transition probability models has contributed to wider literature to codify microsimulation methodology knowledge into open source material. Outlining methods used to improve transition probability goodness of fit and robustness to input parameters in line with larger ATOMIC principles throughout statistics helps to facilitate increased impact of microsimulation by increasing accessibility and reliability of simulated evidence. Case study application of longitudinal transition probability models showed significant improvement in prediction based on root mean square error score. This thesis contributes to a larger portion of microsimulation literature utilising these more complex models to better consistently estimate individual longitudinal trajectories, dynamicism in response to policy interventions, and preservation of population mean and variance.

This output has also contributed to the expansion on existing python microsimulation framework vivarium. This fully open source microsimulation framework can now utilise transition models from the R. programming language that are much easier to implement. A stripped back version of the vivarium framework has also been developed removing unnecessary features and improving runtime. Finally a small package has been made allowing for conversion of R. notebooks into html website pages facilitating automatic creation of online documentation required for large models such as MINOS.

7.2.4 Policy Contributions

Results from the MINOS model have implications for several real-life income and housing support policy. For the Scottish Child Payment, doubling the child payment from £25 to £50 would significantly increase SF-12 MCS score beginning to show diminishing returns. As this policy is being argued in Scottish Parliament and Government budgets continue to be reduced justifying increasing this policy may have substantial long term impact on child poverty and quality of life for Scotland. Similarly increasing the minimum wage to the real living wage also shows improvement in SF-12 MCS scores with low uncertainty but it effects a small percentage of the population with a large cost per capita.

Finally, several housing and income support policies are being proposed in order to counteract high energy bills in the UK. The EPCG and GBIS policies are both implemented in Chapter 5 demonstrating vastly different approaches to reduce energy bills. While the EPCG is cheaper initially by reducing household energy bills the GBIS scheme is cheaper long term as early as 2024 due to high gain in physical well-being from subjective thermal comfort. This result suggests that proactive insulation policy, while very expensive, quickly sees a return on investment. This is supported by response to high energy prices in neighbouring European countries such as Finland that are well insulated showing low vulnerability to fossil fuel pricing and ultimately reduced excess mortality and incidence of NCDs. Major investment in insulation retrofitting is showing mounting evidence to justify the investment both financially, and for improved public health.

7.3 Limitations

There are a number of existing limitations with output from this thesis. First, only associative transition probability models are used in the MINOS model. While associative prediction can produce accurate estimation of future state it is desirable to use causal prediction methods (O'Donoghue and Dekkers, 2018; Arnold et al., 2019) in order to tell more interpretable stories of why a policy enacts change. Causal inference methods including g-formulae (Kühne et al., 2022; Murray et al., 2017) approaches are gradually seeing implementation in microsimulation with existing research groups proposing application to health economics microsimulation (Katikireddi et al., 2022; Arnold et al., 2019). Implementation of causal methods requires substantially more data ideally with exact time between life events (Arnold et al., 2019; Katikireddi et al., 2022) such as disease progression or must be completely simulated and aligned to available administrative data (Ackley et al., 2022).

The MINOS model is highly dependent on available administrative data. If data is not available allowing for the parameterisation of certain causal pathways and policy interventions then these pathways simply cannot be included. As data is also only available in yearly cross sections it is difficult to estimate processes that operate on smaller time scales such as days of unemployment. Extension of the existing MINOS synthetic population is required to include new variables such as housing insulation characteristics (Department of Energy Security and Net Zero, United Kingdom Government, 2024a). Estimating smaller time scale processes again requires full simulation that aligns to available discrete time data (Shewmaker et al., 2022). Application particularly of data assimilation methods has strong potential to incorporate this small time scale modelling into microsimulation (Birkin, 2021; Ackley et al., 2022).

Substantial effort has been made to validate the SIPHER synthetic population against LSOA-level aggregate data (Hoehn et al., 2024) where available but there are still a number of variables within UKHLS data not included. Some validation work has been provided in this thesis assessing energy consumption data but further analysis is required for the wider UK population as well as other pathways variables such as number of cars. This additional validation is required to ensure the synthetic population is representative of UK energy consumption and any emulated policy cost-effectiveness. Additionally the synthetic population has only been generated for the 2020 year population and does not include previous years of observed data (Hoehn et al., 2024). Further work is being undertaken to validate individual trajectories over time and interactions within and between households to further improve the quality of synthetic data but requires substantial computational resources.

The MINOS model is written in the python and R. programming languages. These languages are easier to interpret and develop in for new users but are slow compared to other other microsimulation frameworks in C++ and Java (De Menten et al., 2014; Richiardi and Richardson, 2017). While the MINOS model is fully integrated into Leeds High Performance Computing systems, speed improvements are required for the existing MINOS model for application to many model runs and optimisation methods that can require tens of thousands of iterations. This can be addressed by incoming speed improvements in the python language including static typing and just-in-time packages such as ‘numba’ (Kerr et al., 2021) or migration to other languages notably C++ (Li and O’Donoghue, 2013) and Julia (Lee, 2021) that are designed to combine both speed and high accessibility.

When applying new policies and causal diagrams over 90% of development time is dedicated to validation to ensure model output is sensible and aligns with other available evidence (O’Donoghue and Dekkers, 2018). Application of new policy scenarios and outcomes to the MINOS model then also requires substantial additional development time. Ideally unique causal loop diagrams are developed and implemented for every policy question but this is rarely feasible and the same diagrams are used repeatedly. The MINOS model then requires development time in order to respond to new public health crises with trained staff. To improve response time further codification of validation methods in line with other literature (Burgard et al., 2020) should be supplied in online open source technical documentation facilitating application of the MINOS model beyond this project.

Energy policy applied in this thesis has limited heterogeneity with respect to household characteristics. Detailed housing characteristics are not present in UKHLS data including insulation types, energy efficiency ratings, construction materials, and status such as exterior wall surface areas. Without this information it is difficult to precisely identify how much money a household would save with insulation and what types of insulation can be provided. While this thesis has provided uniform policy interventions in line with real government strategy this data is essential to begin to further optimise energy policy. Future datasets from the Department of Energy Security and Net Zero are providing this information (Department of Energy Security and Net Zero, United Kingdom Government, 2024a) on a household level that will be linked with UKHLS data. However, there is likely to be substantial missing data and a new synthetic population may inevitably be needed. With this data exploration of more rural and urban specific policies such as use of solid wall insulation can be explored in higher detail.

Overall demographic process modelling is limited within the MINOS model. Estimation of long term policy effect is biased is demographics including cohabitation and migration are not incorporated into the microsimulation model structure. These processes have been omitted for various reasons including lack of available data, expensive computational cost, or low relevance to specific income and housing policy but broader application of the MINOS model will likely requires these additional processes. Further inclusion of synthetic population variables and spatial transition probability models such as migration models can better estimate the long term UK population. Additionally social behaviour modelling (Klevmarken, 2022) can provide more dynamic response to policy interventions including uptake and changes in behaviour due to crisis such as lowered energy spending and its effects on other variables.

Additional uncertainty quantification of the MINOS model is also required through sensitivity analysis. Sampling from the synthetic population simulated annealing algorithm (Hoehn et al., 2024), inclusion of further randomisation of coefficients, model selection particularly over time and module ordering, and Monte Carlo noise can all be applied to better test sensitivity of the UK population to income and housing support policy providing more robust, trustworthy results (Burgard et al., 2019; Sharif et al., 2012).

7.4 Future Work

Finally, a review of available work to improve and expand upon this thesis is provided including both methodological improvements and application to further policy scenarios.

7.4.1 Methodological Future Work

There is large scope to expand on the existing policies implemented in this thesis applying more precise policy according to individual subgroups and sub-geographies. Developing additional methods to identify why certain subgroups of the population see limited benefit from existing poverty is required to tailor existing policies. These heterogeneous policies could then be presented to Government bodies as simple extensions of existing strategy that see large cost-effectiveness gains.

Incorporation of further policies will require merging multiple datasets together in order to provide all sufficient variables. One example is the DUKES dataset from the Department of Energy Security and Net Zero (Department of Energy Security and Net Zero, United Kingdom Government, 2024a) providing individual household level information on detailed housing material quality including Energy Performance Certificate ratings and insulation type. Expanded application of simulated annealing will be required in order to merge datasets together (Zaidi and Rake, 2001; Klevmarken, 1997) and ensure this population is representative of the GMCA.

Additionally required datasets may not all be available on the same time and spatial scales. Merging of these datasets then requires incorporation of sensor fusion methods including data assimilation and probabilistic microsimulation (Birkin, 2021; Manoukian et al., 2022). These methods would then allow for estimation of behaviour on short and long time scales seen in pseudo-continuous time microsimulation (Spielauer et al., 2007) including acute disease progression (Archer et al., 2021) and unemployment (Harmon and Miller, 2018) as well as much smaller spatial scales combining sparse neighbourhood information with wider scale Local Authority and national level policy (Burgard et al., 2021; Leknes and Løkken, 2021). This would then allow for efficient simulation of transition processes particularly for Output Area level scales and smaller and may substantially benefit implementation of housing policy providing personalised support. Data assimilation and probabilistic methods also allow for application of an ensemble of microsimulation models, which is already extensively implemented in similar methods such as agent based modelling (Birkin, 2021; Arnold et al., 2019), that can be more computationally efficient using fewer microsimulation model runs as well as quantifying uncertainty due to input data from multiple imputation and simulated annealing methods used to create the synthetic MINOS input population.

Optimisation methods can be also be integrated with dynamic microsimulation in order to provide a data driven approach to improve allocation of public resources (Ballas, Kingston and Stillwell, 2005). Varying policy input parameters down to the household level can be combined with optimisation algorithms to explore specific policies identifying how a fixed amount of money could yield the largest health gain as seen in 'knapsack' problems (Ren et al., 2013).

There are currently limited interactions between individual units within the MINOS model. While there are some household interactions such as sharing housing income and energy expenditure there is significant scope to expand interactions between households. Further demographic processes including internal and external migration (Birkin, Wu and Rees, 2017) as well as long term household formation including cohabitation, marriage, and divorce (Cumpston et al., 2010). These characteristics have been included in multiple existing microsimulation models creating new households using matching methods

(Cumpston et al., 2010) that can be applied to the MINOS model as well to better estimate long term household structure. Additionally there is scope to incorporate variable households interactions with policy interventions (Simshauser, 2023). Policies applied in this thesis have assumed 100% uptake such that all eligible individuals automatically opt-in to any policy and its associated costs and benefits. In reality, logistical issues including lack of awareness due to not having an internet connection and individuals who are outright not interested limit overall uptake. Estimating uptake probability using methods such as propensity score matching (de Vocht et al., 2016) could be very useful in the MINOS model particularly to investigate areas with low participation. Further social behaviour can also be included exploring how individuals respond to other external factors such as market elasticity and energy-saving behaviours (Zapata-Webborn et al., 2024) under high fuel prices exploring how this affects income, lifestyle, and health.

There is also scope to incorporate further interactions between households at the neighbourhood level as part of a general inclusion of spatial effects. This can include spatially oriented policies including variance in how Local Authorities (Bridgen and Robinson, 2023) allocate resources as well as policy interventions such as large communal heat pumps designed to provide shared heating for multiple households (Brown et al., 2024). This can also include more general geographic influences such as how urban or rural an area is providing different employ and transport accessibility or social weather such that meteorological effects can influence well-being (Ballas, 2020). Generally the MINOS model includes limited interactions at this scale similar to other dynamic microsimulation modelling but this approach could be rectified by implementing hybrid ABM-microsimulation methods (Birkin, 2021; Arnold et al., 2019; Richiardi, van de Ven and Bronka, 2023). One immediate potential approach to do this is again using matching or decision support methods providing small groups of households highly personalised support at small areas (de Vocht et al., 2016). Finally policies can be influenced by large scale national and international factors. Further long term nationwide policy changes may occur particularly as electricity becomes cheaper than natural gas. Abolition of existing policy such as the 'clean spark spread', where the price of electricity is directly tied to the price of gas (El Amri et al., 2021), and incorporation of potential future technologies are all feasible within the MINOS model including change in individual energy companies and households. Factoring in new geo-political issues such as further conflict (Simshauser, 2023) and wider energy markets may be beneficial to reduce vulnerability to fossil fuels.

The MINOS model also provides a suitable testing framework for additional testing of best practice in dynamic microsimulation. Further case studies on development of transition probabilities in a full microsimulation allows for exploration within other methodological areas including interaction with other transition dynamics modules and sensitivity analyses exploring how microsimulation models react under new model structure and scenarios.

As multiple research groups are now applying microsimulation to income and housing policies and health outcomes the MINOS model has potential for application in meta-analyses comparing multiple microsimulation outputs. While some comparison has been made in this thesis to existing macro-level simulation studies this further comparison of full dynamic microsimulation model structure will substantially improve the robustness of model results.

Finally, there is scope to improve the scalability and computational speed of the MINOS model. Migration away from the vivarium framework into an R./c++ framework such as OpenMPP (Richiardi, van de Ven and Bronka, 2023) or Julia (Xue et al., 2024) may substantially speed up run time and increase population size facilitating application to new areas and optimisation methods mentioned above.

7.4.2 Applied Future Work

MINOS provides a flexible framework on which to apply more realistic and complex combinations of policy that change over time beyond income and housing support. Income support policies are largely expensive and intervention in other areas particularly further housing and neighbourhood characteristics are desirable for policy partners. Similarly recording of other health outcomes including more modern measures and mental and physical well-being such as Warwick-Edinburgh Mental Well-Being score and specific illnesses such as respiratory disease as they become available in administrative data may provide further interpretable metrics for model users. Application of more realistic policies including combinations of different policies together, policies that vary over time and countries within the UK, policies that adapt to external geopolitics, and a broad range of potentially absurd ideas can all be included in the MINOS model subject to data availability.

In order to facilitate research output beyond the scope of this thesis additional accessibility information is being developed. Full online technical documentation and user tutorials are being provided allowing for continued application both within the SIPHER consortium and potentially in external government bodies. This documentation would then enable discussion with Governments in an iterative process allowing for feedback on existing gaps in policy strategy as part of a complex systems modelling methodology fine tuning based on fiscally and political feasibility of more precise policy and further combination of retrofitting.

Finally, the MINOS model can be used to explore further health outcomes beyond health economics measures particularly including incidence of other non-communicable disease. Exploration of disease such as asthma and obesity may be used combined with other simulated evidence providing more information on policy effect. Application of health economics dynamic microsimulation research is increasing exponentially with new work in Government bodies providing invaluable feedback for expansion to new areas of research. This thesis contributes to this methodology and a promising future applying complex systems modelling to provide personalised dynamic policy within the United Kingdom optimising costs and ultimately protecting public health in the event of future crises. Even solely within the scope of energy poverty policy, electricity prices in the UK are now the highest in the world (Oliver, Matt, 2024) causing widespread further energy poverty where the MINOS model and complex systems modelling can make an immediate impact across all levels of UK government preventing non-communicable disease while facilitating a transition to renewable.

Bibliography

Ackley, S. F., Lessler, J. and Glymour, M. M. (2022). Dynamical modeling as a tool for inferring causation *American journal of epidemiology* **191**(1), 1–6.

Adhikari, S., Lecci, F., Becker, J. T., Junker, B. W., Kuller, L. H., Lopez, O. L. and Tibshirani, R. J. (2019). High-dimensional longitudinal classification with the multinomial fused lasso *Statistics in medicine* **38**(12), 2184–2205.

Aki Tsuchiya, Guoqiang Wu (2021). Sipher-7: a seven-indicator outcome measure to capture wellbeing for economic evaluation https://sipher.ac.uk/wp-content/uploads/2021/10/Sipher-7-report.pdf. Accessed: 2023-06-06.

Alarid-Escudero, F., Gulati, R. and Rutter, C. M. (2020). Validation of microsimulation models used for population health policy *Complex Systems and Population Health* **227**.

Ali, M. A., Alam, K., Taylor, B. and Rafiq, S. (2019). Do income distribution and socio-economic inequality affect ict affordability? evidence from Australian household panel data *Economic Analysis and Policy* **64**, 317–328.

Allendes, V., Carrasco, J. A., Miller, E. J. and Vaughan, J. (2021). Temporal transferability of a microsimulation activity-based model: An application in Chile *SOCHITRAN* .

Alsefri, M., Sudell, M., García-Fiñana, M. and Kolamunnage-Dona, R. (2020). Bayesian joint modelling of longitudinal and time to event data: a methodological *BMC Medical Research Methodology* .

Altinisik, Y. (2023). Addressing overdispersion and zero-inflation for clustered count data via new multilevel heterogenous hurdle models *Journal of Applied Statistics* **50**(2), 408–433.

Amenyah, I. D. and Fletcher, E. A. (2013). Factors determining residential rental prices *Asian Economic and Financial Review* **3**(1), 39.

Andreassen, L., Fredriksen, D., Gjefsen, H. M., Halvorsen, E. and Stølen, N. M. (2020). The dynamic cross-sectional microsimulation model MOSART *International Journal of Microsimulation* **13**(1), 92–113.

Anyaegbu, G. (2010). Using the OECD equivalence scale in taxes and benefits analysis *Economic & Labour Market Review* **4**, 49–54.

Appleyard, L., Packman, C., Lazell, J. and Aslam, H. (2023). The lived experience of financialization at the UK financial fringe *Journal of social policy* **52**(1), 24–45.

Archer, L., Lomax, N., Tysinger, B. et al. (2021). A dynamic microsimulation model for ageing and health in england: The english future elderly model *International Journal of Microsimulation* **14**(3), 2–26.

Ari, M. A., Arregui, M. N., Black, M. S., Celasun, O., Iakova, M. D. M., Mineshima, M. A., Mylonas, V., Parry, I. W., Teodoru, I. and Zhunussova, K. (2022). *Surging energy prices in Europe in the aftermath of the war: How to support the vulnerable and speed up the transition away from fossil fuels* International Monetary Fund.

Arnold, K. F., Harrison, W. J., Heppenstall, A. J. and Gilthorpe, M. S. (2019). Dag-informed regression modelling, agent-based modelling and microsimulation modelling: a critical comparison of methods for causal inference *International journal of epidemiology* **48**(1), 243–253.

Asher, M., Lomax, N., Morrissey, K., Spooner, F. and Malleson, N. (2023). Dynamic calibration with approximate Bayesian computation for a microsimulation of disease spread *Scientific Reports* **13**(1), 8637.

Augustin, J., Andrees, V., Walsh, D., Reintjes, R. and Koller, D. (2023). Spatial aspects of health—developing a conceptual framework *International Journal of Environmental Research and Public Health* **20**(3), 1817.

Axtell, R. L. and Farmer, J. D. (2022). Agent-based modeling in economics and finance: Past, present, and future *Journal of Economic Literature* .

Bae, J. W., Paik, E., Kim, K., Singh, K. and Sajjad, M. (2016). Combining microsimulation and agent-based model for micro-level population dynamics *Procedia Computer Science* **80**, 507–517.

Ballas, D. (2020). The economic geography of happiness *in* K. Zimmermann (ed.)*Handbook of labor, human resources and population economics* Springer pp. 1–24.

Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B. and Rossiter, D. (2005). SimBritain: a spatial microsimulation approach to population dynamics *Population, Space and Place* **11**(1), 13–34.

Ballas, D., Clarke, G., Hynes, S., Lennon, J., Morrissey, K. and O'Donoghue, C. (2013). A review of microsimulation for policy analysis *Spatial microsimulation for rural policy analysis* pp. 35–54.

Ballas, D., Kingston, R. and Stillwell, J. (2005). Using a spatial microsimulation decision support system for policy scenario analysis *in* J. Van Leeuwen (ed.)*Recent advances in design and decision support systems in architecture and urban planning* Springer pp. 177–191.

Ballas, D., Rossiter, D., Thomas, B., Clarke, G. and Dorling, D. (2005). Geography matters: Simulating the local impacts of national social policies *Geography matters: Simulating the local impacts of national social policies* .

Ballesteros-Arjona, V., Oliveras, L., Munoz, J. B., de Labry Lima, A. O., Carrere, J., Ruiz, E. M., Peralta, A., León, A. C., Rodríguez, I. M., Daponte-Codina, A. et al. (2022). What are the effects of energy poverty and interventions to ameliorate it on people's health and well-being?: A scoping review with an equity lens *Energy Research & Social Science* **87**.

Barbrook-Johnson, P. and Penn, A. S. (2022). *Systems Mapping: How to build and use causal models of systems* Springer Nature.

Bates, D. (2014). Computational methods for mixed models *LME4: Mixed-effects modeling with R* pp. 99–118.

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2014). Fitting linear mixed-effects models using lme4 *arXiv preprint arXiv:1406.5823* .

Baum, F. E., Ziersch, A. M., Zhang, G. and Osborne, K. (2009). Do perceived neighbourhood cohesion and safety contribute to neighbourhood differences in health? *Health & place* **15**(4), 925–934.

Benzeval, M., Bollinger, C. R., Burton, J., Crossley, T. F., Lynn, P. et al. (2020). The representativeness of understanding society *Institute for Social and Economic Research* .

Best, N., Stone, D., Ellis, H. and Davis, C. (2023). Spatial planning for climate resilience & net zero—barriers & opportunities for delivering net zero and climate resilience through the local planning system—a report for the climate change committee *Centre for Sustainable Energy & Town & Country Planning Association: Bristol, UK* .

Bhattacharjee, A., Pabst, A., Szendrei, T. and Hewings, G. (2023). Nirems: A regional model at household level combining spatial econometrics with dynamic microsimulation *Technical report* National Institute of Economic and Social Research.

Bibby, P. and Brindley, P. (2013). The 2011 rural-urban classification for small area geographies: a user guide and frequently asked questions (v1. 0) *Office for National Statistics* .

Birkin, M. (2021). Microsimulation *Urban Informatics* Springer pp. 845–864.

Birkin, M., Morris, M., Birkin, T. and Lovelace, R. (2017). Using census data in microsimulation modelling *in* J. Stillwell (ed.) *The Routledge Handbook of Census Resources, Methods and Applications* Routledge pp. 245–260.

Birkin, M. and Wu, B. (2012). A review of microsimulation and hybrid agent-based approaches *Agent-based models of geographical systems* pp. 51–68.

Birkin, M., Wu, B. and Rees, P. (2017). Moses: dynamic spatial microsimulation with demographic interactions *New frontiers in microsimulation modelling* pp. 53–77.

Bolton, P. (2023). Gas and electricity prices under the energy price guarantee and beyond *Technical Report CBP-9714* UK Government Department for Energy Security and Net Zero. Accessed:2024-09-01.

Booth, R. (2023). Pathways, targets and temporalities: Analysing English agriculture's net zero futures *Environment and Planning E: Nature and Space* **6**(1), 617–637.

Boumans, M. J. (2009). Grey-box understanding in economics *Scientific standing: philosophical perspectives, HW de Regt, S. Leonelli, and K. Eigner, eds., University of Pittsburgh Press, Forthcoming* .

Bouzarovski, S. and Simcock, N. (2017). Spatializing energy justice *Energy Policy* **107**, 640–648.

Bradshaw, J. R. and Keung, A. (2022). Fuel poverty estimates for the UK *Poverty* .

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models *Journal of the American statistical Association* **88**(421), 9–25.

Bridgen, P. and Robinson, C. (2023). A decade of fuel poverty in England: A spatio-temporal analysis of needs-based targeting of domestic energy efficiency obligations *Energy Research & Social Science* **101**, 103139.

Broadbent, P., Thomson, R., Kopasker, D., McCartney, G., Meier, P., Richiardi, M., McKee, M. and Katikireddi, S. V. (2023). The public health implications of the cost-of-living crisis: outlining mechanisms and modelling consequences *The Lancet Regional Health–Europe* **27**.

Brown, C., Hampton, S. and Fawcett, T. (2024). Accelerating renewable heat: Overcoming barriers to shared-loop ground source heat pump systems in the United Kingdom *Energy Research & Social Science* **115**, 103644.

Brownson, R. C., Chriqui, J. F. and Stamatakis, K. A. (2009). Understanding evidence-based public health policy *American journal of public health* **99**(9), 1576–1583.

Bu, F., Steptoe, A. and Fancourt, D. (2020). Who is lonely in lockdown? cross-cohort analyses of predictors of loneliness before and during the covid-19 pandemic *Public health* **186**, 31–34.

Burdine, J. N., Felix, M., Abel, A. L., Wiltraut, C. and Musselman, Y. (2000). The sf-12 as a population health measure: an exploratory examination of potential for application. *Health services research* **35**(4), 885.

Burgard, J. P., Dieckmann, H., Krause, J., Merkle, H., Münnich, R., Neufang, K. M. and Schmaus, S. (2020). A generic business process model for conducting microsimulation studies *Statistics in Transition New Series* **21**(4), 191–211.

Burgard, J. P., Krause, J., Merkle, H., Münnich, R. and Schmaus, S. (2019). Conducting a dynamic microsimulation for care research: Data generation, transition probabilities and sensitivity analysis *Stochastic Models, Statistics and Their Applications: Dresden, Germany, March 2019 14* Springer pp. 269–290.

Burgard, J. P., Krause, J. and Schmaus, S. (2021). Estimation of regional transition probabilities for spatial dynamic microsimulations from survey data lacking in regional detail *Computational Statistics & Data Analysis* **154**, 107048.

Burgard, J. P. and Schmaus, S. (2019). Sensitivity analysis for dynamic microsimulation models *Technical report* Research Papers in Economics.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., Bradley, K. A., (ACQUIP, A. C. Q. I. P. et al. (1998). The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking *Archives of internal medicine* **158**(16), 1789–1795.

Cairney, P. (2016). *The politics of evidence-based policy making* Springer.

Cairney, P. (2019). The uk government's imaginative use of evidence to make policy *British Politics* **14**(1), 1–22.

Cairney, P. (2021). The uk government's covid-19 policy: assessing evidence-informed policy analysis in real time *British Politics* **16**(1), 90–116.

Cairney, P. and Oliver, K. (2017). Evidence-based policymaking is not like evidence-based medicine, so how far should you go to bridge the divide between evidence and policy? *Health research policy and systems* **15**, 1–11.

Cairney, P., St Denny, E. and Mitchell, H. (2021). The future of public health policymaking after covid-19: A qualitative systematic review of lessons from health in all policies *Open Research Europe* **1**(23), 23.

Callan, M. J., Kim, H. and Matthews, W. J. (2015). Predicting self-rated mental and physical health: The contributions of subjective socioeconomic status and personal relative deprivation *Frontiers in psychology* **6**, 1415.

Campbell, F., Wong, R., Llewellyen, J., Chambers, D. and Stoniute, A. (2023). P27 evidence gap maps: A policy makers tool for navigating the evidence landscape. a case study of employment and health from the SIPHER (systems science in public health and health economics research) project.

Catapult Energy Systems (2024). Greater Manchester Spatial Energy Plan: Evidence Base Study https://www.greatermanchester-ca.gov.uk/media/1363/spatial_energy_plan_exec_summary.pdf. Accessed: 2024-09-22.

Chai, A., Ratnasiri, S. and Wagner, L. (2021). The impact of rising energy prices on energy poverty in Queensland: A microsimulation exercise *Economic Analysis and Policy* **71**, 57–72.

Chandola, T., Head, J. and Bartley, M. (2004). Socio-demographic predictors of quitting smoking: how important are household factors? *Addiction* **99**(6), 770–777.

Chandola, T. and Jenkinson, C. (2000). The new UK national statistics socio-economic classification (NS-SEC); investigating social class differences in self-reported health status *Journal of Public Health* **22**(2), 182–190.

Charlton, W. and Sana, B. (2023). Simwrapper, an open source web-based platform for interactive visualization of microsimulation outputs and transport data *Procedia Computer Science* **220**, 724–729.

Chatterton, T. J., Anable, J., Barnes, J. and Yeboah, G. (2016). Mapping household direct energy consumption in the United Kingdom to provide a new perspective on energy justice *Energy Research & Social Science* **18**, 71–87.

Chen, A., Machiorlatti, M., Krebs, N. M. and Muscat, J. E. (2019). Socioeconomic differences in nicotine exposure and dependence in adult daily smokers *BMC Public Health* **19**, 1–10.

Chen, T., Reed, H., Parra-Mujica, F., Johnson, E. A., Johnson, M., O'Flaherty, M., Collins, B. and Kypridemos, C. (2023). Quantifying the mental health and economic impacts of prospective universal basic income schemes among young people in the UK: a microsimulation modelling study *BMJ open* **13**(10), e075831.

Chen, Y.-C. (2017). A tutorial on kernel density estimation and recent advances *Biostatistics & Epidemiology* **1**(1), 161–187.

Child Poverty Action Group (2024). Scottish child payment: The gaps https://cpag.org.uk/sites/default/files/2024-07/CPAG%20Strengthening%20Social%20Security%20Scotland%20-%20briefing.pdf. Accessed: 2024-09-04.

Choi, J., Lee, M., Lee, J.-k., Kang, D. and Choi, J.-Y. (2017). Correlates associated with participation in physical activity among adults: a systematic review of reviews and update *BMC public health* **17**, 1–13.

Christensen, R. H. B. (2018). Cumulative link models for ordinal regression with the r package ordinal *Submitted in J. Stat. Software* **35**.

Clay, R., Archer, L., Heppenstall, A. and Lomax, N. (2023). Using the dynamic microsimulation MINOS to evidence the effect of energy crisis income support policy (short paper) *12th International Conference on Geographic Information Science (GIScience 2023)* Schloss Dagstuhl-Leibniz-Zentrum für Informatik ,.

Cockburn, J., Savard, L. and Tiberti, L. (2014). Macro-micro models *in* C. O'Donoghue (ed.)*Handbook of microsimulation modelling* Emerald Group Publishing Limited pp. 275–304.

Collard, S., Collings, D. and Cross, K. (2021). The financial impact of the pandemic.

Cominetti, N. and Murphy, L. (2022). Calculating the real living wage for London and the rest of the UK *London: Resolution Foundation* .

Committee of the Faculty of Public Health in Scotland (2017). Healthier lives: fairer futures. a call to action https://www.fph.org.uk/media/1168/healthy-lives-fairer-futures-final.pdf. Accessed: 2023-04-16.

Congreve, E., Connolly, K., Harrison, J., Kumar, A., McGregor, P. G. and Mitchell, M. (2022). The impact of using an income supplement to meet child poverty targets: evidence from Scotland *Journal of Social Policy* pp. 1–17.

Consortium, S. (2025). SIPHER Inclusive Economies Systems Mapping. https://www.gla.ac.uk/research/az/sipher/systemsmapping/inclusiveeconomies-systemsmap/#collaborativecreationwithsipherpolicypartners. Accessed: 2025-01-03.

Consortium, S. R. (2023). Generated using the SIPHER layered systems map: Energy efficiency and health / income and mental health. https://www.gla.ac.uk/research/az/sipher/systemsmapping/inclusiveeconomies-systemsmap/#policymicrosimulation. Accessed: 2024-08-24.

Craig, P., Barr, B., Baxter, A. J., Brown, H., Cheetham, M., Gibson, M., Katikireddi, S. V., Moffatt, S., Morris, S., Munford, L. A. et al. (2022). Evaluation of the mental health impacts of universal credit: protocol for a mixed methods study *BMJ open* **12**(4), e061340.

Creedy, J., Kalb, G. and Kew, H. (2007). Confidence intervals for policy reforms in behavioural tax microsimulation modelling *Bulletin of Economic Research* **59**(1), 37–65.

Crooks, A., Heppenstall, A., Manley, E. and Malleson, N. (2018). Agent-based modelling and geographical information systems: a practical primer.

Cumpston, J. R. et al. (2010). Alignment and matching in multi-purpose household microsimulations *International Journal of Microsimulation* **3**(2), 34–45.

Cumpston, J. R. et al. (2011). New techniques for household microsimulation, and their application to Australia *Open Research Repository Australia* .

Cushion, S., Soo, N., Kyriakidou, M. and Morani, M. (2020). Different lockdown rules in the four nations are confusing the public *LSE Covid-19 Blog* .

Davillas, A., Burlinson, A. and Liu, H.-H. (2022). Getting warmer: fuel poverty, objective and subjective health and well-being *Energy Economics* **106**, 105794.

Davis, P. and Lay-Yee, R. (2019). Simulating societal change *Counterfactual Modelling for Social and Policy Inquiry. Cham: Springer* .

De Leeuw, E. (2017). Engagement of sectors other than health in integrated health governance, policy, and action *Annual review of public health* **38**, 329–349.

De Menten, G., Dekkers, G., Bryon, G., Liégeois, P. and O'Donoghue, C. (2014). Liam2: a new open source development tool for discrete-time dynamic microsimulation models *Journal of Artificial Societies and Social Simulation* **17**(3), 9.

de Oliveira, C., Matias, M. A. and Jacobs, R. (2024). Microsimulation models on mental health: a critical review of the literature *Value in Health* **27**(2), 226–246.

de Vocht, F., Campbell, R., Brennan, A., Mooney, J., Angus, C. and Hickman, M. (2016). Propensity score matching for selection of local areas as controls for evaluation of effects of alcohol policies in case series and quasi-case-control designs *Public Health* **132**, 40–49.

Dekkers, G. (2015). The simulation properties of microsimulation models with static and dynamic ageing–a brief guide into choosing one type of model over the other *International Journal of Microsimulation* **8**(1), 97–109.

Dekkers, G. and Cumpston, R. (2012). On weights in dynamic-ageing microsimulation models *The International Journal of Microsimulation* **5**(2), 59–65.

Dekkers, G. and Van Den Bosch, K. (2016). Prospective microsimulation of pensions in European member states *Applications of microsimulation modelling* pp. 13–33.

Department for Business, Energy and Industrial Strategy (BEIS) London (2013). National energy efficiency data-framework (NEED).

Department of Energy Security and Net Zero. (2024). International domestic energy prices *Technical report* UK Government. Accessed: 2024-08-02.

Department of Energy Security and Net Zero, United Kingdom Government (2024a). Digest of UK Energy Statistics (DUKES) https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes. Accessed: 2024-09-22.

Department of Energy Security and Net Zero, United Kingdom Government (2024b). March 2024 Subnational Consumption Statistics Methodology and guidance booklet https://assets.publishing.service.gov.uk/media/6604524991a320001a82b125/subnational-methodology-and-guidance-booklet-2024.pdf. Accessed: 2024-09-22.

DeYoreo, M., Rutter, C. M., Ozik, J. and Collier, N. (2022). Sequentially calibrating a Bayesian microsimulation model to incorporate new information and assumptions *BMC medical informatics and decision making* **22**(1), 1–14.

Dhamo, E. and Puka, L. (2010). Using the R.-package to forecast time series: ARIMA models and application *International Conference Economic & Social Challenges and Problems* CRAN pp. 1–14.

Dias, P., Oliveira, A. and Lopes, C. (2011). Social and behavioural determinants of alcohol consumption *Annals of human biology* **38**(3), 337–344.

Dissing, A. S., Gil, A., Keenan, K., McCambridge, J., McKee, M., Oralov, A., Saburova, L. and Leon, D. A. (2013). Alcohol consumption and self-reported (sf 12) physical and mental health among working-aged men in a typical Russian city: a cross-sectional study *Addiction* **108**(11), 1905–1914.

Downward, P., Rasciute, S. and Kumar, H. (2020). Health, subjective financial situation and well-being: a longitudinal observational study *Health and Quality of Life Outcomes* **18**, 1–9.

ECIU Comment of Labours Revised Green Prospeity Plan (n.d.). https://eciu.net/media/press-releases/2024/comment-on-labours-revised-green-prosperity-plan. Accessed: 2024-08-24.

Eichhorst, W., Grienberger-Zingerle, M. and Konle-Seidl, R. (2010). Activating labor market and social policies in Germany: From status protection to basic income support. *German Policy Studies/Politikfeldanalyse* **6**(1).

El Amri, A., Oulfarsi, S., Boutti, R., Sahib Eddine, A. and Hmioui, A. (2021). Carbon financial markets underlying climate change mitigation, pricing and challenges: Technical analysis.

Emden, J. and Murphy, L. (2023). Greengo: Unlocking an energy efficiency and clean heat revolution *Institute for Public Policy Research* .

Essa, M. and Sayed, T. (2015). Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts *Accident Analysis & Prevention* **84**, 41–53.

Essien, A. E., Petrounias, I., Sampaio, P. and Sampaio, S. (2019). Deep-PRESIMM: Integrating deep learning with microsimulation for traffic prediction *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* IEEE pp. 4257–4262.

Eurostat-Glossary (n.d.). Equivalised disposable income https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Equivalised_disposable_income. Accessed: 2025-02-02.

Evans, S. (2022). Analysis: Why UK energy bills are soaring to record highs—and how to cut them *Carbon Brief* **12**.

Fan, L. and Babiarz, P. (2019). The determinants of subjective financial satisfaction and the moderating roles of gender and marital status *Family and Consumer Sciences Research Journal* **47**(3), 237–259.

for Social Justice, T. C. (2023). On target: Protecting vulnerable households from the inflation crisis *Centre for Social Justice, UK* .

French, S. A., Tangney, C. C., Crane, M. M., Wang, Y. and Appelhans, B. M. (2019). Nutrition quality of food purchases varies by household income: the SHoPPER study *BMC public health* **19**, 1–7.

Frick, J. R. and Grabka*, M. M. (2005). Item nonresponse on income questions in panel surveys: Incidence, imputation and the impact on inequality and mobility *Allgemeines Statistisches Archiv* **89**, 49–61.

Frick, J. R., Grabka, M. M. and Groh-Samberg, O. (2009). Dealing with incomplete household panel data in microsimulation models *in* G. Dekkers (ed.)*2nd General Conference of the International Microsimulation Association (IMA): Microsimulation: Bridging Data and Policy. Ottawa, Canada (www. diw. de/documents/dokumentenarchiv/17/diw_01. c. 334116. de/soep_punr_2009. pdf)* International Microsimulation Association ,.

Fumagalli, L., Knies, G. and Buck, N. (2017). Understanding society, the uk household longitudinal study, harmonised british household panel survey (bhps) user guide.

Gabrielli, P., Wüthrich, M., Blume, S. and Sansavini, G. (2022). Data-driven modeling for long-term electricity price forecasting *Energy* **244**, 123107.

Gafni, A. and Birch, S. (2006). Incremental cost-effectiveness ratios (icers): the silence of the lambda *Social science & medicine* **62**(9), 2091–2100.

Gale, C. G., Singleton, A., Bates, A. G. and Longley, P. A. (2016). Creating the 2011 area classification for output areas (2011 oac) *Journal of Spatial Information Science* **12**, 1–27.

Gelman, A. (2005). Analysis of variance—why it is more important than ever.

Gilbert, N. and Troitzsch, K. (2005). *Simulation for the social scientist* McGraw-Hill Education (UK).

Giles, C. and Ridge, M. (1993). The impact on households of the 1993 budget and the council tax *Fiscal Studies* **14**(3), 1–20.

Goedemé, T., Van den Bosch, K., Salanauskaite, L., Verbist, G. et al. (2013). Testing the statistical significance of microsimulation results: A plea *International Journal of Microsimulation* **6**(3), 50–77.

Goodman, A., Guell, C., Panter, J., Jones, N. R. and Ogilvie, D. (2012). Healthy travel and the socio-economic structure of car commuting in Cambridge, UK: a mixed-methods analysis *Social science & medicine* **74**(12), 1929–1938.

Government Finance Function, HM Treasury, United Kingdom Government (2024). The green book (2022) https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2020. Accessed: 2024-09-22.

Graff, M., Carley, S., Konisky, D. M. and Memmott, T. (2021). Which households are energy insecure? an empirical analysis of race, housing conditions, and energy burdens in the United States *Energy Research & Social Science* **79**, 102144.

Gray, D. (2013). *Household finances, well-being and subjective financial situation: An empirical analysis of household survey data* PhD thesis University of Sheffield.

Greszczuk, C. (2019). *Implementing health in all policies* The Health Foundation.

Grey, C. N., Jiang, S., Nascimento, C., Rodgers, S. E., Johnson, R., Lyons, R. A. and Poortinga, W. (2017). The short-term health and psychosocial impacts of domestic energy efficiency investments in low-income areas: a controlled before and after study *BMC public health* **17**, 1–10.

Ham, C. (2021). The uk's poor record on covid-19 is a failure of policy learning *bmj* **372**.

Hancock, P. and Tyler, M. (2024). Live performers' experiences of precarity and recognition during covid-19 and beyond *Work, Employment and Society* p. 09500170231218677.

Hanum, H., Dewi, N. R. and Setyawan, D. (2018). Modeling indonesia inflation data: Gamma-normal or normal *Applied Mathematical Sciences* **12**(6), 271–278.

Haraldsson, H. (2004). *Introduction to system thinking and causal loop diagrams* Department of chemical engineering, Lund University Lund, Sweden.

Harding, A. (2007). Challenges and opportunities of dynamic microsimulation modelling *Plenary paper presented to the 1st General Conference of the International Microsimulation Association, Vienna* Vol. 21 Citeseer ,.

Harding, A. (2017). *New frontiers in microsimulation modelling* Routledge.

Harding, A., Keegan, M. and Kelly, S. (2010a). Validating a dynamic population microsimulation model: Recent experience in Australia *IJM* **3**, 46–64.

Harding, A., Keegan, M. and Kelly, S. (2010b). Validating a dynamic population microsimulation model: Recent experience in Australia *International Journal of Microsimulation* **3**(2), 46–64.

Harmon, A. and Miller, E. J. (2018). Overview of a labour market microsimulation model *Procedia computer science* **130**, 172–179.

Harrell, F. E. et al. (2001). *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis* Vol. 608 Springer.

Harrell Jr, F. E. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis* springer.

Harrison, J., Kumar, A., Congreve, E., Connolly, K. and MacGregor, P. (2022). Modelling packages to meet Scotland's child poverty targets: Scenarios, benefits and trade-offs.

Hartnett, G. S., Vardavas, R., Baker, L., Chaykowsky, M., Gibson, C. B., Girosi, F., Kennedy, D. P. and Osoba, O. A. (2020). Deep generative modeling in network science with applications to public policy research *arXiv preprint arXiv:2010.07870* .

Haslett, S., Jones, G., Noble, A. and Ballas, D. (2010). More for less? comparing small area estimation, spatial microsimulation, and mass imputation *JSM* pp. 1584–1598.

Hill O'Connor, C., Smith, K. and Stewart, E. (2023). Integrating evidence and public engagement in policy work: an empirical examination of three UK policy organisations *Policy & Politics* pp. 1–24.

Hincks, S. (2017). Deprived neighbourhoods in transition: Divergent pathways of change in the Greater Manchester city-region *Urban Studies* **54**(4), 1038–1061.

Hjelmskog, A. (2022). How to combine action on housing retrofit with tackling health inequalities (and other injustices) *UK Collaborative Centre for Housing Evidence* .

Hodgkin, R. and Sasse, T. (2022). Tackling the UK's energy efficiency problem.

Hoehn, A., Lomax, N., Rice, H., Angus, C., Brennan, A., Brown, D., Comrie, E., Elsenbroich, C., Hughes, C. and Katikireddi, S. V. (2023). Op36 quality-adjusted life expectancy (QALE) and its association with economic inclusion: a study of local authorities in England, Scotland, and Wales.

Hoehn, A., Lomax, N., Zia, K., Cox, J., Comrie, E., Fergie, G., Heppenstall, A., Lewis, D., Purshouse, R. and Winterbottom, J. (2024). Op78 SIPHER's synthetic population for individuals in Great Britain 2019-2021: creation, validation, and examples of application.

Höhn, A., Lomax, N., Rice, H., Angus, C., Brennan, A., Brown, D., Cunningham, A., Elsenbroich, C., Hughes, C., Katikireddi, S. V. et al. (2024). Estimating quality-adjusted life expectancy (QALE) for local authorities in Great Britain and its association with indicators of the inclusive economy: a cross-sectional study *BMJ open* **14**(3), e076704.

Höhn, A., Stokes, J., Pollack, R., Boyd, J., Del Cerro, C. C., Elsenbroich, C., Heppenstall, A., Hjelmskog, A., Inyang, E., Kopasker, D. et al. (2023). Systems science methods in public health: what can they contribute to our understanding of and response to the cost-of-living crisis? *J Epidemiol Community Health* .

Hughes, A. and Kumari, M. (2017). Unemployment, underweight, and obesity: Findings from Understanding Society (UKHLS) *Preventive medicine* **97**, 19–25.

Huque, M. H., Carlin, J. B., Simpson, J. A. and Lee, K. J. (2018). A comparison of multiple imputation methods for missing data in longitudinal studies *BMC medical research methodology* **18**(1), 1–16.

Inequalities in London (2023). The rising cost of living.

Institute for Health Metrics and Evaluation, University of Washington, Seattle, USA (2023). Vivarium dynamic microsimulation framework https://github.com/ihmeuw/vivarium. Accessed: 2023-05-22.

Ioannidis, J. P., Cripps, S. and Tanner, M. A. (2022). Forecasting for covid-19 has failed *International journal of forecasting* **38**(2), 423–438.

Iskandar, R. (2019). Adding noise to markov cohort state-transition models *bioRxiv* p. 635177.

Jackman, S., Tahk, A., Zeileis, A., Maimone, C., Fearon, J., Meers, Z., Jackman, M. S. and Imports, M. (2015). Package 'pscl' *Political Science Computational Laboratory* **18**(04.2017).

Jakobsen, J. C., Gluud, C., Wetterslev, J. and Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts *BMC medical research methodology* **17**(1), 1–10.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). *An introduction to statistical learning* Vol. 112 Springer.

Jan, F., Shah, I. and Ali, S. (2022). Short-term electricity prices forecasting using functional time series analysis *Energies* **15**(9), 3423.

Jann, B. (2014). Plotting regression coefficients and other estimates *The Stata Journal* **14**(4), 708–737.

Javed, Z., Valero-Elizondo, J., Maqsood, M. H., Mahajan, S., Taha, M. B., Patel, K. V., Sharma, G., Hagan, K., Blaha, M. J., Blankstein, R. et al. (2022). Social determinants of health and obesity: Findings from a national study of US adults *Obesity* **30**(2), 491–502.

Jensen, M. F. and Blundell, J. (2024). Income effects and labour supply: Evidence from a child benefits reform *Journal of Public Economics* **230**, 105049.

Johnson, E. A., Johnson, M. T., Kypridemos, C., Villadsen, A. and Pickett, K. E. (2023). Designing a generic, adaptive protocol resource for the measurement of health impact in cash transfer pilot and feasibility studies and trials in high-income countries *Pilot and Feasibility Studies* **9**(1), 51.

Johnson, P. (2015). UK consumer price statistics: A review *UK Statistics Authority* **8**.

Jones, P. M. (2017). *A Spatial Microsimulation Analysis of Health Inequalities and Health Resilience* PhD thesis University of Sheffield.

Jones, R. V. and Lomas, K. J. (2015). Determinants of high electrical energy demand in UK homes: Socio-economic and dwelling characteristics *Energy and Buildings* **101**, 24–34.

Josse, J., Mayer, I., Tierney, N. and Vialaneix, N. (2024). Cran task view: Missing data.

Katikireddi, S. V., Kopasker, D., Pearce, A., Leyland, A. H., Rostila, M. and Richiardi, M. (2022). Protocol: Health equity and its economic determinants (heed): protocol for a pan-European microsimulation model for health impacts of income and social security policies *BMJ Open* **12**(7).

Kelly, S. and Percival, R. (2009). Longitudinal benchmarking and alignment of a dynamic microsimulation model *in* G. Dekkers (ed.)*Second IMA General Conference* International Microsimulation Association ,.

Kerr, C. C., Stuart, R. M., Mistry, D., Abeysuriya, R. G., Rosenfeld, K., Hart, G. R., Núñez, R. C., Cohen, J. A., Selvaraj, P., Hagedorn, B. et al. (2021). Covasim: an agent-based model of covid-19 dynamics and interventions *PLOS Computational Biology* **17**(7), e1009149.

Khalil, M. A. and Fatmi, M. R. (2022). How residential energy consumption has changed due to covid-19 pandemic? an agent-based model *Sustainable Cities and Society* **81**, 103832.

Klevmarken, A. (2022). Modeling behavioural response in EUROMOD *International Journal of Microsimulation* **15**(1), 89–96.

Klevmarken, N. A. (1997). Behavioral modeling in micro simulation models. a survey.

Koblyakova, A., Hutchison, N. and Tiwari, P. (2014). Regional differences in mortgage demand and mortgage instrument choice in the UK *Regional Studies* **48**(9), 1499–1513.

Koehler, E., Brown, E. and Haneuse, S. J.-P. (2009). On the assessment of Monte Carlo error in simulation-based statistical analyses *The American Statistician* **63**(2), 155–162.

Kopasker, D., Bronka, P., Thomson, R. M., Khodygo, V., Kromydas, T., Meier, P., Heppenstall, A., Bambra, C., Lomax, N., Craig, P. et al. (2024). Evaluating the influence of taxation and social security policies on psychological distress: A microsimulation study of the UK during the covid-19 economic crisis *Social Science & Medicine* **351**, 116953.

Kopasker, D., Katikireddi, S. V., Santos, J. V., Richiardi, M., Bronka, P., Rostila, M., Cecchini, M., Ali, S., Emmert-Fees, K., Bambra, C. et al. (2023). Microsimulation as a flexible tool to evaluate policies and their impact on socioeconomic inequalities in health *The Lancet Regional Health–Europe* **34**.

Kopec, J. A., Finès, P., Manuel, D. G., Buckeridge, D. L., Flanagan, W. M., Oderkirk, J., Abrahamowicz, M., Harper, S., Sharif, B., Okhmatovskaia, A. et al. (2010). Validation of population-based disease simulation models: a review of concepts and methods *BMC public health* **10**, 1–13.

Kouser, H. N., Barnard-Mayers, R. and Murray, E. (2021). Complex systems models for causal inference in social epidemiology *J Epidemiol Community Health* **75**(7), 702–708.

Krijkamp, E. M., Alarid-Escudero, F., Enns, E. A., Jalal, H. J., Hunink, M. M. and Pechlivanoglou, P. (2018). Microsimulation modeling for health decision sciences using R.: a tutorial *Medical Decision Making* **38**(3), 400–422.

Kuhn, M. (2008). Building predictive models in r using the caret package *Journal of statistical software* **28**, 1–26.

Kühne, F., Schomaker, M., Stojkov, I., Jahn, B., Conrads-Frank, A., Siebert, S., Sroczynski, G., Puntscher, S., Schmid, D., Schnell-Inderst, P. et al. (2022). Causal evidence in health decision making: methodological approaches of causal inference and health decision science *GMS German Medical Science* **20**.

Kum, H. (2010). Imputation techniques in microsimulation *Journal of Policy Studies* **25**(3), 69–83.

Lang, B. M., Kuipers, J., Misselwitz, B. and Beerenwinkel, N. (2020). Predicting colorectal cancer risk from adenoma detection via a two-type branching process model *PLoS computational biology* **16**(2), e1007552.

Lappo, S., Lehtonen, R., Liski, A. and Finland, S. (2015). Uncertainty in microsimulation *Master's Thesis, University of Helsinki* .

Lawrence, W. F. and Fleishman, J. A. (2004). Predicting EuroQoL eq-5d preference scores from the sf-12 health survey in a nationally representative sample *Medical Decision Making* **24**(2), 160–169.

Lee, T. Y. H. (2021). Integrating julia and r (studio) for microsimulation modeling: challenges and rewards https://giabaio.github.io/r-hta/lee/#/. Accessed: 2024-09-30.

Leknes, S. and Løkken, S. A. (2021). Flexible empirical bayes estimation of local fertility schedules: Reducing small area problems and reserving regional variation *Technical report* Discussion Papers.

Lenihan, A. T. (2015). Institutionalising evidence-based policy: International insights into knowledge brokerage *Contemporary social science* **10**(2), 114–125.

Li, J. and O'Donoghue, C. (2012). Evaluating binary alignment methods in microsimulation models.

Li, J. and O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure amd methodology *International Journal of microsimulation* **6**(2), 3–55.

Li, J. and O'Donoghue, C. (2016). An overview of binary alignment methods in microsimulation *New Pathways in Microsimulation* pp. 233–248.

Li, J., O'Donoghue, C. and Dekkers, G. (2014). Dynamic models *in* C. O'Donoghue (ed.)*Handbook of microsimulation modelling* Vol. 293 Emerald Group Publishing Limited pp. 305–343.

Li, L., Shen, C., Li, X. and Robins, J. M. (2013). On weighting approaches for missing data *Statistical methods in medical research* **22**(1), 14–30.

Li, X., Arbabi, H., Bennett, G., Oreszczyn, T. and Tingley, D. D. (2022). Net zero by 2050: Investigating carbon-budget compliant retrofit measures for the English housing stock *Renewable and Sustainable Energy Reviews* **161**, 112384.

Liao, Y., Xiang, Y. and Wang, M. (2021). Health assessment and prognostics based on higher-order hidden semi-markov models *Naval Research Logistics (NRL)* **68**(2), 259–276.

Lidbe, A. D., Hainen, A. M. and Jones, S. L. (2017). Comparative study of simulated annealing, tabu search, and the genetic algorithm for calibration of the microsimulation model *Simulation* **93**(1), 21–33.

Livingston, M., Kearns, A. and Bannister, J. (2014). Neighbourhood structures and crime: The influence of tenure mix and other structural factors upon local crime rates *Housing Studies* **29**(1), 1–25.

Lomax, N., Clay, R., Archer, L., Rice, H. and Heppenstall, A. (2023). Briefing note: A dynamic model of disposable income impacts on mental health *OSF* .

Lomax, N., Hoehn, A., Heppenstall, A. J., Purshouse, R. C., Wu, G., Zia, K. and Meier, P. S. (2024). SIPHER Synthetic Population for Individuals in Great Britain, 2019-2021: Supplementary Material.

Lomax, N. and Norman, P. (2016). Estimating population attribute values in a table:"get me started in" iterative proportional fitting *The Professional Geographer* **68**(3), 451–461.

Lomax, N. and Smith, A. P. (2017a). An introduction to microsimulation for demography *Australian Population Studies* **1**(1), 73–85.

Lomax, N. and Smith, A. P. (2017b). Microsimulation for demography *Australian Population Studies* **1**(1), 73–85.

Lonnie, M. and Johnstone, A. (2023). The cost-of-living crisis is feeding the paradox of obesity and food insecurities in the UK *Obesity* .

Lovelace, R. (2014). Introducing spatial microsimulation with r: a practical.

Lovelace, R. and Dumont, M. (2017). *Spatial microsimulation with R* Chapman and Hall/CRC.

Lüdecke, D. and Lüdecke, M. D. (2015). Package 'sjplot' *R package version* **1**(9).

Lupton, R., Hughes, C. and Lomax, N. (2023). SIPHER inclusive economy indicator set.

Magness, P. W. (2021). The failure of Imperial College modeling is far worse than we knew *American Institute for Economic Research. April* **22**, 2021.

Mallet, L. and Weale, M. (2018). Nowcasting household median income: A comparison of microsimulation and time series approaches *35th IARIW General Conference, Copenhagen* IARIW ,.

Manning, W. (2012). Dealing with skewed data on costs and expenditures *The Elgar Companion to Health Economics, Second Edition* Edward Elgar Publishing ,.

Manoukian, S., Stewart, S., Dancer, S. J., Mason, H., Graves, N., Robertson, C., Leonard, A., Kennedy, S., Kavanagh, K., Parcell, B. et al. (2022). Probabilistic microsimulation to examine the cost-effectiveness of hospital admission screening strategies for carbapenemase-producing enterobacteriaceae (cpe) in the United Kingdom *The European Journal of Health Economics* **23**(7), 1173–1185.

Marois, G. and Aktas, A. (2021). Projecting health-ageing trajectories in Europe using a dynamic microsimulation model *Scientific reports* **11**(1), 1–15.

Marois, G., Potančoková, M. and González-Leonardo, M. (2023). Technical report: QuantMig-Mic microsimulation population projection model *European Union's Horizon* .

Mason, P., Curl, A. and Kearns, A. (2016). Domains and levels of physical activity are linked to adult mental health and wellbeing in deprived neighbourhoods: A cross-sectional study *Mental Health and Physical Activity* **11**, 19–28.

McLay, J., Lay Yee, R., Milne, B. and Davis, P. (2015). Regression-style models for parameter estimation in dynamic microsimulation: An empirical performance assessment.

Meier, P., Purshouse, R., Bain, M., Bambra, C., Bentall, R., Birkin, M., Brazier, J., Brennan, A., Bryan, M., Cox, J. et al. (2019). The sipher consortium: introducing the new uk hub for systems science in public health and health economic research *Wellcome open research* **4**.

Milne, A. (2020). A critical covid-19 economic policy tool: Retrospective insurance *Available at SSRN 3558667* .

Mirzaei, A., Carter, S. R., Patanwala, A. E. and Schneider, C. R. (2022). Missing data in surveys: Key concepts, approaches, and applications *Research in Social and Administrative Pharmacy* **18**(2), 2308–2316.

Morrison, R. (2008). Validation of longitudinal microsimulation models *DYNACAN Practices and Plans* p. 64.

Mosley, M. et al. (2022). Box c: Households savings amid the cost-of-living crisis *National Institute UK Economic Outlook* pp. 34–36.

Müller, K. and Axhausen, K. W. (2010). Population synthesis for microsimulation: State of the art *Arbeitsberichte Verkehrs-und Raumplanung* **638**.

Münnich, R., Schnell, R., Brenzel, H., Dieckmann, H., Dräger, S., Emmenegger, J., Höcker, P., Kopp, J., Merkle, H., Neufang, K. et al. (2021). A population based regional dynamic microsimulation of Germany: the MikroSim model *methods, data, analyses* **15**(2), 24.

Murphy, D. R., Klein, R. W., Smolen, L. J., Klein, T. M. and Roberts, S. D. (2013). Using common random numbers in health care cost-effectiveness simulation modeling *Health Services Research* **48**(4), 1508–1525.

Murray, E. J., Robins, J. M., Seage, G. R., Freedberg, K. A. and Hernán, M. A. (2017). A comparison of agent-based models and the parametric g-formula for causal inference *American journal of epidemiology* **186**(2), 131–142.

Nakagawa, S. (2015). Missing data: mechanisms, methods and messages *Ecological statistics: Contemporary theory and application* pp. 81–105.

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining r2 from generalized linear mixed-effects models *Methods in ecology and evolution* **4**(2), 133–142.

Nash, A. (2019). National population projections: 2018-based *Office for National Statistics* .

National Health Service (NHS) Confederation (2022). Could the energy crisis cause a public health emergency.

Neto-Bradley, A. P., Choudhary, R. and Challenor, P. (2023). A microsimulation of spatial inequality in energy access: A bayesian multi-level modelling approach for urban India *Environment and Planning B: Urban Analytics and City Science* **50**(4), 895–910.

Norman, A. and Corfe, S. (2022). Energy bill support–designing policies to support British households in an age of high prices *Social Market Foundation* .

Odiari, E. and Birkin, M. (2022). Simulating micro-level attributes of railway passengers using big data *Journal of Urban Mobility* **2**, 100027.

O'Donoghue, C., Ballas, D., Clarke, G., Hynes, S. and Morrissey, K. (2013). *Spatial microsimulation for rural policy analysis* Springer.

O'Donoghue, C. and Dekkers, G. (2018). Increasing the impact of dynamic microsimulation modelling *The International Journal of Microsimulation* **11**(1), 61–96.

O'Donoghue, C. and Loughrey, J. (2014). Nowcasting in microsimulation models: a methodological survey *Journal of Artificial Societies and Social Simulation* **17**(4), 12.

O'Donoghue, C., O'Donoghue and Pacey (2017). *Farm-Level Microsimulation Modelling* Springer.

O'Donoghue, C., Sologon, D. M., Kyzyma, I. and McHale, J. (2021). A microsimulation analysis of the distributional impact over the three waves of the covid-19 crisis in ireland *arXiv preprint arXiv:2103.08398* .

Office for Budget Responsibility (2023). The outlook for household income and consumption https://obr.uk/box/the-outlook-for-household-income-and-consumption/. Accessed: 2023-05-01.

Office for Budget Responsibility (2024). The outlook for household income and consumption https://obr.uk/box/the-outlook-for-household-income-and-consumption/. Accessed: 2024-08-22.

Office for National Statistics (2022). Average household income, UK: Financial year ending 2022 https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/householddisposableincomeandinequality/financialyearending2022. Accessed: 2023-08-02.

Office for National Statistics (2024). Lower Super Output Areas (LSOAs) (december 2021) names and codes in ew (v3) http://data.europa.eu/88u/dataset/lower-layer-super-output-areas-december-2021-names-and-codes-in-ew-v3. Accessed: 2024-09-22.

Ogden, K. and Thomas, M. (2024). Higher education spending.

O'Hagan, A., Stevenson, M. and Madan, J. (2007). Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA *Health economics* **16**(10), 1009–1023.

Oksanen, A., Aaltonen, M. and Rantala, K. (2015). Social determinants of debt problems in a Nordic welfare state: A Finnish register-based study *Journal of consumer policy* **38**, 229–246.

Oliver, Matt (2024). Britain paying highest electricity prices in the world https://www.telegraph.co.uk/business/2024/09/26/britain-burdened-most-expensive-electricity-prices-in-world/. Accessed: 2024-09-28.

Orcutt, G. H. (1961). Microanalysis of socioeconomic systems.

Panakaduwa, C., Coates, P. and Munir, M. (n.d.). Evaluation of government actions discouraging housing energy retrofit in the UK: A critical review.

Parham, P. and Hughes, D. (2013). The multimodel ensemble approach to reducing structural uncertainty in decision analytical modelling *Value in Health* **16**(7), A324–A325.

Parra-Mujica, F., Johnson, E., Reed, H., Cookson, R. and Johnson, M. (2023). Understanding the relationship between income and mental health among 16-to 24-year-olds: Analysis of 10 waves (2009–2020) of Understanding Society to enable modelling of income interventions *PLoS One* **18**(2), e0279845.

Pawson, R. (2021). The coronavirus response: Boxed in by models *Evaluation* **27**(2), 149–167.

Payne, R. A. and Abel, G. A. (2012). UK indices of multiple deprivation-a way to make comparisons across constituent countries easier *Health Stat Q* **53**(22), 2015–2016.

Peersman, G. and Wauters, J. (2024). Heterogeneous household responses to energy price shocks *Energy Economics* **132**, 107421.

Pemberton, S. and Humphris, P. (2016). Locality, neighbourhood and health: a literature review *Institute for ReFsearch into Superdiversity* .

Petrik, O., Adnan, M., Basak, K. and Ben-Akiva, M. (2020). Uncertainty analysis of an activity-based microsimulation model for Singapore *Future Generation Computer Systems* **110**, 350–363.

Phillips, D., Simpson, L. and Ahmed, S. (2008). Shifting geographies of minority ethnic settlement: remaking communities in Oldham and Rochdale *Community Cohesion in Crisis?* Policy Press pp. 81–98.

Pintelas, E., Livieris, I. E. and Pintelas, P. (2020). A grey-box ensemble model exploiting black-box accuracy and white-box intrinsic interpretability *Algorithms* **13**(1), 17.

Pirc, E., Pecchia, L., Reberšek, M., Serša, G., Snoj, M., Grošelj, A. and Miklavčič, D. (2018). Study design of a medical device pre-market assessment: a case study on electrochemotherapy *Slovenian Medical Journal* **87**(1-2), 22–40.

Pollack, C. E., Cubbin, C., Ahn, D. and Winkleby, M. (2005). Neighbourhood deprivation and alcohol consumption: does the availability of alcohol play a role? *International journal of epidemiology* **34**(4), 772–780.

Prantner, B. (2011). Visualization of imputed values using the R.-package VIM.

Prowse, P. and Fells, R. (2016). The living wage–policy and practice *Industrial Relations Journal* **47**(2), 144–162.

Purdam, K., Garratt, E. A. and Esmail, A. (2016). Hungry? food insecurity, social stigma and embarrassment in the uk *Sociology* **50**(6), 1072–1088.

Pybus, E. (2023). Priorities for the 2023-24 programme for government.

Qaseem, N., Kausar, R. and Malki, I. (2024). Dynamics of fuel poverty in the UK and the impact of the Energy Company Obligation (ECO) Affordable Warmth Scheme *Available at SSRN 4810917* .

Quarina, Q. (2017). *An analysis of the determinants and scarring effects of economic inactivity and unemployment in the UK* Lancaster University (United Kingdom).

R. Core Team (2021). Team R.: A language and environment for statistical computing, 2015.

Rafique, A. and Williams, A. P. (2021). Reducing household greenhouse gas emissions from space and water heating through low-carbon technology: Identifying cost-effective approaches *Energy and Buildings* **248**, 111162.

Raleigh, V. (2021). What is happening to life expectancy in England *The King's Fund* **585**.

Rees, P. H., Wohland, P., Norman, P., Lomax, N. and Clark, S. D. (2017). Population projections by ethnicity: Challenges and solutions for the United Kingdom *The frontiers of applied demography* pp. 383–408.

Reeves, D. C., Willems, N., Shastry, V. and Rai, V. (2022). Structural effects of agent heterogeneity in agent-based models: lessons from the social spread of covid-19 *Journal of Artificial Societies and Social Simulation* **25**(3).

Regan, N. O. (n.d.). Insulating britain.

Ren, Y., Ordóñez, F. and Wu, S. (2013). Optimal resource allocation response to a smallpox outbreak *Computers & Industrial Engineering* **66**(2), 325–337.

Richiardi, M., Bronka, P., van de Ven, J., Kopasker, D. and Katikireddi, S. V. (2023). SimPaths: an open-source microsimulation model for life course analysis *Available at SSRN 4808042* .

Richiardi, M., Collado, D. and Popova, D. (2021). UKMOD, a new tax-benefit model for the four nations of the UK *International Journal of Microsimulation, forthcoming* .

Richiardi, M. G., Poggi, A. et al. (2014). Imputing individual effects in dynamic microsimulation models. an application to household formation and labour market participation in Italy *The International Journal of Microsimulation* **7**(2), 3–39.

Richiardi, M. G. and Richardson, R. E. (2017). Jas-mine: A new platform for microsimulation and agent-based modelling *International Journal of Microsimulation* **10**(1), 106–134.

Richiardi, M. and Poggi, A. (2012). Imputing individual effects in dynamic microsimulation models. an application of the rank method *Carlo Alberto Notebooks* Vol. 267 Collegio Carlo Alberto ,.

Richiardi, M., van de Ven, J. and Bronka, P. (2023). Back to the future: Agent-based modelling and dynamic microsimulation *Technical report* CeMPA Working Paper Series.

Ripley, B. and Venables, W. (2016). nnet: Feed-forward neural networks and multinomial log-linear models *R package version* **7**(2016).

Rizopoulos, D. (2022). Package 'glmmadaptive'.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* Vol. 81 John Wiley & Sons.

Ruse, J.-L., Stockton, H. and Smith, P. (2019). *Social and health-related indicators of energy poverty: an England case study* Elsevier.

Rutter, C. M., Miglioretti, D. L. and Savarino, J. E. (2009). Bayesian calibration of microsimulation models *Journal of the American Statistical Association* **104**(488), 1338–1350.

Rutter, C. M., Zaslavsky, A. M. and Feuer, E. J. (2011). Dynamic microsimulation models for health outcomes: a review *Medical Decision Making* **31**(1), 10–18.

Salonen, J., Möttönen, J., Tikanmäki, H. and Nummi, T. (2020). Using sequence analysis to visualize and validate model transitions *IJM* .

Salonen, J., Tikanmäki, H. and Nummi, T. (2019). Using trajectory analysis to test and illustrate microsimulation outcomes.

Saltelli, A. and Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? *Futures* **91**, 62–71.

Samarin, M., Sharma, M. and Nagle, N. N. (2024). Geographies and determinants of rent burden: A regional economic analysis of the US metropolitan landscape *Regional Science Policy & Practice* p. 100072.

Sasse, T. and Thomas, A. (2022). Better policy making.

Sawyer, A., Sherriff, N., Bishop, D., Darking, M. and Huber, J. W. (2022). "it's changed my life not to have the continual worry of being warm"–health and wellbeing impacts of a local fuel poverty programme: a mixed-methods evaluation *BMC Public Health* **22**(1), 786.

Schreuder, A., Mets, O. M., Schaefer-Prokop, C. M., Jacobs, C. and Prokop, M. (2021). Microsimulation modeling of extended annual ct screening among lung cancer cases in the national lung screening trial *Lung Cancer* **156**, 5–11.

Schünemann, C., Johanning, S., Reger, E., Herold, H. and Bruckner, T. (2024). Complex system policy modelling approaches for policy advice–comparing systems thinking, system dynamics and agent-based modelling *Political Research Exchange* **6**(1), 2387438.

Scott, A., Evandrou, M., Falkingham, J., Johnson, P. and Zaidi, A. (2003). Beyond 2020: Lessons from the development of the SAGE dynamic microsimulation model for Britain *ESRC-Sage Discussion Paper* **13**.

Scott, I. and Gong, T. (2021). Coordinating government silos: challenges and opportunities *Global Public Policy and Governance* **1**(1), 20–38.

Scottish Government (2017). Poverty act 2017 https://www.legislation.gov.uk/asp/2017/6/contents/enacted. Accessed: 2023-05-26.

Scottish Government (2022). Scottish child payment increased https://www.gov.scot/news/scottish-child-payment-increased. Accessed: 2023-05-26.

Selby, B. F. (2011). *Spatial prediction of AADT in unmeasured locations by universal kriging and microsimulation of vehicle holdings and car-market pricing dynamics* PhD thesis The University of Texas at Austin.

Senedd Cymru and Welsh Parliament (2023). Public health wales consultation response *Public Health Wales* .

Serena, L., Marzolla, M., D'Angelo, G. and Ferretti, S. (2023). A review of multilevel modeling and simulation for human mobility and behavior *Simulation modelling practice and theory* **127**, 102780.

Sharif, B., Kopec, J. A., Wong, H., Fines, P., Sayre, E. C., Liu, R. R. and Wolfson, M. C. (2012). Uncertainty analysis in population-based disease microsimulation models *Epidemiology Research International* **2012**.

Shewmaker, P., Chrysanthopoulou, S. A., Iskandar, R., Lake, D. and Jutkowitz, E. (2022). Microsimulation model calibration with approximate bayesian computation in r: A tutorial *Medical Decision Making* **42**(5), 557–570.

Simshauser, P. (2023). The 2022 energy crisis: Fuel poverty and the impact of policy interventions in Australia's National Electricity Market *Energy Economics* **121**, 106660.

SIPHER Consortium (2025). Participatory systems mapping in a socially distanced world. https://www.gla.ac.uk/research/az/sipher/sharingourevidence/blog/headline_1020102_en.html. Accessed: 2025-01-03.

SIPHER Consortium et al. (2023). Mapping the housing & health system: Methods adopted in participatory systems mapping with Greater Manchester Combined Authority.

Skarda, I., Asaria, M. and Cookson, R. (2021). Lifesim: a lifecourse dynamic microsimulation model of the Millennium birth cohort in England *medRxiv* pp. 2021–02.

Smith, D. M., Heppenstall, A. and Campbell, M. (2021). Estimating health over space and time: A review of spatial microsimulation applied to public health *J* **4**(2), 182–192.

Snell, C. J., Bevan, M. A. and Gillard, R. (2018). Policy pathways to justice in energy efficiency.

Spielauer, M., Horvath, G. T. and Fink, M. (2020). microwelt: A dynamic microsimulation model for the study of welfare transfer flows in ageing societies from a comparative welfare state perspective *Technical report* WIFO Working Papers.

Spielauer, M. et al. (2007). Dynamic microsimulation of health care demand, health care finance and the economic impact of health behaviours: survey and review *International Journal of Microsimulation* **1**(1), 35–53.

Spooner, F., Abrams, J. F., Morrissey, K., Shaddick, G., Batty, M., Milton, R., Dennett, A., Lomax, N., Malleson, N., Nelissen, N. et al. (2021). A dynamic microsimulation model for epidemics *Social Science & Medicine* **291**, 114461.

Statham, R., Smith, C. and Parkes, H. (2022). Tackling child poverty and destitution.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M. and Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls *Bmj* **338**.

Stewart, E., Panel, S. G. M. C., Panel, S. S. C., Panel, S. S. C. and Such, E. (2024). Evaluating participant experiences of community panels to scrutinise policy modelling for health inequalities: the SIPHER Consortium *Research Involvement and Engagement* **10**(1), 4.

Stillwell, J. and Duke-Williams, O. (2003). A new web-based interface to British census of population origin–destination statistics *Environment and Planning A* **35**(1), 113–132.

Stone, J., Robinson, E., Blackwell, C. and Padley, M. (2023). Debt, poverty and living standards in Great Britain.

Sutherland, H. (2018). Quality assessment of microsimulation models: the case of EUROMOD *Technical report* EUROMOD Working Paper.

Taneja, S. and Mandys, F. (2022). Drivers of UK household energy expenditure: Promoting efficiency and curbing emissions *Energy Policy* **167**, 113042.

Tanton, R., Vidyattama, Y. et al. (2010). Pushing it to the edge: extending generalised regression as a spatial microsimulation method *International Journal of Microsimulation* **3**(2), 23–33.

Tanton, R. et al. (2014). A review of spatial microsimulation methods *International Journal of Microsimulation* **7**(1), 4–25.

Taylor, J. and Wren, C. (1997). UK regional policy: an evaluation *Regional Studies* **31**(9), 835–848.

Templ, M., Meindl, B., Kowarik, A. and Dupriez, O. (2017). Simulation of synthetic complex data: The r package simpop *Journal of Statistical Software* **79**, 1–38.

Thomas, G. O., Poortinga, W. and Sautkina, E. (2016). Habit discontinuity, self-activation, and the diminishing influence of context change: evidence from the UK understanding society survey *PLoS One* **11**(4), e0153490.

Thrun, M. C., Gehlert, T. and Ultsch, A. (2020). Analyzing the fine structure of distributions *PloS one* **15**(10), e0238835.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**(1), 267–288.

Tikka, S., Hakanen, J., Saarela, M. and Karvanen, J. (2021). Sima–an open-source simulation framework for realistic large-scale individual-level data generation *International Journal of Microsimulation* **14**(3).

Tomova, G. D., Gilthorpe, M. S., Bruneau, G. C. A. and Tennant, P. W. (2022). P66 defining the transparency, explainability, and interpretability of algorithms: a key step towards fair and just decision-making.

Tsuchiya, A., Wickramasekera, N. and Field, B. (2022). Review of sipher-7 wellbeing indicators.

Tysinger, B. (2021). *Design, Validation, and An Application of the Future Adult Model* PhD thesis Pardee Rand Graduate School.

UK Fuel Poverty Monitor (2022). The hardest hit: Impact of the energy crisis *National Energy Action (NEA) and Energy Action Scotland (EAS)* .

UK Government (2019). New UK numeric education qualification levels http//www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels. Accessed: 2022-12-02.

United Kingdom Government: Department for Energy Security and Net Zero. (2023). Carbon budget delivery plan https://www.gov.uk/government/publications/carbon-budget-delivery-plan. Accessed: 2023-11-22.

Urban, D. and Shrestha, S. (2023). ScotMS–microsimulation model for Scotland: Construction and validation *Rural Development & Food Marketing* .

van de Ven, J., Bronka, P., Richiardi, M. et al. (2022). Dynamic simulation of taxes and welfare benefits by database imputation *Centre for Microsimulation and Policy Analysis Working Paper Series, CEMPA3/22* .

Venables, W. N. and Smith, D. M. (2003). The R. development core team *An Introduction to R, Version* **1**(0).

Vilagut, G., Forero, C. G., Pinto-Meza, A., Haro, J. M., De Graaf, R., Bruffaerts, R., Kovess, V., De Girolamo, G., Matschinger, H., Ferrer, M. et al. (2013). The mental component of the short-form 12 health survey (SF-12) as a measure of depressive disorders in the general population: results with three alternative scoring methods *Value in Health* **16**(4), 564–573.

Waddams, C. (2023). The UK Energy Price Guarantee *Intereconomics* **58**(1), 22–26.

Waddell, P. (2012). Parcel-level microsimulation of land use and transportation: The walking scale of urban sustainability *Travel Behaviour Research in an Evolving World* p. 77.

Ward, C., Robinson, C., Singleton, A. and Rowe, F. (2024). Spatial-temporal dynamics of gas consumption in England and Wales: Assessing the residential sector using sequence analysis *Applied Spatial Analysis and Policy* pp. 1–28.

Warr, D., Feldman, P., Tacticos, T. and Kelaher, M. (2009). Sources of stress in impoverished neighbourhoods: insights into links between neighbourhood environments and health *Australian and New Zealand journal of public health* **33**(1), 25–33.

Wasserstein, R. L., Schirm, A. L. and Lazar, N. A. (2019). Moving to a world beyond “p¡ 0.05”.

Wei, Y., Heun-Johnson, H. and Tysinger, B. (2022). Using dynamic microsimulation to project cognitive function in the elderly population *Plos one* **17**(9), e0274417.

Weimer, D. and Vining, A. (2017). *Policy analysis: Concepts and practice* Routledge.

Weinberg, N., Turner, D., Stansbury, A., Elsden, E. and Balls, E. (2024). A growth policy to close Britain’s regional divides: What needs to be done *M-RCBG Associate Working Paper Series* .

Wetherall, K., Mathew, S., Belkadi, N. and O’Connor, R. (2023). The cost-of-living crisis in the UK: A whole society.

Whelan, C. T. and Maitre, B. (2012). Understanding material deprivation: A comparative European analysis *Research in Social Stratification and Mobility* **30**(4), 489–503.

Williams, C., Lewsey, J. D., Mackay, D. F. and Briggs, A. H. (2017). Estimation of survival probabilities for use in cost-effectiveness analyses: a comparison of a multi-state modeling survival analysis approach with partitioned survival and markov decision-analytic modeling *Medical Decision Making* **37**(4), 427–439.

Wilson, L. E. (2020). *Longitudinal geographies of alcohol consumption: understanding middle-aged adult drinking in the UK.* PhD thesis University of Southampton.

Wolf, D. A. (2001). The role of microsimulation in longitudinal data analysis *Canadian Studies in Population [ARCHIVES]* pp. 313–339.

Xue, L., Liu, G.-P. and Hu, W. (2024). All-in-one framework for design, simulation, and practical implementation of distributed multiagent control systems *IEEE Transactions on Systems, Man, and Cybernetics: Systems* .

Yang, K. (2020). What can covid-19 tell us about evidence-based management? *The American Review of Public Administration* **50**(6-7), 706–712.

Zaidi, A. and Rake, K. (2001). Dynamic microsimulation models: a review and some lessons for SAGE *Simulating Social Policy in an Ageing Society (SAGE) discussion paper* **2**.

Zapata-Webborn, E., Hanmer, C., Oreszczyn, T., Huebner, G., McKenna, E., Few, J., Elam, S., Pullinger, M., Cheshire, C., Friel, D. et al. (2024). Winter demand falls as fuel bills rise: Understanding the energy impacts of the cost-of-living crisis on British households *Energy and Buildings* **305**.

Zhang, Z. (2016). Missing data imputation: focusing on single imputation *Annals of translational medicine* **4**(1).

Zhong, H., Brandeau, M. L., Yazdi, G. E., Wang, J., Nolen, S., Hagan, L., Thompson, W. W., Assoumou, S. A., Linas, B. P. and Salomon, J. A. (2022). Metamodeling for policy simulations with multivariate outcomes *Medical Decision Making* **42**(7), 872–884.

Zhou, M., Li, J., Basu, R. and Ferreira, J. (2022). Creating spatially-detailed heterogeneous synthetic populations for agent-based microsimulation *Computers, Environment and Urban Systems* **91**, 101717.

Zhou, M., Men, Y. and Qing, X. (2021). Current development in microsimulation and experimental innovation method in JUTTA model *Journal of Physics: Conference Series* Vol. 1865 IOP Publishing p. 042136.

Zinn, S. et al. (2014). The MicSim package of R.: An entry-level toolkit for continuous-time microsimulation *International Journal of Microsimulation* **7**(3), 3–32.

Zucchelli, E., Jones, A. M., Rice, N. et al. (2010). The evaluation of health policies through microsimulation methods *Health, Econometrics and Data Group (HEDG) Working Papers* **10**(03), 2.

Chapter 8: Appendices

This remainder of this chapter contains supplementary material used as part of chapters 4 and 5 providing data tables for transition probability models and data preprocessing and imputation strategies.

8.1 Chapter 4 Supplementary Material.

This section contains data tables used to generate MINOS models estimating income and housing support policy described in chapters 4 and 5. These data tables include description of variables used in the MINOS model as well as transition probability model coefficients.

8.1.1 Data Preprocessing Tables

-

age	hh_income	nutrition_quality	SF_12_MCS	
Min.	16.00	-3448	4.00	-8.00
1st Qu.	37.00	1011	12.00	40.54
Median	53.00	1544	16.00	49.36
Mean	51.59	1731	16.84	46.85
3rd Qu.	66.00	2202	20.00	55.93
Max.	90.00	14879	96.00	74.94

Table 8.1: Summary statistics for key continous variables in the MINOS model.

variable	value	1
nkids	0	11512 (72.1%)
	1	2125 (13.3%)
	2	1780 (11.1%)
	3	409 (2.6%)
	4	108 (0.7%)
	5	23 (0.1%)
	6	3 (0.0%)
	7	3 (0.0%)
	8	2 (0.0%)
sex	Female	8922 (55.9%)
	Male	7043 (44.1%)
ethnicity	BAN	293 (1.8%)
	BLA	253 (1.6%)
	BLC	307 (1.9%)
	CHI	64 (0.4%)
	IND	709 (4.4%)
	MIX	315 (2.0%)
	OAS	191 (1.2%)
	OBL	22 (0.1%)
	OTH	56 (0.4%)
	PAK	546 (3.4%)
	WBI	12493 (78.3%)
	WHO	716 (4.5%)
region	East Midlands	1254 (7.9%)
	East of England	1484 (9.3%)
	London	1986 (12.4%)
	North East	613 (3.8%)
	North West	1537 (9.6%)
	Northern Ireland	700 (4.4%)

	Scotland	1258 (7.9%)
	South East	2086 (13.1%)
	South West	1405 (8.8%)
	Wales	842 (5.3%)
	West Midlands	1371 (8.6%)
	Yorkshire and The Humber	1429 (9.0%)
education_state	0	2975 (18.6%)
	1	282 (1.8%)
	2	4032 (25.3%)
	3	1905 (11.9%)
	5	1459 (9.1%)
	6	3182 (19.9%)
	7	2130 (13.3%)
housing_quality	High	5384 (33.7%)
	Low	3868 (24.2%)
	Medium	6713 (42.0%)
neighbourhood_safety	1	5108 (32.0%)
	2	7926 (49.6%)
	3	2931 (18.4%)
marital_status	Partnered	10367 (64.9%)
	Separated	1337 (8.4%)
	Single	3337 (20.9%)
	Widowed	924 (5.8%)
hh_comp	1	2683 (16.8%)
	2	435 (2.7%)
	3	7916 (49.6%)
	4	4931 (30.9%)
job_sector	0	8317 (52.1%)
	1	4784 (30.0%)
	2	2864 (17.9%)
job_sec	0	1622 (10.2%)
	1	620 (3.9%)
	2	1260 (7.9%)
	3	4274 (26.8%)
	4	2112 (13.2%)
	5	1474 (9.2%)
	6	899 (5.6%)
	7	2428 (15.2%)
	8	1276 (8.0%)
loneliness	1	9445 (59.2%)
	2	5208 (32.6%)
	3	1312 (8.2%)

Table 8.2: Summary statistics for key discrete variables in the MINOS model.

8.1.2 Transition Probabilities Coefficients

This section contains all transition probability model coefficients used in transition models for the MINOS microsimulation in chapter 5.

	Houshold Income
(Intercept)	2.8903***
	(0.0009)
previous household income	0.0170***
	(0.0001)
household income difference	−0.0026***
	(0.0001)
age	0.0114***
	(0.0003)
age^2	0.0010***
	(0.0001)
age^3	−0.0031***
	(0.0001)
sex: Male vs Female	0.0018***
	(0.0002)
ethnicity: BAN vs. WBI	−0.0144***
	(0.0010)
ethnicity: BLA vs. WBI	−0.0168***
	(0.0009)
ethnicity: BLC vs. WBI	−0.0132***
	(0.0009)
ethnicity: CHI vs. WBI	−0.0013
	(0.0019)
ethnicity: IND vs. WBI	−0.0046***
	(0.0007)
ethnicity: MIX vs. WBI	−0.0069***
	(0.0009)
ethnicity: OAS vs. WBI	−0.0098***
	(0.0010)
ethnicity: OBL vs. WBI	−0.0101**
	(0.0031)
ethnicity: OTH vs. WBI	−0.0083***
	(0.0020)
ethnicity: PAK vs. WBI	−0.0121***
	(0.0007)
ethnicity: WHO vs. WBI	−0.0031***
	(0.0006)
region: East of England	0.0021***
	(0.0006)
region: London	0.0040***
	(0.0006)

	Household Income
region: North East	−0.0032***
	(0.0008)
region: North West	0.0005
	(0.0006)
region: Northern Ireland	0.0024***
	(0.0007)
region: Scotland	0.0025***
	(0.0006)
region: South East	0.0043***
	(0.0005)
region: South West	0.0003
	(0.0006)
region: Wales	−0.0007
	(0.0006)
region: West Midlands	0.0016**
	(0.0006)
region: Yorkshire and The Humber	−0.0009
	(0.0006)
education state: 0 vs. 1	−0.0008
	(0.0007)
education state: 2 vs. 1	0.0036***
	(0.0007)
education state: 3 vs. 1	0.0054***
	(0.0008)
education state: 5 vs. 1	0.0061***
	(0.0008)
education state: 6 vs. 1	0.0112***
	(0.0008)
education state: 7 vs. 1	
	(0.0008)
NSSEC: 0 vs. 3	−0.0138***
	(0.0003)
NSSEC: 1 vs. 3	0.0068***
	(0.0006)
NSSEC: 2 vs. 3	0.0065***
	(0.0005)
NSSEC: 4 vs. 3	−0.0048***
	(0.0004)
NSSEC: 5 vs. 3	−0.0093***
	(0.0005)
NSSEC: 6 vs. 3	−0.0058***
	(0.0005)
NSSEC: 7 vs. 3	−0.0091***
	(0.0004)
NSSEC: 8 vs. 3	−0.0102***

	Household Income
	(0.0005)
AIC	817036.3989
BIC	817518.1139
Log Likelihood	−408473.1995
Num. obs.	329317
Num. groups: pidp	61110
Var: pidp (Intercept)	0.0005
Var: Residual	0.0017

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.3: Household Income GLMM Coefficients

	Nutrition Quality
(Intercept)	17.2570***
	(0.0863)
household income difference	−0.0292*
	(0.0130)
age	0.5232***
	(0.0347)
sex: Male vs. Female	−1.4293***
	(0.0685)
education state: 0 vs. 3	−0.5823***
	(0.0967)
education state: 1 vs. 3	−0.7983***
	(0.1993)
education state: 2 vs. 3	−0.4280***
	(0.0901)
education state: 5 vs. 3	−0.0112
	(0.1187)
education state: 6 vs. 3	0.7120***
	(0.0985)
education state: 7 vs. 3	1.1512***
	(0.1113)
ethnicity: BAN vs. WBI	−1.9400***
	(0.2728)
ethnicity: BLA vs. WBI	−0.9041***
	(0.2470)
ethnicity: BLC vs. WBI	−0.7002**
	(0.2604)
ethnicity: CHI vs. WBI	1.5264**
	(0.5390)
ethnicity: IND vs. WBI	−0.5573**
	(0.1805)

	Nutrition Quality
ethnicity: MIX vs. WBI	0.1806
	(0.2545)
ethnicity: OAS vs. WBI	0.2370
	(0.3004)
ethnicity: OBL vs. WBI	0.3927
	(0.8541)
ethnicity: OTH vs. WBI	1.5508**
	(0.5848)
ethnicity: PAK vs. WBI	−2.6201***
	(0.1969)
ethnicity: WHO vs. WBI	0.8292***
	(0.1590)
household income	0.2624***
	(0.0223)
ncigs	−0.0805***
	(0.0049)
SF-12 MCS	0.1181***
	(0.0218)
AIC	840495.7799
BIC	840749.2756
Log Likelihood	−420221.8899
Num. obs.	126733
Num. groups: pidp	42937
Var: pidp (Intercept)	38.2830
Var: Residual	25.8728

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.4: Nutrition Quality LMM Coefficients

	SF-12 MCS
(Intercept)	2.5148***
	(0.0027)
previous SF-12 MCS	−0.0363***
	(0.0003)
previous SF-12 MCS ^{2}	−0.0031***
	(0.0002)
sex: Male vs. Female	−0.0093***
	(0.0006)
ethnicity: BLA vs. BAN	−0.0004
	(0.0034)
ethnicity: BLC vs. BAN	0.0028
	(0.0034)
ethnicity: CHI vs. BAN	0.0100$^{\cdot}$

	SF-12 MCS
	(0.0055)
ethnicity: IND vs. BAN	$0.0057^{\cdot}$
	(0.0030)
ethnicity: MIX vs. BAN	0.0139^{***}
	(0.0034)
ethnicity: OAS vs. BAN	0.0052
	(0.0037)
ethnicity: OBL vs. BAN	-0.0001
	(0.0083)
ethnicity: OTH vs. BAN	0.0064
	(0.0059)
ethnicity: PAK vs. BAN	0.0049
	(0.0031)
ethnicity: WBI vs. BAN	0.0063^{*}
	(0.0025)
ethnicity: WHO vs. BAN	$0.0048^{\cdot}$
	(0.0029)
age	-0.0092^{***}
	(0.0003)
age^2	-0.0010^{***}
	(0.0003)
housing quality: Low vs. High	0.0036^{***}
	(0.0008)
housing quality: Medium vs. High	0.0019^{**}
	(0.0006)
household income	-0.0023^{***}
	(0.0003)
neighbourhood safety: 2 vs. 1	-0.0053^{***}
	(0.0006)
neighbourhood safety 3 vs. 1	-0.0105^{***}
	(0.0007)
nutrition quality	-0.0002^{***}
	(0.0000)
smoker (ncigs ¿ 0)	0.0051^{***}
	(0.0008)
AIC	170706.7957
BIC	170947.2324
Log Likelihood	-85327.3979
Num. obs.	76693
Num. groups: pidp	41373
Var: pidp (Intercept)	0.0020
Var: Residual	0.0030

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.5: SF-12 MCS GLMM Coefficients

	Number of cigarettes
Count model: (Intercept)	2.3846***
	(0.0660)
Count model: household income difference	0.0301
	(0.0550)
Count model: previous ncigs	0.0741***
	(0.0099)
Count model: household income	−0.0376
	(0.0631)
Count model: household income ^{2}	0.0073
	(0.0271)
Count model: ethnicity: BAN vs. WBI	−0.3078
	(0.6768)
Count model: ethnicity: BLA vs. WBI	−0.6025
	(0.5267)
Count model: ethnicity: BLC vs. WBI	−0.6546
	(0.4852)
Count model: ethnicity: CHI vs. WBI	−0.1736
	(1.0087)
Count model: ethnicity: IND vs. WBI	−0.1967
	(0.4500)
Count model: ethnicity: MIX vs. WBI	−0.2637
	(0.3066)
Count model: ethnicity: OAS vs. WBI	−0.3789
	(0.6629)
Count model: ethnicity: OBL vs. WBI	−0.4682
	(1.0563)
Count model: ethnicity: OTH vs. WBI	−0.1775
	(0.6546)
Count model: ethnicity: PAK vs. WBI	−0.7241
	(0.5974)
Count model: ethnicity: WHO vs. WBI	−0.0479
	(0.1713)
Count model: age	0.0830$^{\cdot}$
	(0.0483)
Zero model: (Intercept)	4.3198***
	(0.4764)
Zero model: smoker (nigs ¿ 0)	−5.9872***
	(0.6166)
Zero model: ethnicity: BAN vs. WBI	−0.2881
	(3.7984)
Zero model: ethnicity: BLA vs. WBI	−0.0815
	(2.5893)
Zero model: ethnicity: BLC vs. WBI	−1.2670

	Number of cigarettes
	(2.4676)
Zero model: ethnicity: CHI vs. WBI	−0.8585
	(3.9093)
Zero model: ethnicity: IND vs. WBI	0.3048
	(2.3485)
Zero model: ethnicity: MIX vs. WBI	0.1066
	(1.9850)
Zero model: ethnicity: OAS vs. WBI	−0.6433
	(2.4315)
Zero model: ethnicity: OBL vs. WBI	−0.6832
	(7.0662)
Zero model: ethnicity: OTH vs. WBI	1.3262
	(3.2021)
Zero model: ethnicity: PAK vs. WBI	1.0452
	(2.2814)
Zero model: ethnicity: WHO vs. WBI	0.3969
	(1.1639)
Zero model: age	0.0634
	(0.3127)
AIC	559.3286
Log Likelihood	−248.6643
Num. obs.	13380

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.6: Cigarette consumption (ncigs) ZIP Coefficients.

	Housing Quality
previous housing quality: Medium	1.1457***
	(0.3008)
previous housing quality: High	4.3564***
	(0.3988)
age	−0.2788*
	(0.1304)
age^2	−0.2068*
	(0.1016)
sex: Male vs. Female	−0.0004
	(0.2129)
SF-12 MCS	0.0698
	(0.1059)
ethnicity: BAN vs. WBI	−0.5208
	(1.2418)
ethnicity: BLA vs. WBI	−0.7933
	(0.8720)

	Housing Quality
ethnicity: BLC vs. WBI	−0.6761
	(1.0650)
ethnicity: CHI vs. WBI	−0.1933
	(1.4276)
ethnicity: IND vs. WBI	−0.4518
	(0.6543)
ethnicity: MIX vs. WBI	−0.2899
	(0.8181)
ethnicity: OAS vs. WBI	−0.3426
	(0.8318)
ethnicity: OBL vs. WBI	−0.4084
	(2.8566)
ethnicity: OTH vs. WBI	−0.3848
	(1.4676)
ethnicity: PAK vs. WBI	−0.4954
	(0.8067)
ethnicity: WHO vs. WBI	−0.1100
	(0.4628)
household income	$0.2738^{\cdot}$
	(0.1559)
household income^2	−0.0301
	(0.0394)
household income difference	0.0042
	(0.1213)
housing tenure: 2 vs. 1	−0.0291
	(0.2986)
housing tenure: 3 vs. 1	−0.5393
	(0.4232)
housing tenure: 4 vs. 1	$-0.8579^{\cdot}$
	(0.4462)
housing tenure: 5 vs. 1	−0.4010
	(0.9522)
housing tenure: 6 vs. 1	−0.4094
	(0.4411)
housing tenure: 7 vs. 1	−0.3417
	(0.6194)
housing tenure: 8 vs. 1	−0.3768
	(1.7353)
Low—Medium Level CLM Threshhold	−0.4833
	(0.3831)
Medium—High Level CLM Threshhold	2.5411^{***}
	(0.4108)
AIC	700.4484
BIC	818.0973
Log Likelihood	−321.2242

	Housing Quality
Num. obs.	427

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.7: Housing Quality CLM Coefficients

	Neighbourhood Safety
previous neighbourhood safety: 2 vs. 1	1.2512***
	(0.2532)
previous neighbourhood safety: 3 vs. 1	2.3699***
	(0.3328)
age	0.2511*
	(0.1185)
sex: Male	0.1470
	(0.2131)
NSSEC: 1 vs. 0	−0.1271
	(0.6150)
NSSEC: 2 vs. 0	−0.0574
	(0.4701)
NSSEC: 3 vs. 0	−0.0508
	(0.3119)
NSSEC: 4 vs. 0	−0.1168
	(0.4099)
NSSEC: 5 vs. 0	0.0181
	(0.4818)
NSSEC: 6 vs. 0	−0.0269
	(0.5585)
NSSEC: 7 vs. 0	−0.1829
	(0.3924)
NSSEC: 8 vs. 0	−0.2609
	(0.5367)
ethnicity: BAN vs. WBI	−0.6205
	(1.7113)
ethnicity: BLA vs. WBI	0.5466
	(1.0417)
ethnicity: BLC vs. WBI	0.0844
	(1.1703)
ethnicity: CHI vs. WBI	−0.4691
	(1.4930)
ethnicity: IND vs. WBI	−0.0516
	(0.6685)
ethnicity: MIX vs. WBI	−0.2140
	(0.9213)
ethnicity: OAS vs. WBI	−0.1861

	Neighbourhood Safety
	(0.9167)
ethnicity: OBL vs. WBI	1.1373
	(3.2127)
ethnicity: OTH vs. WBI	0.2461
	(1.5249)
ethnicity: PAK vs. WBI	−0.4566
	(1.0045)
ethnicity: WHO vs. WBI	−0.0191
	(0.4969)
household income	0.0576
	(0.1130)
housing quality: Medium vs. Low	0.2778
	(0.3923)
housing quality High vs. Low	0.4570
	(0.4112)
region: East of England	−0.0666
	(0.4974)
region: London	−0.8105
	(0.4961)
region: North East	−0.1283
	(0.6363)
region: North West	−0.3707
	(0.5008)
region: Northern Ireland	0.2753
	(0.8027)
region: Scotland	0.1453
	(0.5307)
region: South East	−0.1144
	(0.4584)
region: South West	0.1723
	(0.5113)
region: Wales	−0.1088
	(0.6658)
region: West Midlands	−0.2671
	(0.5190)
region: Yorkshire and The Humber	−0.1859
	(0.5146)
Level 1—2 CLM threshhold	0.2311
	(0.5664)
Level 2—3 CLM threshold	3.0369***
	(0.5906)
AIC	730.0927
BIC	882.4461
Log Likelihood	−326.0463
Num. obs.	367

	Neighbourhood Safety

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.8: Neighbourhood Safety CLM Coefficients

	Loneliness
factor(loneliness)2	1.7113***
	(0.2522)
factor(loneliness)3	3.2293***
	(0.4693)
scale(age)	−0.1451
	(0.1788)
factor(sex)Male	−0.1880
	(0.2309)
scale(SF_12)	−0.4471***
	(0.1246)
relevel(factor(education_state), ref = "3")0	0.0638
	(0.4321)
relevel(factor(education_state), ref = "3")1	0.2216
	(0.9876)
relevel(factor(education_state), ref = "3")2	0.0695
	(0.3825)
relevel(factor(education_state), ref = "3")5	0.1136
	(0.4921)
relevel(factor(education_state), ref = "3")6	0.0803
	(0.4035)
relevel(factor(education_state), ref = "3")7	0.0348
	(0.4455)
relevel(factor(job_sec), ref = "3")0	0.0751
	(0.3497)
relevel(factor(job_sec), ref = "3")1	0.0958
	(0.6854)
relevel(factor(job_sec), ref = "3")2	−0.2263
	(0.5003)
relevel(factor(job_sec), ref = "3")4	−0.1687
	(0.4520)
relevel(factor(job_sec), ref = "3")5	−0.2732
	(0.5895)
relevel(factor(job_sec), ref = "3")6	−0.1390
	(0.6504)
relevel(factor(job_sec), ref = "3")7	−0.1852
	(0.4434)
relevel(factor(job_sec), ref = "3")8	−0.1069
	(0.5720)

	Loneliness
scale(hh_income)	−0.0200
	(0.1300)
relevel(factor(hh_comp), ref = ”3”)1	0.2440
	(0.4070)
relevel(factor(hh_comp), ref = ”3”)2	0.4249
	(0.6598)
relevel(factor(hh_comp), ref = ”3”)4	0.0244
	(0.2881)
relevel(factor(marital_status), ref = ”Partnered”)Separated	0.3495
	(0.4856)
relevel(factor(marital_status), ref = ”Partnered”)Single	0.3071
	(0.3720)
relevel(factor(marital_status), ref = ”Partnered”)Widowed	0.5801
	(0.5711)
relevel(factor(ethnicity), ref = ”WBI”)BAN	−0.2213
	(1.3531)
relevel(factor(ethnicity), ref = ”WBI”)BLA	−0.0328
	(0.9259)
relevel(factor(ethnicity), ref = ”WBI”)BLC	−0.3179
	(1.1461)
relevel(factor(ethnicity), ref = ”WBI”)CHI	0.2044
	(1.5005)
relevel(factor(ethnicity), ref = ”WBI”)IND	0.0750
	(0.6703)
relevel(factor(ethnicity), ref = ”WBI”)MIX	−0.1594
	(0.8512)
relevel(factor(ethnicity), ref = ”WBI”)OAS	0.2676
	(0.8476)
relevel(factor(ethnicity), ref = ”WBI”)OBL	0.5382
	(3.0551)
relevel(factor(ethnicity), ref = ”WBI”)OTH	−0.2453
	(1.4142)
relevel(factor(ethnicity), ref = ”WBI”)PAK	−0.1915
	(0.8880)
relevel(factor(ethnicity), ref = ”WBI”)WHO	0.1053
	(0.4789)
1—2	1.3434**
	(0.4661)
2—3	4.2287***
	(0.5231)
AIC	660.6849
BIC	818.9024
Log Likelihood	−291.3425

Num. obs.	427

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.9: Loneliness CLM Coefficients

8.2 Chapter 5 Supplementary Material.

8.2.1 Chapter 5 Data Tables

	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
age	16.00	37.00	53.00	51.59	66.00	90.00
net_hh_income	-1148	2526	3989	4547	5823	34968
hh_income	-3448	1011	1544	1731	2202	14879
nutrition_quality	4.00	12.00	16.00	16.84	20.00	96.00
council_tax	0.0	129.7	160.0	163.8	190.3	472.9
SF_12_MCS	-8.00	40.54	49.36	46.85	55.93	74.94
SF_12_PCS	-8.00	43.58	53.06	48.72	56.91	74.73
hh_rent	0.0	0.0	0.0	161.1	0.0	6072.2
hh_mortgage	0.0	0.0	0.0	309.2	545.2	8016.8
hh_comp	1.000	3.000	3.000	2.946	4.000	4.000
ncars	-2.000	1.000	1.000	1.521	2.000	5.000
ncigs	0.000	0.000	0.000	1.308	0.000	50.000
yearly_energy	-33	1875	2503	2689	3275	19969

Table 8.10: Summary statistics for key continous variables in the MINOS model.

variable	value	n	freq	export
nkids	0	11512	0.72	11512 (72.1%)
	1	2125	0.13	2125 (13.3%)
	2	1780	0.11	1780 (11.1%)
	3	409	0.03	409 (2.6%)
	4	108	0.01	108 (0.7%)
	5	23	0.00	23 (0.1%)
	6	3	0.00	3 (0.0%)
	7	3	0.00	3 (0.0%)
	8	2	0.00	2 (0.0%)
sex	Female	8922	0.56	8922 (55.9%)
	Male	7043	0.44	7043 (44.1%)
ethnicity	BAN	293	0.02	293 (1.8%)
	BLA	253	0.02	253 (1.6%)
	BLC	307	0.02	307 (1.9%)
	CHI	64	0.00	64 (0.4%)
	IND	709	0.04	709 (4.4%)
	MIX	315	0.02	315 (2.0%)
	OAS	191	0.01	191 (1.2%)
	OBL	22	0.00	22 (0.1%)
	OTH	56	0.00	56 (0.4%)
	PAK	546	0.03	546 (3.4%)
	WBI	12493	0.78	12493 (78.3%)
	WHO	716	0.04	716 (4.5%)
region	East Midlands	1254	0.08	1254 (7.9%)
	East of England	1484	0.09	1484 (9.3%)
	London	1986	0.12	1986 (12.4%)
	North East	613	0.04	613 (3.8%)
	North West	1537	0.10	1537 (9.6%)
	Northern Ireland	700	0.04	700 (4.4%)
	Scotland	1258	0.08	1258 (7.9%)
	South East	2086	0.13	2086 (13.1%)
	South West	1405	0.09	1405 (8.8%)

	Wales	842	0.05	842 (5.3%)
	West Midlands	1371	0.09	1371 (8.6%)
	Yorkshire and The Humber	1429	0.09	1429 (9.0%)
education_state	0	2975	0.19	2975 (18.6%)
	1	282	0.02	282 (1.8%)
	2	4032	0.25	4032 (25.3%)
	3	1905	0.12	1905 (11.9%)
	5	1459	0.09	1459 (9.1%)
	6	3182	0.20	3182 (19.9%)
	7	2130	0.13	2130 (13.3%)
housing_quality	High	5384	0.34	5384 (33.7%)
	Low	3868	0.24	3868 (24.2%)
	Medium	6713	0.42	6713 (42.0%)
neighbourhood_safety	1	5108	0.32	5108 (32.0%)
	2	7926	0.50	7926 (49.6%)
	3	2931	0.18	2931 (18.4%)
marital_status	Partnered	10367	0.65	10367 (64.9%)
	Separated	1337	0.08	1337 (8.4%)
	Single	3337	0.21	3337 (20.9%)
	Widowed	924	0.06	924 (5.8%)
job_sector	0	8317	0.52	8317 (52.1%)
	1	4784	0.30	4784 (30.0%)
	2	2864	0.18	2864 (17.9%)
job_sec	0	1622	0.10	1622 (10.2%)
	1	620	0.04	620 (3.9%)
	2	1260	0.08	1260 (7.9%)
	3	4274	0.27	4274 (26.8%)
	4	2112	0.13	2112 (13.2%)
	5	1474	0.09	1474 (9.2%)
	6	899	0.06	899 (5.6%)
	7	2428	0.15	2428 (15.2%)
	8	1276	0.08	1276 (8.0%)
loneliness	1	9445	0.59	9445 (59.2%)
	2	5208	0.33	5208 (32.6%)
	3	1312	0.08	1312 (8.2%)
FP10	False	13088	0.82	13088 (82.0%)
	True	2877	0.18	2877 (18.0%)
active	-9	108	0.01	108 (0.7%)
	0	9544	0.60	9544 (59.8%)
	1	6313	0.40	6313 (39.5%)
financial_situation	1	5125	0.32	5125 (32.1%)
	2	7043	0.44	7043 (44.1%)
	3	2852	0.18	2852 (17.9%)
	4	685	0.04	685 (4.3%)
	5	260	0.02	260 (1.6%)
behind_on_bills	-1	33	0.00	33 (0.2%)

	-2	10	0.00	10 (0.1%)
	1	15100	0.95	15100 (94.6%)
	2	767	0.05	767 (4.8%)
	3	55	0.00	55 (0.3%)
	Partnered	10367	0.65	10367 (64.9%)
	Separated	1337	0.08	1337 (8.4%)
	Single	3337	0.21	3337 (20.9%)
	Widowed	924	0.06	924 (5.8%)
number_of_bedrooms	0	31	0.00	31 (0.2%)
	1	779	0.05	779 (4.9%)
	12	2	0.00	2 (0.0%)
	2	3036	0.19	3036 (19.0%)
	3	7198	0.45	7198 (45.1%)
	4	3837	0.24	3837 (24.0%)
	5	882	0.06	882 (5.5%)
	6	157	0.01	157 (1.0%)
	7	33	0.00	33 (0.2%)
	8	7	0.00	7 (0.0%)
	9	3	0.00	3 (0.0%)
auditc	-9	136	0.01	136 (0.9%)
	High Risk	877	0.05	877 (5.5%)
	Increased Risk	2909	0.18	2909 (18.2%)
	Low Risk	7975	0.50	7975 (50.0%)
	Non-drinker	4068	0.25	4068 (25.5%)
housing_tenure	1	6355	0.40	6355 (39.8%)
	2	5768	0.36	5768 (36.1%)
	3	1228	0.08	1228 (7.7%)
	4	968	0.06	968 (6.1%)
	5	187	0.01	187 (1.2%)
	6	1401	0.09	1401 (8.8%)
	7	58	0.00	58 (0.4%)

Table 8.11: Summary statistics for key discrete variables in the MINOS model.

8.2.2 Chapter 5 Transition Model Coefficients

Variable	Coefficient	Standard Error
(Intercept)	4.3255^{***}	(0.0106)
scale(net_hh_income)	0.0080^{***}	(0.0000)
scale(net_hh_income)^2	-0.0004^{***}	(0.0000)
scale(age)	-0.0004^{***}	(0.0000)
scale(age)^2	-0.0003^{***}	(0.0000)
scale(age)^3	-0.0002^{***}	(0.0000)
factor(sex)Male	0.0002^{***}	(0.0000)
relevel(factor(ethnicity), ref = ”WBI”)BAN	0.0001	(0.0002)
relevel(factor(ethnicity), ref = ”WBI”)BLA	-0.0010^{***}	(0.0001)
relevel(factor(ethnicity), ref = ”WBI”)BLC	-0.0013^{***}	(0.0001)
relevel(factor(ethnicity), ref = ”WBI”)CHI	-0.0003	(0.0003)
relevel(factor(ethnicity), ref = ”WBI”)IND	0.0004^{***}	(0.0001)
relevel(factor(ethnicity), ref = ”WBI”)MIX	-0.0006^{***}	(0.0001)
relevel(factor(ethnicity), ref = ”WBI”)OAS	-0.0004^{*}	(0.0002)
relevel(factor(ethnicity), ref = ”WBI”)OBL	-0.0005	(0.0004)
relevel(factor(ethnicity), ref = ”WBI”)OTH	-0.0010^{***}	(0.0003)
relevel(factor(ethnicity), ref = ”WBI”)PAK	0.0001	(0.0001)
relevel(factor(ethnicity), ref = ”WBI”)WHO	-0.0000	(0.0001)
factor(region)East of England	0.0004^{***}	(0.0001)
factor(region)London	0.0012^{***}	(0.0001)
factor(region)North East	-0.0003^{**}	(0.0001)
factor(region)North West	0.0001	(0.0001)
factor(region)Northern Ireland	-0.0002^{*}	(0.0001)
factor(region)Scotland	0.0000	(0.0001)
factor(region)South East	0.0009^{***}	(0.0001)
factor(region)South West	0.0002^{**}	(0.0001)
factor(region)Wales	-0.0001	(0.0001)
factor(region)West Midlands	0.0003^{***}	(0.0001)
factor(region)Yorkshire and The Humber	$-0.0002^{\cdot}$	(0.0001)
relevel(factor(education_state), ref = ”1”)0	$-0.0002^{\cdot}$	(0.0001)
relevel(factor(education_state), ref = ”1”)2	0.0002	(0.0001)
relevel(factor(education_state), ref = ”1”)3	0.0006^{***}	(0.0001)
relevel(factor(education_state), ref = ”1”)5	0.0006^{***}	(0.0001)
relevel(factor(education_state), ref = ”1”)6	0.0011^{***}	(0.0001)
relevel(factor(education_state), ref = ”1”)7	0.0015^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)0	-0.0006^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)1	0.0011^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)2	0.0008^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)4	-0.0005^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)5	-0.0005^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)6	-0.0005^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)7	-0.0008^{***}	(0.0001)
relevel(factor(job_sec), ref = ”3”)8	-0.0008^{***}	(0.0001)
time	0.0001^{***}	(0.0000)
scale(SF_12)	0.0003^{***}	(0.0000)
AIC	937011.8275	
BIC	937507.5236	
Log Likelihood	-468458.9138	
Num. obs.	281173	
Num. groups: pidp	51863	
Var: pidp (Intercept)	0.0000	
Var: Residual	0.0001	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.12: Coefficients for the GLMM Gamma-Normal model predicting household net income.

Variable	Coefficient	Standard Error
(Intercept)	0.0986	(0.3319)
scale(net_hh_income)	0.0197***	(0.0007)
scale(hh_income_diff)	−0.0075***	(0.0005)
scale(age)	−0.0152***	(0.0015)
scale(age)^2	−0.0084***	(0.0009)
scale(age)^3	0.0006	(0.0006)
factor(sex)Male	−0.0059***	(0.0016)
relevel(factor(ethnicity), ref = "WBI")BAN	-0.0037	(0.0048)
relevel(factor(ethnicity), ref = "WBI")BLA	-0.0029	(0.0038)
relevel(factor(ethnicity), ref = "WBI")BLC	−0.0255***	(0.0047)
relevel(factor(ethnicity), ref = "WBI")CHI	0.0298*	(0.0122)
relevel(factor(ethnicity), ref = "WBI")IND	0.0198***	(0.0051)
relevel(factor(ethnicity), ref = "WBI")MIX	-0.0040	(0.0048)
relevel(factor(ethnicity), ref = "WBI")OAS	0.0368***	(0.0054)
relevel(factor(ethnicity), ref = "WBI")OBL	-0.0140	(0.0139)
relevel(factor(ethnicity), ref = "WBI")OTH	−0.0192$^{\cdot}$	(0.0099)
relevel(factor(ethnicity), ref = "WBI")PAK	0.0065	(0.0049)
relevel(factor(ethnicity), ref = "WBI")WHO	0.0249***	(0.0035)
factor(region)East of England	0.0523***	(0.0040)
factor(region)London	0.1268***	(0.0037)
factor(region)North East	−0.0119*	(0.0048)
factor(region)North West	-0.0059	(0.0038)
factor(region)Northern Ireland	−0.0415***	(0.0048)
factor(region)Scotland	−0.0304***	(0.0041)
factor(region)South East	0.0715***	(0.0038)
factor(region)South West	0.0380***	(0.0041)
factor(region)Wales	−0.0139**	(0.0045)
factor(region)West Midlands	0.0096*	(0.0040)
factor(region)Yorkshire and The Humber	−0.0135***	(0.0040)
relevel(factor(education_state), ref = "1")0	−0.0214***	(0.0051)
relevel(factor(education_state), ref = "1")2	−0.0148**	(0.0051)
relevel(factor(education_state), ref = "1")3	-0.0015	(0.0053)
relevel(factor(education_state), ref = "1")5	-0.0042	(0.0055)
relevel(factor(education_state), ref = "1")6	0.0112*	(0.0053)
relevel(factor(education_state), ref = "1")7	0.0158**	(0.0056)
relevel(factor(job_sec), ref = "3")0	−0.0049*	(0.0021)
relevel(factor(job_sec), ref = "3")1	0.0123**	(0.0044)
relevel(factor(job_sec), ref = "3")2	0.0098** z	(0.0031)
relevel(factor(job_sec), ref = "3")4	-0.0020	(0.0022)
relevel(factor(job_sec), ref = "3")5	0.0049$^{\cdot}$	(0.0027)
relevel(factor(job_sec), ref = "3")6	-0.0037	(0.0026)
relevel(factor(job_sec), ref = "3")7	−0.0044*	(0.0019)
relevel(factor(job_sec), ref = "3")8	−0.0072**	(0.0023)
time	0.0008***	(0.0002)
factor(housing_tenure)2	−0.0120*	(0.0051)
factor(housing_tenure)3	−0.0881***	(0.0046)
factor(housing_tenure)4	−0.0731***	(0.0046)
factor(housing_tenure)5	−0.0570***	(0.0056)
factor(housing_tenure)6	0.0162***	(0.0045)
factor(housing_tenure)7	−0.0455***	(0.0119)
scale(SF_12)	0.0040***	(0.0006)
AIC	158844.4151	
BIC	159329.8272	
Log Likelihood	-79369.2076	
Num. obs.	70173	
Num. groups: pidp	17951	
Var: pidp (Intercept)	0.0065	
Var: Residual	0.0129	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.13: Coefficients for the LMM Normal-Normal model predicting yearly household rent bills.

Variable	Coefficient	Standard Error
(Intercept)	-743.4144	(769.5913)
scale(hh_mortgage)	226.0634***	(1.5651)
scale(age)	−43.8486***	(3.5001)
scale(age)^2	−30.2114***	(1.6451)
scale(age)^3	0.2506	(0.9496)
factor(sex)Male	14.9984**	(4.5774)
relevel(factor(ethnicity), ref = "WBI")BAN	−69.0747***	(19.8148)
relevel(factor(ethnicity), ref = "WBI")BLA	78.6062***	(19.1681)
relevel(factor(ethnicity), ref = "WBI")BLC	−71.3269***	(17.8941)
relevel(factor(ethnicity), ref = "WBI")CHI	-17.9907	(30.7321)
relevel(factor(ethnicity), ref = "WBI")IND	1.9513	(10.9253)
relevel(factor(ethnicity), ref = "WBI")MIX	-21.6562	(16.6162)
relevel(factor(ethnicity), ref = "WBI")OAS	20.1698	(18.1873)
relevel(factor(ethnicity), ref = "WBI")OBL	59.7639	(64.9807)
relevel(factor(ethnicity), ref = "WBI")OTH	-38.1401	(37.5667)
relevel(factor(ethnicity), ref = "WBI")PAK	−120.2754***	(12.2650)
relevel(factor(ethnicity), ref = "WBI")WHO	49.2240***	(11.3926)
factor(region)East of England	107.6097***	(10.5983)
factor(region)London	237.2344***	(10.4912)
factor(region)North East	−43.2154**	(13.9723)
factor(region)North West	−20.8226*	(10.3477)
factor(region)Northern Ireland	−96.9094***	(12.4663)
factor(region)Scotland	-9.9938	(11.0845)
factor(region)South East	162.1727***	(9.7221)
factor(region)South West	47.2420***	(10.8494)
factor(region)Wales	−44.1373***	(11.8751)
factor(region)West Midlands	6.4434	(10.6930)
factor(region)Yorkshire and The Humber	−49.6486***	(10.7878)
relevel(factor(education_state), ref = "1")0	-21.8990	(19.4747)
relevel(factor(education_state), ref = "1")2	-30.9707	(19.3183)
relevel(factor(education_state), ref = "1")3	-19.7644	(19.6292)
relevel(factor(education_state), ref = "1")5	-12.9809	(19.9805)
relevel(factor(education_state), ref = "1")6	21.6265	(19.5395)
relevel(factor(education_state), ref = "1")7	73.1569***	(19.8118)
relevel(factor(job_sec), ref = "3")0	−15.1619*	(7.3161)
relevel(factor(job_sec), ref = "3")1	55.6258***	(7.1175)
relevel(factor(job_sec), ref = "3")2	36.0848***	(6.4395)
relevel(factor(job_sec), ref = "3")4	−26.7029***	(5.1184)
relevel(factor(job_sec), ref = "3")5	−24.8846***	(6.5798)
relevel(factor(job_sec), ref = "3")6	−34.9088***	(6.9942)
relevel(factor(job_sec), ref = "3")7	−43.9499***	(5.2622)
relevel(factor(job_sec), ref = "3")8	−56.0638***	(6.8224)
factor(housing_tenure)2	−249.0440***	(9.4308)
factor(housing_tenure)3	−81.4714***	(22.8746)
factor(housing_tenure)4	−264.4725***	(17.3438)
factor(housing_tenure)5	9.1782	(34.1565)
factor(housing_tenure)6	117.0237***	(11.9376)
factor(housing_tenure)7	30.0746	(36.5013)
time	0.8845*	(0.3826)
AIC	1632211.6684	
BIC	1632702.0750	
Log Likelihood	-816054.8342	
Num. obs.	110837	
Num. groups: pidp	23805	
Var: pidp (Intercept)	84236.5658	
Var: Residual	108634.0969	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.14: Coefficients for the LMM Normal-Normal model predicting yearly household mortgage expenditure.

Variable	Coefficient	Standard Error
(Intercept)	-17.1106^{***}	(0.2508)
scale(hh_income)	0.0201^{***}	(0.0006)
scale(hh_income_diff)	-0.0091^{***}	(0.0004)
scale(age)	0.0465^{***}	(0.0017)
scale(age)^2	0.0020^{**}	(0.0007)
scale(age)^3	-0.0137^{***}	(0.0006)
factor(sex)Male	-0.0031	(0.0019)
relevel(factor(ethnicity), ref = "WBI")BAN	0.0003	(0.0071)
relevel(factor(ethnicity), ref = "WBI")BLA	-0.0719^{***}	(0.0063)
relevel(factor(ethnicity), ref = "WBI")BLC	-0.0807^{***}	(0.0068)
relevel(factor(ethnicity), ref = "WBI")CHI	-0.0220	(0.0141)
relevel(factor(ethnicity), ref = "WBI")IND	-0.0397^{***}	(0.0049)
relevel(factor(ethnicity), ref = "WBI")MIX	-0.0293^{***}	(0.0069)
relevel(factor(ethnicity), ref = "WBI")OAS	-0.0120	(0.0076)
relevel(factor(ethnicity), ref = "WBI")OBL	-0.0376	(0.0231)
relevel(factor(ethnicity), ref = "WBI")OTH	-0.0144	(0.0144)
relevel(factor(ethnicity), ref = "WBI")PAK	-0.1013^{***}	(0.0052)
relevel(factor(ethnicity), ref = "WBI")WHO	-0.0232^{***}	(0.0052)
factor(region)East of England	0.0727^{***}	(0.0041)
factor(region)London	0.0494^{***}	(0.0040)
factor(region)North East	-0.0637^{***}	(0.0054)
factor(region)North West	-0.0263^{***}	(0.0040)
factor(region)Northern Ireland	-0.0436	(0.0305)
factor(region)Scotland	-0.6522^{***}	(0.0044)
factor(region)South East	0.1552^{***}	(0.0038)
factor(region)South West	0.0837^{***}	(0.0042)
factor(region)Wales	0.0060	(0.0046)
factor(region)West Midlands	-0.0251^{***}	(0.0042)
factor(region)Yorkshire and The Humber	-0.0980^{***}	(0.0042)
relevel(factor(education_state), ref = "1")0	-0.0172^{**}	(0.0057)
relevel(factor(education_state), ref = "1")2	0.0142^{*}	(0.0057)
relevel(factor(education_state), ref = "1")3	0.0527^{***}	(0.0059)
relevel(factor(education_state), ref = "1")5	0.0395^{***}	(0.0061)
relevel(factor(education_state), ref = "1")6	0.0726^{***}	(0.0059)
relevel(factor(education_state), ref = "1")7	0.0814^{***}	(0.0060)
relevel(factor(job_sec), ref = "3")0	-0.0087^{***}	(0.0018)
relevel(factor(job_sec), ref = "3")1	0.0203^{***}	(0.0028)
relevel(factor(job_sec), ref = "3")2	0.0200^{***}	(0.0024)
relevel(factor(job_sec), ref = "3")4	-0.0073^{***}	(0.0018)
relevel(factor(job_sec), ref = "3")5	-0.0011	(0.0022)
relevel(factor(job_sec), ref = "3")6	-0.0244^{***}	(0.0023)
relevel(factor(job_sec), ref = "3")7	-0.0245^{***}	(0.0017)
relevel(factor(job_sec), ref = "3")8	-0.0293^{***}	(0.0021)
factor(housing_tenure)2	-0.0046^{**}	(0.0016)
factor(housing_tenure)3	-0.2290^{***}	(0.0027)
factor(housing_tenure)4	-0.1862^{***}	(0.0027)
factor(housing_tenure)5	-0.0837^{***}	(0.0049)
factor(housing_tenure)6	-0.0974^{***}	(0.0022)
factor(housing_tenure)7	-0.0383^{***}	(0.0081)
time	0.0091^{***}	(0.0001)
scale(SF_12)	0.0047^{***}	(0.0005)
AIC	422768.8704	
BIC	423324.9649	
Log Likelihood	-211331.4352	
Num. obs.	266292	
Num. groups: pidp	48938	
Var: pidp (Intercept)	0.0363	
Var: Residual	0.0262	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.15: Coefficients for the GLMM Gamma-Normal model predicting houshold yearly council tax bills.

Variable	Coefficient	Standard Error
(Intercept)	7.7170***	(0.0006)
factor(financial_situation)2	0.0000	(0.0000)
factor(financial_situation)3	-0.0000	(0.0000)
factor(financial_situation)4	-0.0000	(0.0000)
factor(financial_situation)5	-0.0000	(0.0000)
scale(yearly_energy)	0.4947***	(0.0006)
factor(housing_tenure)2	-0.0000	(0.0000)
factor(housing_tenure)3	-0.0000	(0.0000)
factor(housing_tenure)4	-0.0000	(0.0000)
factor(housing_tenure)5	-0.0000	(0.0000)
factor(housing_tenure)6	-0.0000	(0.0000)
factor(housing_tenure)7	-0.0000	(0.0000)
scale(age)	-0.0000	(0.0000)
scale(net_hh_income)	0.0000	(0.0000)
factor(heating)1	-0.0000	(0.0000)
factor(housing_quality)Low	-0.0000	(0.0000)
factor(housing_quality)Medium	-0.0000	(0.0000)
factor(region)East of England	-0.0000	(0.0000)
factor(region)London	-0.0000	(0.0000)
factor(region)North East	0.0000	(0.0000)
factor(region)North West	0.0000	(0.0000)
factor(region)Northern Ireland	0.0000	(0.0000)
factor(region)Scotland	-0.0000	(0.0000)
factor(region)South East	-0.0000	(0.0000)
factor(region)South West	-0.0000	(0.0000)
factor(region)Wales	0.0000	(0.0000)
factor(region)West Midlands	0.0000	(0.0000)
factor(region)Yorkshire and The Humber	0.0000	(0.0000)
AIC	-4848827.8872	
BIC	-4848511.1101	
Log Likelihood	2424443.9436	
Num. obs.	284713	
Num. groups: pidp	52001	
Var: pidp (Intercept)	0.0163	
Var: Residual	0.0000	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.16: Coefficients for the GLMM Gamma-Normal model predicting yearly energy expenditure for all combined fuel sources.

Variable	Coefficient	Standard Error
(Intercept)	-140.7250^{***}	(14.1755)
scale(age)	0.4246^{***}	(0.0370)
factor(sex)Male	-1.3784^{***}	(0.0700)
relevel(factor(ethnicity), ref = "WBI")BAN	-2.2057^{***}	(0.2802)
relevel(factor(ethnicity), ref = "WBI")BLA	-1.1313^{***}	(0.2560)
relevel(factor(ethnicity), ref = "WBI")BLC	-1.0677^{***}	(0.2689)
relevel(factor(ethnicity), ref = "WBI")CHI	1.1242^{*}	(0.5433)
relevel(factor(ethnicity), ref = "WBI")IND	-0.8751^{***}	(0.1884)
relevel(factor(ethnicity), ref = "WBI")MIX	-0.3942	(0.2583)
relevel(factor(ethnicity), ref = "WBI")OAS	0.1337	(0.3055)
relevel(factor(ethnicity), ref = "WBI")OBL	0.2266	(0.8329)
relevel(factor(ethnicity), ref = "WBI")OTH	$1.0343^{\cdot}$	(0.5656)
relevel(factor(ethnicity), ref = "WBI")PAK	-2.4149^{***}	(0.2024)
relevel(factor(ethnicity), ref = "WBI")WHO	1.2346^{***}	(0.1744)
factor(region)East of England	0.0952	(0.1672)
factor(region)London	$0.3209^{\cdot}$	(0.1647)
factor(region)North East	-0.3166	(0.2159)
factor(region)North West	-0.4319^{**}	(0.1638)
factor(region)Northern Ireland	-1.7199^{***}	(0.1960)
factor(region)Scotland	-0.4707^{**}	(0.1747)
factor(region)South East	0.3609^{*}	(0.1559)
factor(region)South West	0.5574^{**}	(0.1701)
factor(region)Wales	0.3682^{*}	(0.1874)
factor(region)West Midlands	0.0804	(0.1695)
factor(region)Yorkshire and The Humber	$-0.3033^{\cdot}$	(0.1697)
relevel(factor(education_state), ref = "1")0	0.0187	(0.2366)
relevel(factor(education_state), ref = "1")2	0.1757	(0.2374)
relevel(factor(education_state), ref = "1")3	0.6732^{**}	(0.2453)
relevel(factor(education_state), ref = "1")5	0.7758^{**}	(0.2516)
relevel(factor(education_state), ref = "1")6	1.5830^{***}	(0.2410)
relevel(factor(education_state), ref = "1")7	2.1108^{***}	(0.2475)
scale(hh_income)	0.2502^{***}	(0.0229)
time	0.0779^{***}	(0.0070)
AIC	830879.2472	
BIC	831229.8164	
Log Likelihood	-415403.6236	
Num. obs.	125246	
Num. groups: hidp	86621	
Num. groups: pidp	37994	
Var: hidp (Intercept)	5.4856	
Var: pidp (Intercept)	35.1230	
Var: Residual	22.7925	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.17: Coefficients for the LMM Normal-Normal model predicting nutrition quality.

Variable	Coefficient	Standard Error
(Intercept)	2.9476***	(0.0020)
scale(SF_12_last)	−0.0205***	(0.0002)
scale(SF_12_last)^2	−0.0023***	(0.0001)
scale(net_hh_income)	-0.0002	(0.0004)
factor(behind_on_bills)2	0.0014	(0.0009)
factor(behind_on_bills)3	0.0010	(0.0030)
factor(financial_situation)2	0.0059***	(0.0004)
factor(financial_situation)3	0.0126***	(0.0005)
factor(financial_situation)4	0.0217***	(0.0009)
factor(financial_situation)5	0.0279***	(0.0014)
scale(yearly_energy)	0.0003	(0.0002)
scale(age)	−0.0053***	(0.0003)
factor(heating)1	−0.0096***	(0.0010)
factor(sex)Male	−0.0044***	(0.0004)
relevel(factor(ethnicity), ref = "WBI")BAN	−0.0059***	(0.0018)
relevel(factor(ethnicity), ref = "WBI")BLA	−0.0093***	(0.0016)
relevel(factor(ethnicity), ref = "WBI")BLC	−0.0070***	(0.0016)
relevel(factor(ethnicity), ref = "WBI")CHI	0.0001	(0.0033)
relevel(factor(ethnicity), ref = "WBI")IND	−0.0024*	(0.0011)
relevel(factor(ethnicity), ref = "WBI")MIX	0.0018	(0.0016)
relevel(factor(ethnicity), ref = "WBI")OAS	−0.0033·	(0.0019)
relevel(factor(ethnicity), ref = "WBI")OBL	-0.0022	(0.0053)
relevel(factor(ethnicity), ref = "WBI")OTH	-0.0036	(0.0034)
relevel(factor(ethnicity), ref = "WBI")PAK	−0.0025*	(0.0013)
relevel(factor(ethnicity), ref = "WBI")WHO	−0.0018·	(0.0010)
factor(region)East of England	0.0003	(0.0010)
factor(region)London	0.0019·	(0.0010)
factor(region)North East	0.0040**	(0.0013)
factor(region)North West	0.0019·	(0.0010)
factor(region)Northern Ireland	0.0001	(0.0012)
factor(region)Scotland	0.0004	(0.0010)
factor(region)South East	0.0009	(0.0009)
factor(region)South West	0.0008	(0.0010)
factor(region)Wales	0.0022·	(0.0011)
factor(region)West Midlands	0.0016	(0.0010)
factor(region)Yorkshire and The Humber	0.0006	(0.0010)
relevel(factor(education_state), ref = "1")0	-0.0010	(0.0015)
relevel(factor(education_state), ref = "1")2	-0.0020	(0.0015)
relevel(factor(education_state), ref = "1")3	0.0004	(0.0016)
relevel(factor(education_state), ref = "1")5	-0.0003	(0.0016)
relevel(factor(education_state), ref = "1")6	0.0012	(0.0015)
relevel(factor(education_state), ref = "1")7	0.0021	(0.0016)
scale(hh_income)	0.0006	(0.0004)
factor(housing_quality)Low	−0.0039***	(0.0007)
factor(housing_quality)Medium	0.0007	(0.0004)
factor(neighbourhood_safety)2	−0.0018***	(0.0004)
factor(neighbourhood_safety)3	−0.0044***	(0.0005)
factor(loneliness)2	0.0199***	(0.0005)
factor(loneliness)3	0.0360***	(0.0008)
scale(nutrition_quality)	−0.0009***	(0.0002)
scale(ncigs)	0.0006**	(0.0002)
factor(ncigs > 0)TRUE	0.0002	(0.0008)
AIC	135261.6364	
BIC	135749.9931	
Log Likelihood	-67576.8182	
Num. obs.	62545	
Num. groups: pidp	33960	
Var: pidp (Intercept)	0.0008	
Var: Residual	0.0012	

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ·$p < 0.1$

Table 8.18: Coefficients for the GLMM Gamma-Normal model predicting SF-12 MCS state.

Variable	Coefficient	Standard Error
(Intercept)	2.9809***	(0.8812)
scale(SF_12_PCS_last)	0.1422***	(0.0011)
scale(SF_12_PCS_last)^2	−0.0385***	(0.0006)
factor(financial_situation)2	−0.0128***	(0.0016)
factor(financial_situation)3	−0.0300***	(0.0021)
factor(financial_situation)4	−0.0424***	(0.0035)
factor(financial_situation)5	−0.0619***	(0.0056)
scale(yearly_energy)	-0.0003	(0.0007)
factor(heating)1	0.0118**	(0.0039)
factor(sex)Male	-0.0004	(0.0014)
relevel(factor(ethnicity), ref = "WBI")BAN	0.0135*	(0.0059)
relevel(factor(ethnicity), ref = "WBI")BLA	0.0109*	(0.0053)
relevel(factor(ethnicity), ref = "WBI")BLC	0.0031	(0.0052)
relevel(factor(ethnicity), ref = "WBI")CHI	0.0042	(0.0101)
relevel(factor(ethnicity), ref = "WBI")IND	0.0046	(0.0036)
relevel(factor(ethnicity), ref = "WBI")MIX	0.0048	(0.0050)
relevel(factor(ethnicity), ref = "WBI")OAS	0.0044	(0.0063)
relevel(factor(ethnicity), ref = "WBI")OBL	0.0038	(0.0175)
relevel(factor(ethnicity), ref = "WBI")OTH	0.0134	(0.0112)
relevel(factor(ethnicity), ref = "WBI")PAK	0.0041	(0.0043)
relevel(factor(ethnicity), ref = "WBI")WHO	-0.0017	(0.0032)
scale(age)	−0.0351***	(0.0008)
scale(age)^2	−0.0066***	(0.0007)
factor(region)East of England	0.0010	(0.0032)
factor(region)London	0.0010	(0.0032)
factor(region)North East	−0.0114**	(0.0041)
factor(region)North West	-0.0033	(0.0032)
factor(region)Northern Ireland	−0.0071·	(0.0037)
factor(region)Scotland	-0.0028	(0.0033)
factor(region)South East	0.0021	(0.0030)
factor(region)South West	-0.0019	(0.0032)
factor(region)Wales	−0.0073*	(0.0036)
factor(region)West Midlands	-0.0014	(0.0033)
factor(region)Yorkshire and The Humber	-0.0053	(0.0033)
factor(education_state)1	0.0023	(0.0051)
factor(education_state)2	0.0067**	(0.0021)
factor(education_state)3	0.0142***	(0.0026)
factor(education_state)5	0.0046·	(0.0027)
factor(education_state)6	0.0167***	(0.0022)
factor(education_state)7	0.0165***	(0.0025)
factor(housing_quality)Low	-0.0023	(0.0023)
factor(housing_quality)Medium	-0.0024	(0.0015)
scale(hh_income)	0.0018**	(0.0007)
scale(nutrition_quality)	0.0036***	(0.0006)
factor(ncigs > 0)TRUE	−0.0127***	(0.0022)
factor(loneliness)2	−0.0125***	(0.0015)
factor(loneliness)3	−0.0222***	(0.0027)
factor(active)1	0.0369***	(0.0014)
factor(auditc)Increased Risk	0.0072*	(0.0031)
factor(auditc)Low Risk	0.0026	(0.0030)
factor(auditc)Non-drinker	−0.0234***	(0.0033)
time	0.0005	(0.0004)
AIC	-45415.4190	
BIC	-44918.1730	
Log Likelihood	22761.7095	
Num. obs.	73737	
Num. groups: pidp	29598	
Var: pidp (Intercept)	0.0000	
Var: Residual	0.0314	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.19: Coefficients for the GLMM Gamma-Normal model predicting SF-12 PCS state.

	Estimate	Std.Err	z-value	p-value
(Intercept)	1.9194	0.0669	28.6818	¡ 1e-04
scale(ncigs)	0.0961	0.0036	26.5540	¡ 1e-04
scale(age)	0.0593	0.0101	5.8739	¡ 1e-04
scale(nutrition_quality)	-0.0523	0.0072	-7.2552	¡ 1e-04
I(scale(age)**2)	-0.1041	0.0088	-11.8028	¡ 1e-04
factor(sex)Male	0.0622	0.0167	3.7180	0.00020078
factor(ethnicity)BLA	-0.3481	0.1045	-3.3316	0.00086335
factor(ethnicity)BLC	-0.3325	0.0837	-3.9727	¡ 1e-04
factor(ethnicity)CHI	-0.1077	0.2024	-0.5322	0.59455878
factor(ethnicity)IND	-0.1087	0.0895	-1.2145	0.22454798
factor(ethnicity)MIX	-0.1567	0.0819	-1.9122	0.05584617
factor(ethnicity)OAS	-0.1444	0.1102	-1.3108	0.18993542
factor(ethnicity)OBL	-0.2191	0.1794	-1.2216	0.22185946
factor(ethnicity)OTH	0.1465	0.1418	1.0332	0.30150037
factor(ethnicity)PAK	-0.1220	0.0829	-1.4709	0.14131358
factor(ethnicity)WBI	0.2484	0.0661	3.7578	0.00017140
factor(ethnicity)WHO	0.1261	0.0735	1.7160	0.08616097
scale(time)	-0.0206	0.0050	-4.0884	¡ 1e-04
Zero-part coefficients:				
	Estimate	Std.Err	z-value	p-value
(Intercept)	2.4947	0.1140	21.8909	¡ 1e-04
scale(ncigs)	-2.4952	0.0214	-116.7090	¡ 1e-04
scale(age)	0.3486	0.0165	21.1822	¡ 1e-04
factor(sex)Male	-0.0812	0.0317	-2.5569	0.0105606
factor(ethnicity)BLA	0.2174	0.1605	1.3549	0.1754518
factor(ethnicity)BLC	-0.5890	0.1417	-4.1561	¡ 1e-04
factor(ethnicity)CHI	0.2085	0.2599	0.8024	0.4223353
factor(ethnicity)IND	0.6569	0.1465	4.4850	¡ 1e-04
factor(ethnicity)MIX	-0.2856	0.1424	-2.0052	0.0449421
factor(ethnicity)OAS	0.2378	0.1794	1.3256	0.1849705
factor(ethnicity)OBL	-0.7735	0.2803	-2.7600	0.0057803
factor(ethnicity)OTH	0.0280	0.2675	0.1045	0.9167803
factor(ethnicity)PAK	0.1538	0.1396	1.1020	0.2704521
factor(ethnicity)WBI	0.2875	0.1139	2.5239	0.0116055
factor(ethnicity)WHO	0.0244	0.1308	0.1862	0.8522509

Table 8.20: Mixed ZIP model coefficients for predicting cigarette consumption (ncigs).

	Estimate	Std.Err	z-value	p-value
(Intercept)	0.4136	0.0056	74.4578	¡ 1e-04
scale(ncars)	0.1781	0.0010	180.3037	¡ 1e-04
scale(age)	-0.0396	0.0015	-27.0737	¡ 1e-04
I(scale(age)**2)	-0.0278	0.0012	-23.8676	¡ 1e-04
factor(region)East of England	-0.0052	0.0058	-0.8989	0.3686794
factor(region)London	-0.1361	0.0056	-24.3796	¡ 1e-04
factor(region)North East	-0.0642	0.0077	-8.2844	¡ 1e-04
factor(region)North West	-0.0513	0.0057	-8.9687	¡ 1e-04
factor(region)Northern Ireland	0.0183	0.0068	2.6907	0.0071292
factor(region)Scotland	-0.0579	0.0061	-9.4365	¡ 1e-04
factor(region)South East	0.0057	0.0054	1.0641	0.2872656
factor(region)South West	0.0014	0.0059	0.2319	0.8166224
factor(region)Wales	-0.0109	0.0065	-1.6826	0.0924486
factor(region)West Midlands	-0.0045	0.0059	-0.7671	0.4430007
factor(region)Yorkshire and The Humber	-0.0555	0.0060	-9.3137	¡ 1e-04
scale(hh_income)	0.0239	0.0007	32.9906	¡ 1e-04
factor(S7_labour_state)FT Education	0.0938	0.0049	18.9917	¡ 1e-04
factor(S7_labour_state)FT Employed	0.0490	0.0037	13.2562	¡ 1e-04
factor(S7_labour_state)Job Seeking	0.0257	0.0047	5.4688	¡ 1e-04
factor(S7_labour_state)Not Working	-0.0114	0.0041	-2.8053	0.0050269
factor(S7_labour_state)PT Employed	0.0406	0.0038	10.6820	¡ 1e-04
Zero-part coefficients:				
	Estimate	Std.Err	z-value	p-value
(Intercept)	-5.0250	0.0405	-124.1040	¡ 1e-04
scale(ncars)	-4.7389	0.0202	-234.3816	¡ 1e-04
scale(hh_income)	-0.2723	0.0134	-20.3340	¡ 1e-04
factor(region)East of England	-0.1364	0.0507	-2.6904	0.0071367
factor(region)London	0.3554	0.0432	8.2285	¡ 1e-04
factor(region)North East	0.1650	0.0568	2.9075	0.0036428
factor(region)North West	0.0618	0.0466	1.3272	0.1844295
factor(region)Northern Ireland	0.1127	0.0567	1.9856	0.0470766
factor(region)Scotland	0.0439	0.0486	0.9035	0.3662545
factor(region)South East	-0.0599	0.0477	-1.2569	0.2087952
factor(region)South West	-0.1298	0.0514	-2.5224	0.0116550
factor(region)Wales	-0.0283	0.0548	-0.5166	0.6054390
factor(region)West Midlands	-0.0127	0.0500	-0.2532	0.8001192
factor(region)Yorkshire and The Humber	-0.0075	0.0488	-0.1530	0.8783952
scale(age)	0.1099	0.0090	12.2219	¡ 1e-04

Table 8.21: Mixed ZIP model coefficients for predicting number of cars per household (ncars).

Variable	Coefficient
scale(age)	-0.2248
scale(age)^2	-0.1394
factor(sex)Male	-0.0295
relevel(factor(ethnicity), ref = "WBI")BAN	-0.4817
relevel(factor(ethnicity), ref = "WBI")BLA	0.1386
relevel(factor(ethnicity), ref = "WBI")BLC	-0.5838
relevel(factor(ethnicity), ref = "WBI")CHI	-0.7916
relevel(factor(ethnicity), ref = "WBI")IND	-0.2748
relevel(factor(ethnicity), ref = "WBI")MIX	0.1365
relevel(factor(ethnicity), ref = "WBI")OAS	-0.0129
relevel(factor(ethnicity), ref = "WBI")OBL	0.7082
relevel(factor(ethnicity), ref = "WBI")OTH	-0.6214
relevel(factor(ethnicity), ref = "WBI")PAK	-1.1716
relevel(factor(ethnicity), ref = "WBI")WHO	-0.1305
factor(region)East of England	-0.0959
factor(region)London	-0.2479
factor(region)North East	-0.5194
factor(region)North West	-0.3514
factor(region)Northern Ireland	-1.0225
factor(region)Scotland	-0.2268
factor(region)South East	0.0175
factor(region)South West	-0.1290
factor(region)Wales	-0.3684
factor(region)West Midlands	0.1039
factor(region)Yorkshire and The Humber	-0.2772
relevel(factor(education_state), ref = "1")0	-0.2064
relevel(factor(education_state), ref = "1")2	-0.1555
relevel(factor(education_state), ref = "1")3	-0.1723
relevel(factor(education_state), ref = "1")5	-0.1154
relevel(factor(education_state), ref = "1")6	-0.2117
relevel(factor(education_state), ref = "1")7	-0.1687
factor(housing_quality)Medium	1.1274
factor(housing_quality)High	4.9240
factor(neighbourhood_safety)2	0.1847
factor(neighbourhood_safety)3	-0.0898
factor(loneliness)2	-0.0858
factor(loneliness)3	-0.3622
scale(ncigs)	-0.0160
scale(hh_income)	0.2309
scale(hh_income)^2	-0.0075
scale(hh_income_diff)	-0.1180
factor(housing_tenure)2	-0.1382
factor(housing_tenure)3	-0.2776
factor(housing_tenure)4	-0.0508
factor(housing_tenure)5	-0.4422
factor(housing_tenure)6	-0.2680
factor(housing_tenure)7	-0.8326
scale(yearly_energy)	0.1418
factor(heating)1	22.9731
Low—Medium	21.6068
Medium—High	25.3859
AIC	805.2328
BIC	1028.5727
Log Likelihood	-351.6164
Num. obs.	589

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.22: Coefficients for the CLM model predicting housing quality.

Variable	Coefficient	Standard Error
scale(age)	-0.1708	(0.1248)
factor(sex)Male	-0.2036	(0.2040)
relevel(factor(ethnicity), ref = "WBI")BAN	-0.0012	(1.5184)
relevel(factor(ethnicity), ref = "WBI")BLA	0.3150	(0.8171)
relevel(factor(ethnicity), ref = "WBI")BLC	0.0680	(1.1368)
relevel(factor(ethnicity), ref = "WBI")CHI	-0.6896	(1.4802)
relevel(factor(ethnicity), ref = "WBI")IND	-0.2224	(0.6391)
relevel(factor(ethnicity), ref = "WBI")MIX	-0.0788	(0.7630)
relevel(factor(ethnicity), ref = "WBI")OAS	0.0218	(0.9247)
relevel(factor(ethnicity), ref = "WBI")OBL	0.4664	(3.0298)
relevel(factor(ethnicity), ref = "WBI")OTH	-0.1663	(1.5027)
relevel(factor(ethnicity), ref = "WBI")PAK	-0.1616	(0.8476)
relevel(factor(ethnicity), ref = "WBI")WHO	0.0744	(0.4247)
factor(region)East of England	-0.0294	(0.4560)
factor(region)London	-0.0714	(0.4685)
factor(region)North East	0.1878	(0.5834)
factor(region)North West	-0.1359	(0.4630)
factor(region)Northern Ireland	-0.1897	(0.7167)
factor(region)Scotland	-0.1241	(0.4973)
factor(region)South East	-0.1215	(0.4255)
factor(region)South West	-0.1162	(0.4765)
factor(region)Wales	0.0439	(0.6084)
factor(region)West Midlands	-0.0561	(0.4789)
factor(region)Yorkshire and The Humber	0.0032	(0.4701)
relevel(factor(education_state), ref = "1")0	0.0380	(0.8814)
relevel(factor(education_state), ref = "1")2	0.1516	(0.8755)
relevel(factor(education_state), ref = "1")3	-0.0294	(0.9040)
relevel(factor(education_state), ref = "1")5	0.2046	(0.9096)
relevel(factor(education_state), ref = "1")6	0.1961	(0.8814)
relevel(factor(education_state), ref = "1")7	0.2231	(0.8962)
factor(housing_quality)Medium	-0.0982	(0.2692)
factor(housing_quality)High	-0.2390	(0.2968)
factor(neighbourhood_safety)2	-0.0993	(0.2246)
factor(neighbourhood_safety)3	-0.2060	(0.3075)
factor(loneliness)2	2.3754***	(0.2159)
factor(loneliness)3	3.9475***	(0.3681)
scale(ncigs)	0.0499	(0.0837)
relevel(factor(job_sec), ref = "3")0	0.0558	(0.3959)
relevel(factor(job_sec), ref = "3")1	-0.0535	(0.5315)
relevel(factor(job_sec), ref = "3")2	0.2234	(0.3892)
relevel(factor(job_sec), ref = "3")4	0.0655	(0.3222)
relevel(factor(job_sec), ref = "3")5	0.0026	(0.3899)
relevel(factor(job_sec), ref = "3")6	0.0143	(0.4647)
relevel(factor(job_sec), ref = "3")7	0.0653	(0.3223)
relevel(factor(job_sec), ref = "3")8	0.1982	(0.4012)
scale(hh_income)	-0.0815	(0.1154)
relevel(factor(marital_status), ref = "Partnered")Separated	0.5216	(0.3594)
relevel(factor(marital_status), ref = "Partnered")Single	0.3119	(0.2640)
relevel(factor(marital_status), ref = "Partnered")Widowed	0.5650	(0.4344)
1—2	1.6825˙	(0.9878)
2—3	4.5372***	(1.0089)
AIC	862.6091	
BIC	1085.9490	
Log Likelihood	-380.3045	
Num. obs.	589	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.23: Coefficients for the CLM model predicting loneliness.

Variable	Coefficient	Standard Error
scale(age)	0.2812**	(0.1054)
factor(sex)Male	0.1706	(0.1959)
relevel(factor(ethnicity), ref = "WBI")BAN	-0.7438	(1.6251)
relevel(factor(ethnicity), ref = "WBI")BLA	0.4150	(0.9898)
relevel(factor(ethnicity), ref = "WBI")BLC	0.1459	(1.1231)
relevel(factor(ethnicity), ref = "WBI")CHI	-0.6952	(1.4875)
relevel(factor(ethnicity), ref = "WBI")IND	-0.0314	(0.6394)
relevel(factor(ethnicity), ref = "WBI")MIX	-0.1418	(0.8538)
relevel(factor(ethnicity), ref = "WBI")OAS	-0.1555	(0.8887)
relevel(factor(ethnicity), ref = "WBI")OBL	1.0637	(3.3419)
relevel(factor(ethnicity), ref = "WBI")OTH	0.2350	(1.5297)
relevel(factor(ethnicity), ref = "WBI")PAK	-0.3663	(0.9656)
relevel(factor(ethnicity), ref = "WBI")WHO	-0.0368	(0.4680)
factor(region)East of England	-0.1369	(0.4591)
factor(region)London	$-0.7911^{\cdot}$	(0.4606)
factor(region)North East	-0.1222	(0.5891)
factor(region)North West	-0.3547	(0.4631)
factor(region)Northern Ireland	0.2921	(0.7587)
factor(region)Scotland	0.1308	(0.4933)
factor(region)South East	-0.1219	(0.4234)
factor(region)South West	0.1371	(0.4730)
factor(region)Wales	-0.1150	(0.6246)
factor(region)West Midlands	-0.2965	(0.4784)
factor(region)Yorkshire and The Humber	-0.1242	(0.4758)
relevel(factor(education_state), ref = "1")0	-0.2482	(0.7524)
relevel(factor(education_state), ref = "1")2	-0.2947	(0.7488)
relevel(factor(education_state), ref = "1")3	-0.1637	(0.7799)
relevel(factor(education_state), ref = "1")5	-0.1285	(0.7892)
relevel(factor(education_state), ref = "1")6	-0.1705	(0.7552)
relevel(factor(education_state), ref = "1")7	-0.2064	(0.7704)
factor(housing_quality)Medium	0.2524	(0.3728)
factor(housing_quality)High	0.4124	(0.3912)
factor(neighbourhood_safety)2	1.2465***	(0.2363)
factor(neighbourhood_safety)3	2.3511***	(0.3091)
factor(loneliness)2	-0.1043	(0.2222)
factor(loneliness)3	-0.1387	(0.3953)
scale(nutrition_quality)	0.0298	(0.0990)
scale(ncigs)	-0.0224	(0.0916)
scale(hh_income)	0.0365	(0.1042)
1—2	-0.0245	(0.8898)
2—3	2.7996**	(0.9011)
AIC	838.5405	
BIC	1004.8488	
Log Likelihood	-378.2703	
Num. obs.	427	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.24: Coefficients for the CLM model predicting neighbourhood safety.

Variable	Coefficient	Standard Error
factor(financial_situation)2	2.3668***	(0.2298)
factor(financial_situation)3	4.0825***	(0.3194)
factor(financial_situation)4	5.5143***	(0.4993)
factor(financial_situation)5	6.6110***	(0.8535)
scale(age)	-0.0948	(0.1235)
factor(sex)Male	0.0411	(0.1790)
relevel(factor(ethnicity), ref = "WBI")BAN	0.1003	(1.3689)
relevel(factor(ethnicity), ref = "WBI")BLA	0.6068	(0.7338)
relevel(factor(ethnicity), ref = "WBI")BLC	0.2248	(1.0048)
relevel(factor(ethnicity), ref = "WBI")CHI	0.7749	(1.1435)
relevel(factor(ethnicity), ref = "WBI")IND	0.3383	(0.5472)
relevel(factor(ethnicity), ref = "WBI")MIX	0.1580	(0.7293)
relevel(factor(ethnicity), ref = "WBI")OAS	0.3146	(0.8678)
relevel(factor(ethnicity), ref = "WBI")OBL	1.1446	(2.6456)
relevel(factor(ethnicity), ref = "WBI")OTH	-0.3149	(1.4012)
relevel(factor(ethnicity), ref = "WBI")PAK	0.1683	(0.7279)
relevel(factor(ethnicity), ref = "WBI")WHO	-0.2305	(0.3809)
scale(yearly_energy)	0.0032	(0.0914)
factor(region)East of England	-0.0611	(0.3995)
factor(region)London	-0.2977	(0.4193)
factor(region)North East	0.1319	(0.5219)
factor(region)North West	0.1019	(0.4017)
factor(region)Northern Ireland	-0.3981	(0.6421)
factor(region)Scotland	-0.1743	(0.4378)
factor(region)South East	-0.0122	(0.3711)
factor(region)South West	-0.0742	(0.4154)
factor(region)Wales	0.1574	(0.5345)
factor(region)West Midlands	-0.0341	(0.4226)
factor(region)Yorkshire and The Humber	-0.1104	(0.4123)
relevel(factor(education_state), ref = "1")0	-0.3068	(0.7332)
relevel(factor(education_state), ref = "1")2	-0.1643	(0.7275)
relevel(factor(education_state), ref = "1")3	-0.0885	(0.7528)
relevel(factor(education_state), ref = "1")5	-0.2053	(0.7597)
relevel(factor(education_state), ref = "1")6	-0.2439	(0.7351)
relevel(factor(education_state), ref = "1")7	-0.1852	(0.7490)
factor(housing_quality)Medium	-0.3437	(0.2415)
factor(housing_quality)High	-0.4175	(0.2689)
factor(neighbourhood_safety)2	-0.0404	(0.2028)
factor(neighbourhood_safety)3	-0.1122	(0.2701)
factor(loneliness)2	0.1390	(0.1922)
factor(loneliness)3	0.3427	(0.3221)
relevel(factor(job_sec), ref = "3")0	-0.0119	(0.3561)
relevel(factor(job_sec), ref = "3")1	-0.1754	(0.4589)
relevel(factor(job_sec), ref = "3")2	-0.1786	(0.3460)
relevel(factor(job_sec), ref = "3")4	0.0892	(0.2844)
relevel(factor(job_sec), ref = "3")5	-0.0331	(0.3356)
relevel(factor(job_sec), ref = "3")6	-0.0137	(0.3922)
relevel(factor(job_sec), ref = "3")7	0.1121	(0.2848)
relevel(factor(job_sec), ref = "3")8	-0.0269	(0.3583)
scale(hh_income)	-0.1884	(0.1274)
factor(marital_status)Separated	0.0750	(0.3333)
factor(marital_status)Single	-0.1083	(0.2487)
factor(marital_status)Widowed	-0.2673	(0.4147)
factor(housing_tenure)2	0.2204	(0.2368)
factor(housing_tenure)3	0.5296	(0.3649)
1—2	0.3062	(0.8534)
2—3	3.5655***	(0.8694)
3—4	5.9586***	(0.8956)
4—5	7.6424***	(0.9515)
AIC	1191.6362	
BIC	1476.2852	
Log Likelihood	-530.8181	
Num. obs.	589	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.25: CLM model coefficients estimating financial situation.

Variable	Coefficient
factor(behind_on_bills)2	2.8410
factor(behind_on_bills)3	4.9152
scale(yearly_energy)	0.0729
factor(ethnicity)BLA	0.8506
factor(ethnicity)BLC	0.3087
factor(ethnicity)CHI	-20.4225
factor(ethnicity)IND	-0.6275
factor(ethnicity)MIX	0.2611
factor(ethnicity)OAS	-0.9071
factor(ethnicity)OBL	0.3489
factor(ethnicity)OTH	-0.9696
factor(ethnicity)PAK	-0.2078
factor(ethnicity)WBI	-0.4910
factor(ethnicity)WHO	-0.2663
scale(age)	-0.3074
scale(SF_12)	-0.2621
scale(hh_income)	-0.7284
scale(hh_income)^2	0.0801
factor(financial_situation)2	0.4053
factor(financial_situation)3	1.0077
factor(financial_situation)4	1.7242
factor(financial_situation)5	0.9741
1—2	4.3947
2—3	7.7137
AIC	197.9706
BIC	302.7281
Log Likelihood	-74.9853
Num. obs.	581

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.26: Coefficients for the CLM model predicting if a household feels subjectively behind on bills.

Variable	$Coefficient$
(Intercept)	-26.2290
	(67860.6075)
factor(heating)1	52.4447
	(65066.4820)
scale(yearly_energy)	-0.0017
	(12939.4610)
factor(financial_situation)2	0.0009
	(28656.9391)
factor(financial_situation)3	0.0036
	(38064.7550)
factor(financial_situation)4	0.0054
	(66660.8059)
factor(financial_situation)5	0.0110
	(118895.9082)
factor(behind_on_bills)2	-0.0011
	(67459.9702)
factor(behind_on_bills)3	-0.0154
	(221759.7311)
scale(net_hh_income)	-0.0008
	(13213.7715)
AIC	20.0000
BIC	93.4443
Log Likelihood	-0.0000
Deviance	0.0000
Num. obs.	11319

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.27: Coefficients for the Logistic Regression model predicting subjective thermal comfort (heating).

Variable	Coefficient	Standard Error
(Intercept)	-1.1135	(1.5051)
scale(age)	-0.1328	(0.1134)
factor(sex)Male	0.5014**	(0.1945)
factor(education_state)1	0.3114	(0.8049)
factor(education_state)2	0.1212	(0.3261)
factor(education_state)3	0.2094	(0.3865)
factor(education_state)5	0.2092	(0.4182)
factor(education_state)6	0.2490	(0.3416)
factor(education_state)7	0.4083	(0.3721)
factor(financial_situation)2	0.0518	(0.2343)
factor(financial_situation)3	-0.0534	(0.3092)
factor(financial_situation)4	0.0414	(0.5346)
factor(financial_situation)5	0.3618	(0.8430)
factor(behind_on_bills)2	-0.1119	(0.5248)
factor(behind_on_bills)3	-0.3596	(1.8329)
scale(yearly_energy)	0.0411	(0.0992)
scale(SF_12)	0.1706	(0.1043)
factor(heating)1	-0.3118	(0.5666)
scale(SF_12_PCS)	0.4938***	(0.1262)
scale(hh_income)	0.0900	(0.1094)
factor(ethnicity)BLA	0.5473	(1.5919)
factor(ethnicity)BLC	0.5887	(1.7444)
factor(ethnicity)CHI	-0.6318	(2.0397)
factor(ethnicity)IND	-0.0746	(1.5197)
factor(ethnicity)MIX	0.6944	(1.5711)
factor(ethnicity)OAS	0.4011	(1.5940)
factor(ethnicity)OBL	1.0592	(3.3683)
factor(ethnicity)OTH	0.1917	(1.9554)
factor(ethnicity)PAK	-0.0215	(1.6029)
factor(ethnicity)WBI	0.6128	(1.3822)
factor(ethnicity)WHO	0.7418	(1.4352)
AIC	67.2579	
BIC	296.7721	
Log Likelihood	-2.6290	
Deviance	616.9354	
Num. obs.	11993	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table 8.28: Coefficients for the CLM model predicting physical activity.