

Sonthalia, Shreya (2026) *A system dynamics model of the labour market, gender norms, and mental health in Scotland.* PhD thesis.

https://theses.gla.ac.uk/86380/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

A System Dynamics Model of the Labour Market, Gender Norms, and Mental Health in Scotland

Shreya Sonthalia

BA, MSc

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Health and Wellbeing

College of Medical, Veterinary and Life Sciences

University of Glasgow

July 2026

Abstract

Background: The bidirectional relationship between the labour market and mental health is well established. Yet this evidence derives largely from individual risk factors studied in isolation; how these pathways aggregate into a population's prevalence of poor mental health over decades is far less understood. In Scotland, the prevalence of poor mental health among the working-age population has risen over recent decades, and women's prevalence has remained consistently higher than men's despite progress on gender equality indicators such as the employment rate and the gender pay gap.

Understanding this trajectory, this thesis argues, requires focusing on the population-level interaction between these dynamic conditions: the gendered structure of work (the gender pay gap and unpaid work); the labour market (insecure work and unemployment); and population mental health. This raises three questions for Scotland: how reciprocal feedback between labour market conditions and population mental health has contributed to the rise between 1992 and 2019; why the gender gap has persisted; and which mechanisms dominate the system and can serve as leverage points for policy.

Methods: The hypothesised causal structure, a system of interacting feedback loops, was assembled through an iterative literature review, in which identifying one causal relationship triggered a search for its reverse effect and any intermediate variables enabling the mapping of structure. The synthesis was expressed as a causal loop diagram, a system dynamics tool for capturing feedback mechanisms. It was refined through disconfirmatory interviews with ten interdisciplinary experts, designed to challenge the structure.

Grounded in a complex realist philosophy that pursues explanatory verisimilitude rather than prediction, this conceptual model was then translated into a system dynamics simulation model to test whether the hypothesised structure could reproduce the trends observed between 1992–2019. Parameterisation prioritised published evidence and data, and the model was validated through structural and behavioural tests and sensitivity analyses. It was then analysed to identify which feedback loops drive behaviour over time, key contributors to the gender gap, and the policy combinations that most improve outcomes.

Results: The model reproduced the observed trends endogenously, indicating that the structure can generate the observed trajectory. However, the magnitude of the rise in women's prevalence post-2016 is lower than observed, pointing to forces beyond the modelled structure, notably the welfare system.

The rise in poor mental health owed less to the demand shock of the 2008 recession than to the gradual rise of insecure work that preceded it. This rise quietly eroded the system's capacity to move people out of poor mental health. Consistent with this, counterfactual analysis found that holding insecure work at its 1992 level reduced the increase in poor mental health by more than removing the demand shock of the 2008 recession.

The persistent gender gap is primarily driven by higher incidence among women, arising from their structural position, operationalised in the model through the gender pay gap and unpaid work. Structural decomposition of the gender gap attributed a smaller part to women's slower recovery.

Levers focused on equality narrow the gender gap in prevalence, though childcare and norms levers do so partly by raising men's prevalence. This highlights a redistributive trap. Reducing insecure work lowers prevalence for both women and men but can widen the gap, because, given how the model specifies the effect of insecure work on recovery, men benefit more. The levers therefore act best in combination; expanding childcare alongside reducing insecure work reproduces almost the whole effect that deploying all four levers achieves.

Discussion: The thesis makes four contributions. The first is a dynamic, population-level causal structure that combines health causation and health selection within one feedback structure and decouples incidence from recovery. The second is treating gender as a dynamic structural determinant rather than a static covariate. The third is the model as an explanatory thinking tool that highlights key structural stories such as the buffering role of labour-force exit at the population level. The fourth is methodological: the thesis demonstrates system dynamics applied to the upstream social determinants of mental health, where such models remain rare, and offers the feedback-driven review for future researchers to build on.

For understanding population mental health, the modelling framework offers a structural explanation of how the gender gap can persist despite progress on equality indicators. For decision-making, it suggests asking whether a policy acts on incidence or recovery, and when, in what combination of levers, and in what macroeconomic state. Its policy relevance lies in framing poor mental health as a population-level system of interacting feedbacks: labour market, gender equality and childcare policy are core to preventing poor mental health, and they are better judged together.

Table of Contents

Abstract ii
Table of Contents v
List of Tables viii
List of Figures ix
Acknowledgements xi
Author's Declaration xii
Abbreviations xiii
1 Introduction 1
1.1 Overview of the chapter 1
1.2 The public health challenge in Scotland 2
1.3 The labour market as a critical social determinant 5
1.4 The policy and structural context 6
1.5 The need for a systems perspective 10
1.6 Research aims 11
1.7 Research contributions 12
1.8 Guide to the thesis 13
2 Literature Review: Theoretical and Methodological Foundations 15
2.1 Overview of the chapter 15
2.2 Conceptualising the core components 15
2.3 The bidirectional relationship between the labour market and mental health 21
2.4 The gendered dynamics of work and mental health 26
2.5 Organising the evidence: social determinants of mental health framework 30
2.6 The limits of traditional methods for this complex systems problem 33
2.7 Justification for system dynamics 37
2.8 Research questions and objectives 42
3 Methodology 44
3.1 Overview of the chapter 44
3.2 System dynamics: core concepts and typical process 44
3.3 Philosophical underpinnings: complex realism and structuralism 47
3.4 The research design: the modelling process 51
3.5 Methodological limitations 55
4 Methodology for Conceptual Model Development: Literature Review and Expert Consultation 58
4.1 Introduction 58
4.2 Literature review methodology 58

4.3 Expert engagement 67
4.4 Refinement for narrative clarity 70
4.5 Methodological reflections and limitations 72
5 Conceptual Model: Model Boundary and Causal Loop Diagram 75
5.1 Overview of the chapter 75
5.2 Causal loop diagram 75
5.3 Operationalisation of key concepts 83
5.4 Model boundary 92
5.5 Synthesis of feedback from expert consultation 94
5.6 Summary and interpretation 98
5.7 Strengths and limitations 100
6 Model Development 104
6.1 Introduction 104
6.2 Strategic modelling decisions and architecture 105
6.3 Module 1: The labour market 112
6.4 Module 2: Gender dynamics 122
6.5 Module 3: Mental health 136
6.6 Model validation 146
6.7 Summary of model validation 161
6.8 Strengths and limitations 162
7 Model Analysis 167
7.1 Introduction 167
7.2 Examining gender differences 169
7.3 Structural decomposition of the gender gap 178
7.4 Explaining the historical trends in poor mental health 188
7.5 Counterfactual analysis 193
7.6 Addressing the rising and gendered prevalence of poor mental health 206
7.7 Summary of findings 222
7.8 Strengths and limitations 226
8 Discussion and Conclusion 231
8.1 Introduction and summary of findings 231
8.2 Situating the findings in the literature 234
8.3 Contributions of the thesis 239
8.4 Implications for policy 246
8.5 Limitations 251
8.6 Directions for future research 253
8.7 Concluding remarks 256
Appendices 258

Appendix A Search strategy 258
Appendix B Expert engagement documents 260
Appendix C Outputs from expert engagement 268
Appendix D Simplified causal hypothesis for the simulation 274
Appendix E Parameter tables 275
Appendix F Sensitivity-tested parameters 293
Appendix G Parameters providing exogenous input 296
Appendix H Parameterisation via calculation and constrained calibration 301
Appendix I Structural verification tests for the labour market module 314
Appendix J Structural verification of the gender dynamics module 323
Appendix K Structural verification of the mental health module 331
Appendix L Theil's statistics for isolated modules 340
Appendix M Summary of testing conducted on isolated modules 343
Appendix N Model equations 346
Appendix O Calibration data 352
Appendix P Loop dominance analysis 356
Appendix Q Structural decomposition 364
Appendix R Counterfactual scenario analysis 370
Appendix S Single and combined levers 375
Appendix T Optimisation 384
Glossary 393
List of References 395

List of Tables

Table 4-1: Inclusion, exclusion and prioritisation criteria 63
Table 5-1: Feedback loops in the CLD 77
Table 5-2: Translation of concepts to model variables 84
Table 5-3: Evidence for causal relationships in CLD 86
Table 6-1: Exclusion and aggregation details for simulation model 107
Table 6-2: Sensitivity of model parameters 111
Table 6-3: Summary of parameters in labour market module 118
Table 6-4: Summary of parameters in the gender dynamics module 133
Table 6-5: Summary of parameters in the mental health module 143
Table 6-6: Sensitivity of the 2019 prevalence of poor mental health to the prevalence-recovery mechanism, by gender 145
Table 6-7: Behaviour reproduction statistics 160
Table 7-1: Summary of findings 168
Table 7-2: Guide to the feedback loops analysed in this chapter 172
Table 7-3: Gender difference in loop dominance 176
Table 7-4: Decomposition of the gender gap in poor mental health 187
Table 7-5: Summary of structural decomposition runs (D1 to D4) 188
Table 7-6: Counterfactual analysis: scenarios and results 195
Table 7-7: Summary of counterfactual analysis 205
Table 7-8: Optimisation parameters 210
Table 7-9: Effect of each lever at 2040 relative to no intervention 210
Table 7-10: Each lever's 2040 effect across the nine macroeconomic futures 211
Table 7-11: Difference in gender gap and total prevalence in base case under different optimised objectives 217

List of Figures

Figure 1-1: Trends in poor mental health in Scotland, 1991-2022 3
Figure 3-1: Example of a causal loop diagram (CLD) 45
Figure 3-2: Example stock and flow structure 46
Figure 3-3: Modelling Process 51
Figure 3-4: CLD development process 53
Figure 4-1: Literature review process. 60
Figure 4-2: Screening and prioritisation process 61
Figure 4-3: PRISMA flow chart 62
Figure 5-1: Causal loop diagram (CLD) of the hypothesised labour market–gender norms–mental health system in Scotland (1992–2019) 76
Figure 6-1: Model architecture options: separate and unified sub-systems 105
Figure 6-2: The labour market macroeconomy sector 113
Figure 6-3: The labour market stock-and-flow structure 115
Figure 6-4: Sensitivity analysis within the labour market module 121
Figure 6-5: Loop knockout analysis testing the structural necessity of B1 (demand saturation). 122
Figure 6-6: The gender dynamics module 124
Figure 6-7: Sensitivity analysis: the resistance factor 126
Figure 6-8: The simplified gender dynamics module 127
Figure 6-9: Formulation of the unpaid work burden and gender parity in working hours 128
Figure 6-10: Sensitivity analysis: effect of childcare availability on gender parity in working hours 129
Figure 6-11: Sensitivity analysis: effect of norms on unpaid work burden 130
Figure 6-12: Sensitivity analysis: effect of gender parity in working hours on gender pay gap 131
Figure 6-13: Sensitivity analysis: effect of gender parity in working hours on gendered exposure to insecure work 132
Figure 6-14: Sensitivity analysis: effect of economic signal weights and time for norms to change 135
Figure 6-15: The mental health module 138
Figure 6-16: Sensitivity analysis: effect of gender norms on perceived unpaid work 140
Figure 6-17: Illustrative sensitivity of prevalence of poor mental health to the women's degradation-shape parameter 146
Figure 6-18: Stock conservation 148
Figure 6-19: Integration method robustness 149
Figure 6-20: Recovery dynamics: persistence of poor mental health after a temporary economic shock 150
Figure 6-21: Effect of stable economic stressors on prevalence of poor mental health 151
Figure 6-22: Effect of gender inequality on prevalence of poor mental health 151
Figure 6-23: Loop knockout analysis testing the structural necessity of R1 (unemployment-mental health trap) 152
Figure 6-24: Loop knockout analysis testing the structural necessity of R2 (insecure work-mental health trap) 153
Figure 6-25: Loop knockout analysis testing the structural necessity of R3 (norm reinforcement) 153
Figure 6-26: Loop knockout analysis testing the structural necessity of R4 and B4 (norms–mental health loops) 154

Figure 6-27: Behaviour reproduction graphs 157
Figure 6-28: Effect of exogenous shock in onset and recovery 158
Figure 6-29: Prevalence of poor mental health simulated in the full model compared to the isolated module..... 159
Figure 7-1: Net flow, incidence and recovery among men and women, 1992-2019 171
Figure 7-2: Effect of differential gender sensitivity on the gender gap in the prevalence of poor mental health 179
Figure 7-3: Effect of gender-specific stressors on the gender gap in prevalence of poor mental health 180
Figure 7-4: Prevalence of poor mental health for women: base case vs a counterfactual in which gender inequality does not narrow over 1992–2019 ... 181
Figure 7-5: Overall effect of differential gender effects and additional stressors on the gender gap in prevalence of poor mental health..... 182
Figure 7-6: Women's prevalence of poor mental health across decompositions D1-D3 184
Figure 7-7: Residual gap attributable to differential exposure on the gender gap in prevalence of poor mental health 185
Figure 7-8: Loop dominance by year 189
Figure 7-9: Estimated prevalence of poor mental health under different counterfactual labour market scenarios 196
Figure 7-10: Estimated prevalence of poor mental health under alternative gender pay gaps 198
Figure 7-11: Estimated prevalence of poor mental health under alternative gender norms 199
Figure 7-12: Estimated prevalence of poor mental health under different gender norms and gender pay gap scenarios 200
Figure 7-13: Estimated prevalence of poor mental health under different norms and gender pay gap scenarios 201
Figure 7-14: Estimated prevalence of poor mental health under more egalitarian norms alongside stable job quality (insecure work) 202
Figure 7-15: Estimated prevalence of poor mental health when severing and doubling the effect of prevalence on labour market outcomes 203
Figure 7-16: Results of single objective optimisations 214
Figure 7-17: The structural floor across macroeconomic futures..... 219

Acknowledgements

It is a pleasure to thank the many people who helped me along the way.

I am deeply grateful to my supervisory team for their guidance. To Petra Meier, thank you for being such a grounding presence: generous with your guidance, and unfailingly understanding when life beyond the thesis made the work hard. To Corinna Elsenbroich, thank you for your patience, and for the time you so readily gave to listening and helping me make sense of my own thoughts. To Emma Comrie, thank you for your thoughtful comments and your crucial help in recruiting the experts. And to Mohammad Jalali, your critical and constructive feedback made this research considerably stronger.

My thanks go to my annual reviewers, Vittal Katikireddi and Laurence Moore, whose encouragement and practical advice helped me keep my eyes on the finish. A special thank you to Valerie Wells for refining the search strategy with me, and to Mary-Kate Hannah for her guidance on data management and ethics. I'm grateful, too, to those who offered comments at various points along the way, whose questions and suggestions sharpened the work. To my colleagues on the systems team, thank you for creating a true community despite the distance, and for providing a weekly intellectual home through these years. To Elizabeth Inyang, thank you for the long conversations that untangled my thinking, and for the reassurance of someone who was in it alongside me.

I want to thank Pascal Gambardella for his advice in the early days, and other advisors and mentors I found through the System Dynamics Society for their well-timed steers. Thank you to Jo-An Occhipinti for so warmly welcoming me into her team in Sydney. I am also deeply grateful to the experts who generously gave their time. Though you remain anonymous here, your real-world insight was essential. This research was made possible by the financial support of the Medical Research Council (MRC).

Finally, to Hobbes and Chutki, for keeping life in perspective with their unbothered naps and reassuring purrs. Thank you for always knowing how I was, and for never once asking how the PhD was going.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work. This thesis has not been submitted for any other degree at the University of Glasgow or any other institution.

Shreya Sonthalia

Abbreviations

ASHE	Annual Survey of Hours and Earnings
BHPS	British Household Panel Survey
CIS-R	Revised Clinical Interview Schedule
CLD	Causal Loop Diagram
CMD	Common Mental Disorder
DT	Delta Time (time step for integration)
DWP	Department for Work and Pensions
GHQ-12	General Health Questionnaire, 12-item version
GPG	Gender Pay Gap
IMF	International Monetary Fund
LFS	Labour Force Survey
LMT	Labour Market Tightness
LTM	Loops That Matter (built-in diagnostic tool of Stella Architect that analyses loop dominance)
MH	Mental Health
MMBM	Modified Male Breadwinner Model
OECD	Organisation for Economic Co-operation and Development
ONS	Office for National Statistics
PAF	Population Attributable Fraction
RK4	Runge-Kutta 4th order integration
RMSE	Root Mean Square Error
SD	System Dynamics
SDH	Social Determinants of Health
SF-12 MCS	12-item Short Form Survey, Mental Component Summary
SHeS	Scottish Health Survey
SMTH3	Third-order smoothing function
UKHLS	United Kingdom Household Longitudinal Study
WAP	Working-Age Population
WEGo	Wellbeing Economy Governments
WEMWBS	Warwick–Edinburgh Mental Wellbeing Scale
WHO	World Health Organization
ZHC	Zero-Hours Contract

1 Introduction

1.1 Overview of the chapter

This thesis investigates the rising prevalence of poor mental health among working-age adults in Scotland, and its persistent gender disparity. In the last decade, the share of Scotland's working-age population reporting common mental disorders has increased substantially. The gap between women and men (with women consistently reporting higher prevalence) has not narrowed, despite measurable progress on indicators of gender equality such as the employment rate and the gender pay gap. The aim of this thesis is to develop and test a dynamic causal explanation of the structural mechanisms that have produced this trajectory.

The existing evidence on work, gender, and mental health is substantial but often fragmented: epidemiology has documented the effects of unemployment and precarious employment (which this thesis models as insecure work) on individual mental health and vice-versa; economists have examined the changing structure of the UK labour market; sociology has theorised the role of norms and the gendered distribution of paid and unpaid labour. However, this thesis argues that the trajectory observed in Scotland is not the product of these factors operating in isolation; rather, it emerges from their complex, dynamic interaction at the population level over time. Thus, this thesis adopts a systems perspective and develops a system dynamics model exploring the dynamic relationship between gender norms, the labour market and population mental health in Scotland.

The purpose of this chapter is to establish the motivation for the research. It sets out the specific public health challenge the research focuses on, the policy and structural context within which it has arisen, and the analytical gap the thesis proposes to fill. A detailed review of the theoretical and methodological literature is provided in Chapter 2.

Section 1.2 describes the empirical challenge: the rising prevalence of poor mental health in Scotland, its macroeconomic significance, and the persistent gender disparity that this thesis terms the Scottish "gender paradox". Section 1.3

introduces the labour market as a critical social determinant of mental health, summarising the evidence for bidirectional causation between employment and mental health at both individual and population scales. Section 1.4 situates the problem within Scotland's devolved policy landscape and the UK labour market. Section 1.5 explains why a systems perspective is required to explain the observed trajectory. Section 1.6 states the overarching research aims; Section 1.7 outlines the intended research contributions; and Section 1.8 provides a guide to the remainder of the thesis.

1.2 The public health challenge in Scotland

Poor mental health is a significant global public health challenge associated with profound health, social, and economic costs: it limits individuals' ability to live meaningful lives and engage in fulfilling work (World Health Organization, 2022b). Mental health problems can lead to worse physical health, lower life expectancy, increased risk of social exclusion and poor economic outcomes, and higher likelihood of substance misuse and engaging in crime (World Health Organization, 2004). Cumulatively, these consequences create a substantial population burden; mental health disorders contributed to over 17% of the total years lived with disability worldwide in 2021 (GBD 2021 Diseases and Injuries Collaborators, 2024).

This thesis is concerned with the prevalence of common mental disorders (CMDs) that constitute the largest share of population mental ill-health and for which the evidence base on social and economic determinants is most established (OECD, 2023a; World Health Organization & Calouste Gulbenkian Foundation, 2014). Detailed discussion of measurement, and the rationale for using the General Health Questionnaire (GHQ-12) is provided in Chapter 2 (Section 2.2).

In Scotland, despite a policy environment emphasising mental wellbeing, the prevalence of poor mental health has increased significantly over the last two to three decades, particularly among the working-age population (Vriend, Rocks, & Finch, 2025; Zhang et al., 2023). Figure 1-1 shows the prevalence of poor mental health among men and women in Scotland between 1991 and 2022. Further analysis conducted in England suggests that this represents a real increase in prevalence rather than merely improved reporting (Barr, 2021).

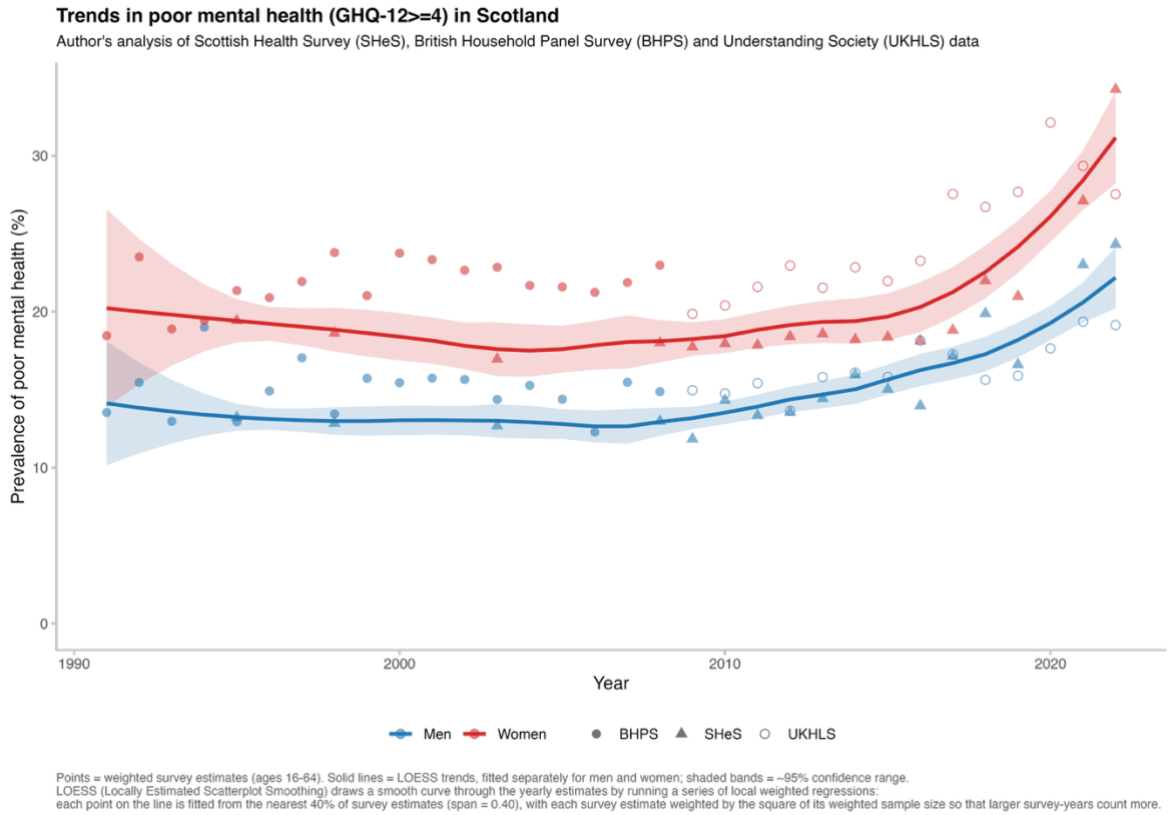


Figure 1-1: Trends in poor mental health in Scotland, 1991-2022

Highlights two patterns among working-age adults (16-64). First, poor mental health has risen since the mid-2010s. Second, poor mental health is consistently more common among women than men, and this gender gap persists throughout. Shaded bands show the approximate 95% confidence range around each trend; the two bands remain separated across the series, indicating the gender difference is unlikely to be an artefact of sampling noise.

This deterioration in population health is not merely a social concern but a macroeconomic issue. An independent review commissioned by the UK Government found that poor mental health costs the UK economy between £74 billion and £99 billion per year; this encompasses direct NHS costs of £3bn and wider costs such as £37-52bn of lost output (Stevenson & Farmer, 2017). Crucially, this deteriorating population mental health trend is increasingly linked to economic inactivity, the share of the working-age population neither employed nor actively seeking work. While economic inactivity comprises several categories (including students, carers, the retired, those with long-term health conditions, and discouraged workers), the UK is currently experiencing a rise in economic inactivity driven by long-term sickness (Powell, 2024). Crucially, over half of those inactive due to illness report experiencing poor mental health (Office for National Statistics, 2023). The International Monetary Fund (IMF) has identified this trend, specifically the post-pandemic decline in labour force participation driven by illness (with the largest increase being due to mental

health), as a "lingering distinctive factor" that sets the UK apart from other advanced economies (Li & Granados, 2023).

It must be noted that official Office for National Statistics (ONS) Labour Force Survey estimates regarding the total scale of economic inactivity are currently the subject of methodological debate. Following the pandemic, survey response rates collapsed, introducing a severe non-response bias that researchers estimate 'lost' nearly a million employed workers from official figures, thereby artificially inflating overall inactivity rates trend (Corlett, 2024). However, while the *overall* aggregate rate may be overstated, based on administrative benefit data, there is general consensus that the *compositional* shift toward inactivity specifically driven by long-term sickness is a genuine, structural trend (Corlett, 2024).

This challenge is particularly acute in Scotland. First, poor health explains a significantly larger share of economic inactivity in Scotland than in the rest of the UK (Randolph, 2024). Second, since 2017 Scotland has recorded higher overall rates of economic inactivity than the UK as a whole; the post-2020 portion of this gap should be read with the LFS non-response caveat noted above, but the pre-pandemic divergence (2017–2019) is robust to that caveat (Randolph, 2024). Reducing economic inactivity is an aim of both the UK and Scottish governments: the UK Government's Get Britain Working White Paper sets a long-term ambition of an 80% employment rate, and the Scottish Government's Programme for Government 2026–2031 commits to reducing inactivity and raising employment (Department for Work and Pensions, 2024; Scottish Government, 2026d).

A critical dimension of this rising inactivity is its unequal distribution between men and women. Long-term sickness has become the most common reason for inactivity among women, underscoring the second dimension of the challenge: the structural persistence of gender disparities (Brett & Macfarlane, 2022). Currently, 29% of inactive women are long-term sick compared to 36% of men (Randolph, 2024).

This gender paradox of mental health is central to this thesis. While Scotland has witnessed significant progress in gender equality, including a narrowing

employment gap (from 10.2 percentage points in 2004 to 5.1 in 2021) and gender pay gap (from 26.5% in 1997 to 11.2% in 2021), gender disparities in poor mental health have remained stubbornly persistent (Scottish Government, 2022c, 2022d, 2024a). The prevalence of poor mental health has been consistently higher among women compared to men in Scotland (Scottish Government, 2023a). This trend is also seen within the rest of the UK and globally (Arias-de la Torre et al., 2021; Global Burden of Disease Mental Disorders Collaborators, 2022). The paradox is that despite substantial advancements in economic status, there has been no corresponding improvement in women's relative wellbeing (Kaiser et al., 2025). This implies that the social determinants responsible for the gap are either insensitive to the particular dimensions on which equality has advanced, or are being counteracted by other dynamics in the labour market. Recent simulation evidence from Canada reinforces this complexity, demonstrating that even if women were assigned men's employment rates and job quality, the gap in clinical distress would narrow only partially while the gap in 'flourishing' mental health would widen (Orchard et al., 2024); the implications of this finding for the structural account of the gender paradox are developed in Section 2.4.

1.3 The labour market as a critical social determinant

Within public health, the Social Determinants of Health (SDH) framework recognises that health outcomes are fundamentally shaped by the "conditions in which people are born, grow, live, work, and age" (World Health Organization, 2008). Among these determinants, the labour market acts as a critical structural factor, and its relationship with mental health has been widely researched (Bambra, 2011; Braveman & Gottlieb, 2014; Frank et al., 2023; Naik et al., 2019).

Evidence highlights the labour market's role in rising mental health problems within the UK. For example, analysis of English local authority attributes approximately one third of the increase in mental health problems since 2009 to increasing unemployment and declining wages (Barr, 2021). This reflects two well-established epidemiological mechanisms of 'health causation' (unemployment and precarious employment as causes of poor mental health) and 'health selection' (poor mental health impacting labour market outcomes),

both of which are reviewed in detail in Chapter 2 (Amiri & Behnezhad, 2019; Paul & Moser, 2009; Pinna Pintor, Fumagalli, & Suhrcke, 2024; Pulford et al., 2022).

Crucially, these impacts extend beyond the individual. Economic conditions, such as local unemployment rates, can influence mental health regardless of an individual's own employment status, and poor mental health can reduce labour productivity through increased presenteeism and absenteeism (Bryan, Bryce, & Roberts, 2021; Bryan, Bryce, & Roberts, 2022; Curtis et al., 2019; Huo et al., 2022). The coexistence of health causation and health selection implies that these mechanisms do not operate in isolation; rather, they aggregate to form dynamic feedback loops where economic conditions and health outcomes reinforce one another over time (O'Halloran & Thomas, 2024). However, there remains a critical gap in understanding how these individual-level pathways aggregate into population-level outcomes. This thesis contends that explaining the persistence of these trends requires treating the labour market and population mental health not as separate, but as interconnected system components whose reciprocal dynamics unfold over time. This framing is the analytical foundation on which the subsequent chapters build.

1.4 The policy and structural context

This section situates the thesis's two empirical concerns — the rising prevalence of poor mental health and its persistent gender disparity — within Scotland's policy landscape. The distinction between reserved and devolved powers is pertinent here, as Scotland's ambitious Wellbeing Economy agenda operates within a UK labour market. Its Fair Work programme and gender equality measures, directed respectively at job quality and the gendered economy, are a distinctive attempt to act on the determinants of these concerns within those constraints, making Scotland a compelling case study.

1.4.1 Scotland’s devolved powers

Scotland has been governed under a devolution settlement since 1999, following the Scotland Act 1998 (Torrance, 2024). The settlement reserves certain policy areas to the UK Parliament and devolves others to the Scottish Parliament. The

Scotland Act 2016 subsequently devolved significant powers over social security, including some Universal Credit-adjacent benefits and disability payments (Mackley, 2020). Most matters directly bearing on the labour market — employment law, the National Minimum Wage, trade-union recognition, immigration, and the Equality Act — remain reserved. Health, education, housing, and most aspects of social care are devolved. This means that the Scottish Government can shape the conditions under which people live, are cared for, and receive benefits, but it cannot legislate directly on the terms under which people are employed. This creates a divergence between the policy ambition and shared UK labour market structure.

1.4.2 The UK labour market: a deregulated backdrop

Over the last four decades, the UK has moved towards one of the least regulated labour markets in the OECD (OECD, n.d.).The collective bargaining system has decentralised and union coverage has fallen significantly (Altunbuken et al., 2022). This power imbalance has fostered a rise in zero-hours and atypical contracts as employers seek to manage fluctuating demand (Slaughter, 2024). Real wage growth has stagnated since the 2008 recession and underemployment has risen (Fraser of Allander Institute, 2022). Insufficient hours have emerged as a key driver of in-work poverty, which remains a significant challenge in Scotland (Birt et al., 2023). These features characterise the labour market that Scotland's working-age population operates within, and which its devolved institutions have limited direct power to reshape.

1.4.3 Scotland's Wellbeing Economy agenda

In parallel with these UK-level trends, the Scottish Government has developed a distinct policy agenda centred on the concept of a Wellbeing Economy, an economic framework that prioritises the collective wellbeing of current and future generations over narrow GDP growth (Scottish Government, 2022a). The agenda builds on the National Performance Framework, first introduced in 2007 and articulates national outcomes including health, equality, and fair work (OECD, 2023b). It was formally cemented when Scotland became a founding member of the Wellbeing Economy Governments partnership (WEGo) in 2018 (Scottish Government, 2025b). The Scottish Government published its National

Strategy for Economic Transformation in 2022 (Scottish Government, 2022a). This ambition was translated into tangible action through a range of policies from Green Industrial Strategy (2024c) to Community Wealth Building Act (2026); the Fair Work Plan (2022b) is the most relevant to this thesis. The Mental Health and Wellbeing Strategy (2023b) subsequently anchored population mental health as a named outcome of this agenda, explicitly linking mental health to fair work and gender equality. More recently, public health and mental health strategy has turned explicitly towards prevention (Scottish Government, 2025a, 2026c). The 2026 Mental Health and Wellbeing Delivery Plan makes prevention and early intervention central and commits to tracking preventative spend (Scottish Government, 2026b).

1.4.4 Fair Work as the delivery framework

The principal vehicle through which the Wellbeing Economy is translated into labour market policy is the Fair Work agenda. The Fair Work Convention, established in 2015, articulated five dimensions of fair work: security, respect, effective voice, opportunity, and fulfilment (Fair Work Convention, 2016). These dimensions correspond closely to the psychosocial functions of work identified in Jahoda's latent deprivation model (Chapter 2), which links them to mental health outcomes. Operationalised through the Fair Work First conditionality regime attached to public procurement and grant funding (Scottish Government, 2024b) and through promotion of the voluntary Real Living Wage (Mignon, 2025), the agenda seeks to improve job quality despite the reservation of employment law. In the same period, fewer employees were paid below the Real Living Wage in Scotland than in the rest of the UK (Mignon, 2025). The Fair Work Action Plan also promotes flexible working as a route to wider participation (Scottish Government, 2022b).

Specific gender-equality interventions complement this agenda. Funded Early Learning and Childcare was expanded to 1,140 hours per year for all 3- and 4-year-olds and for eligible 2-year-olds in 2021, substantially above comparable provision in England at the time (Scottish Government, 2019a). Since then, England's 2025 expansion has extended 38 weeks of thirty funded hours to working parents of children from nine months old, matching the Scottish hours and reaching younger children (Department for Education, 2025). In Scotland,

the Programme for Government 2026–2031 commits to a further expansion providing support 52 weeks a year for children from nine months old to the end of primary school (Scottish Government, 2026d).

The public sector wage floor has been set above the Real Living Wage, with a cash underpin for low-earners (Scottish Government, 2021). These interventions likely contributed to the gender pay gap closing faster in Scotland than in the UK (Blair, 2024; Scottish Government, 2024a).

1.4.5 The gendered economy and the limits of devolved action

Deep-seated mechanisms continue to drive gender inequality in the labour market. A primary driver of the gender pay gap across the UK remains occupational segregation (Andrew et al., 2021). In Scotland, this is exacerbated by occupational downgrading, where women trade skills for the flexibility found in lower-paid part-time work to manage caring responsibilities (Scottish Government, 2019b). Persistent low pay is disproportionately borne by women: 11% of Scottish adults are in persistent low pay, and over 70% of these are women (Birt et al., 2023). The majority of new mothers continue to intend to work part time (Scottish Government, 2019b).

Because employment law is reserved, the Fair Work agenda's enforcement measures are largely voluntary (Irvine, 2021); and because UK-wide drivers of insecure employment (zero-hours contracts, reserved employment-protection legislation) remain outside Scottish policy, the structural sources of insecurity continue to operate even where Scottish policy ambition is high. Within these limits, the Community Wealth Building Act (2026) places duties on local councils and NHS boards to use procurement and employment for local economic development, and the Programme for Government 2026–2031 commits to calling on Westminster to fully devolve employment law (Scottish Government, 2026d).

1.4.6 Scotland as the analytical case

Scotland therefore presents a compelling case for the questions addressed in this thesis. While devolved interventions have yielded measurable progress on certain economic indicators, such as the gender pay gap, the persistence of underlying stressors such as occupational downgrading, unpaid caring burdens,

and insecure employment raises critical questions about the limits of these policies in improving broader wellbeing. It becomes evident that improvements in economic equality do not automatically translate to equivalent gains in mental health. Rather, Scotland's aspirational policy agenda, the structural insecurity of the wider UK labour market, and entrenched gender norms function collectively as a complex, dynamic system. Explaining the trajectory of poor mental health requires an analytical approach capable of capturing how these competing dynamics evolve and interact over time.

1.5 The need for a systems perspective

Existing research has provided robust evidence for individual risk factors connecting employment and health. However, this thesis argues that a new approach is needed to understand the dynamic, population-level feedback loops connecting the labour market, mental health, and gender norms. Contemporary epidemiology offers a range of methods (such as multilevel models, microsimulation, and compartmental approaches) that move beyond single-exposure linear causality. These are powerful for the questions they are designed to answer (typically the average effect of a discrete exposure on an outcome). The argument advanced here, developed fully in Chapter 2 is that the specific question of how the coupled labour market-gender norms–mental health system generates population trajectories over decades is better served by a methodology built around feedback and accumulation than by methods that treat reciprocal causation as a problem of identification.

The thesis therefore adopts a systems perspective that treats the labour market, gender norms, and mental health as interdependent components whose behaviour emerges from their interactions over time, rather than as separate variables. This drives three critical analytical shifts:

First, from individual risk to population dynamics: When the focus shifts to population-level trajectories, an approach that can represent the reciprocal effects between the labour market and mental health is required. These reciprocal effects generate feedback loops in which, for example, deteriorating labour market conditions increase population distress, and rising distress in turn

reduces aggregate labour-force participation, further altering the macroeconomic context.

Second, from gender as a variable to gender as a structure: Traditional approaches often treat gender as a static characteristic whose effect on mental health can be statistically adjusted for. This thesis argues that gender must instead be understood as a structural feature of the economy and society that interacts reciprocally with economic outcomes to disproportionately shape women's mental health trajectories.

Third, from a single timescale to multiple timescales: Labour market conditions change over business cycles or accumulate over longer time periods, whereas gender norms evolve more slowly over generations. Explaining Scotland's trajectory requires representing these differing speeds in a single structure.

Ultimately, this systems perspective addresses a fundamentally different question: rather than 'who is at greater risk?', it asks 'how do the structural conditions that generate risk exposure for the population evolve, why do they lead to different outcomes by gender, and why have they produced the specific trajectory Scotland has observed?' A detailed theoretical justification for this shift in perspective, alongside a comprehensive critique of existing methodological approaches, is provided in Chapter 2.

1.6 Research aims

The overarching aim of this thesis is to develop and test a system dynamics model of how the labour market, gender norms and population mental health interact to produce the rising prevalence of poor mental health, and persistent gender disparity in that prevalence, in Scotland's working-age population between 1992 and 2019. Some of the policies described in Section 1.4, notably the 2021 expansion of funded childcare and the Fair Work Plan, post-date the 1992–2019 calibration window. They form part of the contemporary policy context motivating the research, rather than serving as inputs to the historical model; their relevance is to the forward-looking scenario analysis (Chapter 7) and the policy discussion (Chapter 8). A secondary aim is to use that model to

explore policy-relevant scenarios to 2040 — conditional, exploratory projections of what the estimated structure implies, not forecasts. In doing so, the thesis asks which structural mechanisms have dominated and which offer the greatest leverage for change.

The detailed research questions and objectives that operationalise these aims are developed in Chapter 2, after the empirical evidence and methodological limitations of existing approaches have been reviewed.

1.7 Research contributions

The research makes four contributions:

A dynamic population-level causal structure: By critically integrating fragmented evidence from epidemiology, economics, and sociology into a single testable architecture, this advances a population-level dynamic hypothesis showing how individual-level pathways aggregate into feedback loops at the macroeconomic scale. It demonstrates how these interconnected systems and exogenous economic shocks create self-reinforcing traps that sustain poor population mental health over time. This research advances the field by operationalising complex relationships and time lags that are often theoretically discussed but rarely modelled. In doing so it decouples the incidence of poor mental health from recovery from it, and assigns the system's determinants to distinct timescales, separating slow-moving structural conditions such as gender norms from faster labour market shocks.

Operationalising gender as a dynamic structural determinant: By treating gender norms as a slow-moving structural factor that interacts reciprocally with the labour market, and by modelling how this interaction produces differential exposure and recovery for men and women, this thesis offers a framework that can explain why progress on equality indicators has not translated into reductions in mental health disparity.

The model as an explanatory thinking tool: Rather than forecasting the effect of any single policy, the model elucidates structural mechanisms including the trade-offs that govern how the system responds to intervention - including the

redistributive trap, in which equality gains without job quality improvement relocate the mental health burden rather than removing it.

Methodological contributions: The thesis demonstrates system dynamics applied to the upstream social determinants of mental health, and develops a feedback-driven review architecture suited to mapping structure rather than isolated relationships.

1.8 Guide to the thesis

Chapter 2 provides a critical review of the conceptual and methodological background, identifying research gaps, justifying the research aims and outlining the research questions.

Chapter 3 provides a detailed overview of the methodology and the rationale for its selection. It articulates the philosophical underpinnings of the thesis.

Chapter 4 details the methodology for the development of the conceptual model, including a feedback-driven, iterative literature review approach designed for systems mapping and a disconfirmatory expert-consultation protocol with ten interdisciplinary experts.

Chapter 5 presents the boundary and causal hypothesis for the model. It explains the boundaries adopted while focusing on the labour market. It presents the findings in the form of a causal loop diagram, a visualisation tool within system dynamics to capture feedback mechanisms, and documents the evidential basis for each individual causal link within it.

Chapter 6 describes the development of the simulation model, justifies the model structure and presents results of the testing and refining of the model.

Chapter 7 utilises the calibrated simulation model as a theory-building tool, applying loop-dominance analysis, structural decomposition, counterfactual and scenario analysis and optimisation to decompose the structural mechanisms driving these historical trends, isolating specific feedback loops and leverage points that govern both the rising prevalence of poor mental health and its

persistent gender disparity, and distinguishing the levers that lower overall prevalence from those that narrow the gender gap.

Chapter 8 concludes the thesis by summarising and situating the key findings in the broader literature, and identifying the novel contributions of the research. It discusses the overarching limitations of the research and specifies the implications of the work for policy and future research.

2 Literature Review: Theoretical and Methodological Foundations

2.1 Overview of the chapter

This chapter provides the detailed theoretical, empirical and methodological foundation for the thesis. Section 2.2 defines the core system components (population mental health, the labour market, and gender norms), establishing the conceptual boundaries for the study. Section 2.3 reviews the epidemiological evidence to demonstrate the bidirectional relationship between mental health and the labour market. Section 2.4 examines the gendered nature of this relationship, arguing that gender is not merely a moderator but a structural factor that shapes both the exposure to risk and the psychological impact of risk factors. Section 2.5 situates these findings within the social determinants of mental health framework, organising these structural and intermediate determinants into a causal hierarchy. Section 2.6 critically evaluates the limitations of traditional methods in capturing this dynamic complexity, establishing the necessity of a new approach. Section 2.7 justifies the choice of system dynamics and acknowledges its inherent methodological limitations. Finally, Section 2.8 concludes the chapter by stating the research questions and objectives that operationalise the aims introduced in Chapter 1.

2.2 Conceptualising the core components

To understand the dynamics of the challenge described in Chapter 1, the three core components of the system must be defined: population mental health, the labour market, and gender norms. For each, this section establishes the concept and the conceptual boundaries the thesis adopts. Two choices that shape the thesis: how mental health is measured and why insecure work (rather than wages) is placed at the analytical centre, are critically engaged with, before the interactions between these components are examined in later sections.

2.2.1 Population mental health

Mental health is a complex concept that can be viewed through a binary lens or as a continuum (OECD, 2023a; World Health Organization, 2022b). A binary lens treats it as the absence of illness; a single-continuum lens places wellbeing and

illness on one axis whereas the dual-continua model treats positive wellbeing and mental ill-health as related but empirically distinct dimensions (Keyes, 2002). While wellbeing includes aspects such as life satisfaction, happiness or meaning, ill-health includes common mental disorders such as depression, severe mental illness (e.g. psychosis), and behavioural disorders such as drug addiction or dependence (OECD, 2023a; World Health Organization, 2004). This thesis treats ill-health and wellbeing as analytically separable, following the dual-continua convention adopted by the WHO (2022) and focuses the model on the ill-health dimension for reasons set out below.

This thesis focuses on the prevalence of common mental disorders (CMDs). CMDs are indicative of potential mental health problems and are robust indicators for population mental health (OECD, 2023a). The evidence for the impact of social determinants, particularly economic factors, is most substantial for these conditions (World Health Organization & Calouste Gulbenkian Foundation, 2014). By focusing on mental ill-health or poor mental health, it becomes possible to identify the proportion of the population at risk, as established measures tend to have a cut-off score (Winefield et al., 2012). Prevalence at a population level is understood here as the proportion of individuals at or above a clinically meaningful threshold of poor mental health at a given point in time.

Conceptualising mental health in terms of prevalence establishes a distinction that the remainder of this thesis relies on. Specifically, prevalence is an accumulation governed by two distinct processes: the incidence of new cases (inflow) and recovery from existing cases (outflow). A rise in prevalence may therefore reflect an increase in people moving into poor mental health, a decrease in those recovering, or both. This distinction is critical because different social determinants may act through unique mechanisms, affecting both processes asymmetrically. The operationalisation of prevalence within the simulation model is described in Chapter 6. Because the focus is on prevalence, a caseness-based operationalisation is preferred over a continuous measure.

This thesis uses the General Health Questionnaire-12 (GHQ-12) as its primary measure. The GHQ-12 comprises twelve items asking respondents to rate recent psychological experiences (for example, ability to concentrate, loss of sleep through worry, capacity to make decisions, and feeling constantly under strain)

against their usual state. Total scores range from 0 to 12, with a score of four or more indicating a probable common mental disorder (Goldberg et al., 1997). Although the GHQ-12 can also be scored continuously, it is used here as a dichotomous measure to identify the population with poor mental health at or above the caseness threshold. GHQ-12 is selected for three key reasons: First, it has been administered in the Scottish Health Survey, the UK Household Longitudinal Study (UKHLS) and its predecessor British Household Panel Survey (BHPS) on a consistent basis since 1991, providing the longest calibrated time-series. Second, its caseness threshold makes it suitable for identifying population-level prevalence. Third, the measure has known validity and reliability and is used in Scottish Government reporting on mental health, identified as the appropriate indicator for common mental health problems by Public Health Scotland (Public Health Scotland, 2022). Public Health Scotland includes other indicators for mental health problems such as the Clinical Interview Schedule-Revised (CIS-R) for measuring more specific indicators such as depression or anxiety (Public Health Scotland, 2022).

Two alternative measures were considered and rejected. The Warwick–Edinburgh Mental Wellbeing Scale (WEMWBS), also recommended and used in the Scottish Health Survey since 2008, measures positive wellbeing on a continuous scale and does not support the caseness-based operationalisation required here. The SF-12 Mental Component Summary (SF-12 MCS) provides a caseness-compatible measure of mental functioning, but is not used in the Scottish Health Survey, and is not available across the full calibration period and therefore could not provide the time-series required for the simulation.

The GHQ-12 is a screening instrument for general psychological distress rather than a diagnostic assessment of specific disorders, and its classification of cases carries known measurement properties that warrant explicit acknowledgement. Bell et al. (2005), comparing GHQ-12 cases against the more detailed CIS-R, report an overall false-positive rate of approximately 30% at the threshold of three or more, falling to 26% when the threshold is raised to four. They further demonstrate that this rate is systematically patterned by socio-demographic status: home-owners and those with close confiding relationships are more likely to be classified as false positives, whereas respondents unable to work because of illness or disability are less likely. Within the variables they examined, gender

was not reported as having an independent association with false-positive status after adjustment, supporting the comparability of the measure between men and women. As noted by the authors themselves, false positivity is defined by one self-report instrument against another; the CIS-R itself is not a criterion measure. These properties do not materially undermine the use of the GHQ-12 in this thesis. The analysis does not rely on absolute prevalence as an estimate of true morbidity but on the trend in prevalence over time and the gap between men and women. The three reasons set out above for selecting the GHQ-12 therefore hold, and its known properties are better understood as limitations to be considered in the interpretation of results than as fundamental issues that undermine its use.

2.2.2 The labour market

The labour market is one of the most widely studied and important social determinants of health and mental health (Solar & Irwin, 2010). Traditional economic models conceptualise the labour market as driven by interactions of labour supply and demand, influenced by regulatory, institutional, and policy frameworks (Borjas, 2013).

The psychosocial benefits of employment are most commonly conceptualised through Marie Jahoda's Latent Deprivation Model, which posits that work provides five essential 'latent functions': time structure, social contact, collective purpose, status, and regular activity, alongside the manifest function of income (Paul et al., 2023). On this account, employment is health-protective not solely because it provides earnings, but because it embeds individuals in the structuring features of social life. The labour market is therefore, in this thesis, defined not just by employment rates, but by its capacity to provide these latent functions.

Two further theoretical accounts refine and complicate this view. Warr's Vitamin Model identifies an expanded set of job features — opportunity for control, skill use, interpersonal contact, external goal and task demands, variety, environmental clarity, availability of money, physical security, and valued social position — and adds the important qualification that the relationship between these features and mental health is not monotonic: very high levels of some,

such as task demands, can be as harmful as very low levels (Warr, 1994). Fryer's agency restriction account, by contrast, contests Jahoda's emphasis more directly, arguing that the driver of unemployment-related distress is the erosion of personal agency and the material constraint imposed by income loss, rather than deprivation of the specific latent functions Jahoda identified (Fryer, 1992). For the purposes of this thesis these accounts are read as complementary rather than rival: agency restriction and latent-function deprivation are alternative descriptions of overlapping mechanisms.

These lenses converge in their implications for precarious employment, which may deprive individuals of these latent and manifest functions provided by standard employment (Irvine & Rose, 2022). Where Jahoda's framework highlights what insecurity removes, social-exchange accounts highlight what it withholds: when sustained worker effort is not reciprocated by the expected reward of security, the resulting imbalance generates psychological strain (De Witte, Pienaar, & De Cuyper, 2016). Resource-based models add a third lens, framing insecurity as the depletion of the control resources workers need to manage ongoing job demands (Rugulies et al., 2023). Together, these accounts identify three channels through which precarity bears on mental health: function deprivation, reciprocity failure, and resource depletion. Precarity itself is, in turn, shaped not only by the traditional actors of workers, employers, trade unions, and governments, but also by global economic forces that have reshaped employment relations across the period this thesis examines (Benach et al., 2014; Prosser, 2016).

A legitimate question is why this thesis places precarious employment at the analytical centre rather than other labour market features such as wage levels, occupational segregation, or union coverage. Three considerations justify the choice. First, it was identified as a key factor through exploratory searches, explained in Chapter 4; precarious employment is, in an important sense, the integrative variable through which several other features operate on mental health: low wages, occupational downgrading, and declining union density increase exposure to insecurity, and the psychosocial literature identifies insecurity as the proximate mental health stressor (Benach et al., 2014; Kreshpaj et al., 2020). This is a boundary judgement rather than an empirical finding. Second, precarious employment is an important aspect of Scotland’s Fair

Work agenda discussed in Chapter 1. Third, precarious employment has risen substantially in prevalence across the period this thesis covers (1992-2019), making it a plausible candidate for the observed trend in population mental health. Other labour market features such as occupational segregation, wage levels, union coverage are represented indirectly as drivers of the gender pay gap and insecure work distribution (see Chapter 5).

This thesis operationalises the labour market along four dimensions relevant to mental health production: (i) employment (presence of paid work); (ii) unemployment (involuntary non-employment); (iii) inactivity (long-term sickness, caring, and retirement); and (iv) insecure work. The evidence for the causal relationship between each of these dimensions and mental health, in both directions, is reviewed in Section 2.3. The gendered distribution of these dimensions, and its implications for the persistent gender gap in prevalence, is developed in Section 2.4.

2.2.3 Gender norms

The labour market is also shaped by social norms that govern who is expected to do paid and unpaid work. Societies often ascribe different social roles to men and women, particularly in terms of family responsibilities (Solar & Irwin, 2010). These ascriptions operate as gender norms: shared, sanctioned expectations about appropriate behaviour that are reproduced through institutions and everyday practice, and that have material consequences for labour market participation, the distribution of unpaid work, and ultimately mental health.

A typological tradition organises welfare states by the strength of the male-breadwinner norm, distinguishing "strong", "modified" and "weak" variants (Lewis, 1992). This thesis locates Scotland within the modified male-breadwinner (MMB) regime, in which women are encouraged into paid work but remain culturally coded as secondary earners with primary responsibility for care. The breadwinner model's value here is heuristic rather than taxonomic; it names the specific configuration of norms (women as primary carers, men as principal earners) that the evidence in Section 2.4 shows to be consequential for mental health; in addition to influencing women’s labour market participation, these norms expose women to differential risks, as well as moderate the impact of

unemployment on mental health (Cislaghi et al., 2022; Heise et al., 2019; Moriconi & Rodríguez-Planas, 2021; Strandh et al., 2013).

Gender norms are best conceptualised as slow-moving beliefs that build up and erode over generations (Scott & Clery, 2013). In the model, it is operationalised as the prevalence of norms on gender roles regarding the division of paid and unpaid work, discussed further in Chapter 5.

2.3 The bidirectional relationship between the labour market and mental health

The following discussion concerns four key labour market states: employment, unemployment, inactivity, and insecure work, together with the macroeconomic context of recession, and their relationship to the prevalence of common mental disorders. The evidence is organised around two hypotheses: health causation (the labour market affects mental health) and health selection (mental health affects labour market outcomes). Each section identifies the individual-level mechanism, aggregates it to the population level, and highlights where the existing evidence base is sparse.

2.3.1 Health causation: the impact of labour market on mental health

The association between unemployment and poor mental health is one of the most consistently replicated findings in social epidemiology (Gedikli et al., 2022; Kim & von dem Knesebeck, 2016). The meta-analysis of longitudinal studies by Paul and Moser (2009) remains the strongest synthesis arguing for a causal component, showing that distress rises on job loss and falls on re-employment.

At the individual level, unemployment, like other social determinants, impacts mental health through material, behavioural and psychosocial processes (Compton & Shim, 2015). While employment provides individuals with income, and a part of the negative impact of unemployment on mental health is due to the loss of this income, unemployment also represents a structural loss of the 'latent functions'. By stripping individuals of time structure, status, and purpose, unemployment causes poor mental health (Bahr, Batinic, & Collischon, 2022; Paul et al., 2023; Thomson et al., 2023).

Generous welfare systems can dampen these effects but do not eliminate the impact of unemployment on mental health (Cygan-Rehm, Kuehnle, & Oberfichtner, 2017). This relationship is also seen at the aggregate level; both regional and national (Barbalat & Franck, 2020; Barr, 2021). Furthermore, local unemployment trends can impact an individual's mental health irrespective of their personal employment status (Curtis et al., 2019).

The qualification that any job is better than no job, implicit in the unemployment-distress literature, has been increasingly challenged (Butterworth et al., 2011). The argument that the mental health consequences of insecure work can be comparable in magnitude is central to the framing this thesis adopts. Although Jahoda's theory focused on employment as a binary state, recent research has suggested that precarious employment may also deprive individuals of these latent and manifest functions (Irvine & Rose, 2022).

Precarious employment is a multi-dimensional concept that captures employment situations where the power dynamics between employers and employees are unfavourable for the employee (Benach et al., 2014). It generally encompasses three key aspects: (1) employment insecurity due to duration and terms of contract (temporary contracts, zero-hours contracts or agency work, underemployment), (2) income inadequacy due to low wages or volatility, and (3) lack of rights and protection (Kreshpaj et al., 2020).

Precarious employment therefore acts as a source of uncertainty, stress and low control. As chronic stress can alter biological stress responses and worsen mental health (Fisher & Baum, 2010), there is good evidence to support a causal impact of various aspects of precarious employment on poor mental health at the individual-level, with chronic or persistent precarity having greater effects (Bell & Blanchflower, 2019; De Witte, Pienaar, & De Cuyper, 2016; Pulford et al., 2022; Rönnblad et al., 2019). The impact of precarious employment on mental health can in fact be comparable to that of unemployment (Kim & von dem Knesebeck, 2015). Similarly, underemployment, conceptualised as insufficient hours or lack of skills utilisation, may also not provide all the latent functions of full-time employment (Beck, Warren, & Lyonette, 2024). Studies consistently document the negative impact of underemployment on mental health (Beck, Warren, & Lyonette, 2024; Bell & Blanchflower, 2019; Churchill et al., 2025). As

a chronic stressor, insecure work depletes psychological resources; those transitioning into a poor-quality job from unemployment report higher allostatic load and adverse biomarkers compared to those who remained unemployed (Chandola & Zhang, 2018). This means that insecure work could impact the onset or duration of poor mental health; however, there is a lack of evidence that distinguishes the effects of insecure work in terms of these mechanisms.

Objective precarious work does not universally or deterministically generate a psychological state of uncertainty and powerlessness (Allan, Autin, & Wilkins-Yel, 2021). The adverse effect of insecure work is moderated by the extent to which an individual has the freedom to choose their employment status. When atypical contracts, which can often be classified as precarious in objective terms, provide flexibility through worker choice and control, they can in fact improve health outcomes, whereas if flexibility is involuntary, the consequences are largely negative (Joyce et al., 2010).

While the negative impact of precarious employment on individual mental health has been documented, its impact on population mental health has been rarely investigated directly, leaving an insufficient understanding of how population-level shifts toward precarious employment may influence mental health trends.

Economic inactivity (defined by the ONS as those aged 16-64 years not in employment who have not been seeking work within the last 4 weeks and/or are unable to start work within the next two weeks) could also contribute to poor mental health. Inactivity can be due to various reasons such as education, caring for family, long-term sickness or disability, early retirement or discouragement. These reasons influence latent function deprivation as well as mental health (Paul et al., 2023; Wang & Lu, 2022). Moving from inactivity to employment can improve mental health whereas transitioning from employment to inactivity can worsen it (Kamerade et al., 2019; Thomas, Benzeval, & Stansfeld, 2005).

Finally, at the macro-level, economic downturns negatively impact population health through unemployment and broader labour market conditions such as reduced labour demand (Margerison-Zilko et al., 2016). Recessions, characterised by high unemployment rates, are associated with a rise in population mental distress (Asper et al., 2022; Barr, Kinderman, & Whitehead,

2015; Frasquilho et al., 2016). For instance, the 2008 recession led to a rise in the prevalence of common mental disorders (CMDs) in England (Katikireddi, Niedzwiedz, & Popham, 2012). Critically, the impact extends beyond those directly impacted by job loss (Zivin, Paczkowski, & Galea, 2011). Recessions can also adversely impact mental health through reduced vacancy rates and increased precarious employment (Frasquilho et al., 2016). Paradoxically, some studies suggest that high unemployment rates may protect the unemployed from mental distress, perhaps due to reduced stigma as unemployment is more widespread and attributed to external causes rather than the individual (Frasquilho et al., 2016). This individual-level stigma effect among the unemployed does not, however, override the population-level rise in distress associated with recessions (Zivin, Paczkowski, & Galea, 2011).

In the UK, the policy response to the 2008 recession was a prolonged period of austerity. In Scotland, this retrenchment was driven by Westminster, though the devolved government allocated cuts differently to buffer the impact to an extent (Gray & Barford, 2018). Such policies have been shown to have further impacted population mental health through multiple channels, including cuts to public services, welfare reform, and the growth of insecure work, in addition to the direct effects of unemployment (Stuckler et al., 2017). Reductions in the generosity of welfare benefits, increased conditionality, and stricter eligibility criteria were consistently associated with worsening mental health outcomes (Simpson et al., 2021). Crucially, whether such macroeconomic shocks and changes act as independent drivers, activate other dormant drivers, or magnify existing drivers, is not well-understood.

2.3.2 Health selection: the impact of mental health on labour market outcomes

The system also operates in reverse. The health selection pathway is not a competing account to health causation but a complementary mechanism: both operate simultaneously, and their population-level aggregation produces the reciprocal dynamics that this thesis sets out to model. The health selection pathway describes how poor mental health impacts labour market outcomes through two key mechanisms: reduced employability and reduced productivity.

First, poor mental health significantly increases the likelihood of non-employment, unemployment, early retirement, disability pension and sick leave (Amiri & Behnezhad, 2019; Amiri & Behnezhad, 2021; Campbell et al., 2022; van Rijn et al., 2014). It can reduce the number of hours worked and increase the duration of unemployment (Butterworth et al., 2012; Campbell et al., 2022).

Second, poor mental health increases presenteeism (reduced performance due to working when ill) and absenteeism (sick leave) leading to reduced productivity (Almond & Healey, 2003; Bryan, Bryce, & Roberts, 2021; Bryan, Bryce, & Roberts, 2022; Bubonya, Cobb-Clark, & Wooden, 2017; Curran et al., 2007; Huo et al., 2022; Layard, 2013; Pinna Pintor, Fumagalli, & Suhrcke, 2024). These individual-level consequences aggregate to the population level, resulting in reduced economic growth (Thomas et al., 2023). Evidence suggests that the loss of days due to poor mental health is associated with a reduction in growth rate at the county level in the US (Davlasheridze, Goetz, & Han, 2018).

Two further pathways, job sorting and labour market stigma, are excluded from the model due to uncertainty regarding their aggregate-level impacts. The first is job sorting, which exists at the individual level. Poor mental health can also have negative consequences by sorting individuals into worse jobs, increasing the likelihood of being in insecure employment (Ervin et al., 2023b) or suffering a reduction in income (Campbell et al., 2022). The second is stigma and discrimination in hiring and retention, which operates independently of cognitive impairment and is consequential for how long a spell of poor mental health restricts labour market options after recovery (Brouwers, 2020). Whether this channel is large enough to affect aggregate trajectories in the UK is empirically uncertain; it is flagged here as a pathway the model does not represent endogenously but which conditions the interpretation of results.

2.3.3 A coupled system at the population level

Taken together, the evidence establishes that mental health status is a cause as well as a consequence of labour market position. The separation of these pathways into health causation and health selection is an analytical convenience inherited from study designs focused on individuals; at the population level, it imposes a dichotomy that the aggregate evidence does not support. The

pathways reviewed in Sections 2.3.1 and 2.3.2 do not operate independently of one another. As deteriorating labour market conditions erode psychological resources and raise the incidence of common mental disorders, the resulting psychological morbidity in turn reshapes labour market attachment; the two pathways therefore form a coupled system, in which the consequences of each feed back into the conditions of the other over time.

This coupling has methodological implications that the remainder of the chapter takes up. A research design that holds either side of the relationship fixed, either by treating labour market conditions as exogenous to mental health, or treating mental health as the downstream outcome of fixed labour market status, examines the average effect along one direction of the arrow. It does not uncover the behaviour of the coupled system. Sections 2.6 and 2.7 return to the methodological requirements this implies.

2.4 The gendered dynamics of work and mental health

Research at the individual level often presents a complex picture of gender differences in mental health. While women consistently report higher prevalence of common mental disorders, men are often found to be more psychologically impacted by labour market shocks, such as unemployment and precarious employment (Chum et al., 2022; Gedikli et al., 2022; Kromydas et al., 2021; Padrosa et al., 2022; Paul & Moser, 2009; Thomson, Niedzwiedz, & Katikireddi, 2018). This asymmetry is not universal, however: some US evidence finds unemployment harms the mental health of women and men to a similar degree, suggesting the gendered pattern is specific to the context (Aydiner-Avsar & Piovani, 2019). These findings have led some researchers to treat gender merely as a statistical moderator i.e. a variable that simply dampens or amplifies a relationship. However, this thesis argues that treating gender as a moderator is insufficient. It fails to capture the broader societal context that shapes women's experiences.

Gender operates as a system of “interlocking cogs”, where societal, political, community, interpersonal and family-based components work together to determine exposure to risk and shape norms (Heise et al., 2019, p. 2443). To understand the persistence of mental health inequalities, it is important to

examine how gender norms structure both the labour market and the domestic sphere.

2.4.1 The gender paradox in mental health

Higher female prevalence of common mental disorders is documented across countries (Seedat et al., 2009). Artefactual accounts attribute part of the gap to differential reporting, help-seeking, and diagnostic thresholds; biological accounts include genetic and hormonal differences; structural accounts locate the gap in differential exposure to chronic stressors (Banos & Miragall, 2024).The evidence base is best read as assigning each a partial role (Piccinelli & Wilkinson, 2000). This thesis does not attempt to identify the specific contribution of each; it develops a structural account because it is the one that is actionable at the population level and interacts with the labour market dynamics modelled here.

Within the structural account, a specific puzzle of the "gender paradox" in employment research motivates the argument that follows. Employment is generally protective of mental health, but the protection is asymmetric: women derive less psychological benefit from employment than men. A recent microsimulation study found that assigning women men's employment rates and job quality narrowed the gender gap in clinical distress but widened the gap in positive "flourishing" mental health (Orchard et al., 2024). This indicates that simply equalising labour market participation does not resolve this disparity and that the protective psychosocial mechanisms of employment operate differently for women, a difference that the next sections argue, reflects the suppression of those mechanisms by the structural reality of women's parallel caregiving burden.

To explain this paradox, understanding the underlying structure of gender norms is essential. In the UK context, these norms manifest as the modified male-breadwinner model introduced in Section 2.2.3. This normative structure dictates the impact of economic shocks. In such a society, men's mental health is more impacted by unemployment than women's, whereas no such difference exists in societies where men and women are seen as equal earners (Strandh et al., 2013). More gender-inclusive welfare regimes such as the Nordic states, are

associated with better mental health outcomes for women and smaller gender-related inequalities in mental health compared to those in market-oriented models (McAllister et al., 2018).

Furthermore, these norms actively restructure the labour market for women. The expectation of care responsibilities leads to women working part-time, particularly after childbirth (Harkness, Borkowska, & Pelikh, 2019). Part-time work leads to job downgrading, occupational segregation, wage stagnation and employment precarity (Casarico & Lattanzio, 2025; Florisson & Gable, 2022; Nightingale, 2019; Padrosa et al., 2022). Consequently, structural discrimination, manifesting through lower pay, is a plausible direct driver of worse mental health among women (Platt et al., 2016).

Simultaneously, gender norms dictate the distribution of unpaid labour (McMunn et al., 2020). Despite women's increased participation in the workforce, they continue to perform most of the housework and childcare (Andrew et al., 2021; Wheatley, Lawton, & Hardill, 2018). While economic theories of relative resources or time availability suggest that the partner earning less or with more time does more housework, empirical evidence supports the gender socialisation hypothesis: even when women earn equivalent income, they perform more unpaid labour to conform to normative expectations (Fauser, 2019; Foster & Stratton, 2018; Sahl, 2015). In fact, when women provide more than half the household income, the exchange-based dynamic of "more money buys less housework" breaks down and women report doing more housework to neutralise the gender deviance of their higher earnings (Bittman et al., 2003). However, others have challenged this, arguing that empirical evidence for this hypothesis has not taken into account women's absolute earnings, and that this effect is concentrated in a small group of low-income families with traditional views (Sullivan, 2011). More importantly, as discussed above, the differences in earnings and time availability are themselves shaped by gender roles (Dominguez-Folgueras, 2022).

This normative expectation creates a disproportionate burden on women. In most OECD countries, women spend more total time on work (paid plus unpaid) than men (OECD, 2017). This results in time scarcity, a lack of restorative or leisure time (Strazdins et al., 2016), and imposes a significant mental load

through the cognitive labour of planning and decision-making (Reich-Stiebert, Froehlich, & Voltmer, 2023). Cognitive labour is gendered: women perform more of the anticipation, identification and monitoring components, while decision-making, the influence component, is mostly shared equally (Daminger, 2019). This unequal distribution of unpaid burden creates a gender-specific pathway to distress; unpaid work has a negative impact on the mental health of employed women (Ervin et al., 2022; Piovani & Aydiner-Avsar, 2021).

2.4.2 The interplay of the labour market and domestic sphere

Crucially, these mechanisms do not operate in isolation; they function as a bidirectional relationship. Gender norms drive women into part-time work to accommodate unpaid care. Part-time work, being lower paid and lower status, reinforces the woman's status as the secondary earner and further entrenches the expectation that women are responsible for domestic work.

This dynamic is supported by a systematic review which concluded that it is the specific interplay between the labour market structure and domestic work that drives gender differences in mental health (Cabezas-Rodriguez, Utzet, & Bacigalupe, 2021). Further, this relationship is complex; it is mediated by beliefs. For example, the mental health benefit of transitioning from home-making to paid employment was not seen among women holding traditional gender beliefs when they had children (Lu et al., 2023). This suggests that a potential conflict between the internalised norms and behaviour can cause distress.

Most importantly, this impact also operates over long time horizons. High levels of unpaid work in the present are linked to non-employment and reduced hours for women later in life (Ervin et al., 2023b). Due to its widespread impact on women's lives, unpaid labour has been argued to be a neglected social determinant of health (McMunn, 2023).

2.4.3 Gender as a structural determinant

Understanding the relationship between the labour market and population mental health necessitates a focus on gender as a structural determinant. Crucially, gender norms are structural factors, not distal or less important than

individual-level exposures to risk; in fact, through embodiment, they exert a powerful and pervasive impact (Krieger, 2008).

This shifts the analytical focus from individual-level explanations to the broader societal structures, particularly social norms and inequalities, that generate differential risk distributions between men and women. It is critical to frame these gender differences not merely as statistical observations, but as a health inequality. Drawing on a long tradition of disparities-focused public health scholarship, Braveman (2025) defines a health inequality as a difference that adversely affects a socially disadvantaged group, raising fundamental concerns about fairness and justice. This normative lens justifies the focus on these persistent disparities as a structural problem.

While this thesis primarily operationalises the modified male-breadwinner model through its social, economic and health impacts, it is critical to acknowledge its deeper macroeconomic function. Feminist political economists and theorists of social reproduction argue that this is not merely a lagging cultural artefact, but a structural requirement for cost minimisation within contemporary capitalism (Fraser, 2017). Endogenising the political economy of this norm lies beyond the boundary of this thesis, although the economic function served by these gender norms is important to acknowledge. In this thesis, gender norms are treated as a powerful, slow-moving structural condition that actively configures the labour market and domestic responsibilities, and distributes the risk factors between men and women.

2.5 Organising the evidence: social determinants of mental health framework

The preceding sections have reviewed the bidirectional causation between the labour market and mental health and the gendered structuring of that relationship. To address the theoretical challenge of conceptualising multiple interacting factors, a social determinants framework offers a structured approach by organising these variables into a rigorous causal hierarchy and explicitly clarifying the level at which each determinant operates.

Theoretical models of the social determinants of health (SDH) conceptualise determinants as occurring across distinct layers of social organisation, from the factors most proximal to the individual to those most distal (or “upstream”). Early versions of this approach were most notably described in the “rainbow” model by Dahlgren and Whitehead (1991). The framework has been further developed by the WHO Commission on Social Determinants of Health to more clearly capture the causal hierarchy; from the socio-political context to socioeconomic position (collectively termed as the structural determinants) through to intermediary determinants (Solar & Irwin, 2010). More importantly, this framework incorporates inherent feedback loops by recognising the causal effect of health status back onto socioeconomic position and structural determinants, avoiding a strictly unidirectional assumption. This feedback is not an afterthought to the WHO framework but an explicit element of the Diderichsen model mechanism that underwrites it: “the social and economic consequences of illness…on an aggregate level, feed into the context of society, …and influence aggregate social and economic development" (Solar & Irwin, 2010, p. 24).

Adapting this for mental health, Compton and Shim (2015) further refine the hierarchy into a specific "architecture of risk" for mental illnesses, delineating a clear, progressive pathway from broad societal forces down to individual psychopathology. While their model is largely linear and does not capture the feedback represented in the WHO model, they emphasise the importance of addressing the public policies and social norms that shape the remaining determinants as the most effective way to improve population mental health arguing that "the deeper we go, the greater the overall impact for the population" (Compton & Shim, 2015, p. 422). The focus on upstream determinants is also echoed by Braveman and Gottlieb (2014), who highlight that these macrosocial factors interact, often mediating their effects through complex pathways, and that there are significant time lags before the effects of these determinants on health materialise.

The framework dictates that structural determinants produce health outcomes almost entirely through intermediate determinants, rather than directly. At the apex of the causal hierarchy lies the structural level, which encompasses the socio-political context and the foundational mechanisms that generate social

stratification. In this thesis, the structural level is operationalised by two primary intersecting forces: gender norms and macroeconomic conditions. Gender norms act as a profound and pervasive structural determinant; they are slow-moving, deeply accumulated societal beliefs that configure the social roles, expectations, and power dynamics of men and women. Macroeconomic conditions include the economic context and the broader policy regime, operationalised by the exogenous forces of labour demand and changes in the prevalence of insecure work. The policy regime is also indirectly captured through the manifestation of the interaction between norms and context, specifically the gender pay gap residual not explained by different labour market roles, and childcare availability.

The intermediate level is primarily occupied by the operational realities of the labour market and the domestic sphere. At this level, the analysis observes aggregate unemployment rates and the prevalence of insecure work; here, the structural gender norms translate directly into intermediate material realities such as the persistent gender pay gap and the unequal distribution of the unpaid-work burden. While the proximal level is where mental illness actually manifests biologically and psychologically within the individual, this thesis treats proximal individual processes as aggregated mechanisms rather than as the primary unit of analysis. By measuring population mental health via the aggregated prevalence of common mental disorders, the study observes the collective, macro-level outcome of structural and intermediate forces. Prioritising the aggregate population over the isolated individual is a deliberate methodological commitment, not a limitation; it aligns fundamentally with Geoffrey Rose's (2001) core public health mandate to understand the "causes of incidence" across populations rather than the isolated "causes of cases" within particular individuals. As Pearce (2011) argues, ubiquitous structural risks may not explain population variation but are primary causes of disease, and thus are both causes of incidence and causes of cases. Where a structural determinant is pervasive, between-individual variation in exposure is compressed, and the factor is visible principally at the population level rather than as individual risk.

While these frameworks provide an indispensable conceptual hierarchy for organising evidence, they often lack the capacity to distinguish variables by the speed at which they operate, a limitation Krieger (2008) notes as a problematic

conflation of spatial scale with temporal timelines. The dynamics of population mental health are dictated by mismatched timescales, such as slow-moving gender norms interacting with relatively rapid macroeconomic shocks. A framework that represents only the hierarchy of determinants without their time architecture will systematically misattribute changes in prevalence, typically to whichever determinant has moved most recently, because slow-moving structural conditions are invisible to snapshot analyses (Braveman & Gottlieb, 2014).This, along with the challenges discussed in the next section, suggests the need for a new approach to study population health dynamics and to move toward a more robust understanding of how structural inequities generate mental health disparities over time.

2.6 The limits of traditional methods for this complex systems problem

Despite the substantial evidence regarding the individual links between the labour market, gender, and mental health, our understanding of how these factors interact to produce persistent population-level trends remains limited. This epistemological blind spot is largely attributable to the complex, adaptive nature of this interplay and the inherent constraints of reductionist methodologies in capturing non-linear system behaviours (Heitman, 2017). Population health acts as a complex system where the disease rate is not merely the aggregation of individual risks, but the result of “the dynamic interplay between individuals and population-level properties” (Diez Roux, 2007). While the dominant methodological paradigm has been effective in identifying specific causal associations ("what" matters), it is less suited to capture the dynamic architecture of the system ("how" it functions) that shapes population health (Keyes & Galea, 2015). This section argues that four specific characteristics of this system: circular causality, accumulation (including delays), different timescales and multifactorial causation, require a new approach.

Before proceeding, a clarification of scope is warranted. The argument advanced in the remainder of this section is not that contemporary epidemiology is incapable of engaging with complexity. Social epidemiology has long moved beyond single-exposure linear causality through multilevel hierarchical models that routinely account for spatial and contextual factors such as neighbourhood

deprivation and local unemployment rates alongside individual risk factors. Systems epidemiology has incorporated computational simulations, notably microsimulation, agent-based models and compartmental models, to capture non-linear dynamics, spillover effects, and emergent properties within populations. These methods are powerful for the questions they are designed to answer; typically, whether a discrete policy or exposure produces a measurable change in a health outcome, on average, under identified assumptions. The argument here concerns a different question: how the coupled system of macroeconomic and normative forces generates population mental health trajectories over decades. For this particular problem, system dynamics is more appropriate than the alternatives.

2.6.1 The problem of circularity

The first methodological challenge is that the labour market and mental health form a feedback loop, not a linear chain. As reviewed in Section 2.3, this relationship is bidirectional. These effects cannot be captured with methods that assume a static fixed relationship between variables (Lundberg, 2020). However, research focusing on the impact of macroeconomic factors on health tends to study them as exogenous factors, not accounting for internal dynamics that could have influenced them (Janke et al., 2020).

Traditional models require variables to be specified as either dependent or independent. Mathematically, this forces a linear structure where typically reverse causality (either health selection or health causation) is treated as a bias or confounder to be controlled or adjusted for (Galea, Riddle, & Kaplan, 2010). The distinction here is as much epistemological as it is technical: advanced traditional methods can, in principle, accommodate bidirectionality through instrumental variables, cross-lagged panel designs, or structural equation models, but they do so by treating it as a problem of identification to be resolved. A systems perspective, by contrast, treats circular causation as the defining structural feature to be modelled. From a population health perspective, this circularity is not a statistical nuisance; it is the fundamental engine driving the system's behaviour. To comprehend the persistence of mental health inequalities, it is necessary to deploy methodologies capable of formalising simultaneity i.e. where A causes B and B causes A at the same time

(Diez Roux, 2007). The cost of neglecting this structural feedback is not merely conceptual. In their dynamic panel analysis of macroeconomic conditions and health across Britain, Janke et al. (2020) conclude that ignoring dynamics and feedback would have led their model to underestimate the effects of macroeconomic circumstances by around 50%, indicating that treating the labour market–health relationship as unidirectional systematically compresses its apparent magnitude.

2.6.2 The problem of accumulation

The most consequential limitation for this thesis is the difficulty that traditional linear methods face in capturing the continuous dimension of time, specifically the systemic role of accumulation. In complex systems, the state of the system (e.g., the prevalence of poor mental health or the strength of gender norms) behaves as an accumulation. It is a reservoir of history that builds up and drains slowly. For example, gender norms are not fleeting variables, but accumulations of beliefs that persist over generations. Time-use shows that the domestic division of labour has converged more gradually than women's labour market participation, supporting the slow pace of change (Altintas & Sullivan, 2016; Gershuny & Kan, 2012).

This deep historical accumulation inherently generates path dependence, a phenomenon where early structural disadvantages become “locked in”, creating inertia that resists immediate change (Paina & Peters, 2012). Because this structure causes delays and policy resistance, there is often a significant time lag between an intervention and a visible result at the population level (McAllister et al., 2018). This was demonstrated in the dynamic panel modelling of British employment and health data mentioned above where the lag between employment and health effects is estimated to be over two years (Janke et al., 2020, p. 28). Traditional linear models, which predominantly assume immediate or rigidly fixed-lag effects, are poorly equipped to represent these structural delays. This often leads to the false conclusion that an intervention is ineffective simply because the system has not yet tipped (Diez Roux, 2011).

2.6.3 The problem of different timescales

A further characteristic of the social determinants of mental health that traditional methods struggle to represent is the differential speed at which they operate. Some determinants fluctuate cyclically (for example, unemployment rates shift over the business cycle), others accumulate gradually (for example, erosion of worker bargaining power), or change over generations (for example, cultural beliefs about gender roles).

While social epidemiology has long acknowledged this temporal stratification conceptually, the successful integration of these multiple time-horizons into quantitative, population-level dynamic analysis remains limited. Explaining the trajectory of a population outcome therefore requires a method capable of representing not only which determinants matter, but the different speeds at which they propagate and interact.

2.6.4 The problem of multifactorial and multiplicative causation

Finally, traditional methods focus on isolating single causes and often assume that risk factors are additive (Galea, Riddle, & Kaplan, 2010; Krieger, 2017). However, as Krieger (2017) explicitly demonstrates, forcing the interdependent causes of population health to sum to 100% through metrics like the Population Attributable Fraction (PAF) is a fallacy because multiple factors contribute independently and synergistically to causal pathways. Causes do not act in isolation but rather through a network of causal structures (Keyes & Galea, 2015). This limitation is highlighted by Naik et al. (2019), who conclude that while the economy functions as a “complex interacting system”, the existing evidence base fails to capture the interconnections between economic factors.

In complex systems, causation is multifactorial, interconnected, and multiplicative. The resulting behaviour is not caused by any one causal relationship but by the dynamic, multiple and complex connections between the different elements (Savigny & Adam, 2009). The challenge is not simply to trace long causal chains, but to understand the dynamic interplay across levels of a system that are concurrently active (Diez Roux, 2007).

Risk factor epidemiology has historically prioritised proximal factors (those closest to the outcome), such as individual stress events or health behaviours (Krieger, 2008). While valid, this focus can obscure the distal structural drivers that set the context for those proximal risks (Fink, Keyes, & Cerdá, 2017).

Crucially, the pathway from these distal causes to the outcome is indirect and non-linear (Diez Roux, 2007). A gradual erosion of job security may have little visible impact on population mental health for years until a threshold is reached, after which the system state deteriorates rapidly. Capturing these indirect pathways requires a method capable of simulating the entire causal chain, rather than just the immediate associations between adjacent variables.

2.6.5 The need for a new methodological mandate

The system characteristics outlined above — endogenous feedback loops, slow-moving accumulations, different timescales and non-linear causal pathways — collectively define the gendered relationship between the labour market and population mental health as a classic "wicked problem" (Rittel & Webber, 1973). Because such population health issues are not caused by a single, fixed variable but by evolving, multifaceted structural factors, they should be studied as a complex system (Apostolopoulos, 2020). Continuing to apply reductionist methods to this irreducible complexity will likely yield diminishing returns. In fact, the inability to consider this complexity has been argued as a reason for not being able to effectively tackle health inequalities (Diez Roux, 2011; Henderson, Hammond, & Homer, 2023). Therefore, to answer the research questions posed in this thesis, a method capable of formalising feedback and simulating accumulations over time is required.

2.7 Justification for system dynamics

Systems thinking and systems science offer a necessary alternative to the traditional paradigm capable of addressing the complexities identified in the previous section (Zukowski, Davidson, & Yates, 2019). Systems science can be broadly defined as "analytical approaches that aim to uncover the behaviour of complex systems" (Carey et al., 2015). There is a growing consensus that systems science must be applied to public health challenges to better

understand and address their root causes (Diez Roux, 2011; Hohn et al., 2023; Rutter et al., 2017; Savigny & Adam, 2009). The value of this approach in preventing non-communicable diseases (NCDs), including mental ill-health, is increasingly recognised by major health bodies. For instance, the World Health Organization (2022a) highlights that systems approaches can encompass the complexity of NCD drivers, facilitate consensus-building, and improve policy coherence by identifying synergies or counterproductive effects across government departments.

System dynamics is a specific methodology within the broader field of systems science that uses simulation modelling to operationalise this thinking. It is particularly valuable when systems are governed by feedback loops and when effects are not immediate but unfold non-linearly over time (Langellier et al., 2019). These are crucial characteristics of the complex system described in this chapter and the complementarity of system dynamics in epidemiology has been emphasised (Uleman, Petersen, & Rod, 2025). This section justifies why system dynamics is chosen to investigate the relationship between the labour market, gender, and population mental health.

2.7.1 Focus on feedback and accumulation

System dynamics is a methodology designed specifically for understanding the behaviour of complex systems over time (Sterman, 2010). It is ideal for studying the problem at hand due to its explicit focus on two fundamental structures: feedback loops and accumulations (Homer & Hirsch, 2006).

- Feedback loops (endogenous perspective): System dynamics frames the problem not as a set of external impacts, but as a system of internal feedback. It allows us to model the reciprocal relationship where the labour market impacts mental health, and mental health simultaneously feeds back to alter the labour market. This is not a peripheral convenience of the method but its constitutive feature (Yasarcan, 2023).

- Accumulations (stocks): System dynamics distinguishes between the state of the system (stocks) and the rates of change (flows). This naturally captures the inertia and path dependence described in Section 2.6,

allowing us to model how historical states accumulate to constrain current and future health, capture the delayed effects, and simulate long-term impacts.

This focus on stocks and flows forces operational thinking (Forrester, 1994). While traditional statistical models often focus on correlational relationships, operational thinking requires the researcher to explicitly map the physics of the system and ask how the process works. For example, a correlational view simply links unemployment to prevalence. An operational view forces us to ask: does unemployment increase the incidence of new cases (inflow), or does it suppress the rate of recovery (outflow)? By demanding this level of causal specificity, system dynamics forces the researcher to represent the mechanism through which each relationship operates; not merely that variable A is associated with variable B, but whether A affects B by raising an inflow or reducing an outflow, what the shape of the effect is, and how it interacts with other factors operating on the same mechanism. Whether the resulting specification is supported by evidence is then subject to the model's calibration and validation tests, described in Chapter 6.

2.7.2 Rigorous method to build and test theory

This thesis employs system dynamics as a rigorous method for theory building and testing. The literature review has highlighted that evidence although expansive is fragmented across disciplines. System dynamics acts as a method of triangulation, allowing these diverse evidence sources to be integrated into a single, coherent structure. The developed causal hypothesis can then be tested and refined through simulation.

In system dynamics, a causal hypothesis is a proposed structure of interconnected feedback loops and accumulations that, if operating as described, would generate the observed behaviour over time. Simulation allows us to test this hypothesis for logical consistency and examine if the hypothesised causal structure is sufficient to generate the observed historical trends in the prevalence of common mental disorders and the gender gap in that prevalence. What is assessed is not the significance of individual coefficients but the capacity of the postulated structure to reproduce the observed behaviour of the

system under plausible parameterisation. The structure is considered falsified if it cannot reproduce the reference mode under any parameterisation within the range established by the literature, or if behavioural fit is achieved only through parameter values that contradict the evidence base. The test is therefore the sufficiency of the hypothesised structure; the goal is the demonstration that the postulated mechanisms are compatible with the observed population-level trajectory. This does not establish uniqueness; other structures may generate the same behaviour.

2.7.3 Modelling at the population level

This thesis adopts the system dynamics position that structure generates behaviour: the patterns observed are understood as consequences of the system's underlying feedback and accumulation structure, rather than as the aggregate of individual actions alone. This position aligns with the population perspective set out in Section 2.5: just as Rose's distinction directs attention from the causes of cases to the causes of incidence, the system dynamics premise directs explanation from individual attributes to the feedback and accumulation structure that generates population-level trajectories. This is the basis for the aggregate, population-level representation chosen in this thesis.

Galea, Riddle and Kaplan (2010) observe that system dynamics is most appropriate where the behaviour of interest is driven by feedback between aggregate stocks and the flows that change them, over a timescale sufficiently long that within-cohort heterogeneity averages out in the accumulation. The question this thesis poses maps cleanly: macroeconomic context, gender norms, the distribution of the unpaid work burden, and the prevalence of common mental disorders are aggregate states that evolve in response to feedback from one another over decades. The mechanism of interest operates at the interface between the structural and intermediary tiers of the social determinants of mental health framework. By modelling at the aggregate level, we can focus on the structural forces, such as gender norms or macroeconomic conditions, that shape the health of the population as a whole. This level of abstraction is appropriate to examine how macro-level shifts in the labour market filter down to impact the aggregate accumulation of risk for men and women.

2.7.4 Precedence in public health research

Finally, system dynamics is not a novel method in public health; it is well-established within healthcare and health policy (Darabi & Hosseinichimeh, 2020; Freebairn et al., 2018). The aggregate stock-and-flow representation adopted here is structurally analogous to compartment models familiar in epidemiology (e.g. susceptible–infected–recovered frameworks), differing principally in the focus of endogenous feedback from outcomes back to upstream drivers.

A recent scoping review of system dynamics applications to population-level depression identifies only four models, which cover a diverse range of questions but do not study the social determinants: the comparative effectiveness of general antidepressant use versus targeted prevention among those with recurrent disease; the magnitude of recall error in estimated lifetime prevalence; smoking-related mortality among adults with depressive episodes; and the social transmission of depression through counselling interventions (Graham, Gariepy, & Orpana, 2023).

Despite system dynamics models being well-suited for analysing social determinants of health, there have been limited applications specifically in this area. Three notable examples have successfully demonstrated the value of applying system dynamics to capture dynamic complexity within the social determinants of health, by uncovering underlying mechanisms, testing intervention effectiveness, highlighting trade-offs and simulating long-term impacts: chronic illness and disability in Canada (Mahamoud, Roche, & Homer, 2013), obesity trends in the US (Chen et al., 2018), and mental health service reform in Australia (Occhipinti et al., 2021). This thesis builds upon this established tradition to explicitly model the reciprocal feedback between the labour market and population mental health through a gendered lens.

2.7.5 Limitations of systems dynamics

Adopting system dynamics as the principal analytical method requires an acknowledgement of its limitations. The most fundamental concerns the model boundary. System dynamics models, like all models, represent simplified versions of reality (Sterman, 2002); a model's behaviour is conditional on which

variables are treated as endogenous and which as exogenous. Given the dense interconnections between the social determinants of mental health, focusing on a bounded subset of mechanisms reproduces, in an attenuated form, the reductionism the systems approach seeks to overcome. The boundary chosen here is defended in Chapter 5; any tractable model necessarily selects from a wider causal architecture.

Three further classes of methodological limitation are flagged here and addressed at the points in the thesis where they become tractable: (1) Structural uncertainty (equifinality) i.e. different causal structures can produce the same observed behaviour, so behavioural fit is necessary but not sufficient evidence for the hypothesised structure (discussed in Section 3.5); (2) Operationalisation of qualitative variables such as gender norms within quantitative model equations (addressed in Section 5.7); (3) Parameter uncertainty: the literature underpinning model parameters is heterogeneous in design, population, and measurement (addressed in Section 6.8).

2.8 Research questions and objectives

This chapter has highlighted three critical research gaps in our understanding of the relationship between the labour market, gender norms, and population mental health. First, our understanding of these dynamics at the population level remains limited; individual-level pathways between the labour market and mental health are well established, but their aggregation into population-level feedback structures has not been directly investigated. Second, there is a need for more research examining how gender norms and structural discrimination shape the complex interplay between the labour market and mental health; gender has been extensively studied as a moderator of individual risk but rarely modelled as a slow-moving structural determinant that configures exposure. Third, traditional research methods have not been able to adequately consider the complexity of this system, particularly the feedback mechanisms that sustain these trends; the methodological toolkit of traditional epidemiology, while capable of isolating proximal causal associations, struggles to represent circularity, accumulation, differential timescales, and multiplicative causation simultaneously.

The overarching aim of this thesis, introduced in Chapter 1, is to develop and test a dynamic causal explanation of the structural mechanisms that have produced the rising prevalence of poor mental health, and the persistent gender disparity in that prevalence, in Scotland's working-age population between 1992 and 2019. Building on the gaps above, this aim is operationalised through three research questions:

1. How has reciprocal feedback between labour market conditions and population mental health contributed to the rising prevalence of poor mental health in Scotland between 1992 and 2019?

2. Why has the gender gap in the prevalence of poor mental health persisted despite measurable progress on indicators of gender equality?

3. Which structural mechanisms are dominant within this system, how do their effects differ between the overall prevalence and the gender gap, and which can function as leverage points for policy intervention?

To address these questions, the thesis proceeds through three empirical objectives:

1. **Map the causal structure**: Develop a causal loop diagram representing the dynamic relationships between labour market structures, gender norms, and mental health by synthesising diverse sources of evidence (Chapter 5).

2. **Simulate historical trends**: Translate the causal structure into a validated system dynamics simulation model and test whether the hypothesised structure is sufficient to reproduce Scotland's historical mental health trajectory (Chapter 6).

3. **Explain trends**: Analyse the system dynamics model to uncover how the labour market and gender norms have contributed to the rising prevalence of poor mental health and persisting gender differences (Chapter 7).

3 Methodology

3.1 Overview of the chapter

Chapter 2 established system dynamics as the appropriate methodology to investigate the relationship between the labour market, gender norms, and mental health. This chapter details how it is operationalised in this research. Section 3.2 defines the core vocabulary of the approach. It explains what system dynamics is by introducing its fundamental building blocks, specifically feedback loops and stocks and flows, and describes the typical modelling process. Section 3.3 establishes the philosophical underpinnings, grounding the research in complex realism. Section 3.4 details the specific research design, outlining the five-stage iterative process used to develop, parameterise, and validate the simulation model. Finally, Section 3.5 discusses the methodological limitations specific to this thesis's operationalisation of system dynamics, complementing Section 2.7.5 and the limitations discussed in subsequent chapters.

3.2 System dynamics: core concepts and typical process

Developed by Jay Forrester in 1958, system dynamics is a methodology designed to understand behaviour in complex systems dynamically over time. It is based on the ideas of classical feedback control theory where outputs impact inputs, creating a feedback loop.

3.2.1 Core concepts: feedback loops and accumulation

3.2.1.1 Feedback loops (endogeneity)

Chapter 2 identified the bidirectional relationship between the labour market and mental health. System dynamics formalises this by modelling feedback loops: closed chains of causal influence where variables drive one another in a dynamic process. In such systems, the distinction between cause and consequence becomes artificial.

By focusing on these internal mechanisms, the methodology seeks to identify endogenous causes of behaviour (those arising from the system's own structure) rather than attributing trends solely to external shocks. This principle of

feedback and endogeneity is central to system dynamics. It is this feedback that makes a system dynamic because any observed variable depends not only on the input but also on the past state of the system.

Feedback loops are visually represented through causal loop diagrams (CLDs, Figure 3-1). There are two types of feedback loops: reinforcing and balancing. Reinforcing feedback loops represent mechanisms through which an initial change is magnified. Conversely, balancing feedback loops represent mechanisms that counteract change and maintain system equilibrium. CLDs provide a qualitative representation of the system structure (Sterman, 1994). The strength of feedback loops and shifts in feedback loop dominance lead to dynamic behaviours.

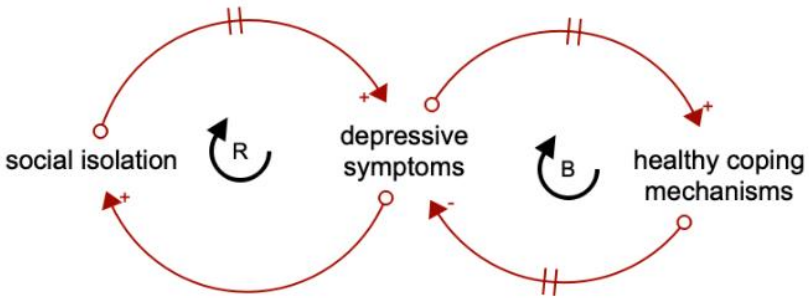


Figure 3-1: Example of a causal loop diagram (CLD)

This illustrates the two fundamental feedback structures in system dynamics.

The reinforcing loop (denoted by R) depicts a reinforcing mechanism where an initial increase in depressive symptoms, leads to a further increase in depressive symptoms, through increasing social isolation

The balancing loop (denoted by B) depicts a balancing mechanism where an initial increase in depressive symptoms leads to a decrease in depressive symptoms, through employing healthy coping mechanisms.

CLDs use arrows to depict causal relationships between variables; a plus sign (+) indicates that a change in the cause variable leads to a change in the effect variable in the same direction (i.e. an increase causes an increase and a decrease causes a decrease), and a minus sign (-) indicates a change in the opposite direction. The parallel lines on the arrow indicate a delay. The polarities are not always straightforward to interpret. For example, a positive link from social isolation to depressive symptoms does not mean that a decrease in social isolation will necessarily lead to a reduction in depressive symptoms, but that symptoms would be lower than what they would be in the absence of decreased social isolation. In other words, depressive symptoms may still increase but at a slower rate.

3.2.1.2 Stocks and flows

While feedback drives change, the memory of the system is captured through stocks and flows. These are essential for feedback loops to operate (Yasarcan, 2023). Stocks represent the state of the system at any given time (e.g. the

number of people with poor mental health), while flows represent the rate at which those states change (e.g. incidence or recovery rates). Stocks, also referred to as state or level variables, are accumulations that increase or decrease over time through flows. Due to the nature of stocks, they do not change instantaneously. As a result, complex systems contain inertia.

A stock and flow structure, often visualised through rectangles and pipes (Figure 3-2), operationalises the CLD and forms the basis of the quantitative simulation model.

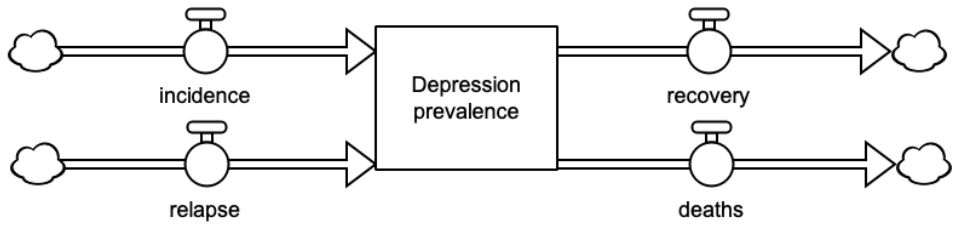


Figure 3-2: Example stock and flow structure

This illustrates the core mechanics of accumulation using the example of depression prevalence.

The stock (rectangle) represents the state of the system at a point in time; here, the total number of people with depression. Flows (pipes with valves) represent the rates of change over a specific time unit (e.g, per month). Clouds mark the model boundary, representing infinite sources or sinks outside the scope of analysis.

Dynamics: Inflows (incidence, relapse) add to the stock, while outflows (recovery, deaths) deplete it. The net change in these flows will determine the change in the prevalence of depression. Put simply, if no one recovers or dies but there are people with new diagnosis of depression in a month, the prevalence of depression will rise. If there are no new diagnoses (all flows are zero), the prevalence will stay constant. Similarly, if there are no new diagnoses or relapses, but people recover or die in a month, the prevalence will reduce. If there are new diagnoses and/or relapses, alongside recoveries and/or deaths in a month, the change in depression prevalence will depend on the net rate of the flows. If the total of new diagnoses and relapses is greater than the total of recoveries and deaths, the prevalence will increase. If it is lower, then the prevalence will decrease, and if it is the same, there will be no change in prevalence.

3.2.2 Typical modelling process

System dynamics modelling is an iterative cycle broadly comprising five stages: problem articulation, dynamic hypothesis, formulation, testing, and policy formulation and evaluation (Sterman, 2010). The process begins with qualitative conceptualisation (stages 1 and 2). This phase defines the scope and identifies the reference mode i.e. the specific dynamic trend the model aims to explain. It then maps the feedback structure, or the dynamic hypothesis, usually

articulated as a CLD. Boundary selection and dynamic hypothesis development are often done in parallel.

While CLDs provide a qualitative map of the system's structure, they cannot determine the net outcome of competing feedback loops. To understand the dynamic behaviour, the stock and flow structure is translated into a quantitative simulation model using differential equations (model formulation). This is followed by testing, often conducted in parallel. This quantitative process is essential to reveal gaps, inconsistencies, and emergent behaviours that are counter-intuitive or invisible in static diagrams alone (Forrester, 1994).

3.3 Philosophical underpinnings: complex realism and structuralism

This research is grounded in complex realism, a theoretical framework that integrates the ontological depth of critical realism with the dynamic principles of complexity theory. This choice is not merely a philosophical stance but a methodological necessity. Complex realism posits an objective reality independent of our perceptions, yet acknowledges that our experience of this reality is context-dependent while our knowledge of it is socially constructed (Byrne, 1998).

3.3.1 Ontological realism

Critical realism posits a stratified reality consisting of three domains: the Empirical (what is observed and experienced), the Actual (events that occur whether observed or not), and the Real (the underlying generative mechanisms and structures that cause events) (Mingers, 2014).

This framework aligns closely with the system dynamics modelling process employed in the thesis.

- The Real: The generative mechanisms operating within Scotland (i.e. feedback loops and stock-and-flow structure) that produce the observed trajectory; the simulation model is treated as a falsifiable hypothesis about this domain, not as a direct representation of it.

- The Actual: Corresponds to the system's behaviour over time (the simulation output).

- The Empirical: Corresponds to the observed data (the specific historical time-series datasets) used to calibrate the model.

To bridge these domains, the research employs the critical realist method of retroduction: reasoning backwards from the observed Empirical trends (the persistent gender gap) to hypothesise the Real generative structures necessary to produce them. This approach integrates two forms of causal inquiry (Campaner, 2011): (1) mechanistic causality: the development of the causal loop diagram represents a search for the web of causes that produces the phenomenon, and (2) counterfactual-manipulative causality: the quantitative simulation allows for the exploration of "what-if" questions, testing how the system's behaviour would shift if certain conditions were varied.

Crucially, the distinction between the Real and the Empirical explains how the same objective reality can generate fundamentally different experiences for different groups (Mingers, 2014). This view validates the focus on gender not just as a variable, but as a lens for understanding how different subgroups uniquely experience the same objective labour market dynamics.

3.3.2 Epistemological relativism

While the system is objectively Real, our understanding of it is rooted in epistemological relativism i.e. our understanding of that reality is socially constructed and fallible. Crucially, this implies that the modelling process is inherently reflexive, influenced by the researcher’s perspectives. Therefore, this research acknowledges that the system dynamics model is a constructed representation; “one plausible model” rather than an absolute truth (Barlas & Carpenter, 1990).

3.3.3 Structuralism and explanatory verisimilitude

Consistent with this view, this thesis adopts a structuralist approach (Clancy, Langarudi, & Zaini, 2023). Since models are glimpses into reality, the goal is not precise prediction but explanatory verisimilitude: a model that closely

approximates the underlying causal structure (Clancy, Langarudi, & Zaini, 2023). This has two critical implications for the research design:

- Theory Building: The model is used primarily as a tool for theory building and testing (Sterman, 2006). Because system dynamics models explicitly articulate cause-and-effect relationships, the method necessitates grounding model development in evidence of causal mechanisms, creating a falsifiable hypothesis to be tested through simulation (Schwaninger & Grösser, 2008).

- Validation: This stance dictates a specific approach to validation which rejects both pure relativism (where validation is purely subjective) and positivist empiricism (where validation is solely predictive accuracy) (Barlas & Carpenter, 1990). Instead, it employs triangulation, establishing validity by integrating theoretical logic, expert critique and behaviour reproduction tests.

3.3.4 Complexity perspective

Complex realism extends critical realism by explicitly incorporating complexity theory, recognising the probabilistic, emergent, and contextually contingent nature of causation (Byrne, 1998).

3.3.4.1 Aggregation

A central challenge in population health modelling is the relationship between individual risks and population outcomes. This research adopts a system dynamics perspective on aggregation. Mathematically, population health statistics are simply the sum of individual health states. However, within this methodology, the persistence of inequalities (such as the gender gap) is hypothesised to be a systemic property: an outcome that arises from the dynamic interaction of feedback loops. The model explicitly links these levels through structural feedback: aggregate stocks (such as prevailing gender norms) determine the risk environment for individuals, altering the rates at which they fall ill or recover. Within a stock, individuals are treated as identical (the perfect-mixing assumption discussed in Section 3.5); the structural feedback operates on the aggregate transition rate rather than on individual

heterogeneity. Crucially, the parameters defining these rates are grounded in robust individual-level evidence (e.g. longitudinal studies). This ensures that the aggregate feedback loops are not merely theoretical abstractions but are built on valid individual mechanisms.

3.3.4.2 System boundary

Complexity theory emphasises the inherent openness of social systems. Consequently, any attempt to model such a system requires imposing an artificial boundary, a decision that is inherently heuristic. This is not a flaw, but a necessary simplification to create a tractable representation of reality.

From this perspective, the model boundary in this research is treated not as an arbitrary cut-off, but as an explicit causal hypothesis. It represents a theoretical claim about which feedback loops are endogenous (essential to generating the problem) and which factors can be safely excluded.

3.3.4.3 Indeterminism

Because complex systems are open, causation is probabilistic. Complex realism sees the social world as structured yet ontologically indeterministic due to the inherent complexity (Williams & Dyer, 2017). Owing to this, complex realism posits that the same outcome can be reached in different ways (equifinality). A related but distinct sense of equifinality arises in simulation calibration where different parameter sets can produce the same historical fit. This has two methodological implications:

- Sensitivity analysis: System dynamics is used here to map the structural tendencies of the system. Through rigorous sensitivity analysis, the research examines how the system behaves across a range of uncertainties.

- Calibration: Because different parameter sets could theoretically produce the same historical data fit, statistical curve fitting is insufficient. Instead, the research employs constrained calibration, where parameters are bounded by theory and external evidence to ensure the model produces the right behaviour for the right reasons.

3.4 The research design: the modelling process

Having established the theoretical and philosophical basis for this research, this section details the specific research design. System dynamics modelling is not a linear process but an iterative cycle of conceptualisation, formulation, and testing. As understanding of the system deepens, earlier stages are revisited and refined.

This research followed the standard five-stage system dynamics process, integrating inductive theory-building with deductive simulation analysis. Figure 3-3 represents the iterative process followed in this research.

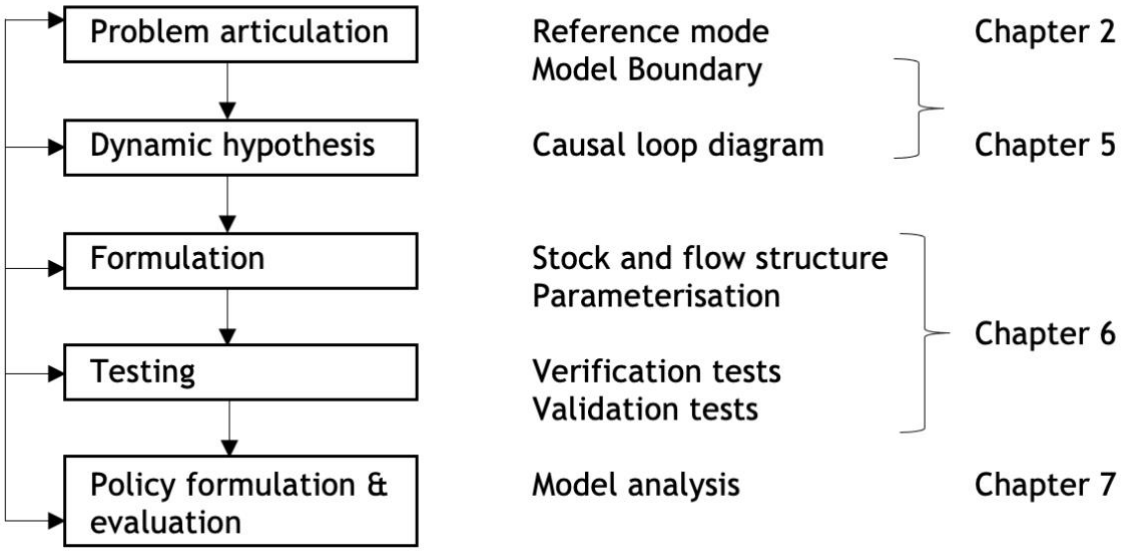


Figure 3-3: Modelling Process

3.4.1 Stage 1: Problem articulation (model boundary)

The first stage defined the scope of the modelling effort by establishing a reference mode and selecting the model boundary.

The reference mode: A reference mode is a behaviour-over-time graph that highlights the historic and expected trends of the key variable of interest. For this thesis, the primary reference mode is the increasing prevalence of poor mental health and persisting gender differences among working-age adults in Scotland. This trend, identified through the analysis of the General Health Questionnaire (GHQ-12) data, serves as the anchor for the model; the aim is to reproduce this dynamic behaviour structurally.

Model boundary selection: The model boundary determines which variables are endogenous (determined within the system), exogenous (external inputs), or excluded entirely. The boundary was drawn to be broad enough to capture the primary endogenous drivers of the problem without becoming intractable.

Initial boundary selection was informed by identified literature (Chapter 2) and focused on unemployment and precarious employment. This boundary was refined iteratively through the literature review and confirmed through expert consultation to ensure it captured the essential dynamics of the Scottish context (Hong et al., 2022). The literature review and expert consultation process are detailed in Chapter 4. The resulting model boundary is presented in Chapter 5.

3.4.2 Stage 2: Dynamic hypothesis (causal loop diagram)

This stage involved conceptualising the underlying feedback structure capable of generating the reference mode in the form of a causal loop diagram (CLD). Literature review was chosen as the primary basis for developing the CLD to ground the causal relationships in good evidence and increase confidence in the structure of the model (Barbrook-Johnson & Penn, 2022; Homer, 2014). An inductive approach was used to develop the CLD through a critical synthesis of existing literature.

Adopting feedback thinking required the researcher to actively search for feedback mechanisms by questioning how variables may be connected. This informed the research design where an iterative process of reviewing was undertaken. The addition of any variable meant that its causal relationship with every other variable in the CLD was investigated, including the possibility of an indirect causal relationship. Figure 3-4 shows this CLD development process. The detailed literature review and synthesis process is described in Chapter 4.

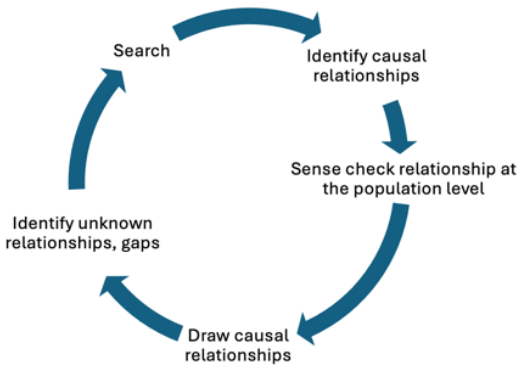


Figure 3-4: CLD development process

Throughout this process, variables were aggregated and reconceptualised to suit a population-level model. For example, the conceptualisation of women's labour force participation evolved from a simple full-time/part-time dichotomy to a more nuanced representation of gender parity in working hours. More importantly, close attention was paid to ensure that the model does not treat population outcomes as a simple sum of independent individuals; instead non-linearities are introduced where individual-level mechanisms imply that thresholds shift with population state (Koopman & Lynch, 1999). For example, the effect of poor mental health on labour market outcomes was theorised as non-linear because, as the prevalence of poor mental health rises, thresholds for exiting employment may shift; at high prevalence, employer tolerance and stigma effects compress the marginal effect of any individual case.

3.4.3 Stage 3: Formulation (simulation model)

The third stage involved translating the dynamic hypothesis into a quantitative stock-and-flow model using the software Stella Architect. This involved describing the feedback structure mathematically and operationalising the causality. For example, the causal link from unemployment to poor mental health is operationalised through the incidence rate, where the proportion of people unemployed impacts the rate of new cases.

The model was initialised using available data from the Office for National Statistics (ONS), published papers and analysis of the UK Household Longitudinal Study (UKHLS). It was parameterised using a multi-method strategy combining literature-derived effect multipliers, primary data analysis, and constrained

calibration for unobservable parameters (detailed in Chapter 6). Where possible, parameters for individual-level risks (e.g. the impact of unemployment on mental health) were derived from meta-analyses of longitudinal studies. For soft variables that are unobservable, such as the adjustment time of gender norms, parameters were estimated based on theoretical reasoning and refined through constrained calibration. Details of the formulation of the model structure and operationalisation are discussed in Chapter 6.

3.4.4 Stage 4: Testing and validation

Testing was conducted in parallel with formulation to ensure both the structural validity and behavioural accuracy of the model. Following the standard system dynamics distinction, verification asks whether the model is a faithful translation of the intended structure, while validation asks whether the model adequately represents the system it is intended to explain.

Model verification was achieved through documenting and checking equations, employing tests such as dimensional consistency (checking unit integrity) and documenting how each link of the CLD was captured in the model. Extreme conditions tests and structural tests of behaviour were also employed to check if the model behaves logically.

Validation was assessed through a combination of expert feedback on the CLD, structural tests to examine the necessity of loops and behaviour reproduction tests to check if the model can endogenously generate the historical trends observed in Scotland.

The methodology for expert input is discussed in Chapter 4, and the feedback and subsequent changes to the model boundary and causal hypothesis are presented in Chapter 5. The specific tests applied and their results are covered in Chapter 6.

3.4.5 Stage 5: Policy formulation and evaluation

The final stage shifted from inductive theory building to retroductive analysis. Scenarios, including changes in labour market conditions and gender norms, were simulated, and the resulting model behaviour was used to reason

backwards from the observed dynamics to the generative mechanisms responsible for them. This analysis, presented in Chapter 7, identifies which dominant feedback loops drive the rising prevalence of poor mental health and the persistence of gender inequalities.

3.5 Methodological limitations

This research employs system dynamics to investigate the complex structural drivers of mental health trends grounded in complex realism. While this approach offers distinct advantages for analysing feedback and accumulation, it is subject to specific methodological limitations. This section outlines the limitations inherent to the philosophical stance and the chosen modelling framework.

Establishing causality in complex systems: This research aims to develop a causal hypothesis of a complex system. It is crucial to acknowledge that establishing causality in complex social systems remains inherently challenging (Kelly-Irving et al., 2023). The causal relationships represented in this model were identified through a comprehensive review of existing literature that prioritised longitudinal studies and temporal analysis to better ascertain causality. Theoretical frameworks, individual-level studies, and aggregate-level analyses were triangulated to strengthen the confidence in the validity of the identified causal links, as relying solely on aggregate data can be less robust for establishing causality. Theories and individual-level studies, grounded in causal inferences, provide insights into the mechanisms and pathways, while aggregate-level studies confirm the relationship at the aggregate-level.

Nonetheless, the causal structure derived from literature is subject to limitations, such as incompleteness, bias, and most importantly an inherent reductionism. The included evidence largely focused on isolated variables, overlooking the complex interplay of factors within the real-world system and potentially mis-specifying the relationship. Furthermore, translating the qualitative causal relationships into the quantitative relationships within the simulation model required simplifications and assumptions.

These are necessary limitations of this research as it attempts to unpack complexity. Consistent with complex realism, system dynamics combines qualitative and quantitative methodologies, facilitating the articulation of the underlying causal structures and the rigorous examination of the developed hypothesis and its implications. The quantitative system dynamics model functions to test the dynamic hypothesis with data. Rigorous iterative model development, sensitivity analysis, and validation procedures increase confidence in the plausibility and robustness of the causal relationships depicted.

Structural uncertainty (equifinality) is a related limitation of simulation-based causal inference: different causal structures can in principle produce the same observed behaviour. Passing the behaviour reproduction tests reported in Chapter 6 is therefore necessary but not sufficient to establish that the hypothesised structure is the one generating the observed trajectory in reality. This limitation is intrinsic to simulation methods; it is mitigated, not eliminated, by requiring behavioural fit to be achieved only under parameterisations consistent with the independent evidence reviewed in Chapter 2, and by subjecting the structure to expert disconfirmation (Chapter 4). Consistent with the complex-realism stance set out in Section 3.3, the model is treated as one plausible structure rather than the unique generative mechanism for the observed phenomena.

Inferential status: A further consequence of this methodological choice is that, unlike quasi-experimental econometric designs, system dynamics does not identify causal effects against an explicit counterfactual. What it offers is a different kind of evidence: a demonstration that a postulated structure is behaviourally sufficient to generate the observed population trajectory. This is a different claim from causal identification, and one appropriate to the question asked: whether a set of coupled mechanisms, operating together at their characteristic timescales, can generate the observed trajectory. The choice of method reflects the shape of the research question, not a claim that system dynamics is inferentially superior to quasi-experimental alternatives for the questions those methods are designed to answer.

Aggregation and homogeneity: A primary limitation stems from the aggregate nature of the model. By choosing to model dynamics at the national level, the

model does not capture heterogeneities at the regional, subgroup, or individual level. Due to the inherent "perfect mixing" assumption of stocks within system dynamics, every individual is treated as identical. For instance, the stock of healthy does not distinguish between people based on age, ethnicity, or socioeconomic background. Consequently, the model assigns the same average probability of developing poor mental health to everyone in that stock, regardless of their individual risk factors. The implication of this limitation is that the model cannot capture how individual-level heterogeneity (outside of gender) influences outcomes. The feedback mechanisms identified operate at the population level; their applicability to specific subgroups (such as by socioeconomic group, age cohort, ethnic group, or geographic region within Scotland) requires further investigation.

However, this aggregate approach is a deliberate choice justified by the research aims to explain the persistence of population-level trends driven by structural forces. By abstracting away from individual idiosyncrasies, the model remains tractable and focused on the dominant feedback loops that shape the risk environment for the population as a whole. Therefore, despite the loss of granular detail, the aggregate approach is the most appropriate and parsimonious method for investigating these high-level structural dynamics.

Reflexivity: The process of model development was inherently influenced by the researcher's choices and perspectives. The methodological stance adopted (Section 3.3) treats this reflexivity as a constitutive feature of the modelling process rather than a defect to be eliminated. To address this, transparency in documenting model assumptions and limitations was prioritised to allow for critical evaluation of the research. Expert consultation was explicitly employed as a validation step, seeking external feedback on the accuracy of the causal hypothesis (against the literature) and sufficiency of the model boundary, to mitigate individual bias. The risks specific to a single-researcher synthesis are discussed in Chapter 4.

4 Methodology for Conceptual Model Development: Literature Review and Expert Consultation

4.1 Introduction

Building on the methodological framework established in Chapter 3, this chapter details the operational process used to develop the conceptual basis of the system dynamics model (in the form of a CLD). The development process was iterative, grounded in a critical review of existing literature and refined through structured expert consultation.

As highlighted in Chapter 2, the complexity of the labour market-mental health nexus requires a departure from linear research protocols. Consequently, this research employed a feedback-driven literature review to refine the model boundary and construct the initial dynamic hypothesis (Section 4.2). This was followed by a rigorous disconfirmatory expert consultation process to validate and refine these (Section 4.3). A subsequent presentational refinement of the resulting CLD, undertaken to prioritise communication of the core dynamic hypothesis, is described in Section 4.4. This chapter outlines how these methods were applied to synthesise diverse evidence into a coherent system structure. It discusses the choices and decisions that were made throughout the process: adding/removing variables and causal relationships, aggregating/disaggregating concepts and changing the visualisation. Reflections on the methodology and its limitations are discussed in Section 4.5.

4.2 Literature review methodology

Grounding the CLD in existing evidence builds model credibility and ensures the causal structure is robust. The aim of the literature review was threefold:

- to refine the model boundary, which required identifying relevant variables;

- to develop the theoretical basis for the system dynamics model, in the form of a CLD, which required identifying the causal relationships between variables; and

- to support the development of the system dynamics model, by providing insight into the functional form and strength of the causal relationship.

4.2.1 Search

The review process departed from traditional linear systematic review protocols to accommodate the feedback perspective. While a standard systematic search can in principle identify any pre-specified relationship, the linear protocol of fixed search strings and fixed inclusion criteria is poorly suited to surfacing causal relationships that were not anticipated at the protocol stage. This means that reciprocal causality operating through indirect pathways might be missed. For example, while an initial search might confirm that precarious employment increases poor mental health, a feedback perspective necessitates a deliberate, targeted search for the reverse pathway: does aggregate mental health impact labour productivity or labour supply? This “closing the loop” strategy meant that the search process was dynamic; the identification of a new variable acted as a trigger for a new, targeted search to map its antecedents and consequences.

This heuristic approach often necessitated a disciplinary migration. For instance, identifying precarious work as a driver in the health literature triggered a shift into labour economics literature to understand the upstream drivers of precarity itself. This cross-disciplinary snowballing was essential to capture the full feedback cycle, preventing the model from treating economic drivers as purely exogenous shocks. Therefore, an iterative, four-stage search strategy was developed (Figure 4-1). The search strategy for stage 2 is in Appendix A.

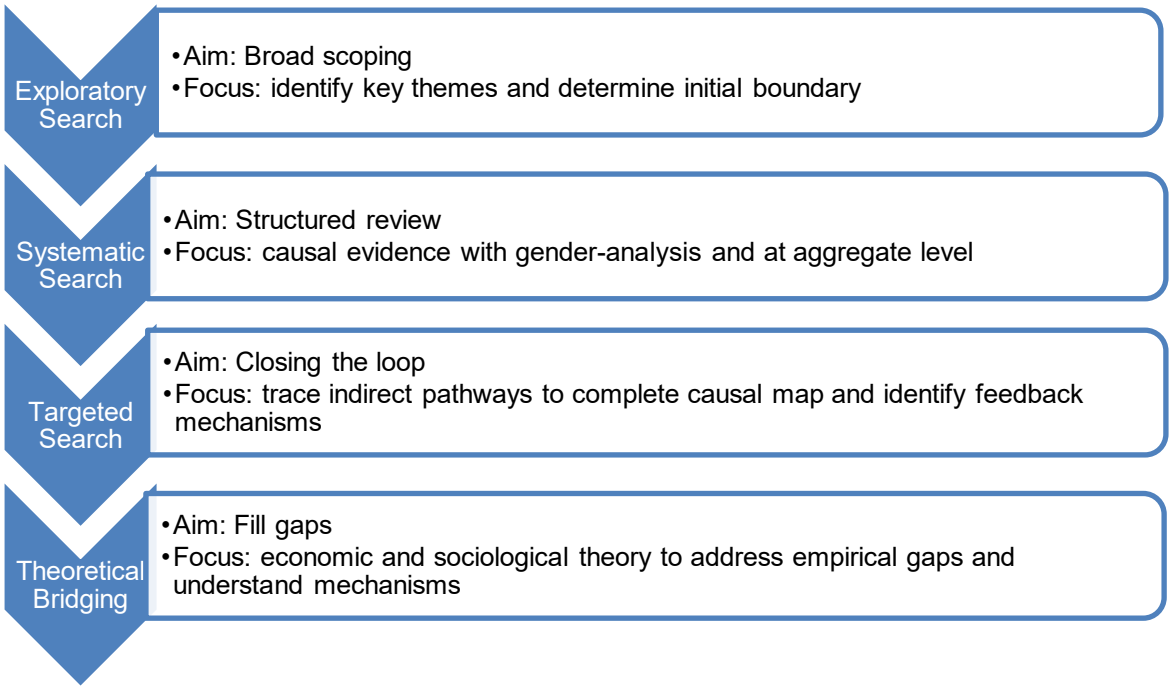


Figure 4-1: Literature review process.

4.2.2 Identification of relevant literature

Given the breadth of the system, a strict set of prioritisation criteria was applied to manage the volume of literature while ensuring model robustness. Studies were prioritised not just on quality, but on their relevance to population-level dynamics and gendered mechanisms. Table 4-1 provides the inclusion, exclusion criteria and prioritisation criteria. In general, meta-analyses and systematic reviews were prioritised. Where possible, causal relationships were evidenced with both individual and aggregate level studies. Figure 4-2 gives an overview of the screening and prioritisation process. Figure 4-3 uses the PRISMA flow chart to show the identification and prioritisation of studies. PRISMA is adopted here for reporting transparency rather than as a search-protocol; the iterative, feedback-driven character of the search means the diagram represents the cumulative inclusion record across multiple search waves rather than a single pre-specified search.

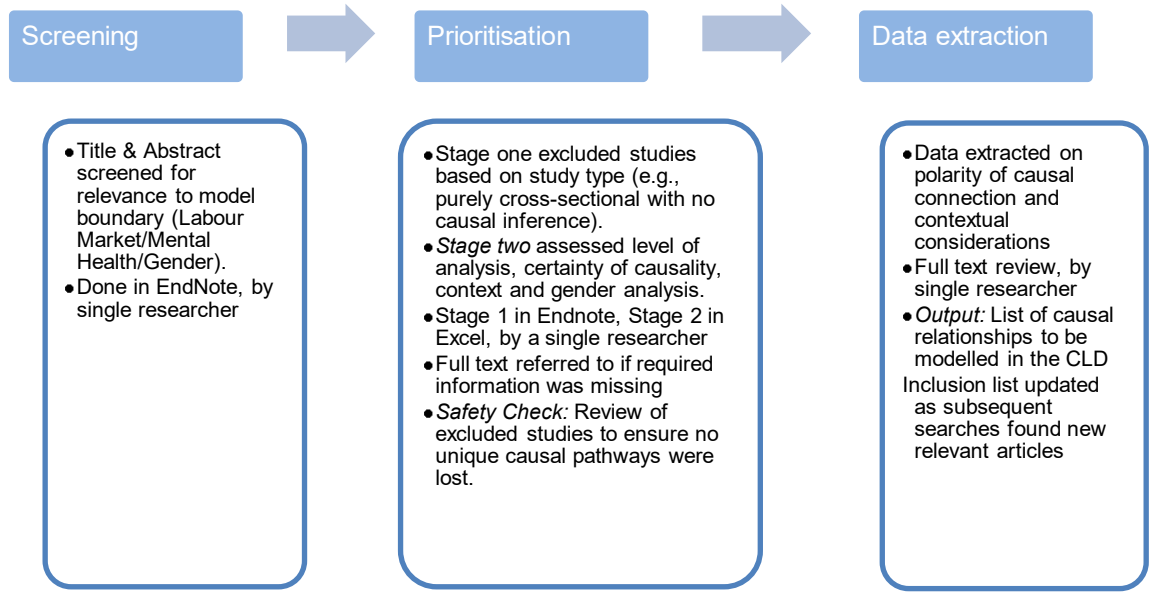


Figure 4-2: Screening and prioritisation process

Figure 4-3: PRISMA flow chart

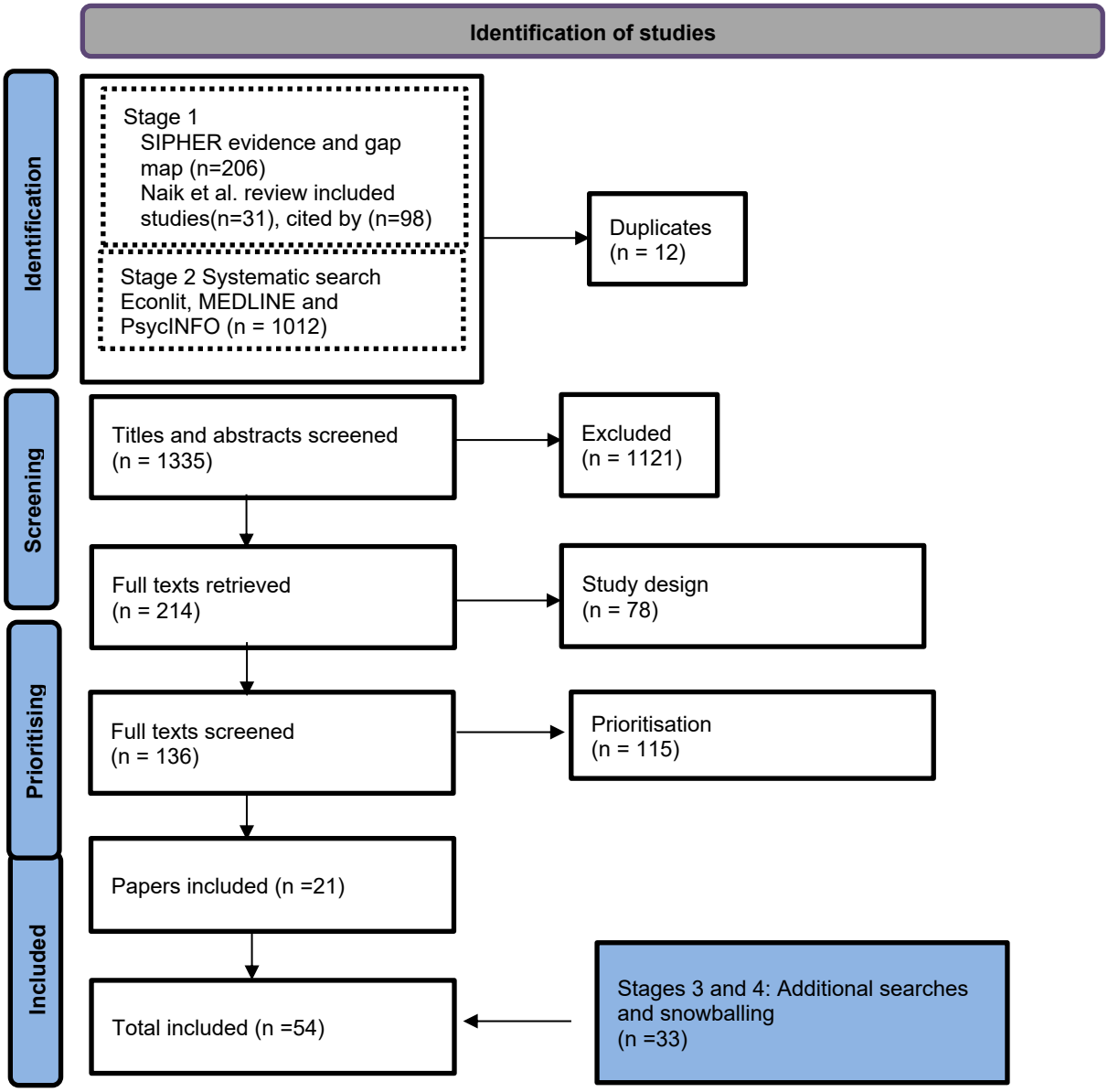


Table 4-1: Inclusion, exclusion and prioritisation criteria

	Include	Exclude	Prioritise
Population-level dynamics			
Study Type	aggregate level studies	trend analysis in only one variable	studies establishing temporality
Variables		variables relevant to smaller spatial scales (such as individuals or firms); e.g. job control, job demands, leadership and management, job performance	
Population		specific subgroups based on occupations (e.g. healthcare workers) or demographics (e.g. single mothers, low-income, ethnicity, refugees)	
Focus on the relationship between mental health and the labour market			
Study Type	longitudinal studies	qualitative and non-empirical studies	meta-analysis given they provide quantitative data required for simulation
Variables	general measures of poor mental health or common mental disorders	postpartum depression (labour market is not an important factor) severe mental illness, suicides and healthcare use (individual level factors are more important) job satisfaction and burnout (not recognised as measures of mental health, focus more on the job specific characteristics)	self-reported measures over clinical diagnosis
	structural variables	interventions and policies	behaviour over intention (e.g. intention to return was excluded given existence of studies measures re-employment)
Population	Working-age population	retired populations (65+)	general population, rather than specific age-groups (e.g. young adults (16-25 years) or older populations (55+))
Gendered lens			
Study Type		studies merely reporting trends by gender without an investigation of contributing factors	studies including a gender analysis
Variables	policies related to childcare	childcare by family members	
Population	women only		
UK based (Given the existence and strength of causal mechanisms can differ across contexts, the aim was to focus on similar contexts.)			
Study Type			
Variables			
Population	high-income countries with similar welfare regimes	low- or middle-income countries culturally different contexts such as Japan and China	studies conducted in the UK

4.2.3 Data extraction

The following information was extracted from the literature: (1) key variables, to identify what is considered within the model boundary and what is excluded, (2) drivers of key variables, to assess whether the variable will be modelled endogenously (i.e. changes dynamically based on other model variables) or exogenously (i.e. variables not modelled to be impacted by other variables in the model but feeding into model variables). For example, the working-age population is an important driver of labour dynamics, however it is driven by factors beyond the scope of the model: birth/death rates and migration. Thus, it was included as an exogenous variable, (3) causal relationships between the variables within the model boundary, to develop the CLD, (4) measures and conceptualisations, to critically assess the relevance of the link at the population level, ways to represent it in the CLD and operationalise it in the model.

4.2.4 Synthesis of the literature

Following the identification of variables and causal relationships through the iterative literature review, the findings were synthesised to construct the preliminary CLD in Stella Architect. Rather than a linear process of compilation, this stage involved an iterative synthesis where variables and relationships were constantly refined, aggregated, or reconceptualised to ensure the final system structure remained both theoretically coherent and tractable. To ensure the model was suitable for population-level simulation, this synthesis was guided by specific formulation principles regarding the translation of evidence, aggregation, and model boundaries.

Translation of evidence to the population level: Causal relationships identified in the literature were critically reviewed based on study design, preponderance of evidence, and relevance to the context. A critical step involved the logical extrapolation of individual-level mechanisms to the population level. Relationships were included only if the individual-level mechanism could be logically aggregated to a population-level dynamic. For example, individual-level evidence suggests that unemployment increases the probability of poor mental health; this implies that an increase in the unemployment rate would increase

the prevalence of poor mental health in a population. Conversely, mechanisms that do not translate directly were excluded. For instance, individual-level evidence suggests those with poor mental health are more likely to be sorted into precarious employment. At the population level, this is captured indirectly in the conceptual model: an increase in the prevalence of poor mental health reduces aggregate productivity, which reduces worker bargaining power and is associated with an increase in the prevalence of precarious work.

Variable conceptualisation: Given the CLD was being developed as a precursor to a simulation model, it was important to conceptualise variables in ways suitable for operationalisation (i.e. refining concepts into measurable units that could be populated with data). This was particularly relevant for gender norms. While the literature discusses gender norms broadly as a sociological construct, the model required a measurable proxy, such as the prevalence of traditional gender role beliefs, that could be operationalised using available longitudinal data. Similarly, the conceptualisation of mental health was refined to ensure that it could capture general levels of distress in the population. This allows the model to focus on the structural relationship between aggregate health states and the labour market, rather than the clinical nuances of specific conditions.

Conceptual aggregation: To maintain model parsimony and focus on structural feedback, complex concepts were aggregated into single variables. This approach was applied to economic inactivity. While the economically inactive population is heterogeneous (comprising students, the long-term sick, and carers), modelling each subgroup individually would introduce unnecessary complexity without altering the fundamental feedback loop between health and labour supply. Therefore, a methodological decision was made to aggregate these groups into a single stock to focus on the net flow of labour withdrawal. Thus, the CLD does not contain a direct causal link from the stock of *Economic Inactivity* to *Prevalence of Poor Mental Health*. The impact on mental health is realised through the flows between these states. For instance, an individual moving from *Economic Inactivity* to *Employment* will experience an improvement in mental health, and this dynamic is captured explicitly in the stock-and-flow formulation of the simulation model (Chapter 6).

Aggregation of pathways: The synthesis process also involved simplifying causal chains to capture the essential dynamic effect without excessive detail. Where multiple intermediate steps led to the same outcome, they were collapsed into a single aggregate pathway. Explicit feedback was sought from experts on such structural simplifications. For example, the literature indicates that lower working hours for women lead to lower earnings through multiple mechanisms: occupational segregation, fewer progression opportunities, and lower returns to experience. Modelling each mechanism separately would add significant structural complexity. Therefore, the chosen structure makes a simplifying assumption that the combined net effect can be effectively represented by two direct pathways: one linking reduced hours to the gender pay gap, and another linking reduced hours to the sorting of women into insecure work. This approach prioritises capturing the essential feedback structures and behavioural dynamics rather than achieving perfect micro-level fidelity.

Criteria for endogenous and exogenous variables: The classification of variables as endogenous or exogenous was determined by the requirement to close feedback loops central to the research aims. For instance, labour demand is driven by global economic forces beyond the scope of a national public health model. It is therefore treated as an exogenous input, modified endogenously by labour productivity in the conceptual model presented in Chapter 5. The productivity-mediated pathway was subsequently excluded from the simulation model on parsimony grounds (Chapter 6); the conceptual treatment described here therefore corresponds to the CLD as a theoretical hypothesis, while the simulation tests a more parsimonious subset of that structure.

Resolving polarity issues: Identified relationships were summarised by polarity to ensure structural consistency. Generally, the review found no contradictions in polarity across the identified studies, apart from the relationship between labour productivity and labour demand, which admits both a displacement effect (negative: higher productivity reduces the need for workers) and a compensation effect (positive: productivity gains lower costs and stimulate expansion) (Calligaris et al., 2023; Chen & Semmler, 2018; Cruz, 2023). However, in the final model this direct link was replaced (see Section 5.5.3) by a link from productivity to insecure work, sidestepping the displacement/compensation

tension by routing the dynamic through worker bargaining power rather than through the employment volume channel.

In summary, the literature review combined an iterative, feedback-driven search strategy with structured prioritisation criteria and explicit synthesis principles governing the translation of individual-level evidence to population-level dynamics, the aggregation of variables and pathways, and the resolution of polarity contestation. The output was a preliminary CLD that was subsequently subjected to disconfirmatory expert consultation discussed in the next section.

4.3 Expert engagement

This section outlines the process of expert engagement undertaken to validate and refine the model boundary and CLD. The objective was not to collect data for qualitative analysis but to identify and correct potential misinterpretations of literature by the researcher, particularly with respect to contextual relevance and aggregation of individual-level effects to model the population-level.

The expert engagement process was designed to invite a holistic critique of the causal hypothesis and the appropriateness of its defined boundaries. This approach aligns with the model's purpose as a theoretical exploration, where the focus of validation is on the plausibility of the core assumptions, not on detailed parameterisation. This approach enabled harnessing domain-specific expertise to seek conceptual validation and to pressure-test the hypothesised causal structure against the real-world knowledge of policy and research experts (Blake et al., 2024).

4.3.1 Expert identification and recruitment

Participants were selected using purposive sampling to ensure diversity across three dimensions of relevance to the model: (i) the Scottish policy and labour market context, (ii) gender as it intersects with employment and mental health, and (iii) macro-level labour market dynamics. Identified experts included epidemiologists, academic researchers focusing on gender and work, and economists and policy analysts from think tanks and governmental organisations including Public Health Scotland, Department for Work and Pensions and

Department of Health and Social Care. Ten experts were recruited; this number was determined by a combination of saturation considerations and the practical aim of triangulating across perspectives in each of the three substantive areas above.

Experts were contacted by email with an information sheet and consent form (Appendix B); written consent was obtained prior to interview. Ethical approval for expert consultation was obtained from the research ethics committee at the College of Medical, Veterinary and Life Sciences, University of Glasgow (application number: 200230153). Interviews were recorded with consent which were stored on secure storage in accordance with the Data Management Plan.

4.3.2 Engagement design and process

4.3.2.1 Interview format

One-to-one, semi-structured interviews of approximately 45-60 minutes each were conducted online via Microsoft Teams between June and July 2025. Interviews were conducted online as participants were located across the UK and the CLD could be easily presented using interactive tools. Interviews were chosen because each expert interview can be seen as an independent validation of the preliminary CLD. Interviews also allow for focusing on specific parts/links based on expertise and on gaps in the validation process, and for talking through the CLD at the interviewee's pace. Additionally some studies have found that it allows experts to more easily voice disagreements compared to group settings, where participants are more likely to build consensus with dominant opinions even when inaccurate (Lindqvist & Carnohan, 2023).

4.3.2.2 Methodological strategy

To mitigate the risk of confirmation bias inherent in a single-researcher synthesis, a disconfirmatory approach was deliberately employed (Andersen et al., 2012). The objective was not merely to seek validation, but to actively uncover the model's flaws and weaknesses. Instead of asking experts if they agreed with the model, the researcher explicitly tasked them with critiquing its accuracy ("do any relationships seem weak or counterintuitive?"), and the overall sufficiency of the model boundary ("are there any important variables... that I

haven't covered?"). Framing critique rather than confirmation as the desired contribution shaped the interview dynamic, eliciting fundamental challenges to the model's assumptions rather than polite agreement.

In developing the model, tentative relationships and variables were strategically included to signal uncertainty and actively seek expert feedback. Representing certain causal pathways as dotted links (such as the impact of the gender pay gap on mental health) or conceptual variables with less certainty (such as the 'growth in median wages') served as an explicit invitation for disconfirmation. This technique was employed to move beyond merely presenting a finished theory and instead engage the experts directly in the process of refining the model's most ambiguous and contested areas.

4.3.2.3 Interview protocol

Orientation and terminology: The session began by introducing the expert to system dynamics concepts (e.g. feedback loops, delays) to ensure a shared language. The methodology and the specific population-level focus of the model were clarified to manage expectations regarding individual-level nuance.

CLD presentation: The CLD was presented systematically, link by link, to make the structure accessible, focus attention on a single links at a time, and facilitate detailed, in-depth discussion of each.

Guiding questions: Experts were asked to provide feedback on three key areas: the accuracy of the causal relationships against their understanding of the evidence; the sufficiency of the model in terms of any critical missing variables; and the overall usefulness of the hypothesised feedback mechanisms. This process was designed to pressure-test the theoretical assumptions of the model against real-world policy and research perspectives, ensuring the final structure was both robust and relevant.

Contextual validation: Discussions explored whether the generic mechanisms accurately reflected the specific Scottish and UK policy context, and if the distinct feedback loops for men and women aligned with their professional understanding.

4.3.3 Synthesis of expert feedback

The feedback gathered from each interview was used to refine and critique the CLD. Data from interviews were extracted and a summary with key feedback from each interview across the CLD's causal relationships and operationalisation of variables was created (and is provided in Appendix C). Feedback was thematically analysed to focus on three specific dimensions:

- **Conceptualisation:** Whether the variable names and aggregation resonated with domain experts (e.g. the terminology of 'precarity' vs 'insecurity').

- **Accuracy:** Identification of missing links or feedback loops that the literature review may have overlooked.

- **Boundary judgments:** Validation of the decision to exclude exogenous factors such as the welfare system.

Where experts offered contradictory views, such as on the strength of the link between poor mental health and economic inactivity, the causal link was retained in the conceptual model but was explicitly marked for rigorous sensitivity analysis in the simulation phase (if there was lack of UK-specific parameters) to test the implications of these different perspectives.

In summary, the structured expert consultation provided a disconfirmatory check on the preliminary CLD across three dimensions: conceptualisation, accuracy, and boundary judgments. Feedback was integrated where multiple experts converged or where it aligned with the literature; contested links were retained in the conceptual model if they aligned with the literature and explicitly flagged for sensitivity analysis in the simulation phase rather than resolved through structural change. The substantive content of this feedback and how it shaped the final CLD is reported in Section 5.5.

4.4 Refinement for narrative clarity

Following the expert consultation process, the CLD was iteratively refined to prioritise the communication of the core dynamic hypothesis. These refinements

involved a process of structural abstraction, where the diagram was simplified to highlight the primary feedback mechanisms without compromising the underlying logic of the simulation model. A link-by-link documentation of the changes to the CLD is reported in Appendix C.

Aggregation: The most significant modification was the consolidation of the three labour market states (Employed, Unemployed, Economically Inactive) into a single variable, Employed Proportion, on the CLD. The simulation model in Chapter 6 retains the full stock-and-flow structure with separate stocks; this consolidation is purely a presentational choice for the qualitative diagram. The choice was justified on three grounds. First, parsimony: in a finite working-age population, the Employed Proportion is the complement of (Unemployed + Inactive), so any loop running through the Employed stock implicitly closes through the corresponding non-employed stock. Second, the central feedback structure of interest in this thesis runs through the employment channel, for its protective effect on mental health (latent functions), its gender difference and for its labour market signalling role. Third, displaying all three states on the CLD without showing the directional flows between them risks misrepresenting bidirectional transitions as static cross-sectional ratios. The implication of this simplification is that the CLD discusses exits to economic inactivity without depicting Inactivity as a separate node; references to these should be read as flows captured explicitly in the simulation model in Chapter 6.

Presentation: Additional refinements included the removal of purely exogenous variables, such as childcare availability. While these factors remain parameterised within the simulation, their inclusion in the qualitative CLD added complexity without contributing to the endogenous feedback loops that drive system behaviour. Finally, variable nomenclature was standardised for brevity (e.g., 'Women's Unpaid Work Burden' was condensed to 'Unpaid Work'), and the colour scheme was updated to provide a visual distinction between general labour market dynamics and those specifically related to gender parity.

These modifications ensure that the CLD provides a parsimonious representation of the system's endogenous drivers, establishing the structural foundation for the detailed causal analysis presented in the following chapter.

4.5 Methodological reflections and limitations

This chapter has outlined a rigorous methodology for developing the conceptual model. General limitations of the system dynamics framework (causality in complex systems, aggregation, reflexivity) are discussed in Section 3.5. This section focuses on limitations specific to the conceptual development methodology described in this chapter: the search strategy and review synthesis, the expert validation process, and boundary selection.

Literature review search strategy and synthesis bias: A primary methodological insight from this research is the tension between traditional systematic review protocols and the requirements of feedback thinking. While the initial research design followed principles of systematic review to ensure rigour, the limitations of this linear protocol ultimately necessitated a shift to an iterative, feedback-driven enquiry. Although it was anticipated that the review would require some iteration, the sheer complexity of the system meant that the initial systematic search was potentially less efficient than a heuristic approach. Systematic protocols prioritise the reproducibility of a pre-defined search string, but identifying feedback loops requires a heuristic approach of "following the thread" where the identification of one variable (e.g. precarious work) triggers a new, targeted search for its reverse impact (e.g. on productivity). Consequently, a significant portion of the final model structure was derived from iterative searches and snowballing rather than the initial systematic search. This research demonstrates that complex-systems mapping requires a search architecture genuinely different from systematic review and offers feedback-driven literature review as one such approach. Future research in this field should reconsider how rigour is defined, prioritising a transparent, iterative search strategy explicitly designed to trace causal chains, rather than adhering to traditional systematic review standards which may be structurally ill-suited for complex systems mapping.

The construction of the model was inherently influenced by researcher reflexivity. The entire process of search identification, data extraction, and synthesis was conducted by a single researcher, introducing a risk of selection bias and subjective interpretation of interdisciplinary evidence. While the expert consultation process was designed to mitigate this by disconfirming the

researcher's assumptions, the process of reconciling expert feedback required the researcher to act as the final theoretical arbiter. Thus, despite the external validation, the final model structure remains a constructed representation inevitably influenced by the researcher's own theoretical lens and decisions during synthesis. A specific risk arising from single-researcher synthesis is that missing variables and missing loops are particularly hard to surface, because the researcher cannot search for what they have not yet conceptualised. This was mitigated by the disconfirmatory expert protocol (see below), but only partly: experts could be asked to identify what was missing, but they could only do so within the variables and frame the researcher had already constructed.

One such frame-level constraint should be stated explicitly: the review of the evidence base, the model, and the gender analysis throughout this thesis treat gender as a binary. The literature search was scoped to studies comparing women and men, and did not seek evidence on non-binary individuals or on the intersection of gender with ethnicity, disability, or socio-economic position. The model therefore cannot speak to these groups. This is stated as a scoping choice that reflects the structure of the available population-level evidence.

Expert validation process. A strength of the conceptual model is that it was not merely literature-derived but expert-validated through a structured, disconfirmatory protocol, a more demanding test than confirmation-seeking review, because experts were asked to falsify the hypothesised structure rather than endorse it. That strength is, however, bounded by three key limitations. First, the sample of ten experts, while disciplinarily diverse, was not exhaustive: validation was strongest for the labour market and gender dimensions and weaker for, for example, welfare-conditionality. Second, where experts disagreed (e.g. on the strength of the mental health → economic inactivity link), the conceptual model retained the one routed in the reviewed evidence, which is methodologically defensible but means that the CLD encodes some links whose magnitude remains contested. Third, expert validation can disconfirm individual links and identify missing variables that experts are positioned to know about; it cannot detect structural blind spots affecting all experts in the sample. Single-researcher synthesis combined with sample-bounded expert review remains an irreducibly partial validation strategy.

Boundary selection: Finally, to create a tractable simulation model, the "open" social system must be artificially "closed" within a defined model boundary. This process of abstraction requires excluding highly relevant factors to maintain analytical focus. The methodology required drawing a "hard" boundary that treats the welfare system as a fixed context rather than an active variable; its implications are discussed in the next chapter. To mitigate the potential for inadequate boundary selection, the boundary was refined iteratively throughout the research process.

5 Conceptual Model: Model Boundary and Causal Loop Diagram

5.1 Overview of the chapter

Building on the methodological framework established in Chapter 4, this chapter presents the conceptual model of the thesis. Whereas Chapter 2 reviewed the evidence for individual causal relationships between the labour market, gender, and mental health, this chapter synthesises the literature into a coherent dynamic hypothesis: a structure of interacting feedback loops that, if operating as specified, would be sufficient to generate the trends discussed in Chapter 1. By mapping the underlying feedback structure, this chapter achieves the first objective of the research.

The primary output of this chapter is the Causal Loop Diagram (CLD). This diagram represents the system structure: the set of feedback loops that is hypothesised to generate the persistent trends of rising distress and gender inequality in Scotland. This chapter bridges the gap between theory and simulation; it defines the qualitative map of the system before Chapter 6 details the quantitative translation of this map into a formal simulation model.

The chapter is structured as follows: Section 5.2 presents the CLD, providing a visual map of the reinforcing and balancing loops that drive the system's behaviour. Section 5.3 then formally operationalises the core variables selected for this model and details the evidentiary basis for each causal link. Section 5.4 defines the model boundary, detailing what is endogenous (modelled) and exogenous (excluded). Section 5.5 outlines how expert consultation refined the model structure. Section 5.6 discusses the theoretical contribution of the CLD, and Section 5.7 concludes the chapter with the strengths and limitations.

5.2 Causal loop diagram

The culmination of the literature review and expert consultation is the CLD presented in Figure 5-1. It consists of variables connected by causal arrows, which form feedback loops that are hypothesised to drive the system's behaviour.

Table 5-1 complements Figure 5-1 by summarising each loop in text.

A note on representation: the CLD in Figure 5-1 represents the three labour market states (*Employed, Unemployed, Economically Inactive*) through the single variable *Employed Proportion*. The full stock-and-flow structure with separate stocks and transitions is retained in the simulation model in Chapter 6. Where the loop narrative below refers to exits to inactivity or transitions between states, these correspond to flows in the simulation model that are not separately depicted on the qualitative diagram.

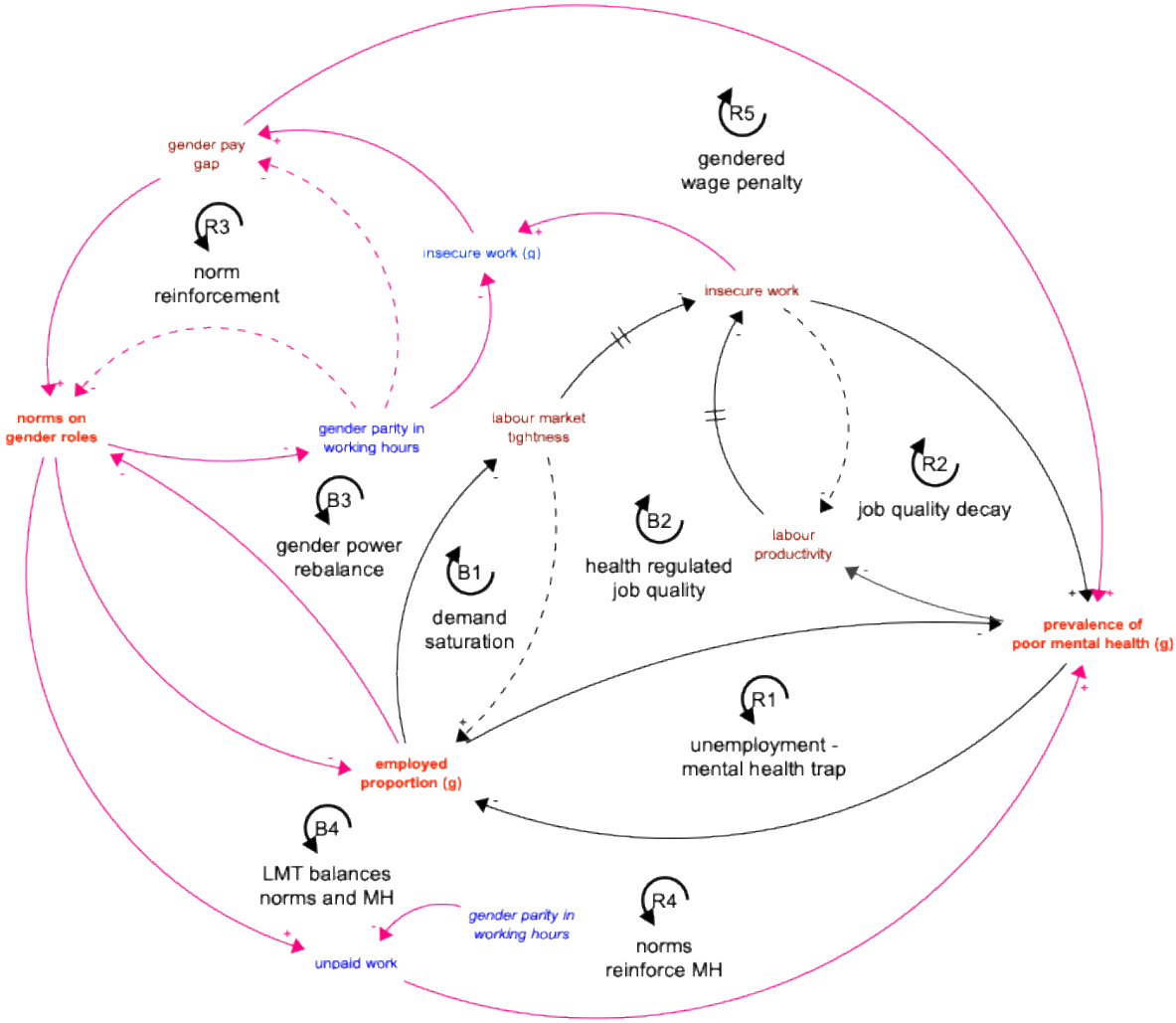


Figure 5-1: Causal loop diagram (CLD) of the hypothesised labour market–gender norms–mental health system in Scotland (1992–2019)

Arrows with a '+' denote a positive relationship (the two variables move in the same direction), while arrows with a '-' denote a negative relationship (the variables move in opposite directions).

Reinforcing loops are denoted by "R" and balancing loop by "B" followed by a number; the direction of the loop around the letter shows whether the feedback mechanism is drawn clockwise or counterclockwise.

The causal links relevant for women only are in pink, whereas those relevant for both men and women are in black. Dotted links are used to denote the existence of an alternative indirect link between two variables.

Variables in red denote stocks. Variables in maroon indicate that the variables have a direct exogenous input. For example, labour market tightness uses the exogenous time series for labour demand to dynamically estimate vacancies.

Table 5-1: Feedback loops in the CLD

Loop	Type	Path
R1 unemployment – mental health trap	Reinforcing	Employed Proportion → (−) → Prevalence of Poor Mental Health (MH) → (−) → Employed Proportion
R2 job quality decay	Reinforcing	Insecure Work → (+) → Prevalence of Poor MH → (−) → Labour Productivity → (−) → Insecure Work Insecure Work → (−) → Labour Productivity → (−) → Insecure Work
R3a norm reinforcement (via hours)	Reinforcing	Norms on Gender Roles → (−) → Gender Parity in Working Hours → (−) → Norms on Gender Roles
R3b norm reinforcement (via gender pay gap)	Reinforcing	Norms on Gender Roles → (−) → Gender Parity in Working Hours → (−) → Gender Pay Gap → (+) → Norms on Gender Roles Norms on Gender Roles → (−) → Gender Parity in Working Hours → (−) → Insecure Work (Women) → (+) → Gender Pay Gap → (+) → Norms on Gender Roles
R3c norm reinforcement (via employment)	Reinforcing	Norms on Gender Roles → (−) → Employed Proportion (Women) → (−) → Norms on Gender Roles *(via Gender Inactivity Gap; see note)*
R4 norms reinforce MH	Reinforcing	Norms on Gender Roles → (+) → Unpaid Work → (+) → Prevalence of Poor MH (Women) → (−) → Employed (Women) → (−) → Norms on Gender Roles
R5 gendered wage penalty	Reinforcing	Gender Pay Gap → (+) → Prevalence of Poor MH (Women) → (−) → Labour Productivity → (−) → Insecure Work → (+) → → Insecure Work among women → (+) → Gender Pay Gap
B1 demand saturation	Balancing	Labour Market Tightness → (+) → Employed Proportion → (−) → Labour Market Tightness
B2 health-regulated job quality	Balancing	Insecure Work → (+) → Prevalence of Poor MH → (−) → Employed Proportion → (−) → Labour Market Tightness → (−) → Insecure Work

Loop	Type	Path
B3 gender power rebalance	Balancing	Norms on Gender Roles → (−) → Employed Proportion (Women) → (−) → Labour Market Tightness → (−) → Insecure Work → (+) → Gender Pay Gap → (+) → Norms on Gender Roles
B4 LMT balances norms and MH	Balancing	Norms on Gender Roles → (+) → Unpaid Work → (+) → Prevalence of Poor MH (Women) → (−) → Employed (Women) → (−) → Labour Market Tightness → (−) → Insecure Work → (+) → Gender Pay Gap → (+) → Norms on Gender Roles

Note on R3c: **The link from Employed Proportion (Women) to Norms on Gender Roles operates through the Gender Inactivity Gap, which is not depicted on the CLD as an intermediate variable. A decrease in women's employment increases the inactivity gap, which reinforces the breadwinner norm.**

The following narrative describes the system behaviour generated by the interaction of these causal relationships, while the formal operationalisation of these variables and the supporting evidence for each individual causal link are documented subsequently in Section 5.3. The CLD comprises five reinforcing loops (R1-R5); R3 itself resolving into three sub-loops (R3a–R3c), and four balancing loops (B1-B4). The reinforcing structure is presented first, then the balancing mechanisms.

Loops (R1 and R2) explain how population mental health and the labour market are locked in a self-reinforcing cycle.

Unemployment-mental health trap (R1) represents the bidirectional health selection/causation relationship in a feedback loop that demonstrates how a change in employment or poor mental health in the population triggers a self-reinforcing cycle.

A decrease in the *Employed Proportion* increases the *Prevalence of Poor Mental Health* through loss of latent and manifest functions, as well as through increased stress about rising unemployment. Crucially, this creates a macro-level selection effect: for those currently unemployed, poor mental health increases the duration of joblessness by reducing the job-finding rate; for those currently employed, it increases the probability of exiting the workforce due to

a reduced ability to work. This can transform a temporary economic shock into a population mental health crisis.

Similarly, a rise in poor mental health due to other reasons also increases the probability of job loss. This increase in unemployment has a negative impact on the population as a whole. This is a vicious cycle where a poor economy and poor population mental health reinforce each other.

Job quality decay (R2): This reinforcing loop extends the logic of R1 from the quantity of jobs to their quality. *Insecure Work* harms *Mental Health*, which reduces *Labour Productivity*. This in turn reduces firms' ability to offer secure, high-quality roles, entrenching a low-skill, low-wellbeing equilibrium. This loop shows how poor job quality can become entrenched. In addition to the indirect link from insecure work to labour productivity via *Prevalence of Poor Mental Health*, *Insecure Work* also reduces *Productivity* due to lack of skills. Higher productivity generates an economic surplus; when productivity declines, firms face tighter margins. In the absence of robust collective bargaining structures to protect employment standards, firms respond to these narrow margins by pivoting toward cost variabilisation: externalising economic risk onto workers via insecure contracts. This further increases insecure work, completing a vicious cycle where diminished economic value and diminished worker power reinforce one another. The simulation model in Chapter 6 excludes labour productivity for parsimony (Section 6.2.3); the structural function of R2 is captured indirectly. The CLD therefore presents the broader theoretical hypothesis, of which the simulation tests a more parsimonious subset.

R1 and R2 operate equivalently for men and women. The next set of loops introduces the gender-specific structure.

Gender norms entrenchment (R3): Three reinforcing loops connect *Norms on Gender Roles* to women's labour market outcomes; together they operationalise the stickiness of gender norms. All the loops are based on the link from *Prevalence of Norms on Gender Roles (Gender Norms)* to *Gender Parity in Working Hours*. Gender norms influence the decisions couples make, particularly after childbirth, leading to lower gender parity as women reduce their paid working hours. This loop captures the idea that part-time work is often a

management strategy to balance gender expectations, rather than necessarily a choice (Tomlinson, 2007).

- R3a (Hours–Norms loop): *Gender Norms* reduce *Gender Parity in Working Hours*; the resulting unequal hours pattern, by virtue of its visibility and prevalence, reinforces the norms.

- R3b (Pay Gap–Norms loop): Lower hours reduce women's human capital accumulation and segregates them into lower-paying roles, widening the *Gender Pay Gap*. The persistent pay gap then makes it economically rational within households for the lower earner to absorb unpaid care, reinforcing the norm.

- R3c (Employment–Norms loop): *Gender Norms* also depress women's *Employed Proportion* directly; lower employment rates further validate the breadwinner model.

Together, these represent the economic entrenchment of gender norms and explicitly mirror the “Success to the Successful” systems archetype (Senge, 1990). As men historically possess higher labour market attachment, households rationally prioritise male employment to maximise income. This allocates men the resource of paid work hours, validating their role as primary earners and justifying further investment in their career capital. Conversely, women are allocated the resource of unpaid domestic labour. In other words, the relative success (gender pay gap) leads to an unequal allocation of resources (working hours) which then determines future success. This structural constraint reinforces the societal norm that women are secondary earners. These three R3 loops operationalise the "interlocking cogs" described by Heise et al. (2019) reviewed in Chapter 2. This represents the model's core theory of structural persistence: institutionalised gender norms directly constrain women's labour market choices.

While R3 entrenches the norm itself, R4 and R5 translate the entrenchment into mental health and economic penalties for women.

Norms reinforce poor mental health (R4): An increase in *Gender Norms* increases the proportion of *Unpaid Work* done by women. The resulting role strain and time poverty negatively affect women's mental health, raising the *Prevalence of Poor Mental Health*. Through the mechanisms outlined in R1, this reduces the *Employment* among women, which through the mechanisms outlined in R3, reinforces *Gender Norms*.

Gendered wage penalty (R5): *Gender Pay Gap* increases the *Prevalence of Poor Mental Health* among women. This reduces *Labour Productivity* and increases *Insecure Work*. Because women are exposed to more insecure work, this widens the *Gender Pay Gap*. R5 differs from R4 in that it operates through the labour market rather than through unpaid work. It captures a baseline eroding structural penalty for women that operates regardless of their personal employment status.

The reinforcing loops above describe the mechanisms of persistence. The balancing loops below describe the structural limits that prevent unbounded behaviour.

Demand saturation (B1): This balancing loop describes the fundamental limit on labour market expansion. As employment rises and approaches total demand, the unfilled vacancy gap closes (saturation). This reduces *Labour Market Tightness (LMT)* and slows the rate of new hiring. This loop acts as a "Limits to Growth" archetype (Senge, 1990), ensuring that employment levels cannot increase indefinitely and will eventually stabilise as they reach the saturation point of the economy.

Health-regulated job quality (B2): An increase in *Insecure Work* increases *Prevalence of Poor Mental Health*. This increases exits to economic inactivity and reduces the *Employed Proportion*. This increases the vacancies and could increase *Labour Market Tightness*. An increase in *LMT* reduces *Insecure Work*. Thus the exit of people from the labour market due to poor mental health can regulate the quality of jobs.

The quantity-quality trade-off: The interaction of loop B2 with **(R1)** reveals a structural tension between employment availability and employment quality. As

shown in the CLD, the direct link from *Employed Proportion* to *Prevalence of Poor Mental Health* is negative, representing the protective effects of employment (via income and latent functions of work). However, this indirect pathway mediated by the labour market creates a counteracting pressure and creates a positive link to *Prevalence of Poor Mental Health*. This creates a plausible situation where efforts to simply maximise employment (quantity) without regulation may inadvertently sustain high levels of distress by shifting the system dominance towards insecure employment (quality), thereby dampening the expected public health benefits of job creation. Whether this counteracting pressure is sufficient to attenuate the public health gains of employment growth depends on which loop dominates, which a static CLD cannot resolve. This generates a specific testable hypothesis: in scenarios that expand employment without regulating job quality, the negative health effects mediated by rising insecure work may partially or fully offset R1's positive effects. The simulation in Chapter 6 represents this trade-off through the labour market tightness pathway (which substitutes for the productivity channel that R2 relies on in the CLD).

Gender power rebalance (B3): Traditional norms (increase in *Gender Norms*) reduce the proportion of women *Employed*. This decreases the overall employed and increases *Labour Market Tightness*. This reduces *Insecure Work*, and thus also reduces *Insecure Work among Women*. As *Insecure Work* reduces, the *Gender Pay Gap* reduces. This then reduces *Gender Norms*. This loop shows that gender norms cannot push women out of the economy entirely without triggering an economic correction that makes those norms unsustainable. At the same time, the labour market itself acts as a brake on social progress. When women enter the workforce and saturate the market, this reduces *Labour Market Tightness*, which in turn increases *Insecure Work*; and this impacts women more than men owing to their part-time participation. Plausibly, the labour market responds to the quantity of labour available by degrading the quality of the jobs.

LMT balances norms and mental health (B4): Traditional *Gender Norms* increase women's *Unpaid Work Burden*, raising their *Prevalence of Poor Mental Health*. This reduces women's *Employment*, which feeds into *Labour Market Tightness* and ultimately into *Insecure Work*, the *Gender Pay Gap* and back into *Gender Norms,* a long balancing path that constrains the behaviour of R4.

Together, these interconnected loops provide a comprehensive explanation for why gender differences in mental health persist and why population-wide distress has risen in the recent Scottish context. For the entire working-age population, the system is anchored by two vicious cycles that transform economic fluctuations into health crises. While men and women are both subject to these macro-economic traps, women face a second, reinforcing layer of disadvantage because societal expectations drive women into part-time roles and justify the continued unequal distribution of paid and unpaid labour. Additionally, even when economic indicators improve, women's baseline mental health is constantly eroded by the burden of unpaid work. Therefore, the persistence of Scotland's gender gap in the prevalence of mental health is hypothesised to arise from this structure: men are primarily exposed to the volatility of the labour market, while women are doubly trapped in a secondary, self-perpetuating cycle that structurally diminishes both their financial rewards and their psychological well-being. Whether this hypothesised structure is sufficient to reproduce the observed trajectory and how much of the gender gap in prevalence of poor mental health can be attributed to each loop is the empirical question addressed in Chapters 6 and 7 respectively.

5.3 Operationalisation of key concepts

This section discusses the operationalisation of the variables in the CLD, translating broad theoretical concepts into model variables and assembling the evidentiary basis for each causal link. Where the chosen operationalisation introduces analytical limitations, these are flagged in Section 5.7. Following the synthesis of the literature and the refinement process, the final conceptual model consists of specific variables, chosen to represent the theoretical constructs of interest, that are documented in Table 5-2. Given the precision required to build a system dynamics model, this table documents the translation of broad theoretical concepts into measurable variables, highlighting the necessary simplifications. The supporting evidence for each individual causal link detailing the polarity, underlying mechanism, and key citations is provided in Table 5-3.

Table 5-2: Translation of concepts to model variables

Theoretical Concepts	Operationalisation and Refinements	Model Variable
Mental ill-health / psychological distress	The variable is conceptualised as a binary state of high psychological distress (GHQ-12 ≥ 4) rather than clinical diagnoses. This aligns with the evidence base and enables exploring how shifts in the systemic risk factors, such as unemployment, affect the entire population. It is worth noting that studies assessing the impact of mental health on productivity often focused specifically on depression although there was at least one study from the UK that used GHQ-12 (Bryan, Bryce, & Roberts, 2022).	Prevalence of Poor Mental Health (Stock)
Gender norms/ attitudes and beliefs regarding women's employment or division of labour	Given the salience of the labour market in this research, this model conceptualises gender norms in terms of the division of labour. This variable was initially named "prevalence of the modified male breadwinner model". However, the operationalisation needed to also include the traditional male breadwinner model. Thus, the concept was later renamed to "norms on gender roles" to capture the broader division of labour.	Prevalence of Norms on Gender Roles (Stock)
Precarious Employment / Low Pay / Perceived insecurity	This is a broad concept that includes contract insecurity, involuntary part-time hours, variable or inadequate pay, limited or no control over hours. It was initially operationalised as perceived job insecurity and inadequate pay. Following expert consultation, it was renamed as insecure work to focus on the secure dimension of Scotland's Fair Work agenda, which captures these dimensions of precarity as well as predictable hours and transparency in pay progression, sick pay and pensions.	Insecure Work Conceptualisation changed from prevalence (during expert consultation) to a weighted index
The "Motherhood penalty"/ "Child penalty"/Part-time work	The disparity in working hours between men and women, especially after childbirth. Using a relative difference rather than absolute levels captures the dynamics of the causal theory and focuses on the magnitude of the difference between men and women.	Gender Parity in Working Hours
Unpaid work / double burden / domestic responsibilities/ second shift	Although unpaid work covers different aspects such as caregiving, the focus is on housework because it has unanimous negative impact on mental health (unlike childcare) and is disproportionately done by women. It is conceptualised as the difference in weekly hours spent on housework between men and women. The method for doing this evolved from absolute difference to percentage difference, to ratio to proportion. Proportion is preferred as there is a theoretical range and this can be tied to the equality and inequality.	Unpaid work burden (Proportion of housework done by women)

Pay gap/ earning differential / undervaluation	While widely reported as an outcome in the literature, the gender pay gap is also conceptualised as a potential driver of gender norms. This relationship was explicitly modelled with uncertainty (as tentative) because, while conceptually sound, it lacked the same level of robust causal evidence present for other connections within the framework.	Gender Pay Gap
Unemployment / exit from labour market / job gain/ job loss	The CLD only represents the stock of employment, but the simulation captures the underlying flows. While the inactive group is heterogeneous, it is aggregated into a single stock to focus on the macroeconomic dynamic of labour supply withdrawal rather than the specific administrative reason for inactivity.	Unemployed / Inactive / Employed (Stock and flow structure)
Macroeconomic Context	To create a dynamic system with feedback loops, the core economic principles of labour supply and labour demand were introduced. Following expert consultation, labour market tightness was added as an intermediate to capture the effect of labour demand on labour market dynamics. Because labour market tightness is based on both employed (numerator is demand – employed) and unemployed, the dynamics of both supply and demand are captured.	Labour Market Tightness (ratio of vacancies to unemployed);
Presenteeism/ Absenteeism/ Productivity / GDP loss	Reduced labour productivity was identified as a key consequence of poor mental health. In addition, precarious work was linked to reduced productivity due to mechanisms such as loss of skill or investment in training.	Labour Productivity

Table 5-3: Evidence for causal relationships in CLD

Causal Link (Cause → Effect)	Polarity	Underlying Mechanism(s)	Key Supporting Evidence
Employed Proportion → Prevalence of Poor Mental Health	-	Captures the benefits of employment and negative effects of unemployment. Individual-level: Employment reduces exposure to the latent and manifest deprivation of unemployment (e.g. loss of status, routine, and social contact). Aggregate-level: High unemployment or rise in unemployment leads to increased population-wide economic anxiety and insecurity, impacting even the employed.	Meta-analyses of longitudinal studies and longitudinal studies in the UK confirm that unemployment leads to poor mental health (Gedikli et al., 2022; Kromydas et al., 2021; Thomson et al., 2023). Aggregate studies show rising unemployment rates are linked to increased mental distress at regional and national levels (Barbalat & Franck, 2020; Barr, Kinderman, & Whitehead, 2015). The impact is further supported by multi-level analyses showing local unemployment trends affect individual mental health regardless of personal employment status (Curtis et al., 2019). Meta-analyses of longitudinal studies confirm that job gain reduces psychological distress, supporting the causal role of employment in improving mental health (Paul & Moser, 2009). Employed individuals report higher levels of both manifest and latent functions compared to the unemployed, and that all of these functions are independently associated with mental health (Paul et al., 2023). Employment improves the prognosis for those with depression (Buckman et al., 2022), highlighting the role of employment in recovery.
Insecure Work → Prevalence of Poor Mental Health	+	Individual-level: Chronic stress from financial instability, uncertainty (unpredictable hours), and lack of control. Deprivation of the same latent functions as unemployment, but often to a lesser degree.	Meta-analyses and systematic reviews of longitudinal studies find a significant causal impact of various dimensions of precarious employment on poor mental health, with effects sometimes comparable to that of unemployment (Kim & von dem Knesebeck, 2016; Pulford et al., 2022; Rönnblad et al., 2019).

Prevalence of Poor Mental Health → Labour Productivity	-	Individual-level: Increased absenteeism (sick leave) and, presenteeism (reduced performance and output while at work). Spillover effects: Negative impact on team productivity due to increased workload for colleagues and loss of experienced staff.	Studies consistently find that depression and poor mental health impact productivity via absenteeism and presenteeism (Evans-Lacko & Knapp, 2016; Lerner & Henke, 2008; Stewart et al., 2003). At a macroeconomic scale, poor population health is estimated to reduce GDP by creating a less productive workforce (Suhrcke et al., 2006). Recent evidence from the UK finds similar results (Bryan, Bryce, & Roberts, 2021; Bryan, Bryce, & Roberts, 2022).
Prevalence of Poor Mental Health → Employed Proportion	-	Individual-level (health selection): Poor mental health causes exit from the labour force into categories such as long-term sick, disabled, or early retirement; poor mental health increases the duration of unemployment spells.	Meta-analysis of longitudinal studies and longitudinal studies in the UK find that poor mental health is associated with an increased probability of transitioning to disability pensions and early retirement (Amiri & Behnezhad, 2019; Amiri & Behnezhad, 2021; Campbell et al., 2022; Linder et al., 2020; van Rijn et al., 2014). A longitudinal study from Australia finds that poor mental health significantly increases the duration of unemployment for individuals (Butterworth et al., 2012).
Employed Proportion → Labour Market Tightness	-	Macro-level: As employment increases, it fills the available labour demand and vacancies decrease, reducing labour market tightness.	Mathematical: Vacancies = Demand - Employment. As the numerator reduces, labour market tightness (vacancies/unemployed) decreases.
Labour Market Tightness → Employed Proportion	+	Macro-level: When there are more vacancies than unemployed (high tightness), there is a higher probability of finding a job, increasing the job gain rate and thus employment. When demand or vacancies are high, there is a decrease in firing, increasing employment.	Analysis of UK data confirms that the job-finding rate is highly pro-cyclical and driven by tightness (Pizzinelli & Speigner, 2017; Schweitzer & Tinsley, 2004). In the context of increasing insecure work, the impact of labour market tightness is greater on insecure work compared to unemployment (Saunders, 2017).

Labour Market Tightness → Insecure Work	-	Scarcity leverage: high vacancies reduce the threat of dismissal, provide more options to workers, allowing them to reject insecure work/demand better terms.	Bank of England analysis suggests that insecure work is likely to be impacted more than unemployment by labour market slack (Saunders, 2017).
Labour Productivity → Insecure Work	-	When productivity is low, firms face tighter margins and pass the risk to employees.	Although there is some decoupling between productivity and wages, it is central to pay prospects (Whittaker, 2019). Low productivity growth has likely to have capped pay (Saunders, 2017).
Insecure Work → Labour Productivity	-	Lack of skill: Insecure work prevents the accumulation of skill due to lack of investment in training and high turnover.	Temporary employment reduces labour productivity growth, with the effect being greater in skilled sectors compared to unskilled (Lisi & Malo, 2017)
Norms on Gender Roles → Gender Parity in Working Hours	-	Socialisation & constrained choice: Societal expectations of women as primary caregivers and secondary earners influence household decisions, leading women to reduce paid work hours, particularly after childbirth	The birth of a first child leads to a large, persistent reduction in women's employment (Kleven, Landais, & Leite-Mariante, 2023). This effect persists beyond childcare years and childcare availability tends to increase part-time rather than full-time employment, indicating a normative rather than purely practical driver (Doan et al., 2022; Zimmert, 2023). A pan-European study finds that non-traditional gender norms among the respondents' grandmothers' cohort is associated with an increase in the probability of mothers' employment, indicating the potential impact of normative change (Moriconi & Rodríguez-Planas, 2021). Kleven, Landais and Leite-Mariante (2023) show a massive and persistent "child penalty" in female labour supply that is not explained by practical barriers alone, but by ingrained norms. UK longitudinal data consistently shows a sharp divergence in employment rates following the birth of the first child (Andrew et al., 2021).
Norms on Gender Roles → Employed Proportion (Women)	-		

Norms on Gender Roles → Unpaid Work	+	Gender socialisation: Societal expectations directly assign the majority of unpaid domestic and care work to women, irrespective of their employment status or relative income.	Evidence shows that unpaid work is highly gendered, with women performing significantly more than men (Kan, Sullivan, & Gershuny, 2011). This disparity is not primarily explained by time availability or relative resources, pointing to gender socialisation as the main driver (Sahl, 2015; Wang & Cheng, 2024).
Gender Parity in Working Hours → Norms on Gender Roles	-	Descriptive norm reinforcement: When women work fewer hours (low parity), it reinforces the societal perception that women are secondary earners.	Grounded in social learning theory: individuals learn behaviours, norms, and attitudes by observing others. This impact of role model has been evidence in a study in the US (Galassi, Koll, & Mayr, 2024). Himmelweit and Sigala (2004) find that identities and behaviours mutually influence each other, but the pathway from behaviour to identity is particularly strong. When mothers are forced by external constraints (e.g. financial necessity) to adopt a behaviour that conflicts with their current gender identity (e.g. a "traditional" mother working full-time), they experience cognitive dissonance. To resolve this, they are more likely to adjust their identity to align with their behaviour than to change their behaviour to match their identity.
Gender Inactivity Gap → Norms on Gender Roles		Social signalling: The visibility of a large population of inactive women validates the traditional breadwinner model as the status quo. An increase in the gender inactivity gap is likely to increase the prevalence of norms on gender roles, but a decrease in the gender inactivity gap may have little to no impact if the hours don’t increase/change.	
Unpaid Work → Prevalence of Poor Mental Health (Women)	+	Role strain & time poverty: The "double burden" of paid and unpaid work creates chronic stress, a significant mental load, and a scarcity of time for rest and health-promoting activities.	A systematic review and a study from Australia conclude that unpaid work contributes to poor mental health among women (Ervin et al., 2022; Ervin et al., 2023a).

Gender Parity in Working Hours → Unpaid Work	-	Time availability: As women work more paid hours, they have less time available for unpaid domestic work; time availability acts as a secondary mechanism in addition to norms.	Evidence finds support for the time availability mechanism, where increased paid work hours for women are associated with some reduction in their unpaid work hours (Sahl, 2015; Wheatley, Lawton, & Hardill, 2018).
Gender Parity in Working Hours → Gender Pay Gap	-	Human capital accumulation: Lower gender parity (i.e. more part-time work) directly inhibits wage progression. Experience gained in part-time roles offers minimal or even negative returns to experience leading to a wage stagnation. Fewer opportunities: Part-time roles may be available in less productive firms with lower pay. Vertical segregation: Part-time work is a major barrier to promotion.	Part-time work is associated with a significant pay penalty (Jones, 2019). Studies have found evidence for lower pay in part-time jobs, discrimination against mothers and occupational segregation (Blau & Kahn, 2017; Budig & England, 2001; Correll, Benard, & Paik, 2007; Jones, Cook, & Connolly, 2023). Differences in full-time work histories explain a large proportion of the UK gender pay gap (Blundell et al., 2016; Olsen et al., 2018). Fewer part-time workers report prospects for advancement compared to full-time workers (Murphy, 2022). Women, particularly after having children, are significantly more likely to work for less productive firms than men (Joyce & Xu, 2019).
Gender Parity in Working Hours → Insecure Work among Women	-	Occupational segregation/ downgrading: Women who switch from full-time to part-time work are often forced to downgrade occupationally, accepting jobs below their skill level to secure the necessary flexibility. Sorting mechanism: The need for part-time work channels women into a segment of the labour market structurally characterised by higher contractual insecurity and low-pay.	A UK study found evidence of occupational downgrading when switching to part-time work (Connolly & Gregory, 2008). A UK study found that mothers are significantly more likely than fathers or women without children to be in poor quality jobs and that this difference was largely explained by their much higher rates of part-time work (Jones, Cook, & Connolly, 2023).

Insecure Work among Women → Gender Pay Gap	+	Structural disadvantage and low pay: Insecure work is associated with low pay. Occupational downgrading is a direct cause of lower hourly pay.	Because the operationalisation of *Insecure Work* includes low pay, a higher prevalence of insecure work among women compared to men directly increases the *Gender Pay Gap*.
Gender Pay Gap → Prevalence of Poor Mental Health (Women)	+	Psychosocial stress from structural inequality: Gender pay gap creates a sense of inequality and perpetuates gender differences in opportunities and resources.	A US study found that after controlling for individual productivity factors, the direction of the wage gap was a significant moderator of mental health. Where female income was less than a matched male counterpart, women's odds of major depression were 2.43 times higher than men's (Platt et al., 2016). *Note: this is a cross-sectional study conducted in US.*
Gender Pay Gap → Norms on Gender Roles	+	Economic reinforcement of norms: A persistent and visible pay gap makes it economically rational within a household for the lower-earning partner (disproportionately the woman) to specialise in unpaid domestic work, thereby reinforcing the gender norm for the next generation.	This link represents a core theoretical hypothesis of the model, explaining the stickiness of gender norms. While direct causal evidence is scarce, it was deemed plausible by experts as a mechanism for structural persistence. An experimental study found that women facing explicit negative gender discrimination supplied less labour than women facing gender-neutral wage inequality (Gagnon, Bosmans, & Riedl, 2025).

5.4 Model boundary

Boundary selection determines what the model can and cannot answer. This section explains the model's spatial scope (Scotland), temporal scope (1992–2040), and what is treated as exogenous or excluded and why. The model boundary was informed by both the research question and the literature review. The research question determined that the boundary must be relevant at the national level and able to investigate trends over the past few decades. Variable selection was crucial and done throughout the literature review.

5.4.1 Spatial and temporal boundary

The model's spatial boundary is Scotland, a choice made to provide a coherent, national-level case study. While the UK economy is highly interconnected, and many of the dynamics explored are UK-wide, significant regional heterogeneity exists. Attempting to model the entire UK would introduce a level of complexity that could obscure the fundamental feedback mechanisms under investigation. Scotland provides a compelling case due to its distinct policy environment emphasising mental wellbeing, fair work, and gender equality, even within the constraints of the broader UK macroeconomic framework. The influence of the wider UK economy is captured exogenously through baseline labour demand. More specifically, the model represents Scotland as a closed system: migration flows, UK-wide policy changes, and cross-border labour market dynamics are excluded. To the extent that these factors influenced mental health trends or labour market conditions during 1992–2019, the model's attribution of effects to internal feedback mechanisms may be overstated. This is a necessary simplification that allows the model to focus on the endogenous feedback structures at the core of this research; the implications are returned to in the limitations of Section 5.7.

The model's time horizon spans from 1992 to 2040, partitioned into a 1992–2019 historical calibration period (used to test whether the hypothesised structure is sufficient to reproduce observed trends (see Chapter 6)) and a 2019–2040 projection horizon (used in Chapter 7 to explore policy-relevant scenarios). The historical start date of 1992 is determined by the availability of consistent, high-quality data on the prevalence of poor mental health from the British Household

Panel Survey (BHPS) and its successor, the UK Household Longitudinal Study (UKHLS), as well as consistent labour market data for Scotland. This timeframe is sufficiently long to capture the dynamics of multiple economic cycles, including the recessions of the early 1990s and 2008, and the subsequent period of austerity. It also allows for the observation of gradual, long-term shifts in key variables, most notably the slow-moving nature of social norms related to gender and work. The projection to 2040 enables the exploration of the long-term implications of the hypothesised system structure and the potential impact of policy scenarios.

5.4.2 Exogenous and excluded variables

As discussed in the exclusion criteria (Section 4.2), factors operating at smaller scales (firm-level) or larger scales (global economy) were excluded to focus on the national labour market dynamics. However, the exclusion of specific variables requires detailed justification.

The welfare system: A significant and valid critique raised during expert consultation (discussed further in Section 5.5) was the exclusion of the welfare system (specifically conditionality) from the model boundary. Experts rightly argued that one cannot fully model gendered health and employment dynamics in the UK without accounting for the structural role of these factors. Recognising the complexity of the UK's institutional framework, the welfare system is considered a fixed feature of the context rather than an endogenous variable. This methodological choice ensures analytical parsimony; however, it implies that the findings are specific to the current welfare regime and must be interpreted within these defined institutional constraints.

Education and skills: Experts also suggested that education is a critical driver of labour market outcomes, with one expert noting that overqualification might be relevant for women returning to work. While education is undoubtedly a driver of earnings and productivity, it was excluded because there are no significant gender differences in educational attainment in Scotland. By excluding it, the overall education level of the workforce is assumed to be a stable background condition that does not drive the specific gendered feedback loops this model investigates.

The business cycle: While the model captures the impact of economic conditions, it does not endogenously generate the business cycle (recessions and booms). Demand for labour is treated as an exogenous input. The model can show how the system reacts to a recession, but it cannot explain why the recession occurred in the first place.

Labour productivity: Productivity is included as an endogenous variable in the conceptual model presented in this chapter. It was subsequently excluded from the simulation model in Chapter 6 (Section 6.2.3) on parsimony grounds, after early iterations found its impact difficult to elucidate when labour demand was treated as exogenous. The conceptual structure is therefore a wider hypothesis than the simulated structure tests; the implications are addressed as a limitation in Sections 5.7 and 6.8.

Poverty: During expert consultation, a valid critique was raised regarding the exclusion of poverty and lone parenthood from the model boundary. Lone parenthood, especially among women, would be an important variable that impacts both their mental health and labour market participation. However, these were not included to prevent adding exogenous drivers that do not refine the core endogenous feedback mechanisms.

5.5 Synthesis of feedback from expert consultation

The methodology of the expert consultation, including expert sample composition, interview protocol, and the disconfirmatory framing, is detailed in Section 4.3. This section synthesises the substantive feedback received and documents how it informed the final CLD. Feedback led to refinements in three areas: conceptualisation, model boundary, and causal structure. The preliminary CLD presented to experts and the detailed feedback summaries are in Appendices B and C.

5.5.1 Conceptualisation

The representation of the gender dynamics was particularly well-received with experts affirming the choice of variables gender norms and gender parity in working hours as well as the operationalisation of the unpaid work burden.

Following expert feedback that modified male breadwinner model (MMBM) was a misnomer, the variable was initially renamed to the modified breadwinner model and ultimately finalised as *Prevalence of Norms on Gender Roles*. The model identifies childbirth as a trigger, but experts stressed the importance of "sandwich care" and the lifelong nature of women's caring responsibilities for parents and spouses, which is a key reason for their continued part-time status. This aligns with the model: the influence of norms is not confined to a fixed period but is set by the commencement of childbearing. The explicit exogenous driver of proportion of women with (dependent) children was later removed, as it was considered a relatively stable feature of the context rather than a driver of the feedback loop. Additionally, experts provided a critical counterpoint to the idea that part-time work is universally negative, highlighting it can be a protective factor for women's mental health by helping them manage the double burden, even if it carries a long-term financial penalty. This is reflected in the emerging insight from the CLD where part-time work is a strategy to manage rather than a choice. There is no direct impact of part-time work on mental health.

Expert feedback helped refine the naming and operationalisation of precarious employment. Explicit feedback was sought on the proposed operationalisation of precarious employment including perceived insecurity (rather than temporary contracts) and low-pay, possibly job autonomy although this may not be possible with certain jobs. Experts challenged the use of narrow terminology for a potentially broader concept of job quality or fair work. Some experts pointed out that uncertainty in work hours could be quite different to low-pay which might be more of a chronic stress rather than the acute stress of precarity. The variable was renamed to *Insecure Work*, as perceived insecurity and low-pay are both elements of the “secure” dimension of the Fair Work framework. It was operationalised as a weighted composite which was better suited to capture the multi-dimensional nature which is not visible in the prevalence.

Although experts highlighted that the distinction between precarity and low-pay is pertinent for the individual as the former is a psychosocial stressor and the latter presents material hardship, these were combined given the high correlation between them and the focus on the macro-level, where the causes and consequences are similar. It becomes conceptually difficult to disentangle

their individual causal chain, attributing a specific weight to each becomes a matter of significant assumption rather than evidence. This strategic aggregation does not permit separate analysis of policies targeting income versus policies targeting employment security. However, it allows for a more robust analysis of the primary feedback loops by avoiding the structural noise of disentangling two highly correlated and interconnected drivers that share common systemic origins.

5.5.2 Boundary refinements

The preliminary CLD presented to experts included a tentative variable *Growth in Median Wages* to capture broader labour market dynamics, such as the impact of wages on labour demand and the feedback from unemployment levels to wages. However, during consultation, economists stressed that the effects of this variable are highly contingent on the broader business cycle or labour market slack, making it difficult to model endogenously without significantly expanding the boundary. Experts also agreed that the impact of wages on mental health is best captured at the lower end of the income distribution (through concepts such as in-work poverty) rather than through the median wage. (This was the reason for marking it as a tentative link.) On probing, experts suggested that the drivers of low pay (e.g. replaceability of workers) were similar to the drivers of precarious employment. Consequently, it was determined that growth in median wages was not a helpful addition to the CLD. Instead, the model boundary (defined in Section 5.4) was refined to exclude median wages and explicitly capture low pay as a component of the insecure work variable described in Section 5.3.

Some experts pointed out that some of the relationships between labour market variables would depend on the macroeconomic context and labour market slack. This would have been captured by the non-linear relationships in the simulation model, but following expert consultation, a new variable of *Labour Market Tightness (LMT)* was added to make this explicit. Labour market tightness is the ratio of the number of vacancies to unemployed people. A higher ratio indicates higher worker bargaining power. This is driven by both labour demand and the unemployment rate. LMT replaced the variables of labour demand and supply as

it captures their dynamics better. In the simulation model, labour demand will be an exogenous time-series impacting LMT.

Significant feedback centred on what is missing from the model and the feedback related to this is discussed under model boundary in Section 5.4.2.

5.5.3 Causal structure

There was broad agreement among the experts on the fundamental relationships depicted in the model. The feedback revealed nuance regarding the strength and nature of specific links. A point of contention is the strength of the link from poor mental health to economic inactivity. Some experts, focusing on Scotland, see this as a significant driver of rising inactivity. In contrast, one expert argued that within the broader UK context, the benefit system acts as a barrier, meaning this link is weaker than in other countries and only applies in cases of severe illness.

The researcher's caution regarding the dotted lines was validated by the experts. The proposed tentative link from *Poor Mental Health* to reduced *Working Hours* was met with the most scepticism, with experts suggesting this is a luxury choice not available to many in the UK, especially those on low incomes. This was not included in the revised model. The link from the *Gender Pay Gap* to *Poor Mental Health* was viewed as conceptually logical but difficult to isolate from broader concepts of workplace discrimination and inequality. One expert suggested that this might operate through intermediate variables. Gender pay gap might contribute to structural discrimination or perceived inequality, which could impact poor mental health. Additionally, gender pay gap might suggest low earnings for women, which is causally linked to poor mental health. This link was included in the revised model. The link from the *Gender Pay Gap* to *Gender Norms* was deemed plausible but lacking strong causal evidence. In line with the commitment to structural accuracy, this was included in the model with the view to specifically test it in the simulation model.

The causal relationship from *Productivity* to *Labour Demand* to *Insecure Work* was conceptualised in the context of worker bargaining power. While experts agreed with this, one expert explicitly emphasised the importance of other

factors such as unions. They also mentioned that the potential decoupling discussed in the context of UK may not hold when the full remuneration package is considered as suggested by some economic analysis. The impact from productivity to labour demand was a mechanism largely explained by firm profitability: if workers are more productive, they generate more profit per head, increasing the firm's desire to hire (assuming non-zero marginal revenue). In the context of insecure work, firm profitability leads to better working conditions for employees. Thus, the relationship from *Productivity* to *Labour Demand* was replaced by a link from *Productivity* to *Insecure Work*, given the impacts on number of jobs can be different in the short and long-term.

A causal relationship was drawn from *Insecure Work* to *Labour Productivity*. Although supported by evidence, this was not initially included to develop a coherent and parsimonious model. This exclusion was not challenged by experts. However, the aggregation of precarity and exclusion of wages made it more important to capture the labour market dynamics, and this causal relationship was added.

Overall, experts showed praise for the model's ambition, comprehensive nature, and structural features, suggesting that general structure is considered a coherent and plausible theory for exploring gender differences in mental health and labour market outcomes. The expert consultation served principally to refine rather than redirect the model: the conceptualisation of *Insecure Work* was sharpened, and *Labour Market Tightness* was added as an explicit intermediate variable. The boundary was held against pressure to extend it (welfare, lone parenthood) on parsimony grounds.

5.6 Summary and interpretation

The causal loop diagram developed in this chapter brings together epidemiological literature and economic theory to capture the interconnectedness between population mental health and the labour market. By treating mental health as a dynamic stock, rather than a static outcome, the model reveals how the labour market can either replenish or erode population mental health over decades. This section summarises the theoretical contributions of the conceptual model. Wider interpretation against the

empirical results, and the policy implications, is reserved for the overarching discussion in Chapter 8. Three contributions are highlighted here.

First, the CLD operationalises the "mutually reinforcing" relationship between population mental health and the macroeconomy (Bloom & Canning, 2008; Thomas et al., 2023) into a formal, dynamic causal hypothesis. Unlike models which treat mental health as static, this model treats it as a dynamic stock that is eroded or replenished by the labour market itself. It also builds on frameworks that describe the connections between macroeconomic variables and between macroeconomy and health, by focusing on the causal impact of health on macroeconomic variables (Naik et al., 2019). A report by the Health Foundation (2020) furthered this framework by highlighting that health impacts the structural determinants of health, but it does not describe the specific variables that are impacted or the causal pathways.

Second, this causal loop diagram positions gender as a structural factor. It first highlights that the benefits of employment for women are systematically eroded through unpaid work, and an economic penalty (gender pay gap) due to constrained career choices. In addition, while the aggregate model cannot depict the individual experience of double burden, it makes the constituent pressures explicit. It shows how women as a population group are subjected to the normative pressures of unpaid work, which when they are employed creates role strain and diminishes the net benefits of employment. Because the overall psychological and financial rewards of employment are structurally lower for women, the relative loss upon becoming unemployed is less severe compared to men, for whom employment is more unequivocally aligned with their primary social role. In this way, the CLD highlights a plausible explanation for a paradox in the literature: why does unemployment often have a less severe impact on women's mental health than men's? It also illustrates that entrenched gender norms exert a constant influence on women's mental health, regardless of their personal employment status. This creates a structural disadvantage that impacts the baseline mental health of women as a population group, providing a systemic explanation for the persistent gender differences observed in Scotland.

Third, the research approach aligns with a growing movement to use systems modelling to inform policy responses to the interconnected social, economic,

and health drivers of mental health. For example, the Mental Wealth Initiative is using a national-level system dynamics model to forecast how cross-sectoral policies in education, the labour market, and mental health systems could enhance national prosperity, conceptualised beyond GDP (Occhipinti et al., 2022). The CLD developed in this thesis contributes to this movement by grounding these mechanisms in the specific institutional and policy context of Scotland, positioning the model as a tool for asking policy-relevant questions.

Central to this is the model's conceptualisation of insecure work as the outcome of shifting worker bargaining power, driven by the interaction of market tightness, productivity, and institutional factors. This structural framing offers a coherent way of reading the historical trajectory in Scotland: the period of 1992–2008 can be understood as a buffered era in which the gradual decay of institutional protections (e.g., declining union density) was offset by productivity growth and high market tightness, so that insecure work prevalence remained low. The post-2008 period, characterised by stagnant productivity and continued institutional decline, removed both buffers, allowing insecure work to entrench. Crucially, this bargaining power is structurally attenuated for women whose choices are constrained by the normative requirement for part-time work.

5.7 Strengths and limitations

The conceptual model has several strengths. Its structure was not settled by a single researcher in isolation but stress-tested through disconfirmatory expert consultation. The boundary itself was refined iteratively throughout the research process rather than fixed at the outset. The evidentiary basis of each causal link is documented, so that the loops resting on thinner evidence are identified rather than obscured. And intangible constructs such as gender norms are represented with the best measure rather than omitted.

These contributions should be read against the evidentiary asymmetry across the loops. Because feedback loops are no stronger than their constituent links, an assessment of the conceptual model requires examining each loop's weakest pathway. R1 (Unemployment–MH Trap) rests on links that are well-supported by multiple meta-analyses and UK longitudinal evidence in both directions. R2 (Job Quality Decay) rests on a weaker chain where the effects of insecure work on

mental health and productivity are well-evidenced, but the productivity to insecure work link rests on Saunders (2017) and Whittaker (2019), which are policy-analytic rather than causal. Similarly, R3 is asymmetric where the norms to hours direction is robust but the reverse is theoretically grounded but empirically thinner. R4 and R5 inherit the asymmetric evidence base of R3. Balancing loops B2 and B3 are inferred largely from the macroeconomic literature which makes their structural form plausible but their empirical magnitudes contested. This evidentiary gradient is a feature of the literature on this complex system, not a flaw of synthesis. Its implication is that some of the loops hypothesised in the CLD rest on a smaller evidence base than others, and should be interpreted with corresponding caution.

Several limitations specific to the CLD should be borne in mind when interpreting its contribution. These should be read alongside the broader methodological limitations of system dynamics discussed in Section 3.5 (causality in complex systems, aggregation, reflexivity), and the limitations of the conceptual model development methodology discussed in Section 4.5 (search strategy and expert validation).

Boundary: The model boundary while necessarily limited omits crucial variables. The feedback from experts regarding the exclusion of the welfare system and poverty is particularly salient. While beyond the current model boundary, these factors represent critical next steps for this line of research. The model boundary was deliberately drawn to maintain parsimony and focus on the central research question. While this means certain dynamics are excluded, this choice was made to ensure the model remains tractable and its core mechanisms understandable.

Context-specific: As with most causal mechanisms, those depicted in the CLD operate in a specific context. The relevant context for the study relates to the political, economic and social structures broadly similar to the neoliberal framework. Different approaches to the economy are associated with different impacts on health (McCartney et al., 2019; Naik et al., 2019), and thus the causal relationships may not hold within an alternative economic approach, making the generalisability of the CLD limited to contexts with similar economic, political and social structures. Several estimates underwriting the CLD also draw

on non-Scottish contexts; while the structural plausibility of the mechanisms is likely to travel, their strengths may not.

Aggregation: The model operates at the national level, aggregating diverse local experiences into population averages. Experts confirmed that significant sub-national variation exists, particularly between metropolitan and post-industrial regions, and that this is also true within Scotland. However, they also acknowledged the pragmatic necessity of the national boundary, given the study's theoretical aims and the practical constraints of obtaining robust, disaggregated data for all model variables, especially social norms. Beyond regional heterogeneity, gender norms are not uniform within a national context: they vary substantially by ethnicity, class, cohort, and region, and an intersectional lens cautions against treating gender norms as a single aggregate (Nandi & Platt, 2024). The aggregate representation adopted here is a modelling choice dictated by the population-level research question and the structure of available time-series data, not a theoretical claim that intra-national heterogeneity is unimportant. The national model is therefore a tool for exploring systemic, structural dynamics, with the acknowledgment that these dynamics will manifest with varying intensity at the local level.

Operationalisation: The operationalisation of *Insecure Work* as a composite index combining low pay, contract insecurity, and other dimensions of precarity is justified by their high correlation and shared upstream drivers at the macro level. However, this aggregation constrains the model’s ability to examine differences in effects between policies targeting income (e.g. minimum wage uplifts) versus policies targeting employment security (e.g. zero-hours contract regulation).

System dynamics requires that every variable in the model, including concepts such as gender norms, which lack a standard physical scale, be represented quantitatively within the model's equations. This carries the risk that model parameters convey a false sense of precision. As Sterman (2002) argues, however, omitting a structurally important variable because it resists measurement is a less defensible position than including it with a transparent best estimate, since omission assigns the determinant an implicit effect of zero.

The operationalisation choices made in this thesis and their parameterisation (Chapter 6) are documented to support critical scrutiny of this trade-off.

Static relationships: The model's core feedback structure is assumed to be time-invariant within the simulation period: the polarity, shape, and effect-multiplier values of each causal link are held constant from 1992 onwards. The strength and even the nature of these links can change as institutional, economic, and normative conditions evolve. A model that allowed link strengths to evolve endogenously is beyond the scope of this thesis but is a natural extension of the structural hypothesis advanced here.

Furthermore, a CLD is a static map and unable to reveal the dynamic behaviour. It cannot tell which loops are dominant. This necessitates the transition from a qualitative conceptual map to a quantitative simulation model, which is the focus of the next chapter.

6 Model Development

6.1 Introduction

Chapter 5 set out a causal hypothesis for Scotland's gender mental health gap: a system of feedback loops linking the labour market, gender norms, and population mental health. A hypothesis of this kind needs a simulation model to be tested. If the model reproduces the rise in poor mental health and the persistent gender gap observed in Scotland, the hypothesis is supported: the hypothesised structure is one plausible explanation. This chapter translates the qualitative conceptual model or causal loop diagram (CLD) into a quantitative model, builds confidence in the simulation through various tests and establishes it as capable of endogenously generating the historical patterns of rising mental distress and persistent gender gap observed in Scotland between 1992 and 2019.

The chapter is organised in three steps. Section 6.2 sets out the strategic choices that shaped the model's architecture: what was included, what was deliberately left out, and how the model was kept small enough to be tractable. Sections 6.3–6.5 then describe the three sub-models that together form the simulation: the labour market (6.3), gender dynamics (6.4), and mental health (6.5). For each sub-model, the structure, key variables, parameterisation strategy, and the tests used to build confidence in the structure are described. Section 6.6 brings the three sub-models together and reports the verification, structural validation through loop knock-out, and behaviour reproduction tests run on the fully integrated model. Section 6.7 summarises the validation evidence and Section 6.8 closes the chapter with strengths and limitations.

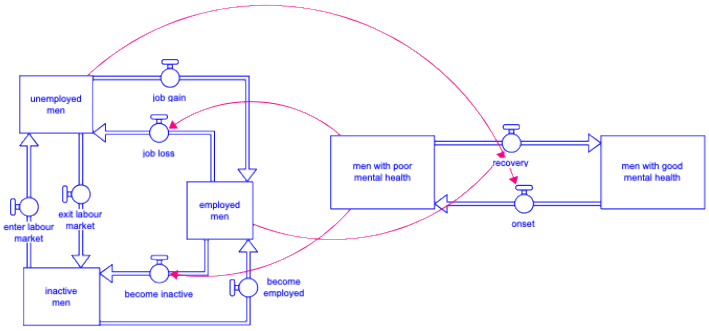
The verification and validation results presented throughout the chapter were generated during an iterative development process. Where the underlying feedback architecture changed, earlier tests were re-run on the revised model, and some module-level parameters were finalised only after the three sub-models had been integrated. The results should therefore be read as a cumulative body of evidence supporting the final model, rather than as a single assessment.

6.2 Strategic modelling decisions and architecture

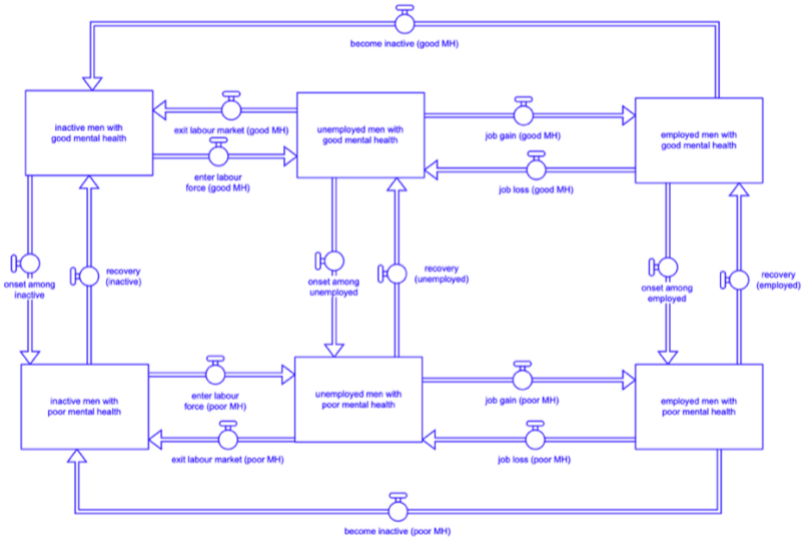
Building a simulation forces choices that the CLD left open. Four choices that shape the model are discussed in this section. First, the labour market and mental health are modelled as two coupled sub-systems (Section 6.2.1). Second, every stock in the model represents a proportion of the working-age population rather than a count of people (Section 6.2.2). Third, the model was deliberately kept parsimonious (Section 6.2.3). Fourth, parameters were anchored in data and literature wherever possible, with formal calibration reserved for variables that cannot be observed directly (Section 6.2.4).

6.2.1 Separate sub-system structure

The labour market and mental health are modelled as separate, interconnected sub-systems rather than a unified, multi-state population structure (Figure 6-1).



Panel A: modelling labour market and mental health as separate sub-systems



Panel B: modelling labour market and mental health as a unified system

Figure 6-1: Model architecture options: separate and unified sub-systems

Panel A (adopted) isolates the population-level feedback loops central to the research question

The separate sub-system architecture (panel A) was chosen because it aligns more closely with the research question, which is concerned with how the overall prevalence of poor mental health in the population affects the overall rates of labour market transitions. By partitioning the model into coupled engines, the model can more easily isolate the primary feedback loops. The sub-systems are linked by causal functions: unemployment modifies the transition from good to poor mental health, while the prevalence of poor mental health impacts flows into and out of employment. This approach provides a clearer representation of population-level feedback, capturing how the structural determinants modify the odds of onset and recovery for the entire cohort.

In a unified structure (panel B) common in epidemiological models, individuals are tracked across combined states (e.g. unemployed with poor mental health). While this offers granular detail, it was deliberately avoided to prevent the complexity that arises from high dimensionality. A unified structure would require 16 flows per gender, compared to 8 in the separate approach. More importantly, the central research questions would be hidden in the differences between flow rates rather than the direct effect multipliers connecting the two sub-models.

In both cases, the third element, gender dynamics would be modelled as a separate stock that exists only at the population level and is represented as a set of norms that exists across genders. The three sub-models — labour market, population mental health, and gender dynamics — are implemented as separate modules in Stella Architect; the first two are the coupled engines, and gender dynamics supplies structural inputs to both, and is impacted by the labour market. The causal links between the labour market, mental health, and gender sub-models operationalise the core dynamic hypothesis of the research.

6.2.2 Prevalence-based stocks

The model adopts a prevalence-based approach to stocks, representing the proportion of the population in a given state (e.g. proportion of working-age men who are unemployed) rather than absolute numbers of people. Normalising the population (where the sum of stocks within a system always equals 1.0) abstracts the model from unnecessary complexity of exogenous demographic

changes, such as births, deaths, and migration, which are not the primary drivers of the central hypothesis. While this boundary choice limits the model's ability to analyse changing age structures, it directly serves the research objective: isolating the endogenous, structural feedback mechanisms that drive gender gaps. Through this, the simulation ensures that any observed divergence in mental health is an emergent property of the modelled socioeconomic forces, rather than an artefact of a growing or shrinking population base.

6.2.3 Parsimony and strategic aggregation

The model boundary was strategically narrowed to avoid over-parameterisation and enhance internal validity. The simulation is a distillation of the CLD, prioritising parsimony. Decisions made for parsimony are summarised in Table 6-1.

Table 6-1: Exclusion and aggregation details for simulation model

Variable / Mechanism	Status	Rationale
Labour Productivity	Excluded	Productivity is often driven by global shocks and factors exogenous to the model. Early iterations found the impact of poor mental health on productivity was difficult to elucidate when labour demand was exogenous. Productivity was therefore removed to maintain parsimony. Consequently, all the loops which pass through productivity are not modelled. The structural function of the productivity-mediated loops in the conceptual model, notably R2 (Job Quality Decay), is captured in the simulation through an alternative route: insecure work affects mental health, which alters labour market states, which feeds back into *Labour Market Tightness (LMT)*[1] and hence into *Insecure Work* (see Section 6.3.1). The simulation therefore realises the dynamic via a different intermediate variable, rather than eliminating the feedback structure entirely.

[1] A note on variable referencing. Throughout this chapter, variable names appear in two forms. In prose, variables are referenced in the same convention used in Chapter 5: capitalised first letter (for stocks), italicised, with spaces between words (e.g., *Employed Proportion*, *Prevalence of Poor Mental Health, Labour Market Tightness, Insecure Work*). The underscored forms (e.g., Employed_Proportion, Prevalence_of_Poor_Mental_Health) appear only in direct quotation of Stella Architect equations, where they are the literal variable names in the simulation software. This convention preserves readability while keeping the model–text correspondence auditable.

Insecure Work to GPG Link	Excluded	While *Insecure Work* is a vital factor, its impact on the *Gender Pay Gap* (GPG) is structurally subsumed within the *Hours Parity* variable in final version. R3b in the CLD has two paths back to the *Gender Pay Gap*: one direct (Hours Parity → GPG) and one indirect (Hours Parity → Insecure Work (Women) → GPG). The simulation retains the direct path (Section 6.4.6) and excludes the indirect one to prevent double-counting and empirical overfitting; R3b is therefore partially captured rather than excluded entirely.
Composite Stressor Indices	Added	Composites consolidate multiple indeterminate factors into single weighted structural drivers. By aggregating causal pathways that the data cannot empirically distinguish or where drivers are highly correlated, the composites reduce the degrees of freedom (i.e. reduces the number of unobservable parameters that require estimation), and enhances the model's robustness. Further, it allows the model to leverage the trend of a collective phenomenon even when time-series data for a specific sub-indicator is fragmented. Finally, it shifts the analytical focus to how the system reacts to aggregated pressure, rather than attributing variance to micro-fluctuations in different stressors. This was done within the mental health and gender modules.
Labour Market Dynamics	Simplified	LMT, in theory, impacts all labour market flows. However, in the simulation it was modelled to impact only three flows: job gain, job loss and movement from inactivity to employment.
Childcare to Inactivity	Excluded	This link was initially modelled but excluded for parsimony, especially in the absence of robust estimates. The impact of childcare availability on gender parity in working hours was retained.

6.2.4 Parameterisation philosophy and methods

To maintain the model's structural integrity, a multi-method parameterisation strategy was employed, prioritising empirical evidence while acknowledging the necessity of theoretical assumptions.

Initialisation: The choice of starting values matters because it determines whether subsequent change is generated by the model's own feedback structure or simply by the system settling away from an arbitrary starting point. The

model uses a partial-equilibrium initialisation strategy in which sub-models are initialised differently:

- Labour market: a mathematical steady state (inflow = outflow) is not forced because Scotland was not in equilibrium in 1992; the system is initially balanced by back-calculating implied baseline rates (by dividing observed 1992 transition data by literature-derived effect multipliers to adjust for the pre-existing structural burden).

- Mental health: initialised in equilibrium to ensure that subsequent dynamics, such as rising mental distress, are driven by internal feedback rather than starting momentum.

- Gender dynamics: norms are not initialised in equilibrium to capture the liberalisation in play; the time for norms to change is the same for reinforcement and erosion.

Simplification: Relationships between variables are often defined by dimensionless table functions. The non-linear shapes of these functions represent complex dynamics such as diminishing returns or critical thresholds, ensuring the model reflects real-world plausibility. Sensitivity analyses were conducted to examine the impact of the table function shapes. Where the following criteria converged: little difference in behaviour across alternative shapes and theoretical plausibility of a linear form, and in some cases, the parameter-reduction benefit of a linear formulation, a linear formulation was selected. This was the case for the labour market and gender dynamics module which constitute the environment or context for poor mental health.

Rejection of trend-line fitting for calibration Early iterations of the model attempted to calibrate against smoothed trend lines for mental health derived from the historical data to avoid annual fluctuations in the data. This approach was rejected because fitting a simulation to a trend line introduces a dual layer of error: the error of the trend line against observed, and the error of the simulation against the trend line. Consequently, all calibration is performed against raw historical data points. This results in a visually noisier fit.

Parameterisation methods: The parameterisation strategy leverages the validated findings of previous research to enhance the model's credibility. The parameterisation strategy for each variable is integrated directly into the module descriptions. This section provides an overview of the five distinct parameterisation methods employed.

- **Literature-based:** Effect multipliers determining causal strength were sourced from high-quality meta-analyses or established theory where quantitative estimates were unavailable. The non-linear functional forms (e.g. sigmoid or concave shapes) are driven by theoretical plausibility. For example, risk ratios for the effect of unemployment and insecure work on poor mental health.

- **Data-driven:** Most model parameters are anchored in ONS data, primary analysis of the harmonised BHPS and UKHLS data for Scotland, or longitudinal survey results. For example, the job gain or job loss rates.

- **Constrained calibration:** For variables that are unobservable (e.g. time for norms to change), formal calibration was used. This is also the case in the absence of any data such as the change in prevalence at which baseline recovery rate halves. This was also used in the labour market to improve empirical fit given the limitations of the available data. To avoid overfitting, search ranges were constrained by theory and only a few parameters were estimated this way. Crucially, only one calibration exercise is used per module.

- **Calculations:** The baseline recovery rate is calculated from data adjusted for equilibrium constraint.

- **Theoretical:** Parameters such as composite weights, the resistance factor and delay parameters are set on theoretical grounds where neither literature nor data were available. These were tested via sensitivity analysis.

Sensitivity analysis was targeted at the parameters whose value or functional form is not directly observable and which therefore carry genuine modelling

uncertainty: parameters estimated by constrained calibration, the shapes of table function relationships, and the theoretically set composite weights. Parameters anchored directly in data or in point estimates from the literature — for example, the baseline labour market flow rates and the effect sizes drawn from meta-analyses — were not separately subjected to sensitivity analysis, as their values are empirically fixed. The results are summarised Table 6-2 and reported in full in Appendix F. Sensitivity analyses for key variables are discussed in Sections 6.3 to 6.5, within their respective modules.

Table 6-2: Sensitivity of model parameters

The parameters subjected to sensitivity analysis, grouped by how strongly model behaviour responds to varying them. Full results are in Appendix F.

Category	Count	Parameters (estimation method)
Critical (alternative parameterisations change the model's qualitative behaviour)	1	effect of gender parity in hours on gendered insecure work (theoretical table function shape)
High (only one parameterisation reproduces the observed trend; alternatives cannot match it)	2	resistance factor (theoretical table function shape), effect of norms on unpaid work (theoretical table function shape)
Medium (magnitude shifts but the qualitative trajectory is unchanged)	7	baseline exit rate (constrained calibration), time for norms to change (constrained calibration), effect of economic signal on norm reinforcement (theoretical table function shape), prevalence threshold (constrained calibration), effect of norms on gender parity in working hours (theoretical table function shape), effect of gender parity in working hours on gender pay gap (theoretical table function shape), effect of insecure work on the effect of employment on recovery (constrained calibration)
Low (no significant difference across the tested shape variations)	8	effect of gender norms on women's inactivity (theoretical table function shape), effect of statutory pension-age change (theoretical table function shape), effect of poor mental health on job gain rate/job loss rate/labour market exit (theoretical table function shape), economic

Category	Count	Parameters (estimation method)
		signal composite weights (theoretical), effect of gender parity in working hours on unpaid work (theoretical table function shape), fraction of gap closed by childcare (theoretical table function shape)

In summary, the simulation is built as three coupled sub-models, operationalised as prevalence-based stocks, deliberately parsimonious, and anchored where possible in data rather than tuned by calibration. With these choices in place, the next three sections describe each sub-model in turn, starting with the labour market.

6.3 Module 1: The labour market

6.3.1 Module structure

Module 1 describes how the working-age population (WAP; adults aged 16-64 years) in Scotland moves between three states: being in work (*Employed*), looking for work (*Unemployed*), and being out of the labour market entirely (*Inactive*). These movements are driven by both exogenous and endogenous factors: from outside the model come *Labour Demand* and trend in *Insecure Work*; from inside the model come the *Prevalence of Poor Mental Health* and the *Prevalence of Norms on Gender Roles (Gender Norms)*, both of which alter the rates at which people enter and leave each state. The module is split into two pieces: a signal-generating sector (Section 6.3.1.1,

Figure 6-2), which turns *Labour Demand* into a measure of how easy it is to find work using the stocks of *Employed* and *Unemployed* from within the model (*Labour Market Tightness, LMT*), and a stock-and-flow sector (Section 6.3.1.2, Figure 6-3), which tracks the population in each labour market state and the transitions between them.

The two components of the labour market interact through the balancing loop B1 (demand saturation). In this structure, increases in the *Employed* and *Unemployed* stocks exert a negative impact on *LMT*. As more people move into employment, *LMT* falls, which slows *job gain* and accelerates *job loss*, returning

the system toward balance. In the simulation, B1 operates simultaneously as a balancing influence on the *Employed* stock and an amplifying influence on the *Unemployed* stock; a richer dynamic than the single balancing arrow in the CLD in Chapter 5, which for visual clarity, only included one labour market state: *Employed*.

6.3.1.1 Labour market macroeconomy

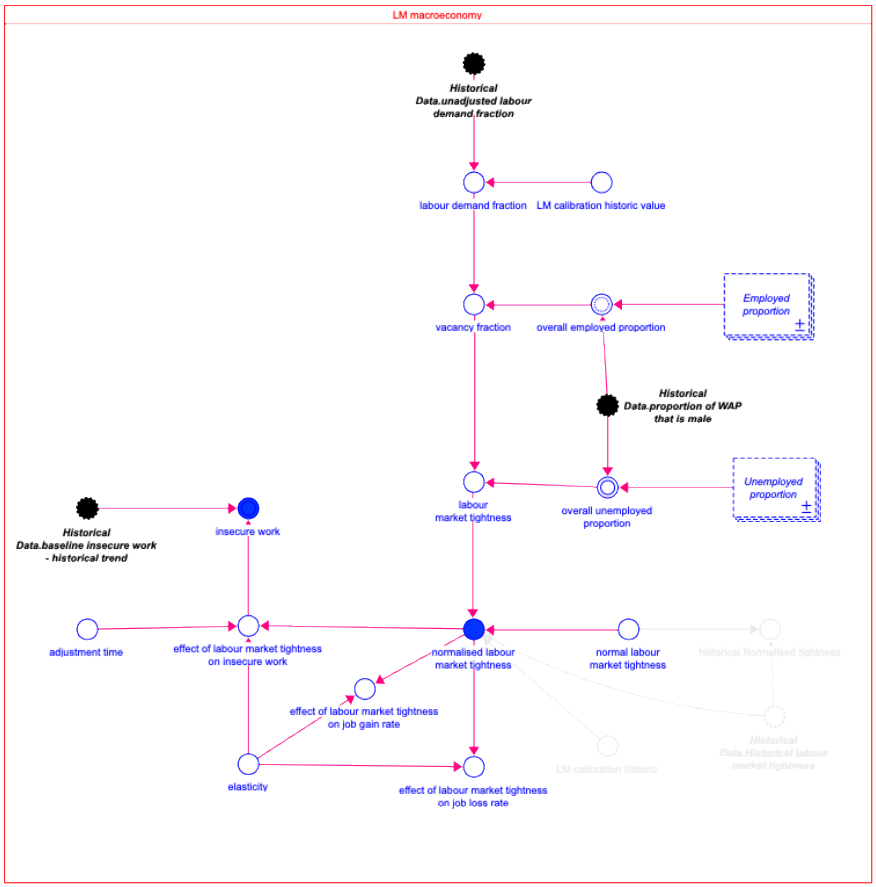


Figure 6-2: The labour market macroeconomy sector

Exogenous labour demand (black) is converted into the vacancy fraction and Labour Market Tightness (blue) — the central endogenous signal that modifies the labour market flows.

The macroeconomic sector takes historical *labour demand* (total jobs plus vacancies, normalised against the working-age population) as an external input. From this, it computes the *vacancy fraction* (the unfilled portion of labour demand) and from that, *LMT*[2], defined as the ratio of vacancies to *unemployed*. High tightness signals a sellers' market for labour: workers have more bargaining power, hiring picks up, and insecure work tends to fall. LMT is the central

[2] The *proportion of working-age population (WAP) that is male* was added as an exogenous auxiliary variable to calculate the overall proportion of employed and unemployed that is used to estimate labour market tightness, alongside the vacancy fraction.

endogenous signal in the labour market; it modifies the level of insecure work and three labour market flows: job gain rate, the job loss rate, the rate of moving from inactivity into employment. These three labour market flows were selected because they are most directly mediated by cyclical employer demand and worker bargaining power. The equation defining the vacancy fraction is formulated so that even under deep downturns a small non-zero level of vacancies remains, reflecting persistent frictional churn.

Equation for vacancy fraction

The vacancy fraction is the difference between the *labour demand* and *employed population*, representing the unfilled labour demand. It is given by:

IF labour_demand_fraction = 0 THEN 0 ELSE ((labour_demand_fraction-overall_employed_proportion) + SQRT((labour_demand_fraction-overall_employed_proportion)^2) + 0.0001) / 2

At zero gap, the equation creates a smooth asymptotic curve that maintains a non-zero frictional floor for vacancies, reflecting the reality that some turnover and hiring persist even in deep downturns due to mismatch and churn.

The effects of *LMT* are based on *normalised tightness*. *LMT* is normalised over the historical average (1992–2019). When *LMT* equals its historical average, the effect multiplier equals one, ensuring the endogenous impact matches the implicit impact already present in the exogenous baseline. The translation of this tightness signal into effect multipliers is governed by elasticity equations rather than table functions. This aligns with the Cobb-Douglas function established in economic theory, ensuring the amplitude of cyclical fluctuations is theoretically grounded.

Insecure Work is treated as an exogenous trend modified endogenously by *LMT*. The endogenous effect has a three-year delay reflecting a lag between macroeconomic conditions and contractual change as employment terms are renegotiated. To capture the multidimensional and changing nature of precarity, insecure work is a weighted index of low pay, underemployment, involuntary temporary work, and zero-hours contracts (ZHC); ZHC carry a double weighting because they capture both income inadequacy and temporal unpredictability. The elasticity of *Insecure Work* to *LMT* is assumed to be equal to that of labour market flows, for parsimony.

6.3.1.2 Labour market states

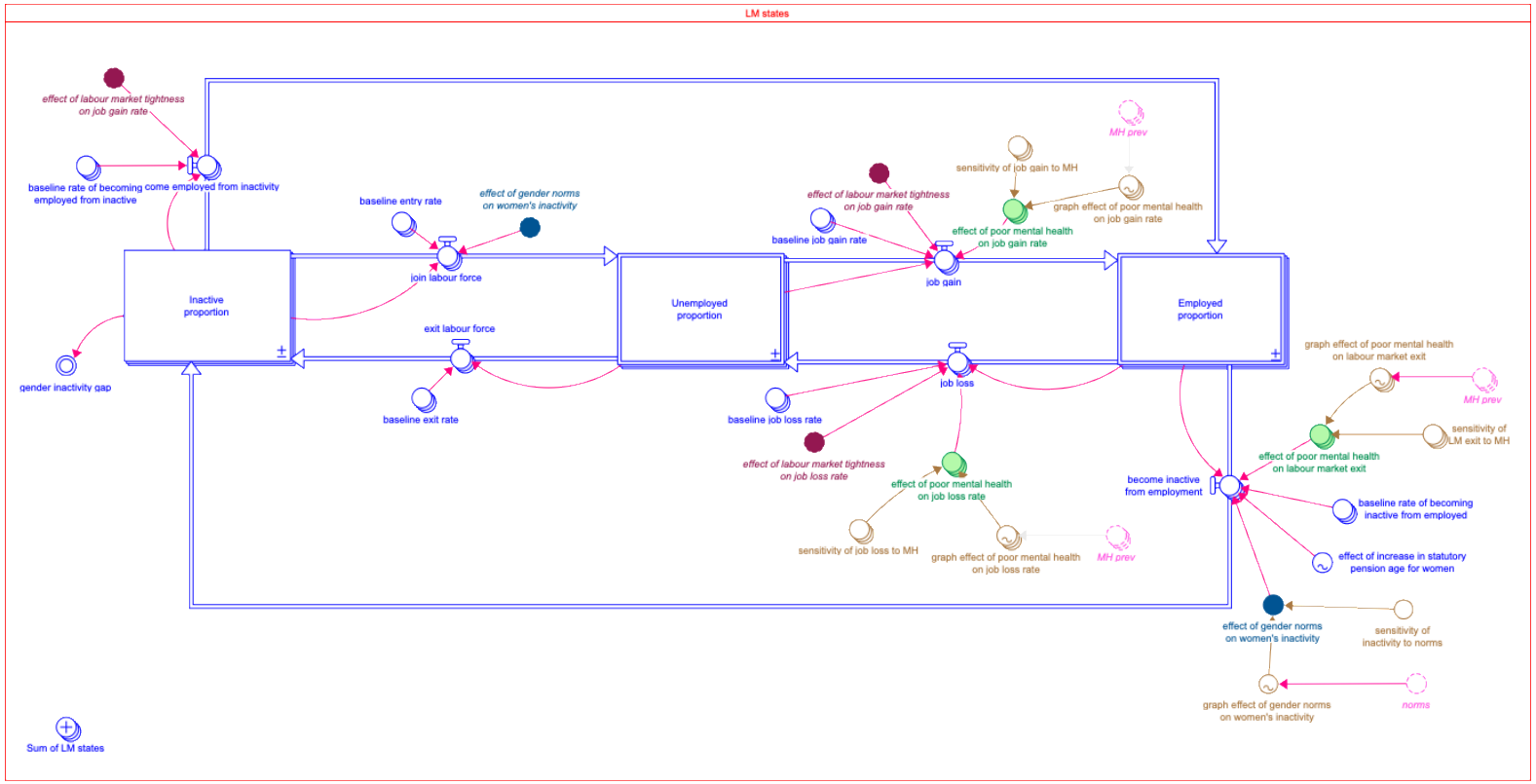


Figure 6-3: The labour market stock-and-flow structure

Three mutually exclusive states (employed, unemployed, inactive) arrayed by gender, and six flows; effects feed in from gender dynamics module (blue), mental health module (green) and within labour market module (maroon). Brown variables aid in the sensitivity analysis and model testing.

All stocks and flows within the labour market are implemented using gender-based arrays [Men, Women]. This means that while the visual structure remains parsimonious, the model tracks separate values for men and women across every transition. This allows the model to maintain identical flow logic while applying gender-specific parameter values and norm-based multipliers, ensuring structural consistency across the genders.

Stocks: The working-age population is divided into three mutually exclusive states, where the sum for each gender always equals 1. This allows the model to focus on relative rates of change and gender disparities independent of absolute population shifts.

- *Employed Proportion (Employed)*: Individuals currently in the workforce undertaking paid work.

- *Unemployed Proportion (Unemployed)*: Individuals actively looking for employment.

- Economically *Inactive Proportion (Inactive)*: Individuals outside the labour force i.e. not employed and not seeking employment (e.g. due to care responsibilities or long-term illness).

Flows: The movements between these stocks are governed by six flows.

- *job gain*: This represents the transitions from *Unemployed* to *Employed*. It is endogenously modified by *Prevalence of Poor Mental Health* and *LMT*. Baseline rates reflect the historical annual rate of finding employment for the population without poor mental health.

- *job loss*: This represents the transitions from *Employed* to *Unemployed*. Similar to *job gain*, this is modified by *Prevalence of Poor Mental Health* and *LMT* and uses a baseline rate for population without poor mental health.

- *become inactive from employed*: This represents the transition from *Employed* to *Inactive*. While *Prevalence of Poor Mental Health* impacts

both genders, women's flows are uniquely modified by *Gender Norms* and *Statutory Pension Age Change* (for women).

- *become employed from inactivity*: There is a direct flow from *Inactive* to *Employed*. An annual model cannot capture short periods of unemployment. Many people who re-enter the labour market after a period of inactivity (e.g. students after graduation, parents after childcare years, or carers) may secure a job within a year. From the perspective of an annual snapshot, they transition directly from being inactive to being employed. This flow is endogenously influenced by *LMT*.

- *join labour force*: This represents the transitions from *Inactive* to *Unemployed*. This rate is impacted by *Gender Norms*, with women's rates lower than men's. To maintain parsimony, the model uses the same effect multiplier as that for the movement from employment to inactivity. Given the reverse effect on this flow, it uses a reciprocal formulation where the effect of gender norms divides the rate for women.

- *exit labour force*: This represents transitions from *Unemployed* to *Inactive*. This rate is estimated by calibration as explained in the next section.

6.3.2 Parameterisation

The parameters for this module are largely anchored in ONS data. A summary is reported in Table 6-3 and full details are in Appendix E. Baseline flow rates were calculated using quarterly hazard rates from ONS.[3] For flows sensitive to *Labour Market Tightness*, regression analyses were performed to estimate elasticities. Observed rates were back-adjusted for effect of multipliers to get the implied baseline. For parsimony, all baseline flow rates within the labour market were assumed to be the same for men and women apart from the exit rate.

The baseline rates for the *exit from the labour force* (unemployment to inactivity) were disaggregated by gender and estimated using **constrained**

[3] Quarterly rates were converted to annual rates using the formula: 1 - (1 - quarterly rate)^4

calibration. To account for differences in scale of the stocks, the stocks were weighted by 1/average value of historical data. Without this weighting, variables with large absolute values (e.g., Employment) would dominate the calibration. This calibration followed a three-stage process. A symmetric search around the UK mean (±50% of 0.552) failed to identify a viable value for women. An empirically-guided search then revealed a gendered divergence: men's rates clustered near the UK mean, while women's rates required materially higher values. Asymmetric bounds were therefore adopted for the final calibration: men capped at +25%, women at +100%. The departure from the UK mean is acceptable because this is the only flow being calibrated, the only rate hypothesised to differ between men and women, and the Scottish data are not fully captured by UK-wide averages. Full calibration logs are in Appendix H2.

A sensitivity analysis revealed that these revised rates significantly alter the system's resting state (Appendix H2). When 1992 values are held constant, the model settles at a lower level of female unemployment and a higher level of female inactivity. Critically, these calibrated values corrected a structural misalignment in the women's sub-sector. Without this calibration, the female inactivity stock declined steadily, failing to reflect historical persistence. For the remaining stocks, this parameter does not alter the direction of the settled state, only the magnitude. To evaluate the impact of parameter uncertainty on the model's core findings, a multivariate sensitivity analysis was conducted. This demonstrated that while the magnitude of labour market stocks varies slightly, the longitudinal trajectory of rising psychological distress and the persistence of the gender mental health gap remain structurally robust.

Table 6-3: Summary of parameters in labour market module

Parameter	Source	Method	Confidence; Issue
Labour demand	ONS	Data series, estimation	Moderate; pre-1996 jobs estimated using UK data; vacancy data is for UK and pre-2001 is estimated.
Insecure work (baseline trend)	ONS	Data-driven	Moderate; composite created; linear trend is derived after partialling out effect of LMT

Parameter	Source	Method	Confidence; Issue
baseline job gain rate baseline job loss rate	ONS	Data-driven, regression for LMT, adjusted for mental health effect	High; UK-level
baseline rate of becoming inactive (employed → inactive)	ONS	Data-driven, adjusted for mental health, state pension, and gender norms	Medium; UK-level each of three adjustments introduce uncertainty.
baseline rate of becoming employed (inactive → employed)	ONS	Data-driven, regression-adjusted for LMT	High; UK-level
baseline entry rate (inactive → unemployed)	ONS	Data-driven, adjusted for gender norms	Medium; UK-level and adjusted by gender norms
baseline exit rate (unemployed → inactive), by gender	ONS + constrained calibration	Calibrated within asymmetric bounds (men +25%, women +100%)	Low–Medium (women); Medium (men)
elasticity of job gain/job loss/becoming employed from inactivity/insecure work to LMT	ONS	Data-driven, regression	High; UK-level
effect of state-pension change on women	ONS + Jenkins & Chandler (2010)	Empirical ratio	Medium; UK level; comparison window is short
effect of gender norms on women's inactivity		Theoretical	Low; anchor points are theoretical
effect of poor mental health on job gain rate	Butterworth (2012); Germinario et al. (2022)	Literature-anchored	High; not specific to Scotland

Parameter	Source	Method	Confidence; Issue
effect of poor mental health on job loss rate	van Rijn (2014)	Literature-anchored	High; not specific to Scotland
effect of poor mental health on labour market exit	Amiri (2019), Campbell et al. (2022)	Literature-anchored	High; not specific to Scotland

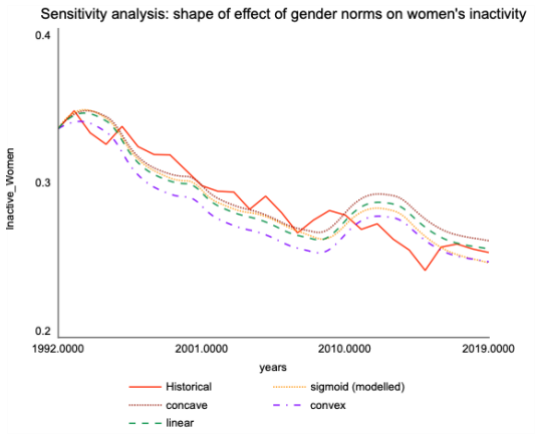
6.3.3 Structural Verification

The labour market module was verified through four categories of tests: extreme-conditions tests (inputs pushed to theoretical extremes), boundary-adequacy tests (exogenous inputs varied by ±20%), behavioural-integrity tests (transient shocks and loop knock-outs), and sensitivity analyses (the influence of non-linear table-function shapes). Full results are in Appendix I, with a summary below. One behavioural-integrity test, examining the necessity of a feedback loop, is reported in full below.

- The module behaved logically under zero, doubled, and shocked labour demand.

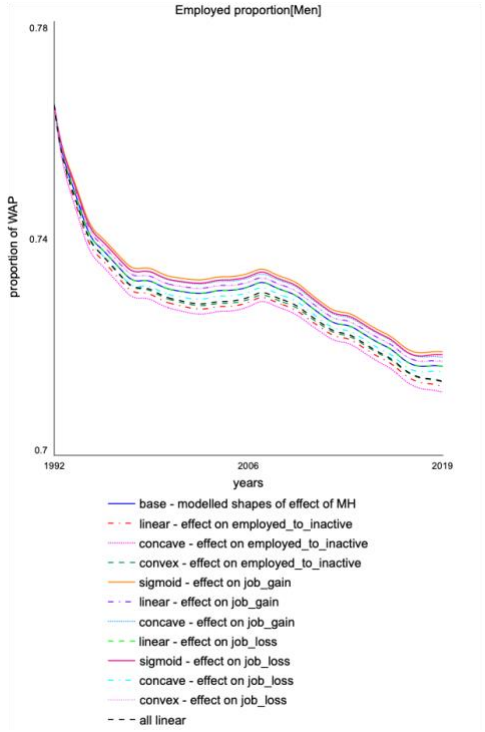
- The trajectory of poor mental health was robust to ±20% variation in exogenous inputs.

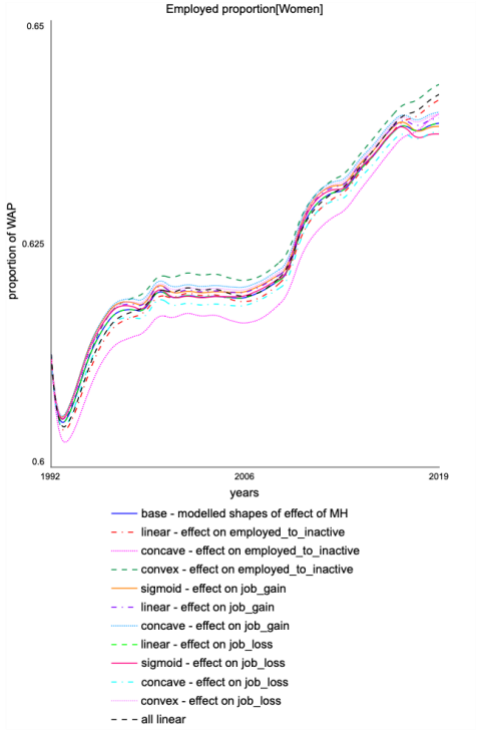
- Across sigmoid, concave, convex and linear alternatives the three table functions (effect of gender norms on women's inactivity, effect of poor mental health on the labour market flows, and effect of statutory pension-age change) produced trajectories identical in direction and timing, linear forms were adopted for parsimony (Figure 6-4).





Panel A





Panel B

Figure 6-4: Sensitivity analysis within the labour market module

Panel A shows that all four shapes (sigmoid, concave, convex, and linear) produce trajectories that are functionally identical in terms of direction and timing for two multipliers (1) the effect of gender norms on women's inactivity, and (2) the effect of change in statutory pension age on the rate of becoming inactive from employed (among women). Panel B shows different formulations for the effect of prevalence of poor mental health on labour market flows for men and women.

Structural necessity of balancing loop b1 (demand saturation). The contribution of loop B1 was tested by comparing a base case against a step increase in demand with and without the elasticity link between *Labour Market Tightness* and the labour market flows (Figure 6-5). In the active configuration, LMT stabilised as rising employment exerted a saturation effect on the market. In the knock-out configuration, LMT rose to an unregulated steady state and the unemployment stocks failed to respond to the demand shock at all, remaining

flat at their 1995 levels. The failure of unemployment to react when the LMT link is severed provides direct evidence that loop B1 is the primary structural engine driving macroeconomic adjustment in the labour market module.

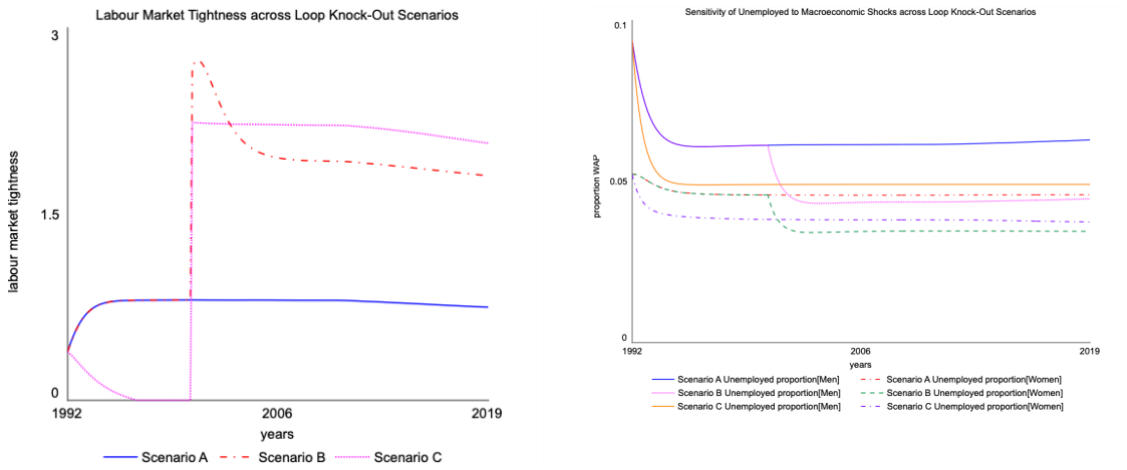


Figure 6-5: Loop knockout analysis testing the structural necessity of B1 (demand saturation).

The base case (A) was compared against a step increase in demand with (B) and without (C) the elasticity link between Labour Market Tightness (LMT) and labour flows. In Scenario B, the active balancing loop regulated the system, allowing LMT to stabilise at a value of 1.8 by 2019 as the increased employment levels exerted a satiation effect on the market. Conversely, in Scenario C, where the LMT elasticity was set to zero, the system became effectively decoupled. In this knock-out state, the LMT rose to a higher, unregulated steady state of 2.0, while the unemployment stocks exhibited zero response to the demand shock, remaining flat at their 1995 levels.

In summary, the labour market module turns external labour demand into the annual flows of people between employment, unemployment, and inactivity. These are also influenced by *Gender Norms* and *Prevalence of Poor Mental Health* described in Sections 6.4 and 6.5. The calibration revealed that the exit rate from unemployment to inactivity is markedly higher for women in Scotland than UK-wide benchmarks would suggest. Sensitivity testing confirmed that the model's trajectory is robust to plausible parameter variation.

6.4 Module 2: Gender dynamics

6.4.1 Module structure

This sub-model operationalises the evolution of the structural forces of gender norms that create and perpetuate gender inequality.

Stocks: It contains one intangible stock representing the accumulation of norms over time, the *Prevalence of Norms on Gender Roles (Gender Norms)*. It is represented as the proportion of heterosexual couples in which the male

breadwinner or the modified male breadwinner norm is held. This stock is not disaggregated by gender as it is an aggregate measure of gender inequality.

Flows: The stock is driven by two flows:

- *norm erosion* (outflow): *Gender Norms* are eroded continuously as demographic cohorts are replaced. The mechanism is therefore generational cohort replacement i.e. the gradual turnover of the working-age population rather than a change of view within individuals over a single working life. The length of a working career provides an empirical proxy for this turnover timescale: it ranges from approximately 35 years (a career resumed after childbirth) to 45–50 years (a full working life, given women's retirement age of 60 before 2009 rising to 65). Using constrained calibration within this range, the time constant was initially estimated at 40 years and refined to 45.

- *norm reinforcement* (inflow): This represents the maintenance and strengthening of traditional norms driven endogenously by economic signals and norms themselves.

Figure 6-6 provides the visualisation of the accumulation of *Gender Norms*. The next sections describes the inflow (Section 6.4.2) and its components (Sections 6.4.3 –6.4.4).

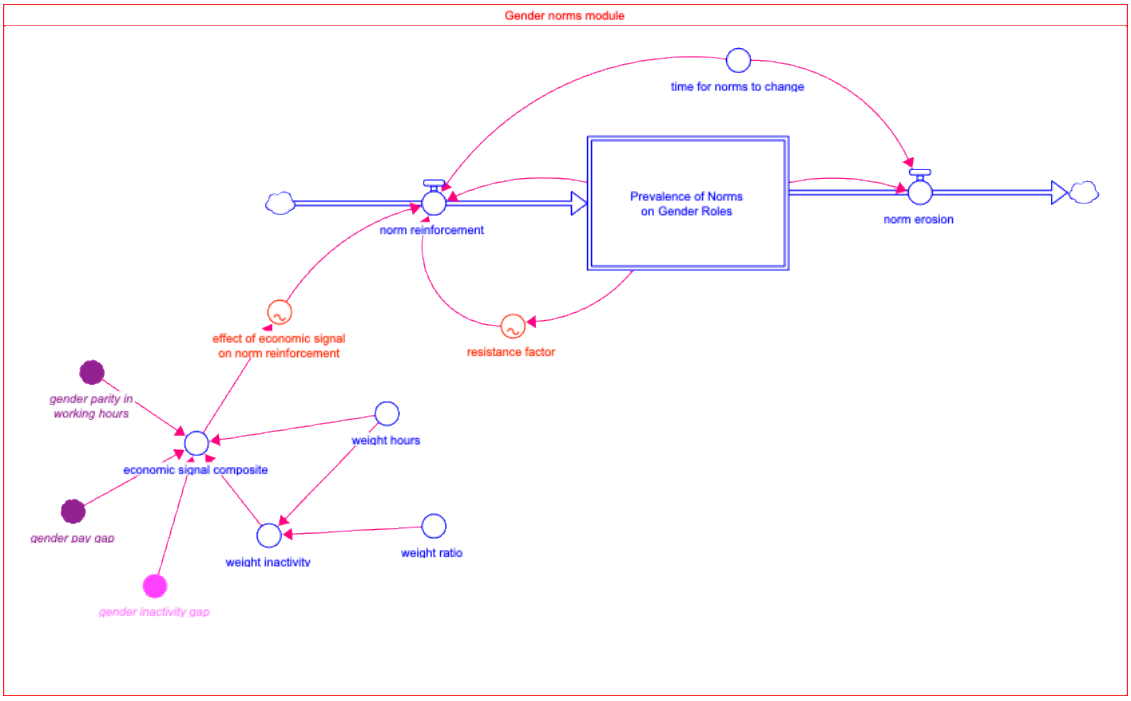


Figure 6-6: The gender dynamics module

A single gender norms stock with reinforcement and erosion flows; it takes inputs from other modules (pink); red variables highlight table-function multipliers discussed in Sections 6.4.3-6.4.5

6.4.2 Persistence of norms

The inflow that maintains traditional gender norms, *norm reinforcement*, captures the idea that, as long as the economic environment continues to reward gendered specialisation, traditional norms keep being passed on. The flow is constructed from four elements:

1. *Gender Norms:* Reinforcement is proportional to how widely the norm is already held, because the pressure to conform comes from the people currently upholding it. This contrasts with standard models of social diffusion, which dampen the inflow as the population becomes saturated.

2. *Economic signal:* A composite of three indicators of gender inequality in the labour market (described in Section 6.4.3) that measure how strongly current conditions reward gendered specialisation.

3. *Resistance factor:* Even when economic conditions favour change, norms shift slowly because of cultural stickiness of gender norms; the resistance factor (Section 6.4.4) captures this.

4. *A time constant: Time for norms to change* sets how long the inflow takes to act. It represents the average time required for a shift in ideology to manifest at the population level. The same time constant governs erosion.

Formulation for norm reinforcement

Norm reinforcement is formulated as: MIN((Prevalence_of_Norms_on_Gender_Roles* (effect_of_economic_signal_on_norm_reinforcement + (1-effect_of_economic_signal_on_norm_reinforcement)*resistance_factor))/ time_for_norms_to_change, (1-Prevalence_of_Norms_on_Gender_Roles)/DT)

MIN(X, (1-Prevalence_of_Norms_on_Gender_Roles)/DT): The use of a minimum function ensures that the norms do not exceed the maximum value of 1 and that it operates at full capacity until the population is 100% saturated.

effect_of_economic_signal_on_norm_reinforcement: This captures the effect of the composite economic signal.

resistance_factor: To capture the structural stickiness of gender norms, the reinforcement inflow is modified by an endogenous resistance factor. This captures a resistance to change, where the realised effect of the economic signals is more reinforcing than the implied effect of economic signals.

6.4.3 Economic signal composite

The economic signal that drives norm reinforcement is a weighted composite of three indicators: (i) *Gender Pay Gap (GPG),* (ii) *Gender Parity in Working Hours (Hours)* expressed as the equality gap (1-gender parity in working hours), and (iii) the economic *Inactivity Gap*. Given each variable can only range from 0 to 1, the range of the composite is also 0 to 1.

Hours and *GPG* are assigned 0.4 weight, while inactivity is 0.2. The lower weight on inactivity is theoretically motivated: a reduction in the inactivity gap is a weaker signal of egalitarianism when women are moving into part-time work. *Hours* and *GPG* are weighted equally because no obvious reason favours one over the other. Sensitivity analysis (reported in Appendix J) found that the trajectory is robust to the weighting choice.

The composite signal is translated into a multiplier acting on norm reinforcement using a concave function anchored at (0,0) and (1,1), indicating that there is no reinforcement due to economic signals at full parity and full reinforcement at maximum inequality. Concavity captures hypersensitivity to even marginal levels of inequality: minor disparities serve as sufficient proof to validate traditional roles and the steepest decline in norm reinforcement occurs only when the system approaches near-perfect parity. Sensitivity analysis

(reported in Appendix J) shows that *Gender Norms* is sensitive to the non-linearity of this multiplier: concave function shows a slower decline in norms compared to the linear, convex or sigmoid alternatives.

6.4.4 Resistance Factor

The *resistance factor* captures the cultural stickiness of gender norms: even when economic conditions favour change, norms shift slowly. It is implemented as a multiplier on the reinforcement inflow and based on *Gender Norms*. It was originally formulated as a sigmoid function bounded between 0.1 (low prevalence of traditional norms) and 0.8 (high prevalence), reflecting a state of cultural lock-in in which the system becomes effectively decoupled from economic counter-signals under traditional norms.

Sensitivity analysis (Figure 6-7) confirmed that resistance is structurally necessary: without it, norms deteriorate faster than observed even when the time constant is set to its theoretical maximum of 50 years.

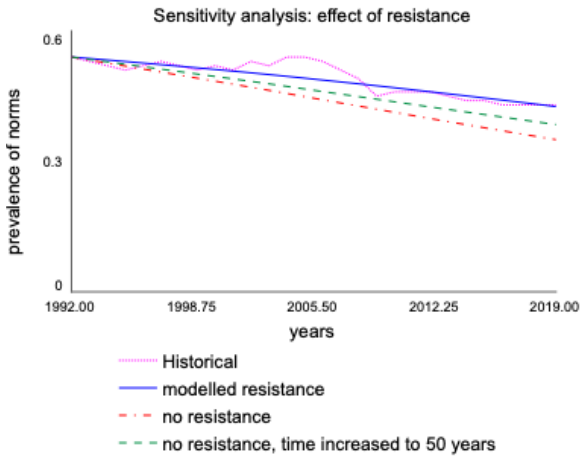


Figure 6-7: Sensitivity analysis: the resistance factor

Without resistance, the prevalence of norms erodes faster than observed even at the 50-year maximum time constant, confirming that resistance is structurally necessary.

Sensitivity analysis on the *shape* of the resistance factor (Appendix J) found that a linear formulation with the range simplified to (0, 1), so that resistance equals *Gender Norms*, reproduces the historical trend as well as the original sigmoid while removing one calibrated parameter. The linear form was therefore adopted. Section 6.4.5 brings this together with a parallel simplification of the economic signal effect.

6.4.5 Simplifying the norm reinforcement structure

Two simplifications, both supported by sensitivity analysis, were made to the norm reinforcement inflow. First, the resistance factor was reformulated from a sigmoid to a linear function equal to *Gender Norms* (Section 6.4.4). Second, the effect of the economic signal on norm reinforcement was reformulated from a concave table function to a linear one, so that the effect equals the economic signal composite itself. Together these removed two table-function parameters, reducing the model's degrees of freedom. The final formulation of norm reinforcement therefore rests on just three parameters: the time for norms to change, the weight on hours in the economic signal composite, and the weight ratio for inactivity.

The simplified norm reinforcement equation is:

norm reinforcement = MIN

[gender norms × (economic signal composite + (1 − economic-signal composite) × gender norms)]/time for norms to change ,

(1 – gender norms)/DT)

The simplified module structure, with these three parameters highlighted (in green), is shown in Figure 6-8.

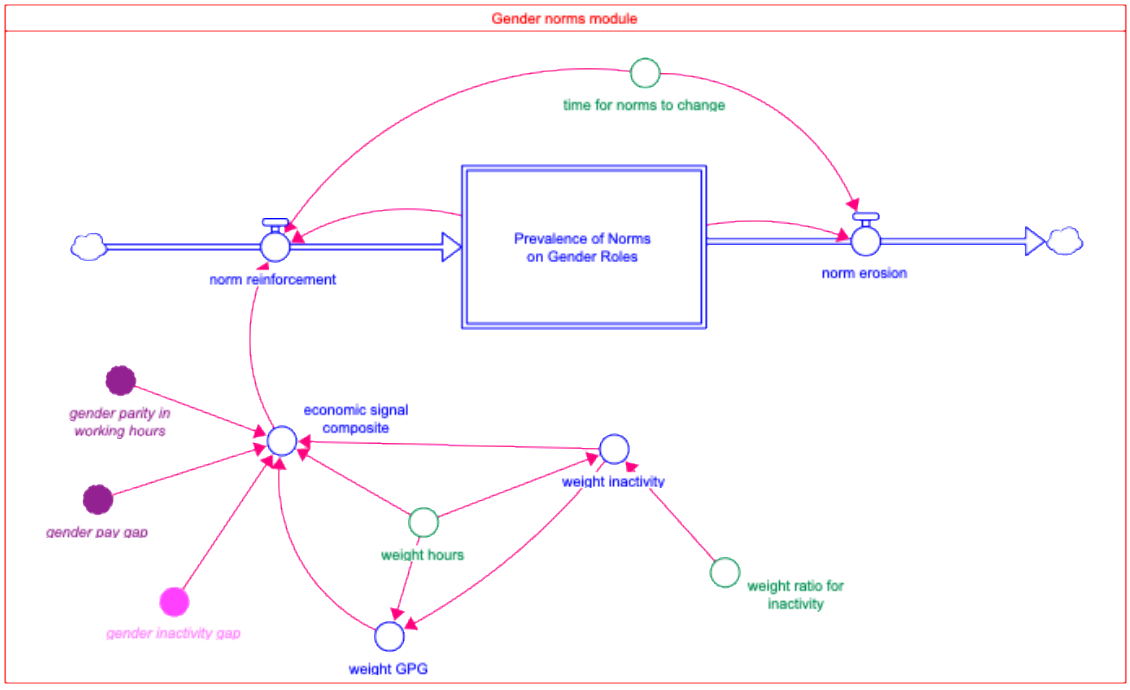


Figure 6-8: The simplified gender dynamics module

After simplification the norm-reinforcement structure rests on three parameters — time for norms to change, weight on hours, and weight ratio for inactivity — highlighted in green.

6.4.6 Gender inequality

This section describes the downstream effects of norms on gender inequality. *Gender Norms* directly impacts the *Unpaid Work Burden*, *Gender Parity in Working Hours*, women's *Inactivity* levels, and indirectly impacts the exposure to *Insecure Work* among women and the *Gender Pay Gap* through the *Gender Parity in Working Hours*. Figure 6-9 shows the formulation of Unpaid Work Burden and Gender Parity in Working Hours.

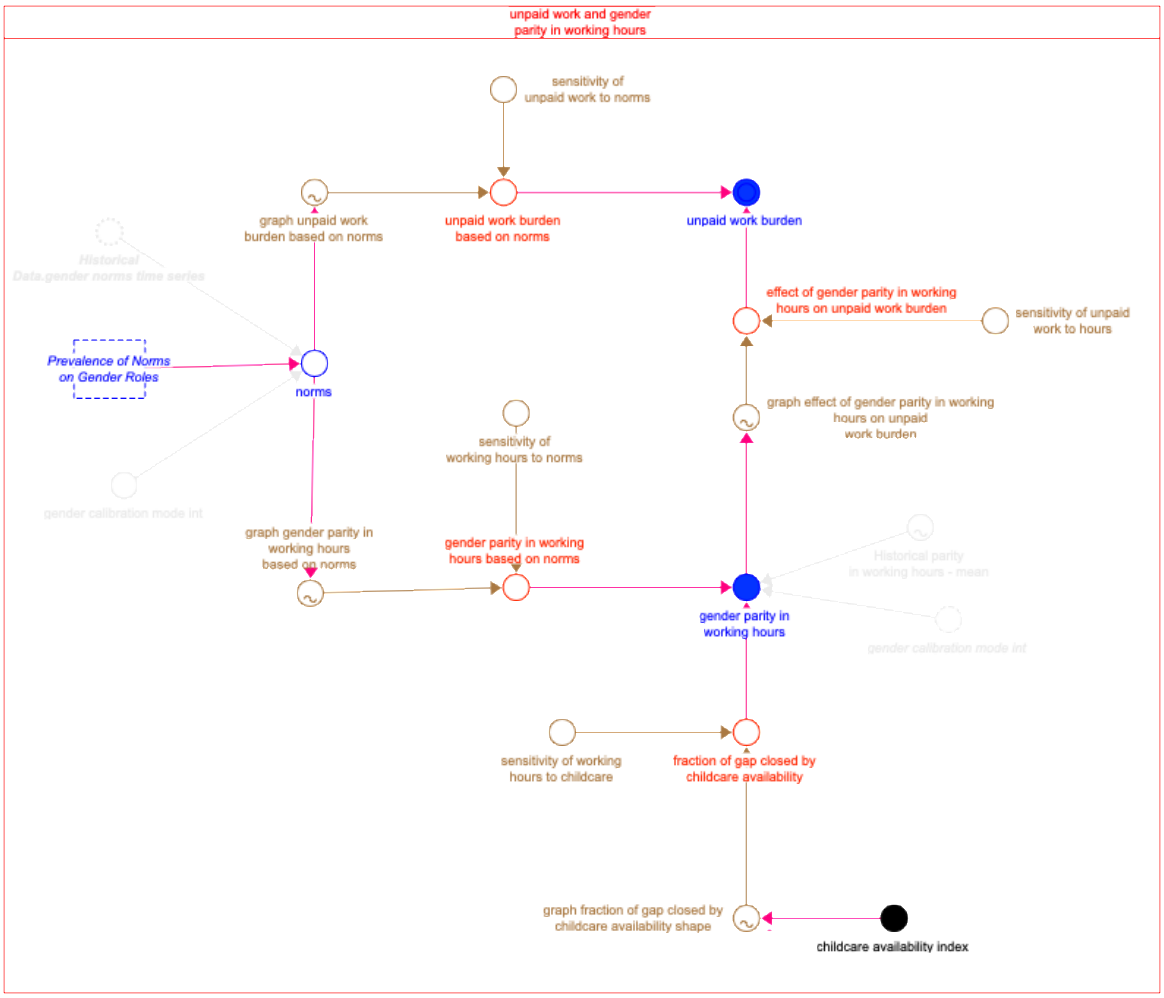


Figure 6-9: Formulation of the unpaid work burden and gender parity in working hours

Shows how norms drive both and how childcare availability modulates working hours parity.

6.4.6.1 Gender Parity in Working Hours

Gender Parity in Working Hours (Hours) is operationalised as the ratio of women's to men's average weekly hours. It is determined largely by *Gender Norms*, but can be reduced by childcare availability. *Childcare availability* is included as an explicit exogenous variable, modelled as a childcare availability index. Unlike the relatively stable proportion of mothers in the workforce,

childcare provision is a key policy lever that has changed over the model's time horizon and including it allows for the exploration of policy scenarios. However, to maintain parsimony, it is not modelled with its own feedback structures. Instead, it acts as a dynamic modulator on the relationship between *Gender Norms* and *Hours*, by closing the gap between men and women's working hours.

The functional form was hypothesised to be convex: small amounts of childcare (a few hours per week) cannot meaningfully change a woman's labour market status from part-time to full-time. Sensitivity analysis (Figure 6-10) showed that the childcare availability did not produce a significant difference in model behaviour; however it was retained given the potential intervention point. Although the historical effect is small, probably due to the low availability, the potential for impact remains high. The concave and linear functions overestimate the gains, thus the theoretically-based convex shape was retained. Leverage points were tested against different formulations of the effect of childcare, as reported in Chapter 7.

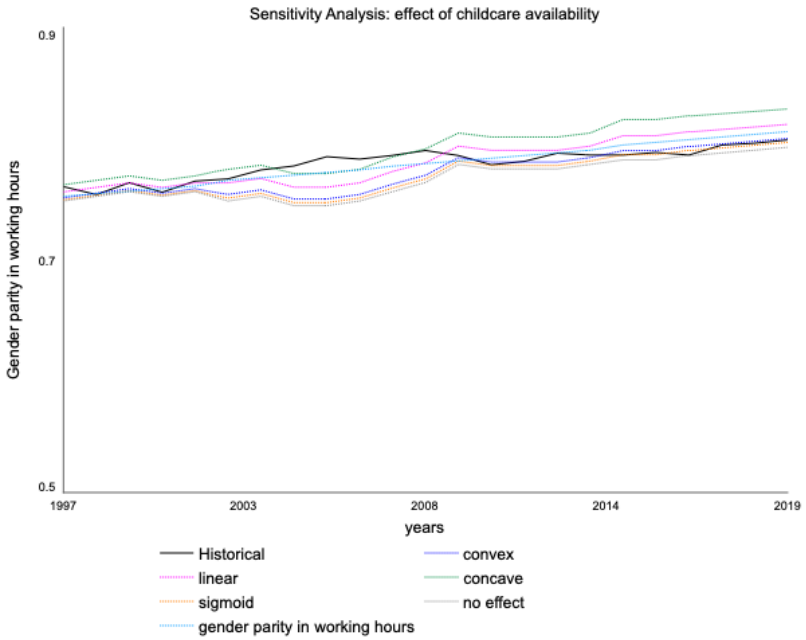


Figure 6-10: Sensitivity analysis: effect of childcare availability on gender parity in working hours

Childcare produced little change in historical behaviour given low historical availability; the convex shape was retained.

6.4.6.2 Unpaid work burden

Unpaid Work Burden (*Unpaid Work*) is the proportion of housework performed by women: 0.5 is 50% indicating parity and 1 is 100% indicating maximum disparity.

It is primarily driven by *Gender Norms*, with *Hours* reducing but not eliminating the inequality. To ensure that the proportion of unpaid work done by women has a floor of 0.5, the equation is formulated such that gender parity in working hours acts only on the gap from current proportion to equal proportion. The *effect of gender parity in working hours on unpaid work* is a function ranging from (0.5, 1) to (1, 0.5), which suggests that at maximum disparity in working hours (0.5), unpaid work burden is fully determined by norms. At full parity in working hours (1), the disparity in unpaid work halves. Following sensitivity analysis (reported in Appendix J) the *effect of gender parity in working hours* was reformulated to a simple linear function of the form (1.5 − parity).

The shape of the *effect of norms on unpaid work* was tested across concave, convex, and linear alternatives (Figure 6-11). Only a concave function could reproduce the observed magnitude; others fell short even when their range was extended to the theoretical maximum. The concave form was therefore retained.

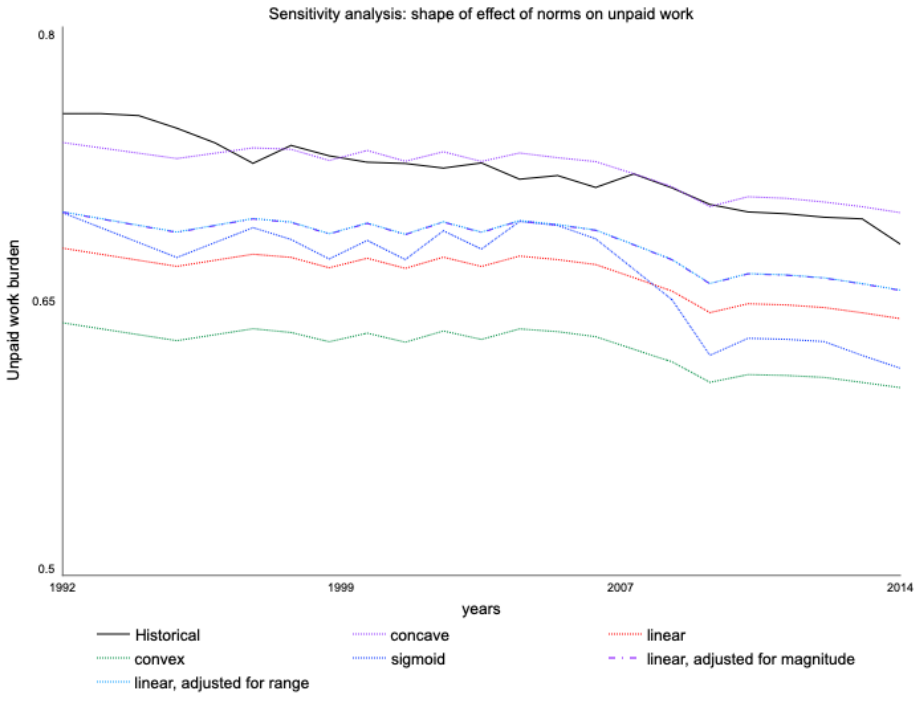


Figure 6-11: Sensitivity analysis: effect of norms on unpaid work burden

Unpaid work was sensitive to the shape of the effect of norms. While the trajectories were similar, only concave was able to replicate the magnitude.

6.4.6.3 Gender Pay Gap

The *Gender Pay Gap* combines an exogenous baseline (the gender pay gap for full-time employees) with an endogenous penalty driven by *Hours*. The gender pay gap for full-time employees is chosen as the baseline to represent the factors not captured in the model. Unlike most other variables impacted by multiple modelled factors in the model which are formed multiplicatively, the gender pay gap is formed additively, capturing the cumulative nature of disadvantage: the structural penalty for reduced hours adds to whatever pay disparity already exists among full-time workers, rather than scaling it.

Sensitivity analysis showed that the magnitude of the *Gender Pay Gap* is dependent on the shape of the *effect of gender parity in working hours*, with the best fit with a steep concave function. This shape slightly overestimates the *gender pay gap* in the early years but matches the long-run trajectory.



Figure 6-12: Sensitivity analysis: effect of gender parity in working hours on gender pay gap

The magnitude of the GPG depends on the shape; a steep concave function gave the best fit.

6.4.6.4 Gendered exposure to insecure work

To represent the relationship between gendered working patterns (i.e. *gender parity in working hours*) and *insecure work*, this model operationalises a sorting mechanism. A higher prevalence of insecure work among women is not primarily because their jobs are individually made precarious, but because they are systematically sorted into a pre-existing segment of the labour market that

offers part-time flexibility. Therefore, the model assumes that while macroeconomic factors determine the overall level of insecure work in the economy, gendered factors act as the primary mechanism that dictates the distribution of men and women into these roles. This is modelled by making women's prevalence of insecure work sensitive to their reduced working hours, reflecting a trade-off between flexibility and job quality. As *Gender Parity in Working Hours* decreases, women bear a disproportionately higher share of *Insecure Work*.

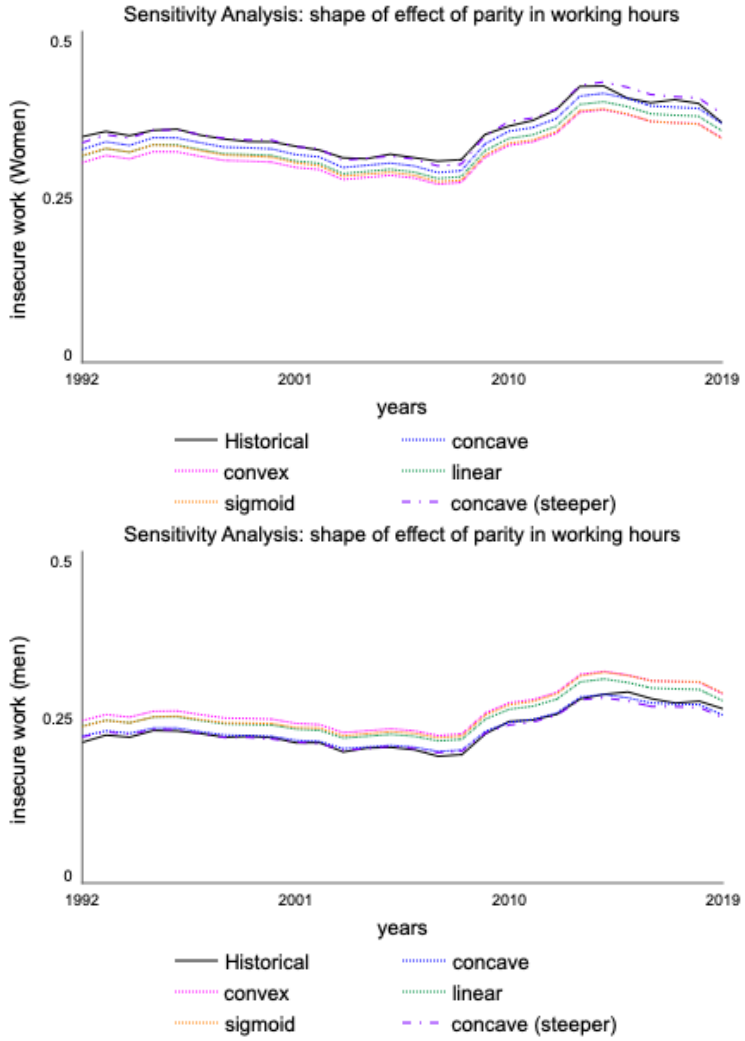


Figure 6-13: Sensitivity analysis: effect of gender parity in working hours on gendered exposure to insecure work

Only a steep concave shape reproduces the observed gender disparity; other shapes reverse the gender ordering.

Sensitivity analysis (Figure 6-13) showed that the concave curve was the only suitable shape as the other shapes resulted in higher insecure work among men and lower insecure work among women. The reverse effects implied that they

were unsuitable to capture the gender disparity. The best fit was obtained with a steeper concave curve.

6.4.7 Parameterisation

Table 6-4 details the data sources and parameterisation of the variables in this module.

Table 6-4: Summary of parameters in the gender dynamics module

Parameter	Source	Method	Confidence
time for norms to change	cohort-replacement theory (Perales, Lersch, & Baxter, 2019) with UK career length as an empirical proxy for the turnover timescale	Constrained calibration	Medium; 45 years is empirically defensible but not uniquely identified by data
gender parity in working hours based on norms	Andrew et al. (2020), ONS hours data	Theoretically-driven shape + one literature-driven anchor point	Medium; mid-range anchor uses post-childbirth average hours from a single UK study
unpaid work burden based on norms	Kan, Sullivan and Gershuny (2011); Zhou and Kan (2019)	Theoretically-driven shape + literature-driven anchor points	Medium-high; anchored at three points sourced from longitudinal data; alternative shapes cannot reproduce the observed magnitude.
effect of gender parity on norms		Theoretical	Low
effect of gender parity in working hours on insecure work	Florisson and Gable (2022)	Theoretically-driven shape + literature-driven anchor point	Medium; anchored on data at the current parity value; alternative shapes cannot reproduce the observed gender disparity.

Parameter	Source	Method	Confidence
baseline gender pay gap	ASHE (1997 onwards); New Earnings Survey extrapolated for pre-1997	time-series + trend-extrapolation	High; trend-extrapolation join produces a near-perfect numerical match but the linear projection cannot capture non-linear dynamics of the early-1990s recession.
effect of gender parity in working hours on gender pay gap	Manning and Petrongolo (2006)	Theoretically-driven shape + literature-driven anchor point	Medium; anchored on a single high-quality source;
childcare availability index	Scottish policy documents	Theoretically-constructed index	Medium; multiplicative combination for intensity and coverage is a choice
fraction of gap closed by childcare		Theoretical	Low

6.4.8 Structural Verification

The gender dynamics module was verified through the same four categories (as the labour market module): extreme-conditions, boundary-adequacy, behavioural-integrity and sensitivity tests. Full results are in Appendix J, with a summary below. One extreme-conditions test is reported in full, as it corrected a structural error.

- *Gender Norms* was robust to the ±20% boundary test on the baseline *Gender Pay Gap*.

- Sensitivity analyses across the economic signal composite weights and the time for norms to change (40, 45 and 50 years) showed difference in slope but no difference in trajectory of *Gender Norms*.

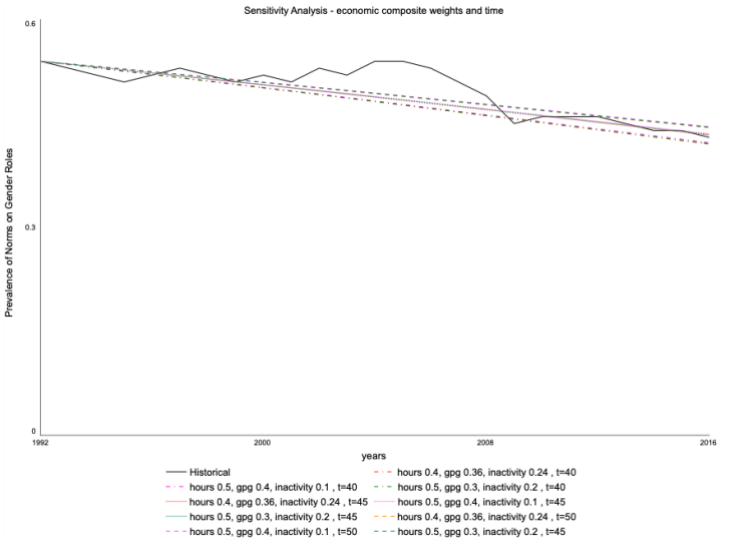


Figure 6-14: Sensitivity analysis: effect of economic signal weights and time for norms to change.

Slope varies slightly across 40, 45 or 50 years and different weights to gender parity in working hours, gender pay gap and inactivity gap; trajectory remains unchanged.

- To demonstrate the iterative nature of model development, one test is reported here that led to an error identification and subsequent correction. Extreme-conditions testing under perfectly egalitarian norms (*Gender Norms* = 0) revealed that the original multiplicative formulation drove the *Unpaid Work Burden* below its theoretical floor of 0.5. The cause was the application of the time-availability multiplier to the entire burden rather than only to the inequality component. The equation was reformulated to isolate the norm-driven inequality before applying the multiplier. This produces logical behaviour across all conditions: at full parity in working hours, the inequality halves but never falls below 0.5, and at maximum inequality the burden tends to the norm-driven ceiling. The correction was made pre-integration and was preserved in all testing.

In summary, the gender dynamics module represents traditional gender roles as a slowly-changing stock, reinforced by ongoing economic signals and eroded by cohort replacement. The chosen parameterisation gives this stock a time constant of around 45 years, broadly the length of a working career, and reproduces the gradual liberalisation of norms observed in Scotland between 1992 and 2019. The downstream variables (Gender Parity in Working Hours, Unpaid Work Burden, Gender Pay Gap, and women's exposure to Insecure Work)

provide the channels through which *Gender Norms* feed into both the labour market and mental health modules.

6.5 Module 3: Mental health

6.5.1 Module structure

The mental health sub-model tracks the share of the working-age population in each of two states of good and poor mental health, separately for men and women. Each year, some people in good mental health become distressed (the incidence or onset flow), and some people with poor mental health recover (the recovery flow). The rates at which these flows operate are not fixed: they respond endogenously to economic and gender conditions generated by the other two sub-models. This module is therefore the part of the simulation where the dynamic hypothesis that the labour market and gender norms together drive the observed rise in distress and the persistence of the gender gap is directly tested. Figure 6-15 provides the visualisation of this.

Stocks: Like the labour market module, the mental health module is disaggregated by gender using arrays, and stocks represent proportions of the gender-disaggregated working-age population, and thus equal 1 for both men and women. There are two prevalence-based stocks:

- *Prevalence of Good Mental Health*: The proportion of the gender-disaggregated working-age population currently below the threshold for psychological distress.

- *Prevalence of Poor Mental Health*: The proportion of the population experiencing psychological distress based on caseness on the General Health Questionnaire-12 (GHQ-12).

A simpler version of this module was tested first. In it, both flows operated at constant rates fixed at their 1992 values. This simplest possible version failed to reproduce the observed rise in distress even with an endogenous recovery rate; the model could not generate the non-linear inflection points in the historical data. This failure provided the empirical mandate to formalise the drivers of

onset and recovery as endogenous feedback mechanisms i.e. to make the *rate* of onset itself respond to the labour market and gender variables.

Flows: The transition between these two states is governed by two flows:

- *incidence or onset*: This flow represents the rate at which individuals transition from good to poor mental health. It is influenced by economic stressors (both genders) and gendered inequalities (women only) generated by other modules. The *baseline rate* represents the incidence rate among those in secure jobs. Incidence follows a multiplicative cumulative risk formulation; baseline rate is multiplied by the effects of the two stressors. Two composites are constructed that reduce the model's dependence on multiple unobservable table functions, moving from five separate incidence multipliers (unemployment, change in unemployment rate, insecure work, gender pay gap and unpaid work) to two integrated structural drivers (economic stressors, women's inequality). These are described in the next section.

- *recovery*: This flow represents the transition from poor to good mental health. This rate is positively influenced by those in employment (adjusted for insecure work), operationalising the positive impact of employment as well as the negative impact of inactivity. It is also endogenously modified by the *Prevalence of Poor Mental Health* where a high community prevalence of distress can diminish the resources available for individual recovery and reduce the rate of recovery.

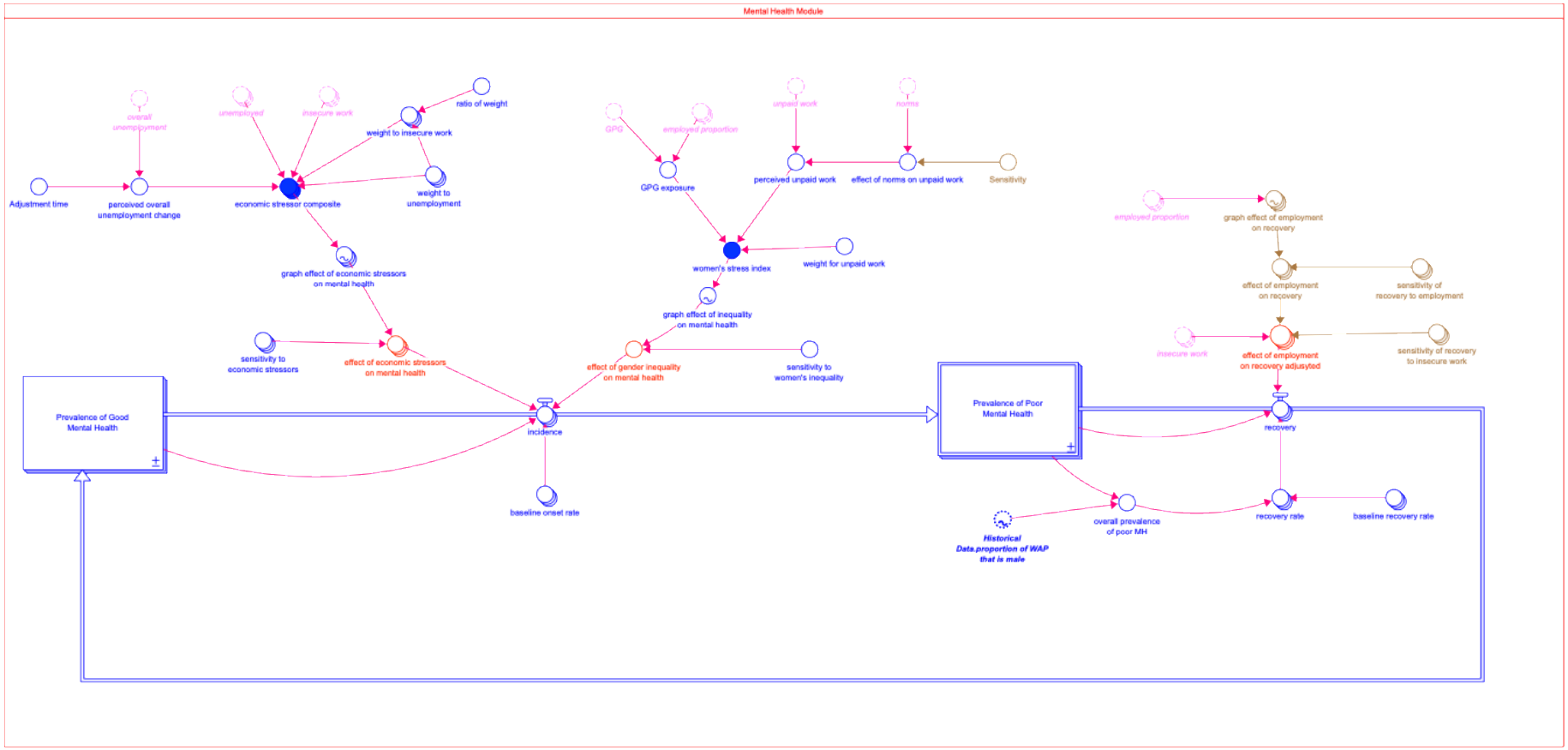


Figure 6-15: The mental health module

Two prevalence stocks with incidence and recovery flows; blue marks the structural composites, red their effects on the flows; inputs from other modules are highlighted in pink.

6.5.2 Structural composites

To mitigate the risk of over-parameterisation, the module consolidates five separate causal links into two integrated structural composites:

- **Economic stressors:** A weighted index aggregating the individual-level impact of unemployment, the macroeconomic volatility (change in unemployment rate), and insecure work.

- **Women's inequality:** A composite representing the *Gender Pay Gap* and *Unpaid Work Burden*.

6.5.2.1 Composite 1: Economic stressors

The weighting is grounded in evidence where the effect sizes for impact of unemployment are generally larger than that of insecure work. Given that the unemployment and change in unemployment rate consider the same underlying variable, the total weight assigned to these is 0.6 to ensure insecure work has a sufficient impact. Further, given the focus on the population effect, both pathways of unemployment are weighted equally. The resulting weights are 0.4 insecure work, 0.3 unemployed proportion, 0.3 change in unemployment rate. Whereas the first two are gender-disaggregated, the change in unemployment is at the aggregate level. The change in unemployment rate enters through a third-order smoothing function with an estimated two-quarter perceptual delay, capturing the lag between economic volatility and its psychological toll. The effect multiplier resulting from the composite equals 1 when the composite is 0. It represents that the baseline rate is not altered when there is no economic stress.

6.5.2.2 Composite 2: Women's inequality

The *Women's Inequality* composite captures two structural penalties faced by women: the *Gender Pay Gap* (*GPG*) and the *Unpaid Work Burden*. Both are converted into "felt" stressors that respond to context. The *GPG* is multiplied by women's employment share, so that as more women enter paid work, the population-level exposure to the pay penalty rises even when the gap itself is shrinking. The *Unpaid Work Burden* is expressed as the departure from parity

(i.e. anything above a 50/50 split) and then amplified by a normative multiplier: as traditional norms erode, the same unequal share of domestic work feels more unjust and acts as a sharper stressor. The two components are combined with a 60/40 weighting in favour of unpaid work, reflecting the stronger evidence base for the unpaid work–mental health link.

Sensitivity analysis was employed to test the impact of the normative sensitivity - while varying sensitivity altered the magnitude of the composite, the trajectory of the stressor remained stable across scenarios except when it had no impact (Figure 6-16).

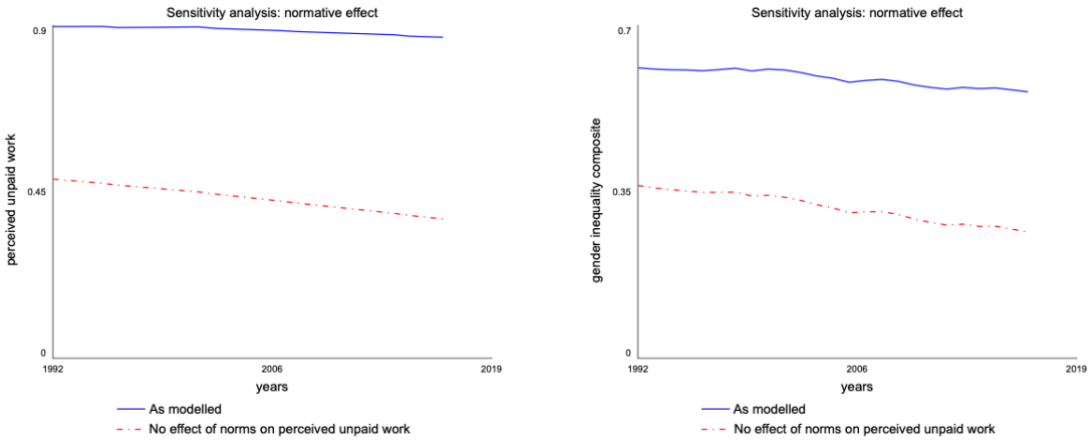


Figure 6-16: Sensitivity analysis: effect of gender norms on perceived unpaid work

Varying the normative multiplier changes the magnitude of the women's-inequality composite but not the trajectory of the stressor.

6.5.3 Effect of employment on recovery

The *effect of employment on recovery* accounts for the Insecure Work. By modelling the benefits lost from employment, insecure work is captured as a structural barrier that keeps people in poor mental health, rather than just a stressor that causes it. This modelling choice is consistent with UK longitudinal evidence that re-employment into poor-quality work confers no mental health advantage over remaining unemployed and is accompanied by worse physiological (allostatic-load) markers (Chandola & Zhang, 2018).

Given the lack of empirical data with which to parameterise this effect, this adjustment was formulated as relative. An increase in insecure work from 1992 reduces the positive impact of employment. At levels equal or below that in 1992, insecure work has no effect on recovery. At a certain level, it reduces the

positive effect of employment to 0 i.e. the effect multiplier is 1 (baseline recovery rate). Crucially, reaching this neutralisation point does not imply that secure employment ceases to provide individual recovery benefits. Rather, it signifies a point of net structural neutralisation. This neutralisation threshold was estimated separately for men and women using constrained calibration. The effect is modelled as a linear table function and applied to the benefit component of the *effect of employment on recovery*.

Because the specific points at which *Insecure Work and Prevalence of Poor Mental Health* (described in Section 6.5.4) neutralise recovery are not directly observable, a phased **constrained calibration** strategy was employed. This involved a four-stage process (see Appendix H for full logs). By setting plausible floors for parameter values, the model was prevented from over-fitting to noise and was forced to find the most robust structural explanation for the gendered divergence in recovery.

The calibrated threshold at which insecure work neutralises the recovery benefits of employment differ by gender, illustrating path dependency. Because the model anchors the effect of insecure work relative to a 1992 baseline, women's structurally higher historical exposure to insecure work (0.37 compared to 0.24 for men) is reflected in the divergence: their recovery benefit is neutralised at a smaller relative increase in insecure work albeit at a higher absolute level. Calibration indicates this neutralisation occurs when absolute insecure work reaches 0.44 for women (a roughly 20% relative increase) versus 0.34 for men (a roughly 40% relative increase).

6.5.4 Endogenous recovery rate

This subsection describes the feedback mechanism that turns the natural balancing dynamic of recovery into a reinforcing one when prevalence is rising. This parameter, referred to as the endogenous recovery rate, underwent a necessary and rigorous iterative evolution.

The mechanism. In a standard stock-and-flow structure, recovery flows grow proportionally with the stock of poor mental health: the more people in distress, the more people recovering, all else equal. This self-correction has limits. As

community prevalence rises, healthcare and informal support resources become strained, waiting times lengthen, and the rate of recovery falls (Clark et al., 2018; Siciliani, 2025). This creates a reinforcing mechanism that works against the natural balancing tendency: rising prevalence slows recovery, which allows prevalence to rise further. Thus, the recovery rate is modified by the change in overall prevalence relative to 1992, operating as a multiplier on the baseline recovery rate.

Why the initial formulation failed. The relationship was first specified as a sigmoid function with a floor at half the baseline recovery rate. Loop-dominance analysis showed that this formulation contributed essentially nothing to model dynamics during 1992–2019: the mechanism was switched off. A sigmoid relationship also turned out to be theoretically inappropriate, because it assumes a buffer zone in which small increases in demand have no effect on service capacity, whereas in practice a system operating near capacity feels pressure from the first additional patient.

The revised formulation. Following focused constrained calibration (Appendix H), a divergent gendered formulation was adopted: a concave function for men, and a linear function for women, both with an implied halving point at 150% of 1992 prevalence. This reformulation modestly activated the previously dormant loop and improved the fit for women (R^2 rising from 0.08 to 0.14). Substantively, it generates an asymmetric recovery trap: any marginal rise in distress proportionally degrades recovery capacity for women, while men retain a buffer at low and moderate stress levels. Sensitivity analysis (Section 6.5.6) shows that the integrated model is robust to the halving point and to any linear-or-concave degradation shape; a materially higher prevalence emerges only under a strongly convex women's shape, which the calibration does not support. The mechanism is therefore a bounded structural uncertainty, dormant in the historical range, retained so that it can activate under the more extreme conditions.

6.5.5 Parameterisation

Stocks are parameterised based on primary analysis of harmonised BHPS and UKHLS data for Scotland. *Baseline onset rate* is estimated to remove

initialisation bias[4]. The gender-disaggregated *baseline recovery rates*[5] are calculated such that the model initialises in equilibrium.

Following the integration of the three sub-modules, the baseline mental health recovery rates were recalculated (Men: 0.4096; Women: 0.348) to ensure the system initialised in a state of partial equilibrium. This adjustment was necessary because the resting state of the integrated model differs slightly from the isolated historical proxies used during module development. Specifically, the endogenously simulated values for variables such as *Unpaid Work Burden* and the *Gender Pay Gap* deviated from their raw 1992 time-series anchors. Table 6-5 provides the details of parameterisation for all variables in this module.

Table 6-5: Summary of parameters in the mental health module

Parameter	Source	Method	Confidence
baseline onset rate	Primary analysis of BHPS/UKHLS Scottish sub-sample	Back-calculated from observed 1992 onset, dividing by multipliers	High
baseline recovery rate	BHPS/UKHLS	Calculated from 1992 equilibrium constraint	Medium
effect of economic stressors on mental health	Curtis et al. (2019); Rönnblad et al. (2019); Thomson et al. (2023)	Theoretically-driven shape; Literature-driven anchor points	High; odds ratios come from meta-analysis
effect of gender inequality on mental health	Piovani and Aydiner-Avsar (2021); Platt et al. (2016); Wheatley, Lawton and Hardill (2018)	Theoretically-driven shape; literature-driven anchor points	Medium; studies rather than meta-analysis; single anchors for each component

[4] This baseline onset is calculated as follows: Implied baseline = observed rate in 1992/effect of multipliers in 1992. It was initially estimated using the available UK flow data (Salis et al., 2021) and a weighted average of effect multipliers (across 1992-2019; weighted by the share of men and women in the at-risk population). The estimate was subsequently refined using primary data analysis of the Scottish sub-sample in BHPS and UKHLS from 1992 onwards.

[5] This was calculated using the baseline onset rate, the prevalence of poor/good mental health and the effect multipliers for 1992, such that inflow to and outflow from prevalence of poor mental health are equal.

Parameter	Source	Method	Confidence
effect of employment on recovery	Salis et al. (2021)	Linear; literature-driven anchor points	Medium; UK-level discrete-time hazard
effect of prevalence on recovery (by gender)		Constrained calibration	Low; calibration hit search bounds without uniquely identifying the parameters; the gender-divergent shapes and the halving point of 1.5 are retained on theoretical grounds.
Insecure work threshold (effect on recovery)		Constrained calibration	Low; calibrated within plausible bounds but not uniquely identified by historical data.

6.5.6 Structural verification

Verification of the mental health module included equilibrium initialisation, extreme conditions, and parameter sensitivity. Detailed results and graphs are reported in the Appendix K.

- With all stressors held constant at their 1992 values, the mental health stocks remain flat throughout the simulation (equilibrium initialisation), confirming the implied baseline onset and recovery rates are mathematically sound. An initialisation sensitivity test was also conducted, artificially inflating 1992 prevalence to 0.25 for men and 0.3 for women. The system sought a clear equilibrium at 0.152 and 0.234 providing evidence that the implied baseline calculations are mathematically sound.

- The extreme-conditions tests (conservation of flow, zero recovery rate, and maximum stressors) all produced the expected behaviour, with no stock exceeding its population boundary

- Sensitivity analyses on composite weighting for both economic stressors and gender inequality, and different shape choices for the table functions of effect multipliers indicated that the trajectories are robust to these.

Given the effect multipliers on effect of employment adjusted for insecure work were modelled as linear, sensitivity analysis for these were not conducted.

- The prevalence-recovery mechanism, including both the halving point and the gender-specific degradation shapes, was subjected to one-way and multivariate sensitivity analysis (Table 6-6). The model is robust to the halving point and to any linear-or-concave degradation shape: with all four parameters varied together, the median 2019 prevalence sits essentially at the retained base (men 15.9%, women 23.5% against a base of 23.6%) and the central 95% of runs lie within 15.8–20.1% (men) and 23.4–28.3% (women). A materially higher prevalence emerges only under a strongly convex women's shape (Figure 6-17); the full range then reaches 27.1% (men) and 37.4% (women), but only 6 of 200 women's runs exceed 28%.

Table 6-6: Sensitivity of the 2019 prevalence of poor mental health to the prevalence-recovery mechanism, by gender

Base (retained parameters): men 15.9%, women 23.6%; halving point 1.5, men's shape 1.5 (concave), women's shape 1 (linear)

A. Univariate (each parameter swept across its range; others held at retained values):

Parameter varied	Range	Men 2019	Women 2019
Halving point (men)	1.2–2.0	15.8–16.0%	23.6–23.7%
Halving point (women)	1.2–2.0	15.8–15.9%	23.5–24.9%
Degradation shape (men)	0.5–2.0	15.8–18.8%	23.6–24.9%
Degradation shape (women)	0.5–2.0	15.8–18.2%	23.4–31.1%

B. Multivariate (200-run Latin-hypercube; halving points 1.2–2.0, shapes 0.5–2.0):

	Men 2019	Women 2019
Median	15.9%	23.5%
Central 95% of runs	15.8–20.1%	23.4–28.3%
Full range	15.8–27.1%	23.4–37.4%

The base sits at the low end of every range: the mechanism raises prevalence only when the degradation shape is more convex than the retained value (women's shape < 1). Under the full multivariate sweep the median stays at the base and only 6 of 200 women's runs exceed 28%; the high tail (women to 37.4%, men to 27.1%) requires a strongly convex shape.

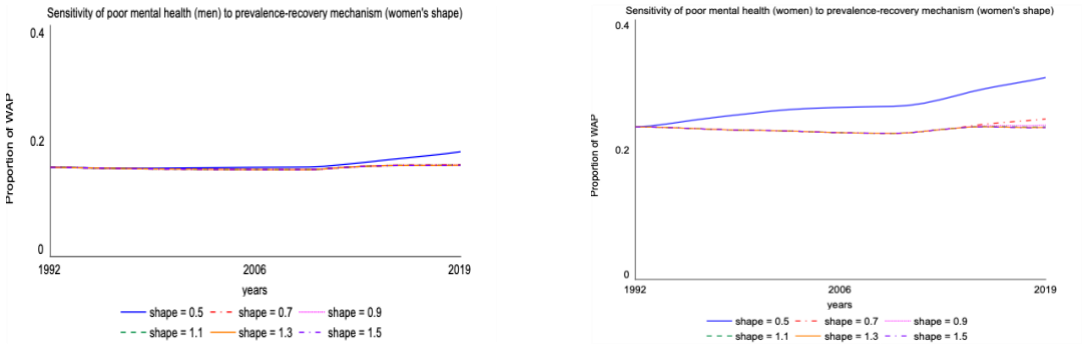


Figure 6-17: Illustrative sensitivity of prevalence of poor mental health to the women's degradation-shape parameter

Women's shape swept 0.5→1.5 (0.5, 0.7, 0.9, 1.0 (retained), 1.2, 1.5); men's parameters and both halving points held at retained values (halving point 1.5; men's shape 1.5). Only a strongly convex shape (0.5) lifts the trajectory

In summary, the mental health module turns labour market and gender variables into annual rates at which people enter and leave poor mental health, separately for men and women. Most of the explanatory load is carried by the two structural composites (economic stressors and women's inequality), which compress many partly-correlated drivers into stable signals. The endogenous prevalence-recovery mechanism adds an asymmetric recovery trap for women, but is largely dormant across 1992-2019. The next section brings these mechanisms together with the labour market and gender dynamics modules and asks whether the integrated system can reproduce the observed history.

6.6 Model validation

Validation in system dynamics is a sequence of structural checks designed to establish that the model is generating the right behaviour *for the right reasons* i.e. that any historical fit is the product of the hypothesised mechanisms rather than overfitting.

Tests in this thesis were performed at three points in the development cycle. Some tests were run on each sub-module in isolation (with historical inputs fed in for variables that would normally come from other sub-models). These are reported within each module's verification subsection (Sections 6.3.3, 6.4.8, 6.5.6). After the modules were coupled, foundational tests were re-run and further behavioural tests were conducted on the fully integrated system; the results of these tests, are reported in this section. Several parameters — notably the baseline recovery rates, the prevalence threshold governing endogenous

recovery, and the insecure work effect on recovery — were finalised only after integration, because the resting state of the coupled system differs slightly from the isolated case. Where this applies, the relevant module section flags it explicitly.

The remainder of this section reports the integrated-model results in four steps: foundational verification of mathematical and logical integrity (Section 6.6.1), behavioural tests of structural integrity that probe whether specific feedback loops generate the dynamics they are meant to (Section 6.6.2), loop knock-out experiments that test the structural necessity of each major loop (Section 6.6.3), and behaviour reproduction against historical Scottish data using Theil's inequality statistics (Section 6.6.4).

6.6.1 Foundational model verification

This initial phase ensured that the simulation model is a technically correct and faithful translation of the CLD. These tests were done for each module and then repeated with the fully integrated model.

Unit consistency: This is a crucial test of logical consistency that focuses on the resolution of any dimensional inconsistencies, ensuring that the units of all variables and flows were logically coherent (e.g. a stock of proportion of working-age population/WAP must have flows with units of proportion of WAP/year). Stella Architect's built-in "Units Check" feature was used to conduct an automated scan of all model variables and their corresponding equations. The model was confirmed to be dimensionally consistent, with no errors in the mathematical formulation of the relationships between stocks, flows, and auxiliaries. This test verifies that the model adheres to fundamental principles of physical and logical conservation.

Stock conservation: These were conducted to verify the conservation of mass within the model's closed systems. A verification run was executed with all exogenous time-series inputs held constant at their initial 1992 values, thereby isolating the model's internal structure from external influences. The sums of proportions for the gender-disaggregated stocks in both the labour market module and the mental health module were monitored throughout the

simulation. In all cases, these sums remained exactly 1.0 for the entire simulation period. This result confirms that the stock-and-flow structures are correctly specified and that no population is being erroneously created or destroyed due to flaws in the model's logic. The test was also repeated with exogenous inputs as time-varying.

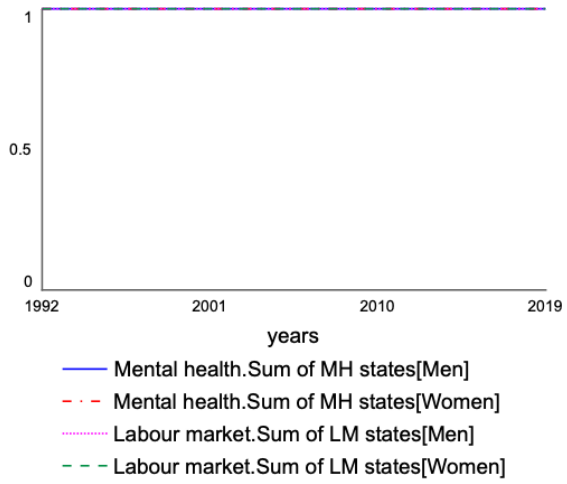


Figure 6-18: Stock conservation

The sum of stocks in both the labour market and mental health modules, for both men and women is 1 throughout the simulation period.

Integration stability and numerical verification: The model was tested for integration stability by varying the time step (DT) across three intervals: 0.25, 0.125, and 0.0625 years. Additionally, the results were cross-verified using both Euler and fourth-order Runge-Kutta (RK4) integration methods to check for sensitivity to the integration algorithm. No difference was observed in values of stock. Figure 6-19 provides an example graph for *Prevalence of Poor Mental Health* among women.

Euler integration was ultimately selected as the primary method to maintain conceptual alignment with the operational stock-and-flow logic of the system. For all final simulations and policy testing, a DT of 0.0625 (representing 1/16th of a year) was adopted. This value is substantially smaller than the shortest time constants in the model, specifically the mental health recovery lags and labour market transition delays, thereby ensuring that the integration error is negligible and the model's physics are robust. The fastest relative flow in the system was identified as the labour market exit rate for women (1.1 /year), corresponding to a time constant of approximately 0.91 years (or 47 weeks).

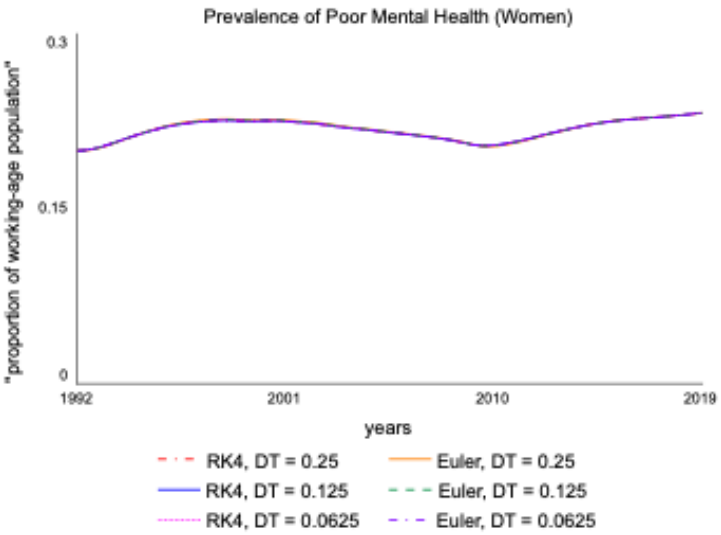


Figure 6-19: Integration method robustness

Women's prevalence of poor mental health (used as an example) under DT = 0.25, 0.125, 0.0625 yr and Euler vs RK4; trajectories indistinguishable.

Fidelity of formulation to the dynamic hypothesis: A systematic mapping exercise was conducted to trace every causal link and feedback loop articulated in the final CLD to its corresponding variable, equation, and functional form in the Stella model. An audit of all equations was conducted. One error was identified and corrected: the causal relationship from gender norms to economic inactivity among women was erroneously not represented in the model. This test was repeated as the model was iterated upon, and any errors identified were rectified.

6.6.2 Behavioural tests of structural integrity

To verify that the model's behaviour is generated by the hypothesised feedback structures rather than exogenous forcing, specific structural tests were conducted. These tests isolate specific feedback loops to verify their polarity and dominance.

Verification of recovery dynamics. The first test was to verify that the *unemployment-mental health trap* (R1) generates persistent effects following a temporary shock. A temporary 50% increase in economic stressors was applied for a 5-year duration (using a STEP function) at three distinct intervals (1995, 2005, 2015). If the reinforcing loops are active, the *Prevalence of Poor Mental Health* should not immediately return to baseline when the stressor is removed; it should exhibit a delay. As seen in Figure 6-20, the *Prevalence of Poor Mental*

Health returns to original levels after 15 years. This confirms that the model captures the scarring effect of economic recessions, where health consequences persist after the economic shock has subsided.

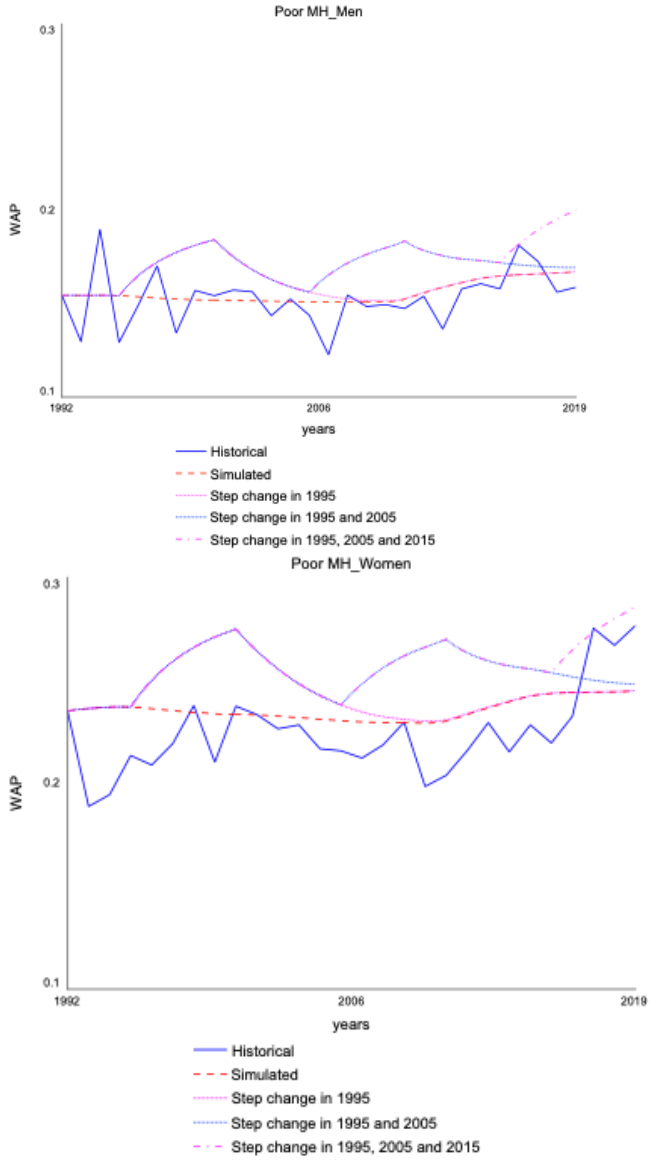


Figure 6-20: Recovery dynamics: persistence of poor mental health after a temporary economic shock

A temporary 5-year 50% rise in economic stressor in 1995, 2005 and 2015 found that return takes 15 years.

Conversely, the control run confirmed internal stability. If there is no change in the economic stressors over time, there would have been no rise in *Prevalence of Poor Mental Health*. There is a slight decrease for women because of the relative progress in gender inequality, particularly in the absence of rising exposure to *Gender Pay Gap* (employed proportion was kept constant along with the economic stressors).

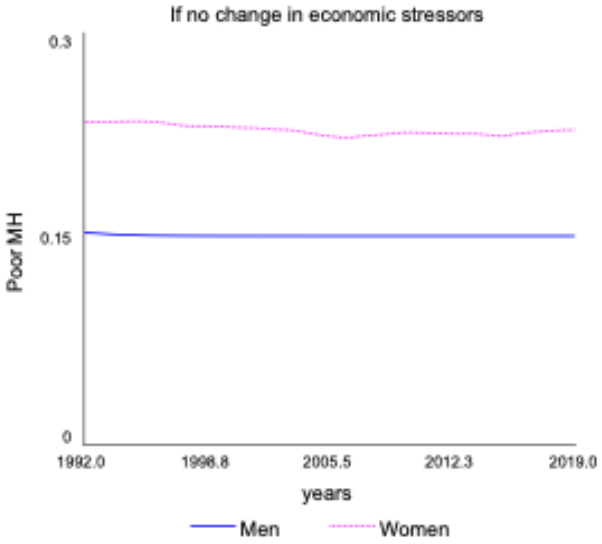


Figure 6-21: Effect of stable economic stressors on prevalence of poor mental health

The rise in prevalence among men can be attributed to economic stressors within the model as in the absence of a change in stressors, there is no change. For women, there is a slight decrease and subsequent small rise.

Isolation of the gender inequality feedback. To examine the structural contribution of gender-specific stressors, a zero-inequality test was conducted. The sensitivity of women's mental health onset to gender inequality was set to zero after 1992. When the effect was removed, the *Prevalence of Poor Mental Health* among women was lower and close to that of men. This structurally validates the hypothesis that the gender gap in the model is not hard-coded but is endogenous and confirms that the model is capable of generating equality if structural barriers are removed.

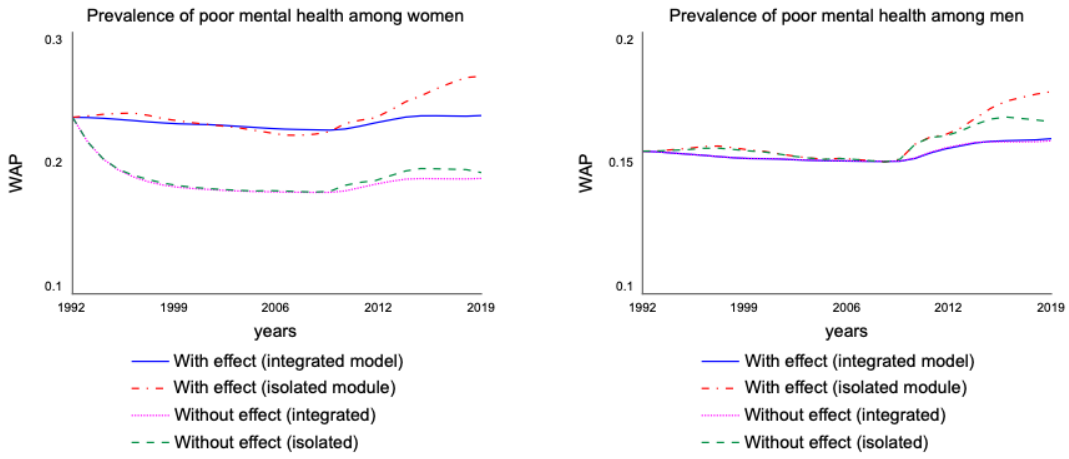


Figure 6-22: Effect of gender inequality on prevalence of poor mental health

Among women, there is a significant reduction in prevalence when there is no effect of gender inequality on onset (within both the isolated module and integrated model). There is little difference in men's prevalence with or without the effect. For this analysis, the baseline onset and recovery rates are not recalculated, but instead the sensitivity to gender inequality is modified after 1992 to be 0 i.e. to have no effect on onset.

6.6.3 Loop knock out tests

6.6.3.1 R1: unemployment-mental health trap

To evaluate the structural dominance of the link between population health and economic participation, a knock-out analysis was performed on loop R1. This test involved severing the endogenous feedback from the mental health module to the labour market by setting the multipliers for the *effect of poor mental health on job gain, job loss,* and *labour market exit* to a constant 1.0 (neutral). When the loop is inactive, unemployment is lower, with the gap between active and inactive being greater for women. This is expected given the *Prevalence of Poor Mental Health* is higher among women. Crucially, there is a significant divergence in the *Prevalence of Poor Mental Health* after 2010 in active and inactive loop configurations. The divergence is driven by recovery rather than incidence – higher unemployment/reduced employment slows recovery by reducing the benefit of employment on recovery. This is the reason the effect is later in the simulation rather than from the start.

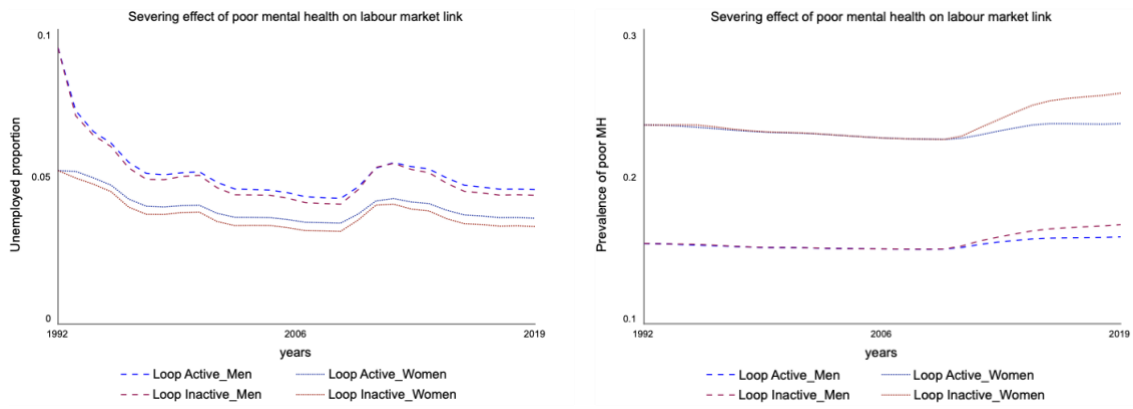


Figure 6-23: Loop knockout analysis testing the structural necessity of R1 (unemployment-mental health trap)

Prevalence of poor mental health and unemployed proportion among men and women were compared when the loop was active and when it was not. Severing the effect of poor mental health on labour market reduced the unemployed proportion and increased the prevalence of poor mental health.

6.6.3.2 R2: insecure work-mental health trap

The reinforcing mechanism between insecure work and poor mental health was severed by (1) removing its effect on onset (setting the weight to 0), (2) removing its effect on recovery (effect of employment not adjusted for insecure work), and (3) both. In scenarios (2) and (3), prevalence reduces throughout the

simulation. In scenario (1), there is a rise in prevalence at the same time as the base scenario of active loop but the level remains lower.

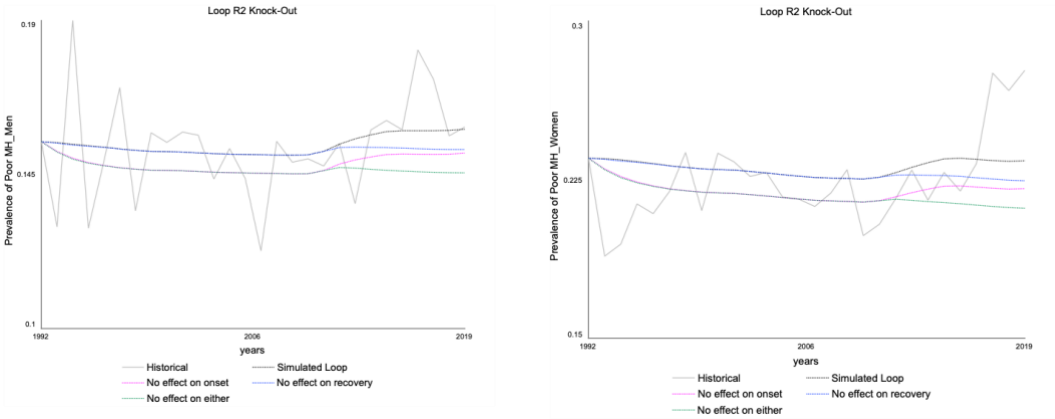


Figure 6-24: Loop knockout analysis testing the structural necessity of R2 (insecure work-mental health trap)

R2 severed three ways — (1) insecure work effect on onset set to zero; (2) employment's recovery benefit not adjusted for insecure work (no effect on recovery); (3) both compared to the active loop. Removing the recovery pathway (2, 3) lowers prevalence throughout; removing only onset (1) keeps the timing of the rise but at a lower level.

6.6.3.3 R3: norm reinforcement

Gender Norms was lower when Loop R3 was severed, either by preventing economic signals from reinforcing norms, or by severing the link between *Gender Norms* and *Gender Parity in Working Hours*, or between *Gender Parity in Working Hours* and *Gender Pay Gap*.

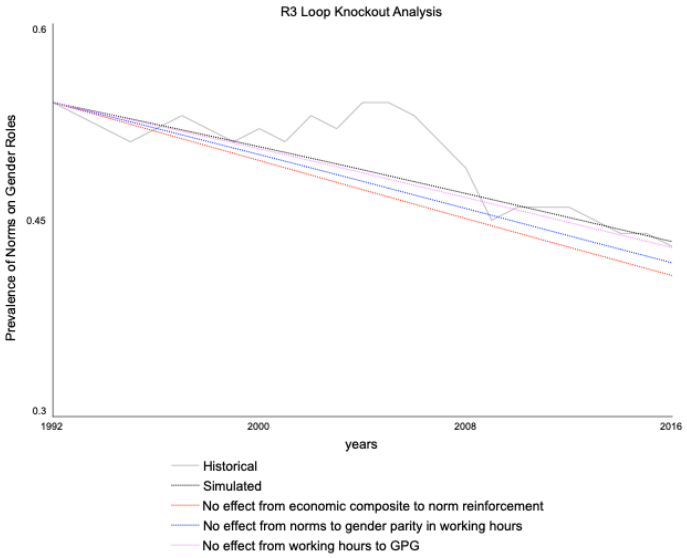


Figure 6-25: Loop knockout analysis testing the structural necessity of R3 (norm reinforcement)

R3 severed by (i) preventing economic signals from reinforcing norms, (ii) cutting norms → working-hours parity, or (iii) cutting working hours parity → gender pay gap. In each case the norms stock erodes further than with the loop active.

6.6.3.4 R4 and B4: norms-mental health loops

Loop R4 and B4 function in parallel and removal of links in this loop could have reinforcing or balancing effects. As seen, if the link from norms to poor mental health (via *Unpaid Work Burden* and *Gender Pay Gap*) is removed, the *Prevalence of Poor Mental Health* drops significantly, highlighting their importance in women's prevalence. If the relationship from poor mental health to labour market is severed for women, the prevalence rises slightly, suggesting that the exit from labour market functions as a balancing mechanism. As expected, there is almost no effect on the *Prevalence of Poor Mental Health* among men.

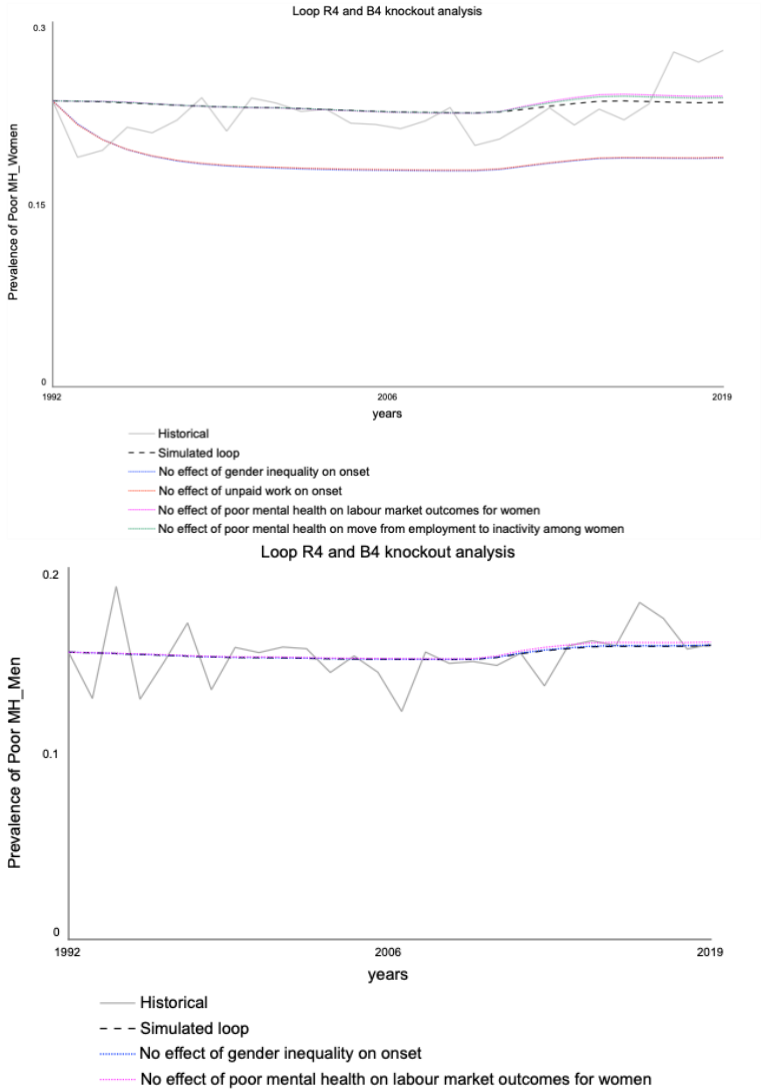


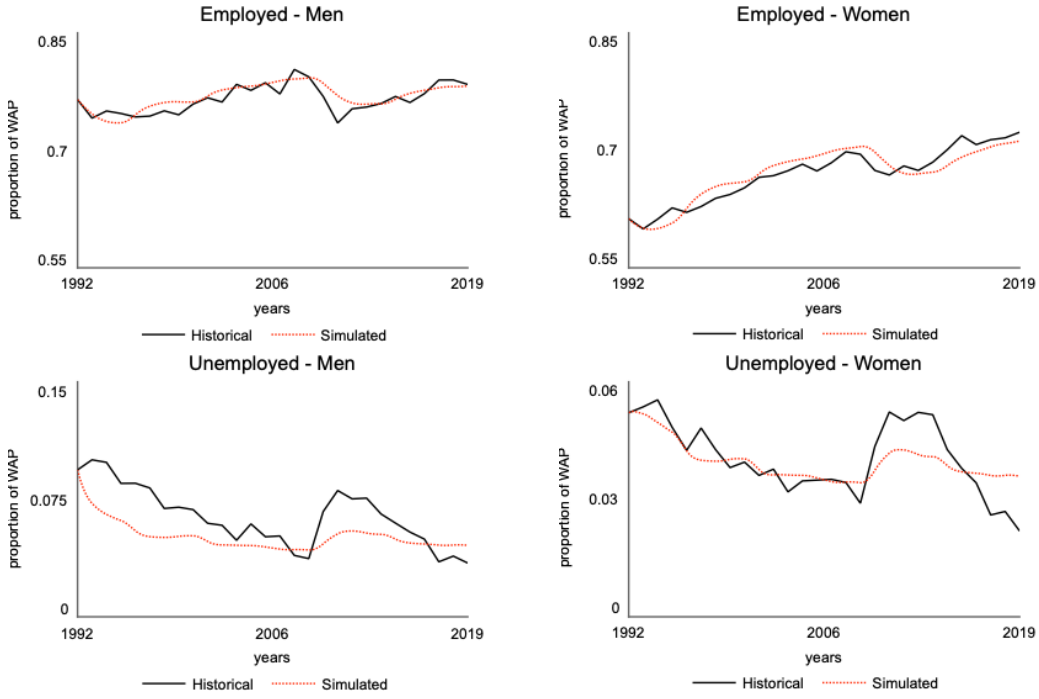
Figure 6-26: Loop knockout analysis testing the structural necessity of R4 and B4 (norms–mental health loops)

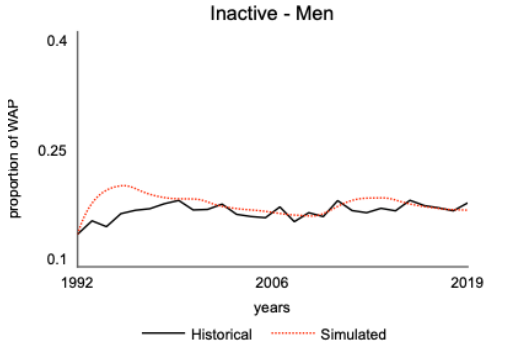
Removing the norms → poor mental health path (via unpaid work and the gender pay gap) lowers women's prevalence; severing the poor mental health → labour market path for women raises it slightly (labour market exit acts as a balancing outflow). Negligible effect for men.

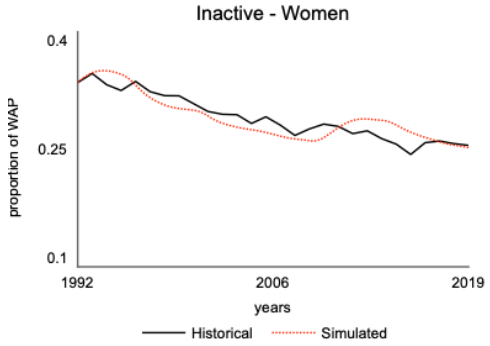
6.6.4 Behaviour reproduction

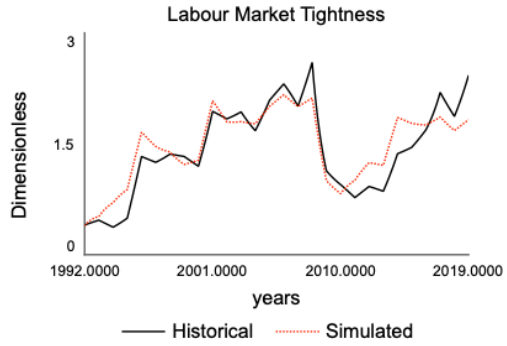
The integrated model reproduces the long-term structural trends with low bias but does not reproduce year-to-year volatility, which it is not designed to do. Specifically, the rise in *Prevalence of Poor Mental Health* in both genders, the persistence of the gender gap in prevalence, and the gradual liberalisation of gender norms are all reproduced. The model under-simulates the sharp rise in women's prevalence after 2016. This is consistent with the model's purpose (long-run structural mechanism rather than point prediction) and with its known boundary exclusions (for example, welfare reform is outside the boundary). The remainder of this section reports the underlying Theil's statistics and what they imply about the model boundary.

Results in this section are based on the final model after all changes discussed above had been made. The primary calibration period was strictly defined as 1992 to 2019. This is because 2020-2022 was a period characterised by a massive, external structural shock (COVID-19), not by the system's endogenous feedback dynamics. Figure 6-27 compares the historical time-series data with the simulated trends for all key model variables.




Inactive - Men
proportion of WAP
0.4
0.25
0.1
1992
2006
2019
years
Historical
Simulated

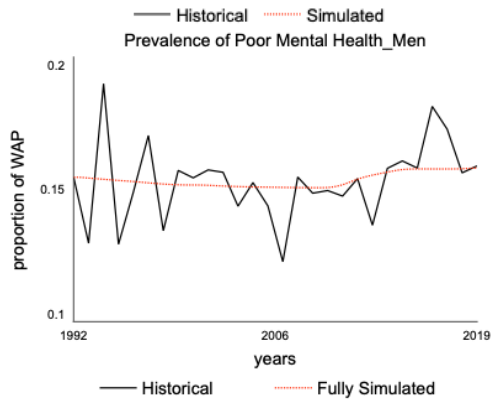

Inactive - Women
proportion of WAP
0.4
0.25
0.1
1992
2006
2019
years
Historical
Simulated

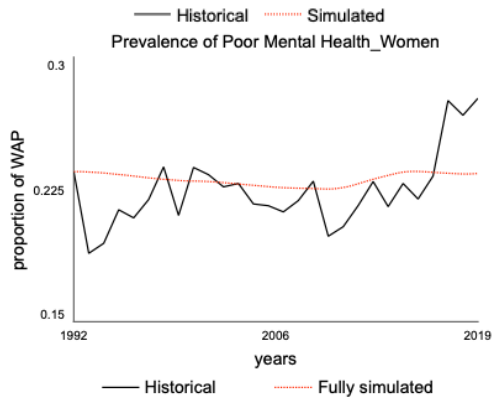

Labour Market Tightness
Dimensionless
3
1.5
0
1992.0000
2001.0000
2010.0000
2019.0000
years
Historical
Simulated

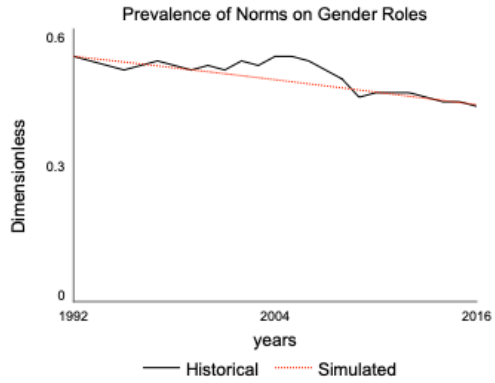

Insecure Work
Dimensionless
0.4
0.2
0
1992.0
1998.8
2005.5
2012.3
2019.0
years
Historical
Simulated

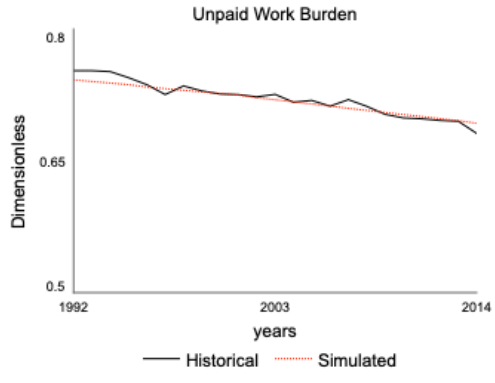

Insecure work - Men
0.45
0.225
0
1992.0000
2001.0000
2010.0000
2019.0000
years
Historical
Simulated


Insecure work - Women
0.45
0.225
0
1992.0000
2001.0000
2010.0000
2019.0000
years
Historical
Simulated


Prevalence of Poor Mental Health_Men
proportion of WAP
0.2
0.15
0.1
1992
2006
2019
years
Historical
Fully Simulated


Prevalence of Poor Mental Health_Women
proportion of WAP
0.3
0.225
0.15
1992
2006
2019
years
Historical
Fully simulated


Prevalence of Norms on Gender Roles
Dimensionless
0.6
0.3
0
1992
2004
2016
years
Historical
Simulated


Unpaid Work Burden
Dimensionless
0.8
0.65
0.5
1992
2003
2014
years
Historical
Simulated

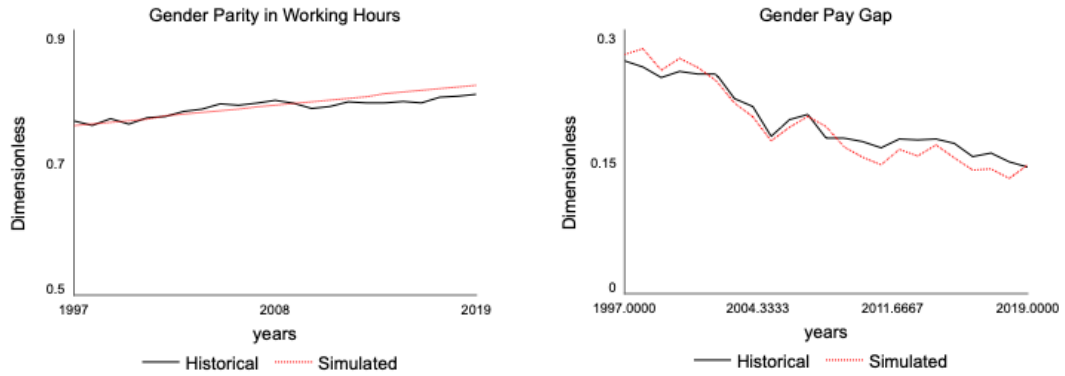


Figure 6-27: Behaviour reproduction graphs

Simulated trajectories (dotted red lines) against historical data (solid black lines) for all key variables, 1992–2019.

To rigorously evaluate the model's structural fidelity, behavioural reproduction was assessed using Theil's Inequality Statistics, following system dynamics best practices (Sterman, 2010). This provides a structured way to assess how a simulation deviates from observed data. Theil's statistics decompose model error into three components: bias (indicating the model consistently over- or under-shoots), variance (indicating the model mismatches the amplitude of fluctuations), and covariance (random noise the model should not be expected to capture). A structurally sound model should concentrate its error in the unsystematic component (covariance) and minimise the systematic components (bias and variance). The analysis below was conducted in two phases: first on the isolated sub-modules to verify independent mechanical accuracy, and subsequently on the fully integrated system to evaluate coupled feedback dynamics.

6.6.4.1 Isolated sub-module performance

When isolated from endogenous systemic feedback, the fundamental engines of the model demonstrate robust capability in capturing historical trends (Appendix L). The labour market and gender dynamics modules reliably capture historical means, with error dominated by covariation rather than bias. The Theil's statistics for the isolated labour market and gender dynamics modules and the fully integrated module are very similar.

Mental health module. For men, the error is well-structured: bias is very low (0.3%), indicating the model accurately tracks the historical mean, and the majority of error is concentrated in the variance component (73%) — the model

underestimates the magnitude of year-to-year fluctuations. The RMSE is 0.014 — meaning the model's predictions deviate from the historical data by approximately 1.4 percentage points on average, against a historical prevalence that ranges from roughly 14% to 18%. The model captures the structural trajectory while failing to match the amplitude of short-term variability, precisely the pattern expected from a model designed to reproduce long-run feedback dynamics rather than annual data points. This failure to match the amplitude of short-term survey variability is reflected in the low R-square.

For women, bias is higher than men but still low (9.4%), with variance accounting for 68% of the error. The RMSE is 0.021 — a deviation of approximately 2.1 percentage points against a range of roughly 20% to 28%. However, the pattern breaks sharply in 2017. Women's historical prevalence jumps from 0.233 (2016) to 0.275 (2017) — a single-year increase of 0.043 — and sustains this higher level through 2019 (0.267, 0.277). This historical deterioration observed for women from 2017 onward cannot be explained by amplifying the model's existing endogenous pathways, suggesting that forces outside the model boundary might be important. To test this, a 3-year temporary shock to baseline onset and recovery rates was examined. This captures the increase but suggests a return to equilibrium conditions over time.

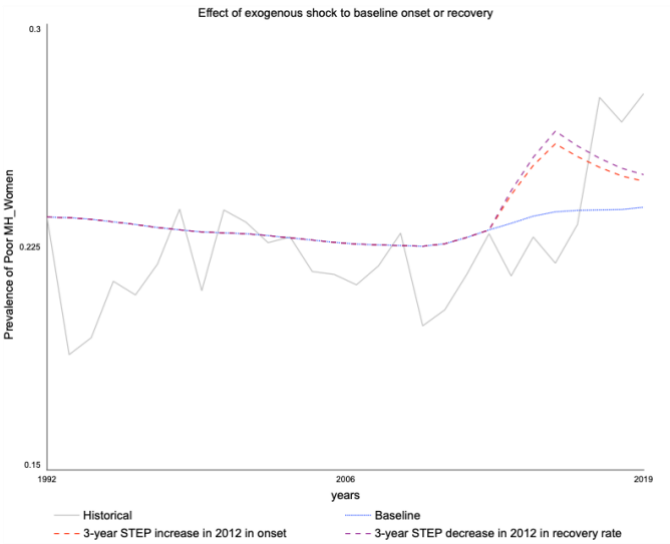


Figure 6-28: Effect of exogenous shock in onset and recovery

A temporary 3-year rise in baseline onset and recovery rates reproduces the post-2016 jump in women's prevalence but decays back to equilibrium.

The fully integrated model performs notably worse than the isolated mental health module alone (Figure 6-29).

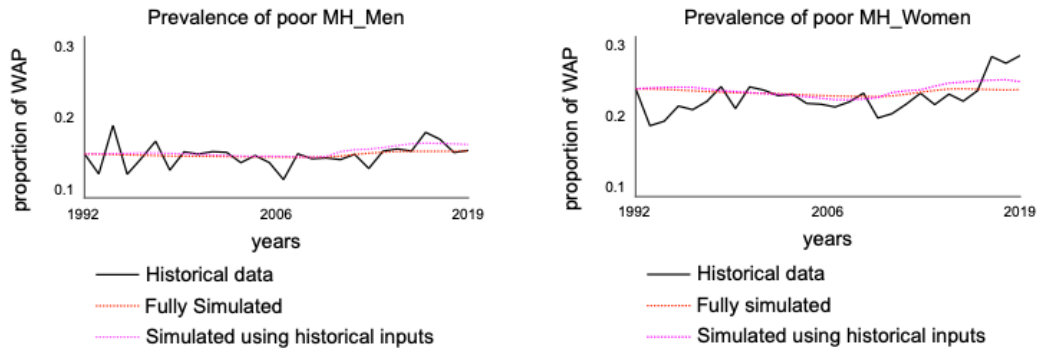


Figure 6-29: Prevalence of poor mental health simulated in the full model compared to the isolated module

The isolated module tracks the rise more closely; the integrated model runs below it because the endogenous labour market smooths the recession signal — a gap of ~25% of the total divergence from the historical data by 2019.

To diagnose the issue further, the Theil's statistics from three configurations were compared. In the fully isolated configuration, the mental health module receives historical values for all labour market and gender variables, isolating its internal dynamics from any errors generated upstream. The gender module's contribution to error is negligible; the gender-incorporated and fully isolated configurations produce virtually identical results. This comparison suggests that when given historical labour market inputs, the isolated module shows a clear inflection around 2009-2010, with prevalence rising thereafter. The model's internal feedback structure is responsive to worsening conditions. However, the slight optimism in the simulated labour market (compared to historical inputs) has cumulative consequences.

The integrated labour market module produces marginally more favourable employment conditions than the historical data; smoother recession response, slightly higher employment recovery. These small annual differences accumulate over the simulation period, resulting in lower stressor signals to the mental health module and marginally higher employment-based recovery support. The net effect: the integrated model suppresses the prevalence rise that the mental health module would otherwise produce. By 2019, the gap between the isolated and integrated simulations represents roughly 25% of the total divergence between the historical trajectory and the integrated simulation.

The remaining divergence is structural i.e. the mechanisms modelled do not, on their own, explain the full rise in the *Prevalence of Poor Mental Health*. The divergence between the model's trajectory and the historical data after 2016 identifies the boundary of the modelled mechanism. The 2017 increase in the historical data i.e. a single-year jump of approximately four times the standard deviation of annual changes represents a departure that the labour market-gender norms mechanism alone cannot generate. This points to forces operating outside the model's boundary.

6.6.4.2 Fully Integrated System Dynamics

When the modules are fully integrated, endogenous variables replace historical data inputs, allowing the full feedback architecture to operate autonomously. Table 6-7 demonstrates that the model successfully captures the fundamental structural trends, evidenced by low bias error. This confirms that the simulation replicates the long-term rise in poor mental health and the persistent gender gap.

Table 6-7: Behaviour reproduction statistics

Theil's inequality decomposition shows low bias across all variables and confirms the model reproduces the structural trajectory; low R^2 reflects year-to-year survey volatility the model does not aim to capture.

	R square	MAE/ Mean	RMSE	Bias	Variation	Covariation
Prevalence of Poor Mental Health[Men]	0.16	0.06	0.01	0.00	0.73	0.27
Prevalence of Poor Mental Health[Women]	0.14	0.07	0.02	0.09	0.68	0.22
Unemployed Proportion[Men]	0.59	0.23	0.02	0.48	0.22	0.30
Unemployed Proportion[Women]	0.63	0.11	0.01	0.02	0.48	0.51
Inactive Proportion[Men]	0.11	0.07	0.02	0.31	0.02	0.67
Inactive Proportion[Women]	0.78	0.04	0.01	0.00	0.00	1.00
Employed Proportion[Men]	0.64	0.01	0.01	0.10	0.04	0.86
Employed Proportion[Women]	0.86	0.02	0.01	0.01	0.01	0.98
Labour Market Tightness	0.84	0.14	0.26	0.02	0.24	0.74
Insecure Work	0.85	0.04	0.01	0.03	0.05	0.92
Insecure Work[Men]	0.84	0.04	0.01	0.01	0.13	0.87

Insecure Work [Women]	0.83	0.03	0.01	0.07	0.00	0.93
Prevalence of Norms on Gender Roles	0.69	0.03	0.02	0.21	0.03	0.75
Gender Parity in Working Hours	0.82	0.01	0.01	0.13	0.33	0.54
Unpaid Work Burden	0.94	0.01	0.01	0.11	0.46	0.43
Gender Pay Gap	0.97	0.06	0.01	0.16	0.47	0.37

6.7 Summary of model validation

A behavioural fit is necessary but not sufficient evidence for a structural hypothesis: a simulation can sometimes reproduce a historical trajectory under more than one parameter set (the calibration form of equifinality). The structural tests are the primary evidence that the fit reported here is generated by the hypothesised feedback architecture rather than by parameter tuning.

The persistence of the gender gap, despite measurable improvements in women's employment and the gender pay gap, is replicated by the model and provides the primary support for the dynamic hypothesis advanced in Chapter 5.

The integrated model successfully reproduces the long-run rise in poor mental health and the persistent gender gap. However, it under-simulates the magnitude of the post-2016 rise in women's prevalence, with 2019 simulated prevalence being approximately 15% lower than the observed. This residual divergence is structural rather than parametric: it reflects forces operating outside the model boundary. The four-year lag between austerity onset and the inflection in women's distress is consistent with cumulative policy exposure and the stock-and-flow structure of accumulation of prevalence. The model's quantitative estimates for the most recent years should therefore be interpreted as conservative.

In summary, the integrated model passes foundational verification, reproduces the historical trajectory of distress and gender gap with low bias, and the structural tests confirm that the dynamic hypothesis articulated in Chapter 5 is empirically defensible within the model boundary. The model's underestimation of the post-2016 surge in women's distress marks the limits of what the modelled mechanisms can explain.

6.8 Strengths and limitations

This section reflects on the strengths and limitations of the choices made during the translation of the CLD to the simulation model.

Transparency: A key strength of this model is the transparency of its iterative development process. The transition from a highly complex descriptive structure to a parsimonious simulation model was driven by systematic testing. By documenting how early structural tests informed final simplifications, such as adopting linear formulations when behaviour modes remained unchanged, the model's simplicity is empirically derived rather than assumed.

Parsimony: Building on this transparency, a primary strength of this model is its empirically grounded parsimony. By distilling complex constructs into transparent, mathematically bounded composites, the model avoids the trap of over-parameterisation and equifinality. The targeted use of constrained calibration for unobservable variables, coupled with rigorous sensitivity testing, ensures that the model's behaviour is a genuine emergent property of its feedback architecture.

Robust parameterisation: This research actively mitigates the risk of equifinality through a dual strategy. First, rather than relying on calibration, the majority of parameters are anchored in data and literature. Formal calibration was reserved primarily for unobservable variables, such as the decay of gender norms or specific recovery thresholds, and was constrained within theoretically plausible bounds. The only exception was the baseline exit rate, which was adjusted to account for the use of UK data and reflect gendered labour dynamics. Second, where parameters remain poorly identified — most notably the prevalence-recovery halving point and degradation shape governing the endogenous recovery rate — they are isolated as explicit structural uncertainties. Sensitivity analysis was used to confirm that the model’s core findings remain stable despite these uncertainties.

Temporal resolution: The simulation operates on an annual time step. This is appropriate to the multi-year feedback and accumulation processes the model is built to represent, but it means the model cannot resolve within-year dynamics

such as a sub-annual response to a job loss, or seasonal variation in labour market flows and it sets a floor on the temporal precision of any claim about the timing of feedback effects.

Conceptual–simulation divergence: Loops R2 (job quality decay) and R5 (gendered wage penalty) pass through labour productivity, which was subsequently excluded from the simulation model on parsimony grounds; the structural function of these loops is captured indirectly through labour market tightness. Similarly, the CLD includes two paths from *Gender Parity in Working Hours* to *Gender Pay Gap*: a direct link, and an indirect link via *Insecure Work*. The simulation retains the direct link and excludes the indirect one to prevent double-counting and empirical overfitting. R3b is therefore partially captured rather than excluded entirely; the simulation cannot adjudicate the relative weight of the two paths. On the contrary, in the simulation, reinforcement is driven by the economic signal composite (the *Gender Pay Gap*, *Gender Parity in Working Hours*, and the economic *Inactivity Gap*) and, in addition, by a direct dependence on the norms stock itself, through the resistance factor. The interpretation of R3's role in loop dominance analysis reported in Chapter 7 should be read with this operationalisation in mind.

Boundary exclusions: To achieve structural clarity, necessary boundary exclusions were made. A notable boundary choice in this model is the aggregation of the working-age population into a single 16–64 cohort, rather than employing an age-disaggregated structure. While this was necessary to maintain the parsimony required for a stable iterative development and testing process, it represents a limitation in capturing heterogeneous lifecycle dynamics. Evidence from the literature, such as Zhang et al. (2023), highlights that young adults (aged 16–24) have experienced significantly higher prevalence and distinct longitudinal trends of mental distress compared to older cohorts, particularly during periods of economic disruption. By treating the population as homogenous, the model may mask the specific vulnerabilities of younger workers who face increasing difficulties in transitioning into their first substantive employment. Furthermore, there is increasing policy interest in how labour market transitions vary at the extremes of the working-age spectrum, specifically the premature exit of older workers due to chronic health conditions versus the delayed entry of younger adults. More broadly, the aggregate stocks

cannot capture differential vulnerability across socioeconomic groups, ethnic groups, or geographic regions within Scotland. The feedback mechanisms identified operate at the population level; their applicability to specific subgroups requires further investigation.

The model normalises to the working-age population and represents women's unpaid work burden as a function of gender norms and the division of paid work, applied uniformly across all working-age women. This burden is operationalised as housework, however, and does not include caregiving. Two demographic shifts over 1992–2019 are therefore relevant. The average age at first birth rose, and population ageing raised the demand for informal elder care, which falls disproportionately on women. Because caregiving lies outside the model's operationalisation of unpaid work, any contribution these shifts made to the gendered trend in mental health is not represented by the mechanisms the model contains.

Similarly, a key finding from the iterative development and testing process was the weakness of the macroeconomic feedback loop. Tests for this reinforcing loop demonstrated that the system returns to its historical baseline after a shock, indicating a lack of loop dominance. This is a deliberate consequence of the model boundary, which treats baseline *Insecure Work* as exogenous time-series drivers. The model is fundamentally structured as a systemic reaction model, designed to simulate the amplification and volatility of the cycle caused by health and gender factors. The lack of a permanent stock in this loop, means the model cannot settle at a new, permanently shifted equilibrium. This is a necessary simplification to maintain model parsimony and focus, acknowledging that the broader macroeconomic context remains unmodelled.

The integrated model under-simulates the rise in women's prevalence after 2016. This represents the limits of the chosen model boundary and points to the role of other factors: austerity-era welfare reform, the staged Universal Credit rollout, and progressive NHS mental health service degradation are the most plausible candidates. The model's quantitative magnitude estimates for the most recent years should therefore be interpreted as conservative.

While the model successfully explains the long-term structural rise in prevalence of poor mental health and the persistent gender gap, it cannot, by design, predict the specific timing of short-term sampled fluctuations. For instance, the low R-squared values observed for the prevalence of poor mental health reflect this stochastic gap, which is expected given the year-to-year fluctuations in the historical data. Focusing on the temporal change across decades, the model intrinsically smooths real-world changes over shorter time periods.

Structural shielding: The reliance on linear formulations for the labour market and gender dynamics introduces a degree of structural shielding. Because the macroeconomic and societal engines respond proportionally and gradually (linearly) to changes, they are somewhat insulated from the sudden, cascading crises that occur when population mental health crosses a critical threshold. Consequently, the model likely yields conservative estimates regarding the ultimate severity of the feedback strength between deteriorating mental health and economic participation.

Operationalisation and parameterisation: Gender norms are operationalised as a single stock that erodes through demographic replacement over a 45-year timescale. This is a defensible representation of the cohort-replacement mechanism through which deeply held norms typically change, but it cannot capture cultural backlash, the influence of social media, or rapid norm shifts driven by social movements.

Several baseline transition rates (notably the labour market flow rates) are calculated from UK-wide samples and applied to the Scottish population. Where Scottish-specific rates differ systematically from UK-wide averages, model behaviour may be sensitive to this extrapolation. The constrained calibration of the female exit rate is a deliberate response to one such case; other parameters retain UK-wide values without re-calibration.

This spatial boundary choice (of testing a UK-parameterised model against a Scottish trajectory) also carries a policy dimension that the boundary justification acknowledges. Much of the economy and policy architecture relevant to the modelled mechanisms (employment law, the minimum wage, and most of the welfare system) operates at the reserved, UK-wide level; yet some

of the levers closest to the model's mechanisms (such as childcare provision) are devolved to Scotland. The model does not represent the interaction between devolved Scottish policy and reserved UK-wide dynamics, in particular the possibility that devolved interventions are overshadowed by UK-level forces; the policy implications of this devolved/reserved split are developed in Chapter 8.

The insecure work composite combines low pay, involuntary temporary employment, underemployment, and zero-hours contracts and the weights are theoretically motivated rather than empirically estimated, and alternative weightings might produce different magnitudes (though sensitivity analyses suggest the qualitative trajectory is robust). The composite cannot adjudicate between policies targeting income (e.g., minimum-wage uplifts) and policies targeting employment security (e.g., zero-hours-contract regulation).

The incidence composites combine multiplicatively: the baseline onset rate is scaled by the economic stressor and women's inequality multipliers. This is the standard epidemiological cumulative-risk form, and it is the form implied by the double burden and role-strain literatures, in which domestic and labour market demands interact rather than sum. It is, however, consequential: it is part of why a given labour market shock produces a larger systemic response for women than for men (Chapter 7). While the shapes of the individual composites were varied in sensitivity analysis, their multiplicative formulation was not tested against an additive formulation. The gendered loop dominance asymmetry should therefore be read as conditional on a theoretically-grounded multiplicative specification; a direct test against longitudinal data on whether women's labour market and domestic stressors combine additively or multiplicatively would help refine this.

7 Model Analysis

7.1 Introduction

Chapter 6 described the model, discussed the verification and validation tests employed and established the model as a useful tool to examine the historical persistence of the gender difference and rise in the prevalence of poor mental health. This chapter uses the model as a theory-building tool to analyse the causal hypothesis and draw insights about the observed trends as well as potential interventions.

It is important to emphasise at the outset that the model generates structural insights, not forecasts. The specific percentage contributions and gap reduction figures reported in this chapter describe the relative importance of mechanisms *within the model's representation of the system*. They should be read as evidence about which structural mechanisms are load-bearing and which are dormant.

Specifically, the model is used to answer the research questions outlined in Chapter 2:

First, **why does the gender gap in poor mental health persist?** The analysis indicates that men's mental health is dominated by recovery mechanisms, whereas women's mental health is dominated by incidence and the competing forces of reinforcing traps and balancing recovery. Structural decomposition reveals that the majority of this gap is driven by women’s differential social position (gender-specific incidence pressures), with the remainder driven by differences in recovery sensitivity, differential exposure and historical path dependence.

Second, **why did poor mental health rise for both men and women between 1992 and 2019?** The findings indicate that the mental health escalation was not solely caused by the immediate demand shock of the 2008 recession. Instead, the gradual rise of insecure work silently degraded the system's underlying recovery capacity. When the recession hit, this diminished capacity allowed the

unemployment mechanisms to overwhelm the system, triggering a vicious, self-perpetuating cycle.

Third, **what are the structural leverage points?** The analysis demonstrates that the system's policy levers are decoupled: interventions that relieve gender-specific incidence pressures act as equality levers required to close the gender gap, whereas interventions that improve job quality (reduce insecure work) act as population levers required to lower overall mental distress. These levers are not distributionally neutral, however: in low-precarity futures (S4, S8) the equality levers narrow the gap partly by raising men's prevalence, while reducing insecure work lowers the total but widens the gap.

The chapter applies four complementary analytical techniques to the simulation: loop-dominance, structural decomposition through what-if scenarios, counterfactual analysis of historical conditions, and experimental scenario runs on potential levers and optimisation across alternative futures. Each method asks a different question of the same model, and the confidence in these findings rests on their convergence. Table 7-1 maps each technique to the research question it addresses and the section in which it is developed.

Table 7-1: Summary of findings

Maps the research questions to their corresponding chapter sections, key findings and analytical methods.

Question	Key finding	Analytical Method	Section
Why does the gap persist?	Men’s mental health dominated by recovery; women’s by competing reinforcing traps and recovery.	Flow analysis; Loop dominance	7.2
	Gender gap: 61% additional stressors (women-specific pressures of gender pay gap and unpaid work burden acting on incidence), 8% differential recovery (67% when combined), 3% differential exposure (women’s heavier sorting into insecure work); 30% residual (path dependence in the 1992 starting position).	Structural decomposition	7.3
Why did the prevalence rise?	The crisis activated pre-existing vulnerability built through the gradual rise of insecure work.	Loop dominance	7.4

Question	Key finding	Analytical Method	Section
	No-recession ≈ 4% reduction in total prevalence; stable insecure work ≈ 7%; combined adds nothing beyond insecure work; Health selection is protective at system level.	Counterfactual analysis	7.5
What can be done?	Equality and job quality are complements; equality alone can redistribute the prevalence burden from women to men.	Counterfactual analysis	7.5
	Reducing insecure work lowers the total prevalence but widens the gap in low-precarity futures; Norms and gender pay gap do not add much more than childcare and insecure work reduction only.	Optimisation Experimental scenario runs	7.6

7.2 Examining gender differences

To examine why the gender gap persists, this section systematically isolates the forces driving the system's underlying architecture. First, Section 7.2.1 examines the mechanical accumulation of poor mental health, distinguishing between the flows of incidence and recovery for each gender. Section 7.2.2 uses loop-dominance analysis to diagnose which feedback loops are most powerful in each gender's subsystem and how the balance between reinforcing and balancing forces differs.

Three terms recur throughout this chapter. Incidence is the rate at which people develop poor mental health (the inflow); recovery is the rate at which they leave it (the outflow); and a reinforcing trap is a self-perpetuating cycle in which poor mental health and adverse labour market conditions each worsen the other. Recovery serves as a balancing force keeping the prevalence in check. These govern the gender difference. For men, prevalence of poor mental health is governed largely by recovery. For women, recovery is contested by reinforcing traps kept alive by pressures men do not face — the pay gap, the unpaid-work burden, and heavier exposure to insecure work. The difference between the genders is therefore not which mechanisms are present but the balance of power between the reinforcing and the balancing ones.

7.2.1 Incidence and recovery dynamics

Before isolating the generative feedback loops driving the Scottish gender paradox, it is necessary to examine the mechanical plumbing of the system and differentiate between inflow and outflow. By analysing historical net flows, this section finds that within the model's representation, the persistence of women's prevalence is primarily driven by a simulated failure to recover; a function of a systemic clogging of the drain, rather than merely a spike in onset.

The basic analysis of stocks requires looking at the inflows, outflows and net flow (incidence – recovery). As Figure 7-1 shows, net flow sits slightly below zero until 2008 for both men and women. After 2008, it rises sharply: for men, it stays positive to the end of the calibration period (in 2019), while for women, it briefly returns to negative before becoming positive again. The flow data reflect this: incidence is relatively flat for men compared with women until 2008, consistent with stable economic conditions but shifting gender inequality.

The incidence and recovery flows data also highlight that there was a larger reduction in recovery for women compared to men. This reflects the underlying model assumption – women’s recovery rate is more sensitive to the prevalence of poor mental health.

To isolate whether the subsequent accumulation in the stock is driven by a surge in onset or a failure to recover, the two flows are compared directly. Incidence does not rise to sustain the accumulation: men's incidence stabilises from around 2012, while women's continues to decline across the later period — yet women's prevalence does not fall, remaining elevated to the end of the calibration period. Because onset is easing rather than intensifying, the persistence of women's elevated prevalence cannot be an onset effect; it is sustained by a failure of recovery. The recovery flows for both genders, and for women especially, do not return to their pre-2008 levels, and once prevalence has risen the reduced recovery rate becomes self-perpetuating. The loop dominance analysis in Section 7.2.2 is consistent with this: recovery (drain) holds only 22% of the relative dominance for women against 59% for men. Section 7.3 shows the same sequencing, with gender-specific incidence pressures governing the early decades and recovery vulnerability activating later.

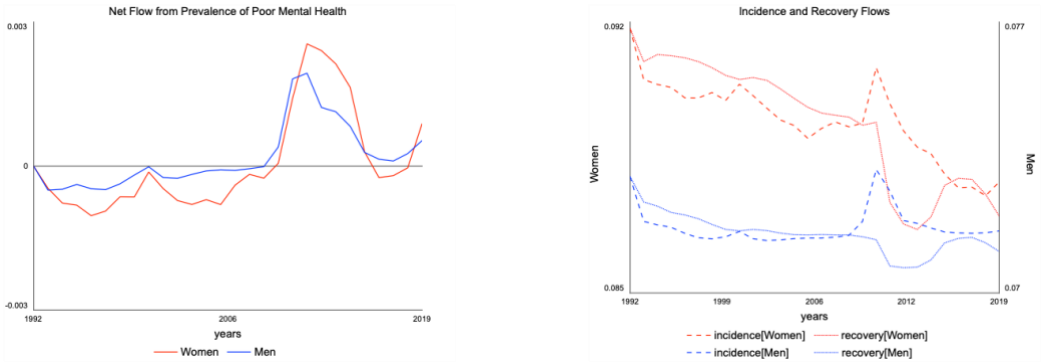


Figure 7-1: Net flow, incidence and recovery among men and women, 1992-2019

Incidence (inflow) and recovery (outflow) for each gender, and their difference (net flow). Net flow sits just below zero for both genders until 2008, then rises; women's recovery falls further and for longer than men's, which is what sustains women's elevated prevalence.

7.2.2 Loop dominance

This section reports results from loop dominance analysis that focus on loops that pass directly through the prevalence of poor mental health stocks to examine the gender difference; a broader unfiltered analysis covering the full system is presented in Section 7.4. The feedback loops referred to in this section and in Section 7.4 are summarised in Table 7-2. This analysis was conducted using the Stella Architect's Loops That Matter (LTM) algorithm. The full LTM methodology, including the methodological adaptations required for the gender-arrayed structure of this model, is documented in Text Box 1.

Text Box 1: Loops that Matter (LTM) Analysis Methodology

The 'Loops That Matter' (LTM) algorithm is an analytical tool that tracks causal momentum through a complex network over time (Schoenberg, Davidsen, & Eberlein, 2020). It interrogates the system at every unit of time to answer: which specific feedback mechanism is actively driving the observed changes in the system right now? It achieves this by identifying all active loops in the model and by calculating the variance transmitted along every link in the model (i.e. its behavioural impact). If a change in variable A consistently produces a major change in variable B, that pathway is assigned a high link variance score. By multiplying the variances of all the links that form a closed feedback loop, the algorithm calculates the total behavioural dominance of that loop. A loop that transmits large changes scores highly; a loop where the signal dissipates scores low. This process essentially identifies the underlying mechanisms that are forcing the system to behave the way it does at any given time, shifting the analytical focus from what is happening to why it is happening. However, it is important to highlight that a loop's dominance score reflects rate-of-change of variables, not their absolute levels. A variable that is high but stable contributes little; a variable that is rising rapidly contributes a great deal.

Approach

Loops with minor mathematical permutations were aggregated into loop families defined by the variable pair. Rather than collapsing the sub-loops into a single metric, their absolute dominance scores were summed independently by polarity. This yielded total reinforcing dominance and total balancing dominance for each loop. This approach quantifies not only the total causal burden transmitted through the pathway, but also the internal structural tension and directionality of that transmission.

Methodological caveats
(1) Magnitude bias: Because the algorithm's variance calculation is intrinsically tied to the absolute magnitude of the stocks, populations with a higher baseline volume artificially inflate the LTM scores. This is the case for the stock of prevalence of poor mental health where there is a gender difference: women's average prevalence is 0.23 compared to 0.15 among men (approximately 1.5 times). Raw cross-gender magnitude comparisons are therefore invalid.

To address this, all results are therefore reported as within-gender relative proportions: asking "what share of the forces acting on women's mental health does this loop represent?" rather than "how strong is this loop compared to the equivalent loop for men?" Cross-gender comparison is based on the ranked ordering and proportional structure of loop dominance, not on raw scores; absolute dominance was replaced by localised relative proportionality, calculated as the loop family's dominance divided by the total informational dominance of that specific gender's subsystem. Loops that are cross-gender were excluded from gender analysis because loops crossing the gender boundary absorb the mathematically larger variance of the female subsystem. Evaluating these against the much smaller baseline of the endogenous male subsystem artificially inflates their relative dominance for men.

(2) Material churn conflation: System dynamics models track both behavioural changes (informational signals) and basic accounting (e.g. proportional movements between *Employed, Unemployed,* and *Inactive* states). LTM assigns massive dominance to these high-volume demographic transitions, conflating structural plumbing with societal mechanisms and thereby masking true behavioural traps.

To isolate the mechanisms of interest, the analysis distinguishes between these. Informational loops are feedback mechanisms driven by causal mechanisms highlighted in the CLD; for example, unemployment worsening mental health, which then reduces employment prospects, which further worsens mental health. Material loops are accounting identities: people moving proportionally between employment states because they are part of a closed population (churn) or recovering from poor mental health at a rate proportional to the stock size (drain).

(3) Relative Measure: LTM percentages are relative dominance scores i.e. the loop's share as a percentage of the total. A loop's dominance score can rise either because its own activity grew or because competing loops weakened. Distinguishing these requires examining either the absolute link scores or model variables.

Table 7-2: Guide to the feedback loops analysed in this chapter

Summarises the reinforcing and balancing loops referred to in Sections 7.2 and 7.4.

Loop name	ID	Polarity	Mechanism
Unemployment–mental health (MH) trap	R1	Reinforcing	unemployment worsens MH, and poor MH in turn reduces employment
Insecure work–MH (job-quality decay)	R2	Reinforcing	insecure work worsens MH, which lowers employment and labour market tightness, feeding back into more insecure work
Norms entrenchment	R3	Reinforcing	traditional norms lower working hours parity and widen the gender pay gap, which feed back to reinforce norms
Norms–MH	R4 / B4	Reinforcing (+ balancing arm)	traditional norms raise the unpaid work burden, worsening women's MH and lowering their employment, which reinforces norms

Loop name	ID	Polarity	Mechanism
Prevalence–recovery degradation		Reinforcing	higher overall prevalence lowers the recovery rate, which raises prevalence further
Demand saturation	B1	Balancing	higher labour market tightness raises employment, which lowers tightness
Health-regulated job quality	B2	Balancing	the balancing arm of the insecure work–MH structure: worse MH lowers employment and tightness, easing insecure work
Gender power rebalance	B3	Balancing	traditional norms lower women's employment and tightness, reducing insecure work and the pay gap, which feeds back to norms
Recovery from poor MH ("drain")		Balancing	a larger poor MH stock means more people recovering (self-correcting outflow)
Churn		Reinforcing	proportional movement between stocks
other drain		Balancing	the draining of stock other than poor MH

7.2.2.1 Findings

This section is based on the relative loop dominance scores for each gender reported in Table 7-3.

Men's mental health is dominated by recovery; women's by competition between recovery and reinforcing traps. Within the model, men's prevalence is governed predominantly by material loops; the drain loop alone holds 59%. Whereas for women, no single mechanism dominates; women's prevalence shows substantial influence from active informational feedback loops.

This means men's mental health stock is primarily governed by its own natural outflow dynamics; the modelled system self-corrects and its dominant tendency is toward returning to baseline. However, the 2008–13 crisis period shows that even men's drain dominance can be overwhelmed by sufficiently strong external shocks (men's prevalence does rise). The feedback loops identified in the CLD

are structurally present for men but behaviourally subordinate to recovery under the historical conditions observed.

Women's recovery does not dominate the way it does for men; the drain (recovery) loop holds only 22% of relative dominance for women and is rivalled by the unemployment-MH reinforcing loop (22%). Other loops exert both reinforcing and balancing pressures. The fundamental asymmetry is not primarily about which loops are active; it is about the balance of power between reinforcing and balancing forces. Compounding pressures acting on women (the gender pay gap, the unpaid work burden, and heavier sorting into insecure work) keep women's reinforcing traps live enough to rival recovery, whereas for men the same loops produce only transient perturbations because recovery dominates.

Given there is a gender difference in the threshold for the effect of insecure work on recovery, further LTM analyses were conducted to test the sensitivity of the findings to this, by changing women's threshold to equal that of men's. The finding of gender asymmetry in the drain dominance was robust to this change. Appendix P provides the results for this sensitivity analysis run.

The 2008–13 crisis amplifies the gendered asymmetry rather than creating it. During 2008–13, women's informational reinforcing loops (unemployment-MH + insecure work-MH + norms-MH) sum to 52% of relative dominance. Men's equivalent sum is 5%. Within the model, the crisis does not create new feedback mechanisms; it amplifies existing ones selectively by gender. The unemployment-MH and insecure work-MH loops were present throughout the simulation, but for women they operated at low dominance before 2008 and surged during the crisis. For men, these loops remained structurally subordinate to drain throughout. The crisis period exposes the system's latent gender asymmetry: the balance of power between reinforcing and balancing forces shifts for women under economic stress while remaining stable for men.

The unemployment-MH reinforcing loop is the dominant informational loop for women. Among men, it has only a 3% relative dominance, active only 5 years, whereas for women, it has 22% relative dominance, active all 27 years, peaking in 2011 (during the crisis). During the period of greatest system change (2008–

13), this loop escalates from 5% to 35% for women while remaining at 5% for men. This gendered divergence appears counterintuitive. As discussed in chapter 2, much of the literature finds unemployment particularly damaging for men in societies where breadwinner identity is central. The model structure provides a distinct interpretation: the unemployment-MH loop exhibits higher relative dominance for women because interacting factors are operationalised to compound multiplicatively. Women's unemployment occurs in the context of additional structural pressures. In the model, the same unemployment shock propagates through a richer set of feedback pathways for women, producing larger systemic effects even if the direct psychological impact of unemployment per se were identical.

Norms-MH loops are persistently active for women but individually weak. The norms-MH reinforcing and balancing loops each contribute only 2% overall (4% during 2008–13). They are active for 27 years but never achieve high dominance. Four interpretations are possible and the model cannot distinguish between them. First, modest effect: norms may genuinely operate as a persistent background drag rather than a dominant driver: their effect is real but structurally modest relative to economic mechanisms. Second, indirect effect: norms may exert their primary influence indirectly, transmitting through other loops rather than acting directly on mental health: norms increase insecure work sorting, which feeds into the insecure work-MH loop that does show high dominance. If this is the case, the 2% direct contribution understates the total causal role of norms. Third, model limitation: the model's operationalisation of norms may be too simple to capture the full mechanism; mental health is not modelled as a direct driver of norms. Fourth, analysis limitation: LTM analysis reflects rates of change, so a variable like norms that is a high but stable variable would register low dominance regardless of its true importance.

Table 7-3: Gender difference in loop dominance

Reports the localised relative loop dominance scores for each gender. It quantifies what proportion of the total causal forces acting on mental health is attributable to a specific feedback loop. It reveals that men's mental health is dominated by recovery (drain) whereas women's by reinforcing loops.

Loop Family	Component	Men	Women	2008-13 [M]	2008-13 [W]	Yrs Active [M]	Yrs Active [W]	Yrs Active 2008-10 [M]	Yrs Active 2008-10 [W]
Unemployment-MH	Reinforcing	3%	22%	5%	35%	5	27	2	3
Unemployment-MH	Balancing	0%	4%	0%	7%	0	27	0	3
Insecure work-MH	Reinforcing	4%	7%	0%	13%	27	27	3	3
Insecure work-MH	Balancing	3%	13%	0%	0%	12	11	1	0
Norms-MH	Reinforcing	0%	2%	0%	4%	0	27	0	27
Norms-MH	Balancing	0%	2%	0%	4%	0	27	0	27
Prevalence-recovery	Reinforcing	4%	1%	4%	1%	7	7	0	0
Prevalence-recovery	Balancing	2%	1%	2%	1%	7	7	0	0
Total of informational loops		**16%**	**52%**						
Churn	Reinforcing	25%	25%				27	3	3
Drain	Balancing	59%	22%				27	3	3

- **Localised relative dominance scores (Rel Dom):** each value shows what proportion of the total causal forces acting on one gender's prevalence of poor mental health is attributable to that loop family. A drain dominance of 59% for men [M] means that 59% of all the causal influence on men's poor mental health stock comes from the recovery outflow. These percentages sum to 100% within each gender (including drain and churn). Critically, 59% for men and 22% for women cannot be directly compared as magnitudes; the normalisation is within each gender's own subsystem. What can be compared is the structural architecture: how the balance of power between reinforcing and balancing forces differs.
- **Excluded loops:** Two loops passing through both men's and women's prevalence stocks were excluded from both genders' analyses, as they inflated men's relative dominance (norms-MH reinforcing was artificially reported at 23%).
- **Magnitude bias:** Raw LTM absolute dominance scores are not comparable across genders due to women's larger base stock. Within-gender relative dominance and ordinal ranking are used instead.
- **Filtering:** This analysis reports on a subset of loops that are relevant to the prevalence of poor mental health.
- **Note on insecure work-MH:** This loop's high score is inflated by demographic accounting flows (churn) that the loop pathway passes through. The dominance score reflects the volume of demographic transitions rather than a genuine behavioural feedback mechanism. It is retained in the table for transparency but excluded from the narrative interpretation. Net of this churn-inflated cell, the informational totals are approximately 13% (men) and 39% (women); the direction and scale of the gender asymmetry are unchanged.

7.2.2.2 Linking LTM findings to the CLD

This subsection links the loop-dominance findings to the loops named in the CLD in Chapter 5.

R1 (the unemployment-MH trap) is identified as the dominant informational loop for both genders, but with dramatically different magnitudes: 22% relative dominance for women (active all 27 years) versus 3% for men (active 5 years). R2 (the job quality decay) is partially confirmed through the insecure work-MH loops, though the productivity channel hypothesised in the CLD was excluded from the simulation model to maintain parsimony, so the full R2 loop cannot be traced. R3 (Norms Entrenchment) operates through the gender dynamics sub-model but does not pass directly through the poor mental health stock, so it does not appear in this filtered analysis; its contribution is examined in Section 7.4. The norms-MH loops capture the norms to mental health pathway, contributing 2% each; structurally present but behaviourally weak as direct pathways, consistent with the interpretation that norms operate primarily through indirect transmission.

The LTM analysis results support the CLD's core hypothesis: men and women are embedded in the same macroeconomic environment, but women do not simply experience that shared system with different settings. Their subsystem adds a structural layer of gender-specific stressors (the gender pay gap and the unpaid work burden) that men do not face, channels heavier sorting into insecure work that affects both onset and recovery, and applies some different table functions to the shared loops (notably the effect of prevalence on recovery). It is, in this sense, a genuinely different system for women; which is why the same macroeconomic conditions produce different dynamics by gender.

7.2.3 Summary

The mechanical flow analysis and the loop dominance analysis expose the same gender asymmetry: the mechanical view highlights that the later period higher prevalence is due recovery failure; the loop dominance view sees women's recovery loop holding only 22% of relative dominance against men’s 59%, rivalled by reinforcing traps that hold 22% on their own. The implication is that the

persistent gender gap is consistent with an architecture where balancing forces are weaker for women.

7.3 Structural decomposition of the gender gap

This section conducts a structural decomposition to isolate where the gender gap in poor mental health originates within the model structure. The analysis proceeds through a series of "what-if" experiments. Each test modifies one aspect of the gender difference built in the model's structure and re-runs the simulation over the full 1992–2019 period. By comparing the resulting gender gap against the base case, the contribution of each structural feature is isolated. The gender gap in prevalence in 2019 (women's prevalence minus men's) provides a comparable summary figure. The base case produces a gender gap of 0.077 in 2019 (women: 0.236, men: 0.159).

This section decomposes the 0.077 gender gap in 2019 prevalence as a sequence of nested runs (decompositions D1–D4). Women's greater recovery sensitivity (D1) accounts for 8% whereas gender-specific incidence pressures (D2) account for 61%; because the two are slightly sub-additive, their standalone sum (69%) is marginally larger than their combined effect (67%, D3). Equalising the implied baseline recovery rate (D4) — the path-dependence residual — then closes a further 30%, leaving 3% unexplained, consistent with differential exposure. Because the implied recovery rate already absorbs women's baseline exposure, the 30% and the 3% are not cleanly separable; this final split is indicative rather than exact.

7.3.1 What if women had the same effect parameters as men? (decomposition D1)

By construction, the model applies an identical function for the effect of economic stressor on onset to both genders; gender differences enter through recovery composite. This test asks how much of the gap comes from women's greater sensitivity to the same pressures, and is therefore operationalised through the recovery pathway; the two table functions governing women's recovery are replaced with their male equivalents:

- **Effect of insecure work on recovery**: The table function for women, which reaches its minimum degradation at a ratio of 1.2 (current-to-initial insecure work), is replaced with the table function for men, which reaches the same minimum at a ratio of 1.4. This means women now tolerate a larger increase in insecure work before their recovery capacity is fully degraded.

- **Effect of prevalence on recovery rate:** Women's linear function, where each unit increase in prevalence produces a proportional decline in recovery capacity, is replaced with men's concave function, where recovery capacity remains resilient until prevalence reaches a high threshold before declining meaningfully.

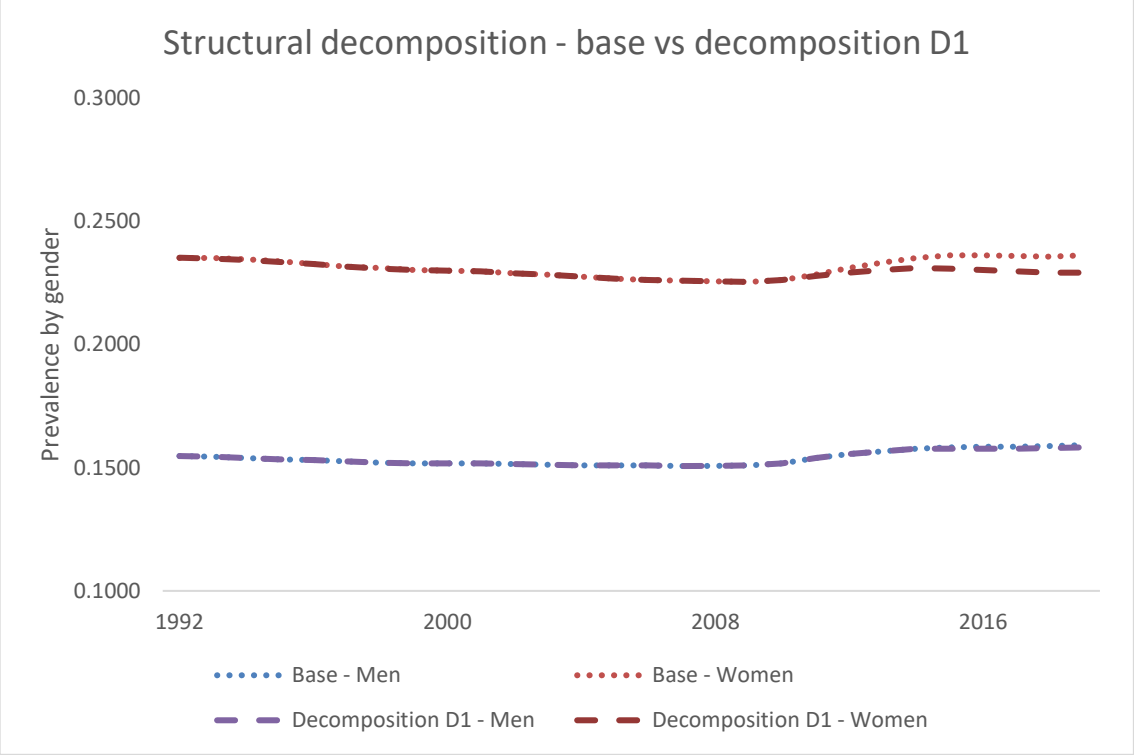


Figure 7-2: Effect of differential gender sensitivity on the gender gap in the prevalence of poor mental health

Visualises mental health trajectories when equalising recovery sensitivity parameters between genders. It shows that changing the sensitivity parameters alone closes only a modest portion of the gap (8%).

Equalising recovery sensitivity parameters closes 8% of the gender gap. This is a modest contribution given the recovery asymmetry (concave vs linear) and the large difference in drain loop's relative dominance highlighted in Section 7.2. However, this does not mean the recovery asymmetry is unimportant. Rather, it

reflects how the sensitivity parameters interact with the system's starting conditions. In 1992, the model initialises in equilibrium. The baseline recovery rate (calibrated to achieve this equilibrium) already compensates for women's greater sensitivity. This experiment shows that, given the model's equilibrium initialisation, changing the *shape* of the sensitivity function alone does not substantially drive the *changing* of the gap over the simulation period.

7.3.2 What if women did not face additional stressors? (decomposition D2)

The gender inequality effect captures the direct impact of structural inequality (the gender pay gap, the unpaid work burden) on women's rate of developing poor mental health. This test asks: how much of the gender gap is attributable to this direct incidence pathway? Gender inequality is set to zero from 1995 onwards, resulting in the effect of gender inequality on incidence being 1 (no effect). The three-year lag from the simulation start (1992) allows the model to settle from initialisation before the intervention takes effect.

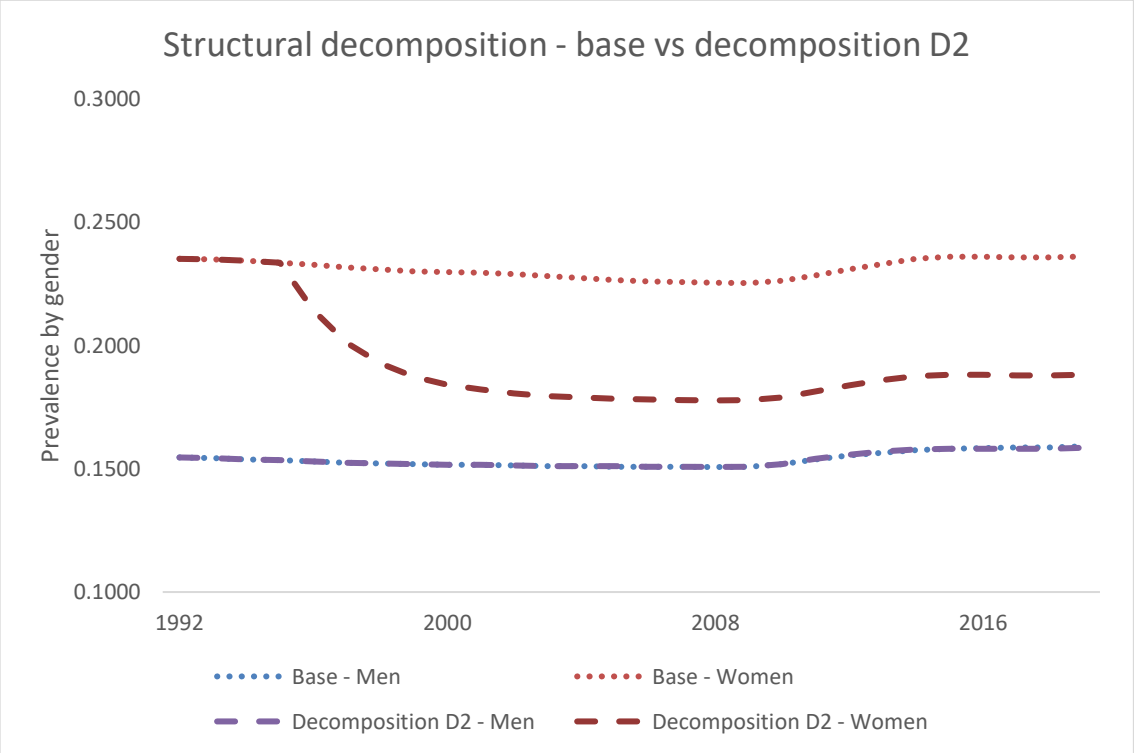


Figure 7-3: Effect of gender-specific stressors on the gender gap in prevalence of poor mental health

Visualises mental health trajectories when removing additional women-specific incidence pressures (from 1995). It shows that when the effect of gender inequality on mental health is removed, the gender prevalence gap closes substantially (61%).

Women's prevalence reduces significantly (0.236 to 0.188). Men's prevalence is essentially unchanged (0.159 to 0.158). Because the model channels all gender-specific incidence pressures through a single inequality variable, removing the variable tests their collective contribution rather than attributing the effect to any single dimension of inequality. Under this test, removing the gender inequality effect closes 61% of the gap, the largest contributor identified in the decomposition.

The gender inequality effect is the *only* gender-specific variable on the incidence side of the model. In back-calculating the implied baseline rate as common for men and women, all gender differences in incidence are channelled through this single multiplier. In a more disaggregated model with multiple gender-specific incidence pathways, this 61% would be distributed across them. The large contribution therefore reflects both the importance of gender inequality as an incidence driver *and* the structural parsimony of the model, which concentrates all gender-specific incidence effects into one variable. Figure 7-4 illustrates this dynamic, showing the extent to which historical narrowing of gender inequality has already reduced women's prevalence of poor mental health below what it would otherwise have been by visualising a counterfactual.

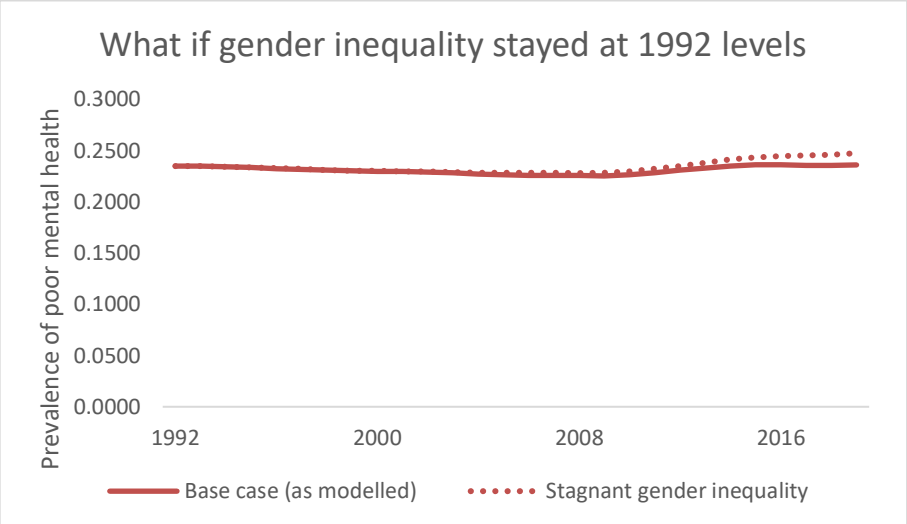


Figure 7-4: Prevalence of poor mental health for women: base case vs a counterfactual in which gender inequality does not narrow over 1992–2019

Charts a counterfactual scenario demonstrating how historical progress in gender equality structurally shielded women; Withholding the historical narrowing of the gender pay gap and the erosion of norms leaves women's prevalence rising to higher peaks, showing that the historical improvement in gender equality already reduced women's prevalence below what it would otherwise have been.

7.3.3 What if women had the same parameters as men and did not face additional stressors? (decomposition D3)

To examine the interaction of the two mechanisms, the above experiments are run simultaneously - the gender inequality effect on incidence is removed from 1995, and women's recovery sensitivity parameters are replaced with men's. Mechanisms can compound when combined producing a combined effect greater than its parts, or mechanisms may overlap, producing a combined effect smaller than the sum of parts.

The combined intervention closes 67% of the gap, slightly less than the sum of the individual contributions (61% + 8% = 69%). Within the model's structure, the incidence and recovery pathways, therefore overlap only marginally: the gender inequality effect governs the inflow of women into poor mental health and the recovery sensitivity governs how quickly they can exit. Removing one does not meaningfully change the dynamics of the other.

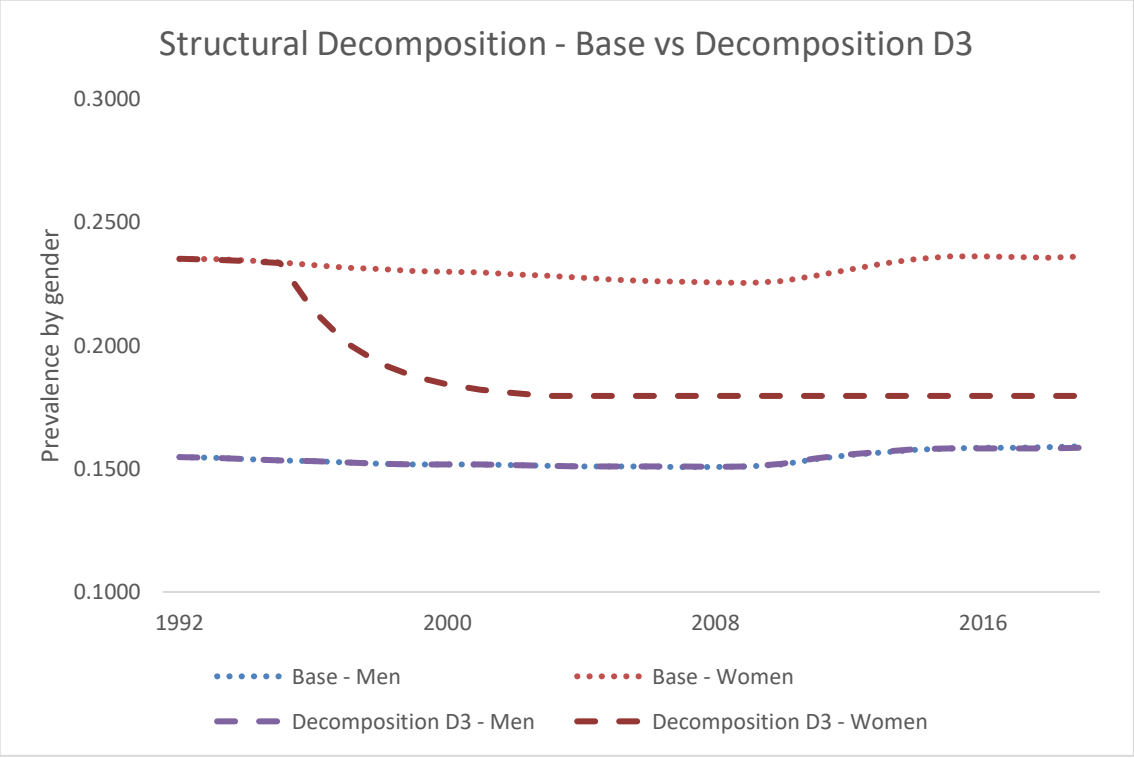


Figure 7-5: Overall effect of differential gender effects and additional stressors on the gender gap in prevalence of poor mental health

Shows the combined impact on prevalence trajectories when equalising recovery sensitivity and removing incidence stressors (D1 + D2) closes 67% of the 2019 gender gap.

The near-additivity at the 2019 endpoint (67% combined vs 69% summed) is confirmed by the full time-series trajectories which show that the two pathways are not merely non-interacting but temporally sequenced. Decomposition D2 (removing the gender inequality effect on incidence) produces an immediate decline in women's prevalence from 1995, stabilising around 2000 and then reflecting the slight post-2012 rise visible in the base case. Decomposition D1 (equalising recovery sensitivity) tracks the base case closely until approximately 2012, then diverges; women's prevalence remains flat rather than continuing to rise. The incidence pathway operates early, reducing the inflow of women into poor mental health from the moment the intervention takes effect. The recovery pathway only becomes visible late in the simulation, once prevalence has risen enough during the post-2008 period to push women's recovery function into the range where the linear-versus-concave distinction has a material effect. The two pathways produce additive effects at the endpoint because they operate in sequence rather than in competition: the incidence pathway governs the first two decades, the recovery pathway governs the last. This temporal sequencing suggests that interventions targeting gender-specific incidence pressures would produce earlier benefits, while interventions addressing recovery capacity would become important only after prevalence has already risen. Figure 7-6 visualises this sequencing: structural inequality drives women into poor mental health early in the timeline, while degraded recovery capacity traps them there during later periods of economic stress.

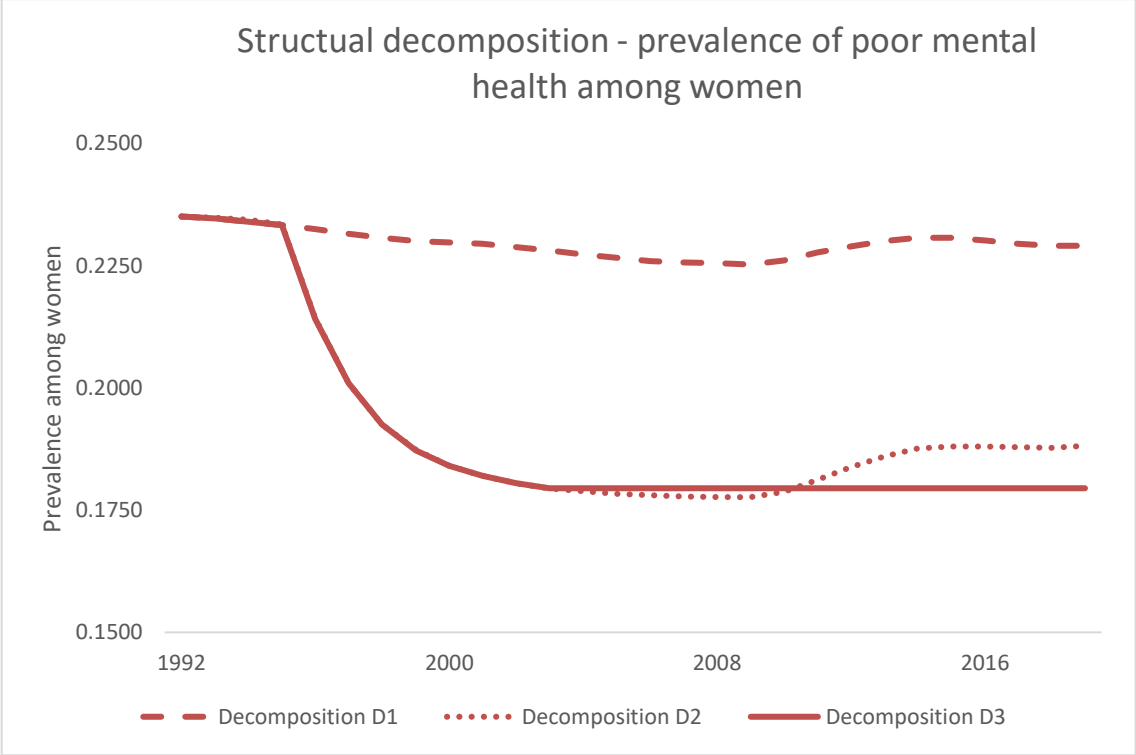


Figure 7-6: Women's prevalence of poor mental health across decompositions D1-D3

Charts women's mental health trajectories across scenarios to visually highlight temporal sequencing, show that the two mechanisms (equalising recovery D1 and removing additional stressors D2) operate independently and additively and that incidence stressors dominate the early decades while recovery vulnerability activates later.

7.3.4 What is the estimated contribution of differences in exposure? (decomposition D4)

After simultaneously removing the gender inequality effect on incidence and equalising recovery sensitivity parameters (D3), a gap of 0.025 or approximately a third of the original gap remains. Decomposition D4 equalises the implied baseline recovery rate (women given men's rate from 1995) brings this gap almost to zero (0.002). It reflects the entangled contribution of two factors:

1. **Path dependence and calibration.** Women enter the simulation in 1992 with a higher prevalence of poor mental health than men, reflecting accumulated historical disadvantage prior to the simulation period. The implied baseline recovery rate is lower for women (0.348) than for men (0.410). This rate is not an independently estimated parameter: it is a calibration residual that absorbs women's additional incidence pressures and lower recovery multipliers at the point of initialisation. In other words, it is set to whatever value balances every structural pressure women already faced; including the accumulated disadvantage that

predates the simulation and any factors the model does not represent explicitly. When men's and women's starting conditions are equalised directly, the gap almost vanishes, confirming that this residual is path dependence embedded in the starting position. Because the implied recovery rate is held constant, the model cannot distinguish disadvantage accumulated before 1992 from pressures that continue at a stable level outside the boundary. These scenarios are reported in Appendix Q and find that the calibration residual contributes 30% of the gap.

2. **Residual differential exposure.** The ~3% that remains after D3 and D4 (Base − D3 − D4) captures women's disproportionate sorting into insecure work: even after equalising sensitivity and removing the inequality multiplier, the same effect multiplier applied to women's higher exposure produces a marginally larger effect. Because D4 already absorbs baseline exposure, this remainder is not cleanly separable from the 30% and is reported as indicative rather than as a distinct run.

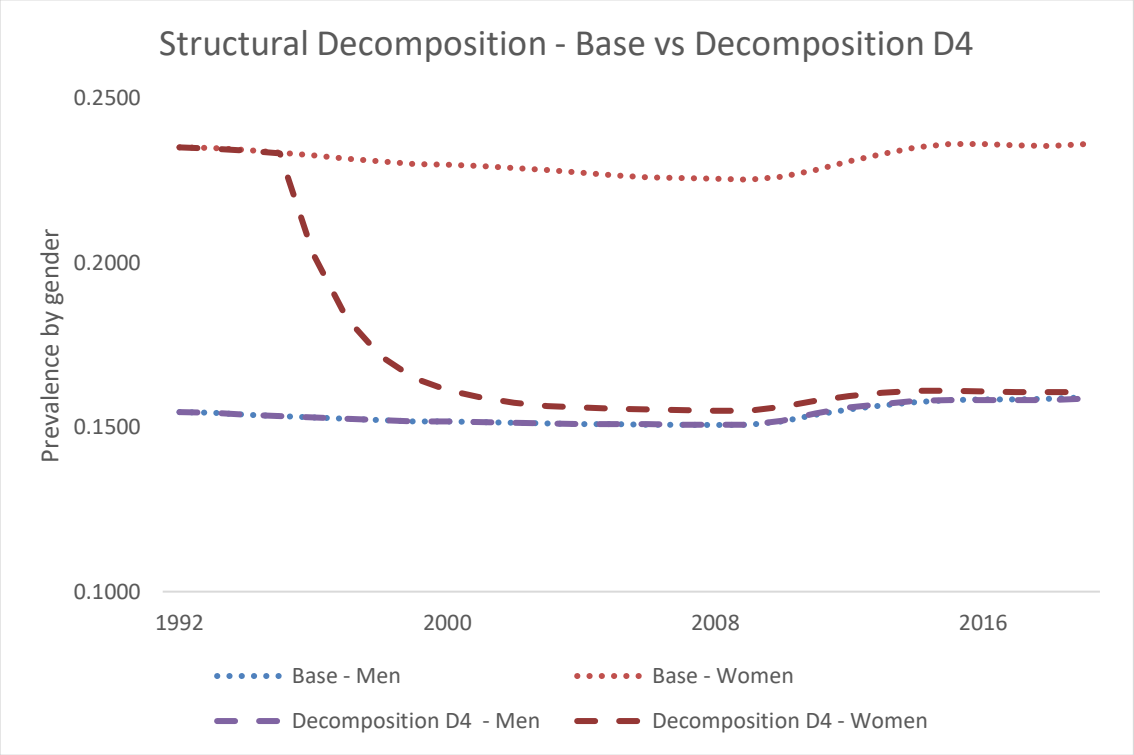


Figure 7-7: Residual gap attributable to differential exposure on the gender gap in prevalence of poor mental health

Illustrates the small residual gap that persists purely due to differential labour market exposure to risk

The 3% residual should not be misread as evidence that women's differential exposure is unimportant. The earlier runs have already removed some effects of that exposure: Decomposition D2 removed the gender inequality multiplier on incidence (which amplifies the impact of unemployment and insecure work), and Decomposition D1 equalised the recovery sensitivity that makes women's recovery more vulnerable to insecure work conditions. What remains after these removals is the residual sorting effect: the small additional gap attributable to women being exposed to higher levels of risk factors even when the gendered amplification of that exposure has been neutralised.

7.3.5 Robustness of the decomposition

Three robustness checks support the decomposition. First, the gender asymmetry in the effect of prevalence on recovery — operationalised in the base case as a concave function for men and a linear function for women — was tested across four alternative specifications spanning the full range from linear to concave for both genders. The 2019 gender gap varied by less than 0.0011 (under 1.5% of the base case gap), confirming that the decomposition does not depend on the precise functional form of this mechanism. Second, the 30% residual finding is corroborated by an equilibrium-recalibration test in which men's and women's starting stocks are equalised; the gender gap disappears in both directions (≈0.002 and ≈0.006), strengthening the path-dependence reading. Third, the gender direction encoded in the implied recovery rate (men 0.410, women 0.348; ratio 1.18) is supported by data-adjusted recovery rates calculated from Scottish Health Survey transitions (ratio 1.11–1.12), confirming that the asymmetry is empirically grounded and not a calibration artefact. Full results for all three checks are reported in Appendix Q. These results operate partly through the endogenous recovery mechanism, whose halving point (prevalence threshold = 1.5) Chapter 6 identifies as a critical and not uniquely identified parameter; the recovery channel magnitudes here should be read as conditional on that value.

7.3.6 Summary

The structural decomposition analysis isolates the specific generative mechanisms driving the simulated gender gap in poor mental health. The

decomposition delivers two findings the loop dominance view alone could not. First, the gap divides into four components: 61% gender-specific incidence pressures, 8% recovery sensitivity (together comprises 67%), 3% differential exposure, and a 30% path-dependent calibration residual (Table 7-4). Second, and less visible from the percentages, the incidence and recovery channels are not simultaneous but temporally sequenced: the incidence pathway governs women's prevalence trajectory through the first two decades; the recovery pathway only becomes material after 2008, once aggregate prevalence rises into the region where the linear-versus-concave distinction comes into play.

Table 7-4: Decomposition of the gender gap in poor mental health

Summarises the percentage breakdown of the structural components driving the gender gap in poor mental health in 2019. Decomposition D1 and D2 together (D3) close 67% of the gap, their standalone contributions sum to 69% because the incidence (D2) and recovery pathways (D1) overlap slightly, so their combined effect is a little less than the sum and this leads to a total contribution greater than 100%. Equalising the implied recovery rate (D4) closes a further 30%, leaving a 3% remainder consistent with residual differential exposure.

Component	What it captures	Contribution %
Differential vulnerability (D1)	Gender difference in recovery sensitivity parameters	8%
Differential social position (D2)	All gender-specific incidence pressures, channelled through the model's single inequality multiplier	61%
Differential exposure (Base – D3 – D4)	Women's disproportionate sorting into insecure work; differences in unemployment	3%
Residual (D4)	Path dependence, unmodelled factors, and starting conditions	30%

Table 7-5 reports the scenario-by-scenario reductions in the gap underlying these proportions. The decomposition provides critical context for interpreting the loop-dominance analysis. The high relative dominance of the unemployment-MH and insecure work-MH reinforcing loops for women during 2008–13 should not be read as evidence that economic shocks alone are doing the work. Within the model, those acute shocks compound multiplicatively with the chronic incidence pressure captured by the 61% component; women's reactivity in the LTM analysis is a symptom of that chronic pressure.

Table 7-5: Summary of structural decomposition runs (D1 to D4)

Details the resulting prevalence for men and women, the new gender gap, and the percentage reduction in the gap to quantify how much each structural feature contributes to the overall gender gap.

Decomposition	Reduction %	Gap reduction	Gap	Prevalence (Women)	Prevalence (Men)
Base			0.0770	0.2361	0.1591
D1	~8%	0.0061	0.0709	0.2291	0.1582
D2	~61%	0.0473	0.0297	0.1882	0.1585
D3 (D1 + D2)	~67%	0.0517	0.0253	0.1838	0.1585
D4	~30%	0.0234 (vs D3)	0.0019	0.1607	0.1588
Base – D3 – D4	~3%				

7.4 Explaining the historical trends in poor mental health

The loop dominance analysis in Section 7.2 examined which feedback mechanisms govern the prevalence of poor mental health, using a filtered analysis that included only loops passing through the mental health stocks. This section broadens the analytical lens to the full system, including feedback mechanisms that operate within the labour market and gender dynamics sub-models without directly passing through poor mental health. The purpose is to identify when and why the structural conditions for the mental health crisis developed. Where the filtered analysis in Section 7.2 asked "what governs mental health?", this analysis asks "what governs the system?"

Loop dominance analysis traces the rise through three-phases: a quiet build-up in which structural vulnerability accumulated largely unseen, the crisis years when it discharged and a changed system afterwards. It finds that the recession activated dormant loops rather than create new ones.

7.4.1 A three-phase chronology

Figure 7-8 shows the result: the relative dominance of each loop family across 1993–2019. Most years are quiet and the system is dominated by routine movement of people between employment states and mental health states (churn and drain). Against that steady baseline there are four short periods when

a feedback loop briefly takes over: 1999, 2006, 2010–11 and 2017. At each of these, the specific drain for poor mental health i.e. natural recovery is almost squeezed to zero. The recovery loop Section 7.2 calls 'drain' appears here, in the full-system view, as the separately labelled 'drain-recovery'; the larger 'drain (other)' band is the draining of the model's other stocks.

The chronology is described below in three key phases, delimited by loop dominance. The crisis comprises the years in which a mental health reinforcing loop is visible (2009-11); 2008, when no such loop is visible, falls in the preceding build-up. Year-level mechanism detail that supports these is reported in Text Box 2, and the full LTM dominance time series is in Appendix P.

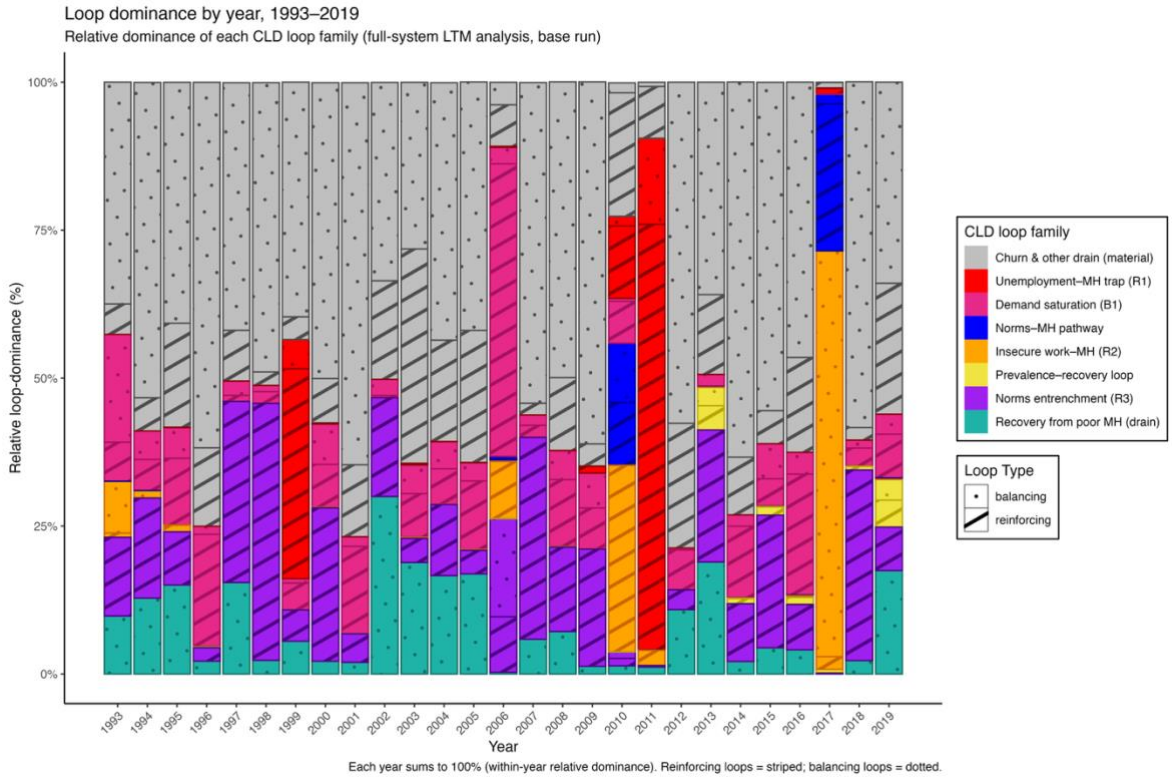


Figure 7-8: Loop dominance by year

Maps the changing relative dominance of the system's underlying feedback loops across 1993-2019, differentiating between reinforcing and balancing mechanism of the loop sets. The bands are ordered to keep a stable baseline: the recovery band — drain-recovery, the outflow by which people recover from poor mental health — is anchored at the bottom; the reinforcing informational loops are stacked above it from least to most volatile (the unemployment–MH trap R1, the insecure work–MH loop R2, the norms–MH and norms-entrenchment R3 channels, the demand saturation loop B1, and the prevalence–recovery degradation loop); and the material accounting flows — churn & other drain (movement between employment states and the draining of the model's other stocks) — form the residual band at the top. This figure shows the persistent presence of drain and churn through most years; the four punctuated departures (1999, 2006, 2010–11, 2017) when informational loops surge; the drain-recovery band collapses to near zero in these years whereas it has a steady presence in calm years.

A reading aid: The model's feedback loops run on three different clocks, set by the delays built into the structure. A fast clock (roughly one to two years) carries ordinary recovery and the unemployment–MH vicious cycle. A medium clock (about three years) carries the slow erosion of job quality, as labour market conditions feed through into insecure work. A slow clock (45 years, the time for one generation of workers to be replaced) carries the entrenchment of gender norms. These three speeds are a design feature, not a finding; the finding is which clock is driving the system at any given moment, and that is what the chronology below traces.

Phase 1: The quiet build-up (1992-2008): Prevalence is roughly flat for both genders and the system is largely governed by natural churn and drain. The recovery of poor mental health (drain-recovery) is itself an active balancing force in these calm years (for example 30% of the system's behaviour in 2002 and around 17–19% in 2003–05) confirming that, before the crisis, the system's own recovery was doing real work.

The only informational loop that stays consistently active is norm entrenchment (R3). R3 maintains the structural conditions for gender inequality through two channels: first, persistent gender pay gaps and working hours inequality generate economic signals that reinforce traditional norms, which in turn constrain women's labour market participation and perpetuate the conditions that sustain the pay gap; second, norms reinforce themselves by resisting change: the more prevalent the norms, the less effective changes in economic signals are in reducing norm reinforcement. The structural conditions for gender inequality are being actively maintained.

A brief, early activation of the unemployment–MH cycle appears around 1999. Its trigger is the small rise in unemployment of 1996–97, which surfaces in the loop a year or two later. It confirms that this mechanism existed long before 2008; the later crisis activated an existing vulnerability rather than creating a new one.

The crucial development of this phase is invisible in the prevalence data. By 2006, labour market tightness has reached its highest value in the simulation and the demand saturation loop (B1) peaks, so the labour market can absorb little

further employment. Around the same time, the exogenous rise in insecure work crosses the level at which it begins to affect recovery, becoming material for the first time. The two coincide rather than being causally linked: in the model a tighter market strengthens workers' bargaining power and thus tends to reduce insecure work. So the emerging insecure work effect reflects the exogenous trend, not the tightness peak. Simultaneously, the norms balancing loop (B3 - gender power rebalance) peaks as women's increased participation tightens the labour market, widens the gender inactivity gap, and ultimately reinforces traditional norms. These conditions create a structural vulnerability: a system with increasing insecure work. The 2008 recession provides a shock, but the mental health crisis does not begin with the recession itself.

Phase 2: The Crisis (2009–2011): The unemployment–MH cycle (R1) is a vicious circle that needs both of its arms to fire: unemployment worsening mental health (health causation), *and* worse mental health then feeding back to reduce employment (health selection), and this loop takes two to three years. So the cycle is barely detectable in 2009, when unemployment has risen but prevalence has not yet moved enough to close the loop. What dominates first, in 2010, is the *job quality* channel: insecure work is changing faster in 2010 than in any other year of the simulation, and because it acts through recovery (the faster channel) as well as onset, the model picks up its signal first. Only in 2011, once both arms of the unemployment cycle have fired, does R1 reach its peak - the single highest loop score in the whole simulation.

Phase 3: Recovery under a new constraint (2012–2019): Through 2009–2011 the recovery of poor mental health is essentially switched off, with dominance falling to 1.3%, 1.4% and 1.2% i.e. nothing is damping the rise while the unemployment cycle builds. Only in 2012, with unemployment falling, does recovery re-engage (drain-recovery returns to 11%). The mental health crisis, in other words, was a delayed and indirect consequence of the recession, carried mostly through accumulated insecure work rather than the demand shock itself.

Recovery re-establishes itself after the crisis but it now operates alongside the prevalence-recovery loop that appears for the first time in 2013. The system stabilises again, but the recovery from poor mental health is now subject to a feedback that did not exist before: when prevalence is high, recovery is

degraded. Another structural vulnerability is exposed in 2017 when the insecure work-mental health balancing loop surges. When recovery is degraded by insecure work, the norms-mental health channel becomes visible. This finding refines the filtered analysis of Section 7.2 that norms-mental health contributed only 2% of relative dominance. The full unfiltered analysis shows that this 2% average masks a punctuated pattern in which the norms-mental health loop is dormant for most years and surges to meaningful levels only under specific structural conditions.

Text Box 2: Year-level loop dominance mechanisms

The chronology turns on a few years when one feedback loop briefly took over the system. The scores are relative: each is a loop's share of the total change driving the system that year. They reflect how fast variables are changing, not how high they are, so a large but steady variable scores low. Full year-by-year scores are in Appendix P.

1999 — R1 (unemployment-mental health trap) rises to over 35%. The trigger is the small rise in unemployment in 1996–97, whose effect surfaces with the loop's one-to-two-year lag; the spike is amplified in relative terms by the simultaneous collapse of the norms-reinforcing loop (from over 43% in 1998 to about 5%), while the absolute link scores confirm R1 itself became materially more active. It then collapses to near zero for almost a decade.

2006 — Labour market tightness peaks. With the labour market at its tightest, it can absorb no more employment; the demand saturation loop (B1) surges to about 50%. Insecure work registers meaningfully for the first time (about 10%), and the recovery of poor mental health is momentarily squeezed almost to nothing (0.3%). No crisis is visible yet, but the ground is now prepared. The model's tightening through the mid-2000s is consistent with the data, but it understates the magnitude recorded in 2008; whether the loop dominance pattern around 2008–09 would differ under a fuller reproduction of labour market tightness therefore remains open.

2009 — The recession pushes unemployment up sharply (labour market tightness falls by more than half, the largest single-year move in the simulation), yet prevalence has barely moved. The unemployment cycle (R1) is still only just detectable (about 1%): it has completed only half a turn.

2010 — The system starts to compound. Three reinforcing loops fire together — insecure work–mental health (about 32%), unemployment-MH (about 12%) and norms–MH (about 11%) — while recovery collapsed to around 1%. The job quality channel led, because insecure work was changing faster this year than in any other and acts through the quicker recovery route (in addition to the slower incidence route). This reduces recovery for men for the first time.

2011 — The poor mental health stocks are now visibly rising. The unemployment-MH cycle (R1) reaches complete dominance at about 72%, the single highest loop score in the whole simulation. This is the culmination of a two-to-three-year

build, not an immediate effect of the 2008 shock: both halves of the vicious circle have finally fired.

2012 — The system begins to settle. As unemployment falls, the recovery of poor mental health switches back on (11%) and the ordinary background drains reassert themselves (about 58%).

2013 — A new vulnerability appears. The prevalence-recovery degradation loop (the endogenous recovery rate) became active for the first time (about 4% reinforcing, 3% balancing). Dormant before the crisis, it switches on once prevalence is high enough for high distress to start slowing recovery, a legacy the system now carries forward.

2017 — System self-corrects. With most variables stable, a jump in labour market tightness sets off the health-regulated job quality loop (B2) at about 69%; it is a self-limiting response as conditions improved. Once insecure work had degraded recovery, the usually hidden norms-MH loop also became briefly visible (about 25%). This self-correction should be read with a caveat: the late 2010s is a period in which the model under-represents the continued rise in women's distress. The system's settling here is therefore consistent with the model providing a conservative, lower-bound account of the recent period.

7.4.2 Summary

The chronology corroborates the temporal sequencing identified in Section 7.3 with historical specificity. Norms entrenchment is the most persistently active informational loop across the entire simulation, but the loops directly governing mental health (R1, R2) remain dormant until the medium-delay insecure work channel has accumulated to threshold (visible at the 2006 capacity peak) and the recession provides a trigger. The 2008 demand shock did not generate the crisis; it discharged a vulnerability the system had been quietly assembling since 2006. After 2012, drain reasserts but the system carries forward a new vulnerability — endogenous degradation of recovery rate (prevalence-recovery loop), visible at the 2017 punctuation. What the analysis cannot adjudicate is whether different exogenous trends in insecure work would have prevented the rise. Section 7.5 tests this counterfactually.

7.5 Counterfactual analysis

This section uses counterfactual analysis to interrogate the causal logic of the system. By asking "what would have happened if historical conditions were different?", the counterfactuals reveal which mechanisms matter and through which pathway.

Sixteen scenarios systematically alter the model's exogenous inputs and structural parameters; each was examined over the full 1992–2019 period. The scenarios are organised into four groups: those acting on the labour market (recession and insecure work), those acting on gender dynamics (gender pay gap (GPG), gender norms), those acting on both and those that test the mental health–labour market feedback mechanism. For each scenario, two metrics are reported: the prevalence gap in 2019 and the change in each gender's prevalence relative to the base case to identify scenarios that close the gap by making one gender worse rather than by making both better. All scenarios are compared against the base case (total prevalence of 0.198 and a gap of 0.077 at 2019, women: 0.236, men: 0.159). Table 7-6 provides an overview of all the scenarios.

The counterfactual scenarios collectively yield three findings. First, removing the 2008 recession reduces the total prevalence of poor mental health by 4% and the gender gap by 7%, whereas stabilising insecure work reduces the total by 7% and the gap by 9%, and combining the two adds nothing beyond insecure work alone — the cyclical shock is not the mental health lever; structural job quality is. Second, accelerating norms erosion closes about 40% of the gap but redistributes the burden of poor mental health from women to men: as women's working hours parity rises, a fixed pool of insecure work is reallocated toward men, so women's prevalence falls while men's rises and the total changes little; equality and quality must be pursued jointly. Third, severing the mental health to labour market feedback paradoxically worsens both genders' outcomes, because labour market exit is partially protective at the system level

Table 7-6: Counterfactual analysis: scenarios and results

Details how the 16 different counterfactual scenarios impact the prevalence and gender gap in 2019.

Scenario	Description	Men 2019	Women 2019	Total 2019	%change in total	Gender Gap 2019	%change in gap	Both improve?
Base		0.159	0.236	0.198		0.077		
Labour Market Scenarios								
A	No recession	0.154	0.226	0.191	-3.8%	0.072	-6.5%	Yes
B	Stable insecure work	0.149	0.219	0.185	-6.8%	0.070	-9.1%	Yes
N	No recession + stable insecure work	0.149	0.219	0.185	-6.8%	0.070	-9.1%	Yes
Gender Dynamics Scenarios								
C	GPG frozen at 0.30	0.161	0.246	0.204	3.1%	0.085	10.4%	Both worse
D	GPG frozen at historical	0.161	0.245	0.204	2.8%	0.084	9.1%	Both worse
E	GPG fixed at 0.10	0.159	0.232	0.196	-1.0%	0.073	-5.2%	Marginal
F	GPG baseline fixed at 0.05	0.159	0.234	0.197	-0.5%	0.075	-2.6%	Marginal
G	Accelerated norms erosion	0.171	0.217	0.195	-1.9%	0.046	-40.3%	Men worsen
H	Low starting norms (0.25)	0.165	0.232	0.199	0.4%	0.067	-13.0%	Men worsen
I	High starting norms (0.75)	0.154	0.234	0.195	-1.7%	0.080	3.9%	Marginal
J	High norms + GPG historical	0.155	0.240	0.198	0.0%	0.085	10.4%	No
K	Low norms + GPG 0.10	0.165	0.232	0.199	0.4%	0.067	-13.0%	Men worsen
Combining Labour Market and Gender Dynamics								
L	Low norms + stable insecure work	0.151	0.210	0.181	-8.7%	0.059	-23.4%	Yes
M	Low norms + stable insecure work + GPG 0.10	0.151	0.211	0.182	-8.4%	0.060	-22.1%	Yes
Mental health - Labour Market Feedback								
O	MH→LM feedback severed	0.167	0.257	0.213	7.4%	0.090	16.9%	Both worse
P	MH→LM feedback doubled	0.156	0.229	0.193	-2.5%	0.073	-5.2%	Yes

Scenario N (no recession + stable insecure work) is grouped with the labour market scenarios A and B though lettered later.

7.5.1 Labour market scenarios: no recession (A), stable insecure work (B), both (N)

Three scenarios alter the labour market: removing the 2008 recession (A), holding insecure work at its 1992 level (B), and combining the two (N), to isolate which channel carries the mental health consequences.

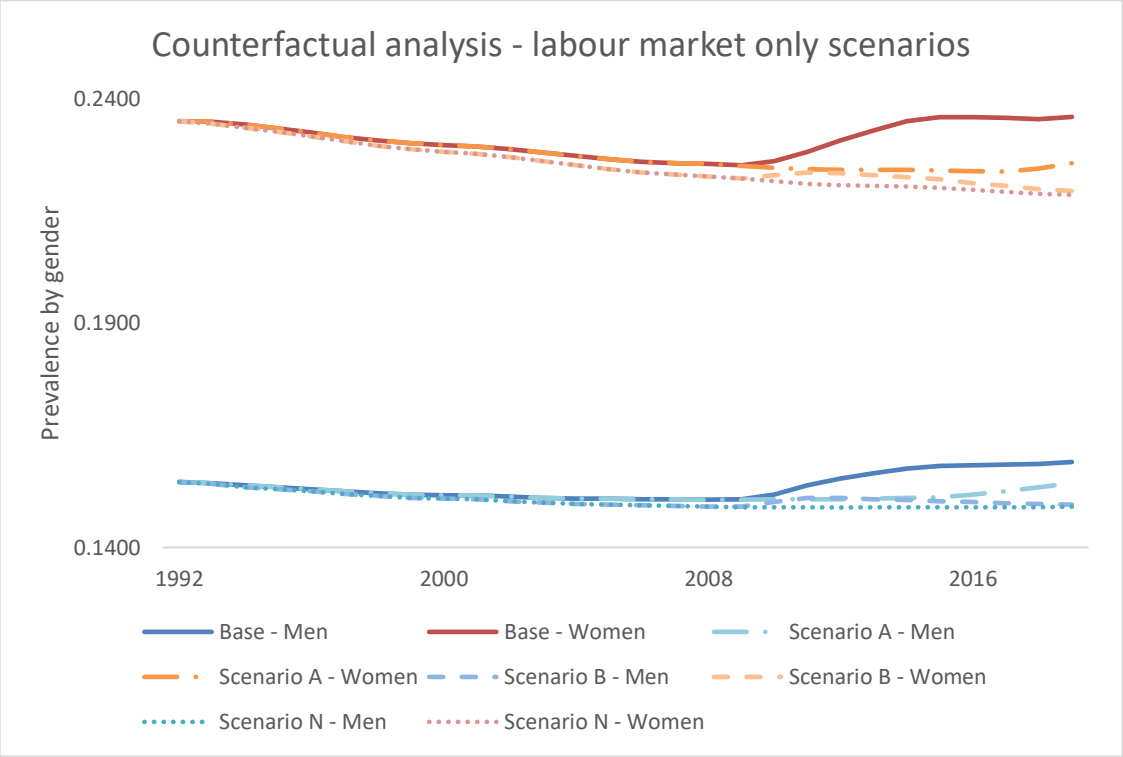


Figure 7-9: Estimated prevalence of poor mental health under different counterfactual labour market scenarios

Shows the results of Scenario A where the 2008 macroeconomic demand shock was removed, Scenario B where the exogenous input to insecure work is held constant at its 1992 value, and when both are combined (Scenario N). It shows that all scenarios would have slightly reduced the prevalence and that stabilising job quality would have had a greater effect than preventing the recession.

1. Removing the recession has a modest effect. Holding labour demand at its 2008 level from 2008 onwards, the 2019 total falls from 0.198 to 0.191 and the gap falls from 0.077 to 0.072. Both genders improve, women more than men (gap reduces by 6.5% whereas the total by 3.8%), but prevalence still rises in the final years of the simulation.

2. Holding the baseline trend in insecure work at its 1992 level (while allowing the endogenous dynamics (labour market tightness effects) to operate) does slightly more, reducing the total to 0.185 and gap to 0.070.

3. Combining both interventions (Scenario N) gives a total of 0.185 and gap of 0.070 — identical to stable insecure work alone (Scenario B). Removing the recession adds nothing once insecure work is already stable.

In scenario B, the benefit comes almost entirely through recovery: in the base case, rising insecure work crosses a threshold that erodes the protective effect of employment on recovery, and holding it flat preserves that benefit. The incidence channel barely moves, because the economic stress composite operates in a narrow, relatively flat range. Two parameterisation sensitivities — a heavier weight on insecure work and a steeper effect of economic stressors on onset curve — leave the 9% result unchanged (Appendix R).

The near-equivalence of B and N is informative; it reveals which labour market channel carries the mental health consequences. Scenarios A and N both remove the recession, so labour market tightness is virtually identical and they differ only in whether the exogenous baseline insecure work trend is rising (Scenario A) or is frozen (Scenario N). This difference creates a large divergence in insecure work exposure by 2019. In Scenario A, the rising insecure work eventually crosses the recovery degradation threshold which reduces the benefit that employment provides for recovery whereas in Scenario N, the full employment benefit is preserved. Scenarios B and N, by contrast, isolate the recession's direct effect. The recession's spike in B creates an additional incidence multiplier that is too small and transient to materially affect prevalence trajectories, resulting in a small difference in incidence between B and N.

The implication is that within the model's labour market channel, it is structural job quality and not the cyclical shock that carries the mental health consequences. Stabilising the insecure work parameter produces significantly larger effects than merely neutralising the simulated demand shock. This does not diminish the recession's real economic costs, but it suggests that from a mental health perspective, the slow deterioration in employment quality was more consequential than the acute demand shock.

A caveat is necessary here. Labour demand and the baseline insecure work trend are exogenous inputs to the model. In Scenario A, the rise in insecure work continues because the exogenous baseline trend is unchanged; only the cyclical

shock is removed. This means the test isolates the recession as a discrete event from the structural trend in job quality. It cannot tell whether the insecure work trend itself would have continued without the recession, since that trend enters the model from outside. What the test does show is that the recession, as defined by a reduction in labour demand, was not the primary driver. The loop analysis identified the unemployment-MH reinforcing loop (R1) as the dominant mechanism during 2010–11. Scenario A removes the trigger, yet the gap barely changes. The structural vulnerability of rising insecure work, persistent gender inequality, slow norms erosion were already building before 2008.

7.5.2 Gender dynamics scenarios: gender pay gap (C to F), norms (G to I) or both (J, K)

The next set varies the gender dynamics inputs: the gender pay gap (C–F) and the level and speed of norm change (G–I), with two combinations (J, K).

Gender Pay Gap (GPG): These scenarios show that its historical narrowing (from 0.30 in 1992 to 0.15 by 2019) was protective. Freezing GPG at its 1992 level widens the 2019 gender gap by about 10% (C, D); pushing it lower closes the gender gap by a similar margin (E, F). The effect runs solely through women's incidence, so men are essentially unaffected.

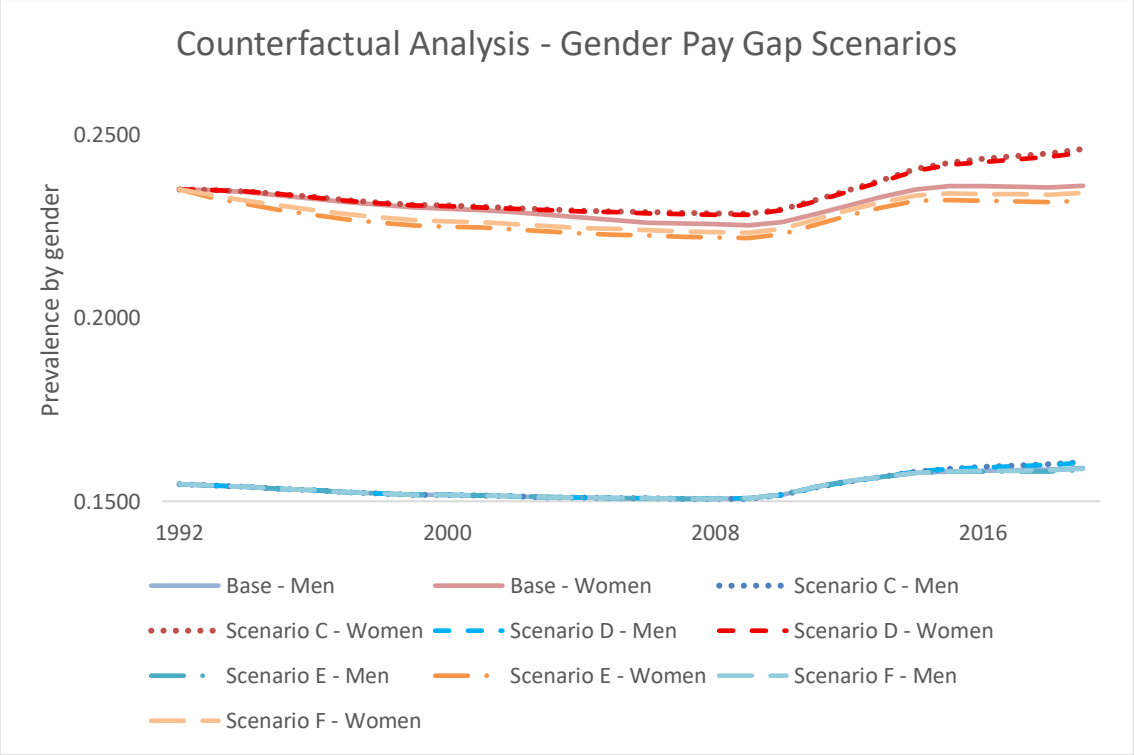


Figure 7-10: Estimated prevalence of poor mental health under alternative gender pay gaps

Compares four scenarios: C and D freeze the GPG at its 1992 level; C holds it at the 1992 constant value (0.3), turning off the endogenous multipliers operating on it; scenario D holds the exogenous baseline GPG at its historical 1992 value, allowing the endogenous multipliers to operate and causing a slow decline to 0.28. Scenarios E and F set the GPG to lower values (E fixes the GPG at 0.10; F fixes the baseline GPG at 0.05, allowing a decline from 0.14 to 0.12). Scenarios C and D produce a ~10% widening of the prevalence gap whereas E and F close the 2019 gap by ~10%; the historical narrowing was protective.

Gender norms: These scenarios show that more egalitarian norms redistribute the mental health burden; they reduce prevalence of poor mental health among women but increase it among men. In the base case, traditional norms sort men into more secure work, and when norms equalise rapidly in a labour market where insecure work is rising, men lose that shield. This does not mean gender equality harms men in general; it means the benefits of equality depend on the quality of the labour market into which both genders are equalised.

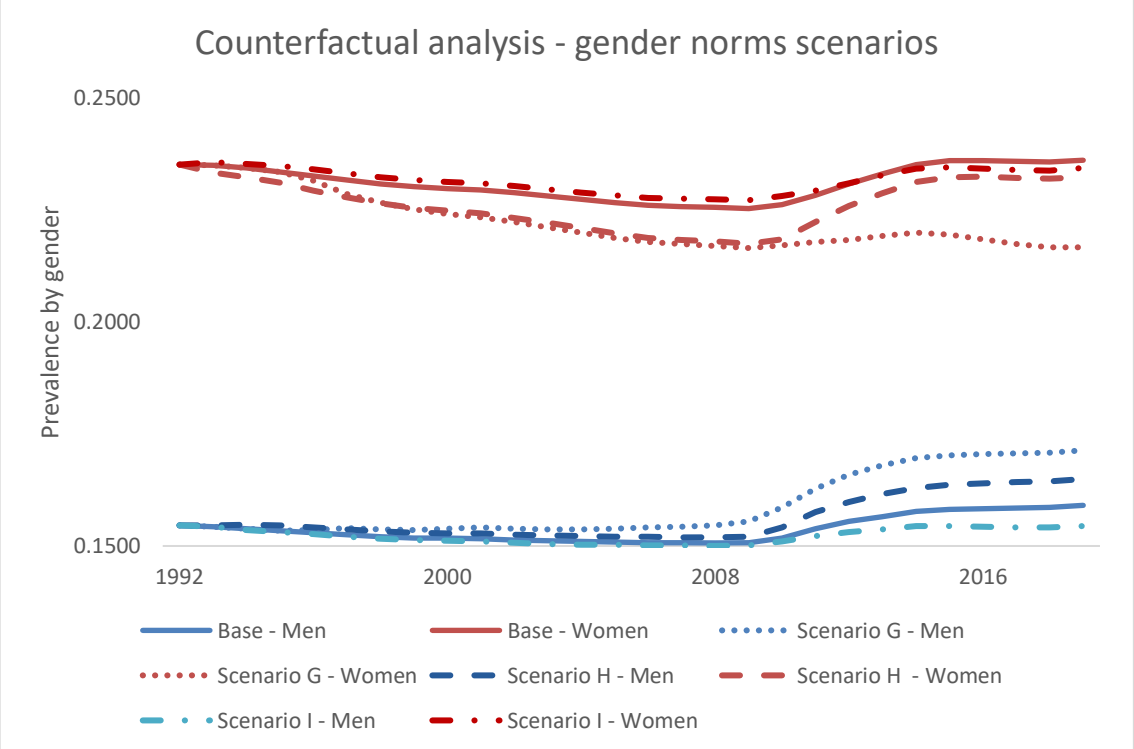


Figure 7-11: Estimated prevalence of poor mental health under alternative gender norms

Scenario G accelerates norm erosion (operationalised through a one-time STEP increase in norm erosion in 1995), H starts the model at lower norms (more egalitarian) whereas I at higher norms (more traditional). Scenarios G and H highlight a redistributive trap where rapid equality in a deteriorating labour market inadvertently pushes men into higher insecure work and worsens their prevalence.

Scenario I tests more traditional gender norms. With lower hours parity, men are sorted more strongly into secure work, but women's exposure to insecure work does not increase as strongly since their employment drops under more traditional norms. This leads to a marginal improvement on the effect of

employment on recovery among women. Unemployment also reduces among women leading to almost no difference in the economic stress affecting incidence. Stress due to gender inequality increases due to higher unpaid work and higher GPG (due to lower hours parity). However, the gains in recovery outweigh these, leading to a lower prevalence by 2019. Men's prevalence is even lower in 2019 (due to higher employment), resulting in almost no gender difference compared to the base case.

Combining both: Gender norms and GPG operate through overlapping channels (both feed the gender inequality index), and reducing one makes the marginal effect of reducing the other negligible, leading to almost no difference in prevalence among women in 2019 (scenario K). Crucially, without changing norms, there is no change in men's prevalence. Thus, the gap reduces more by changing norms. Pairing traditional norms with a stagnant GPG modestly worsens women's outcomes (scenario J).

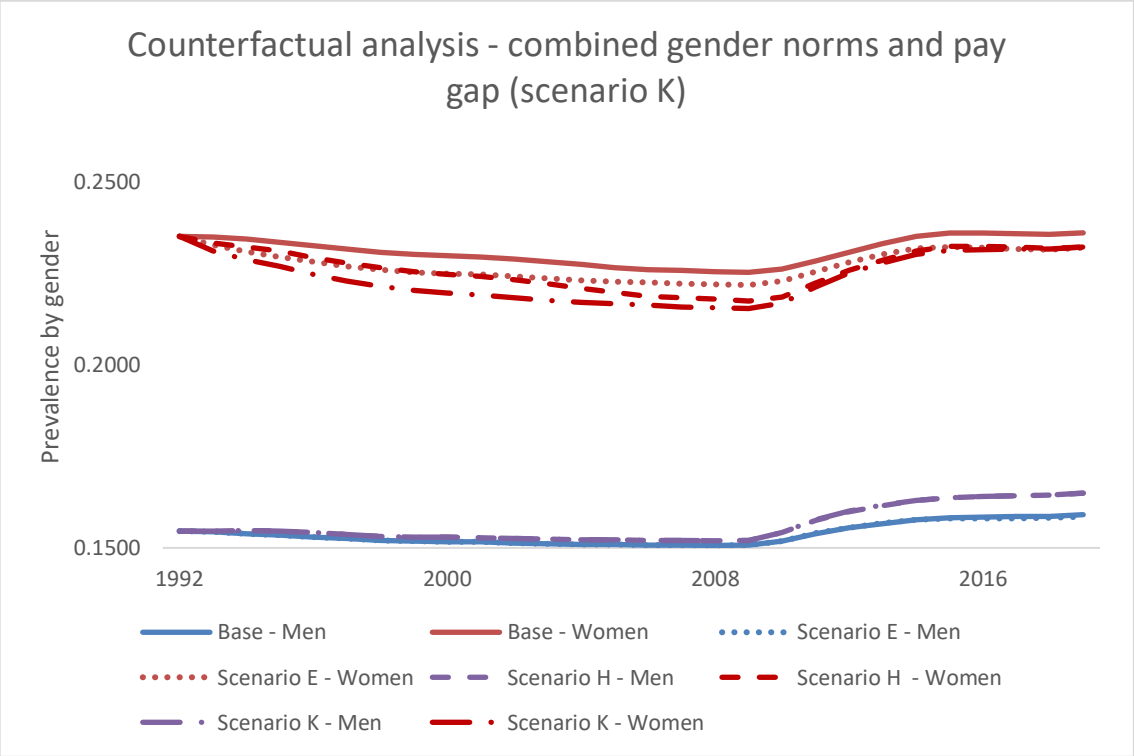


Figure 7-12: Estimated prevalence of poor mental health under different gender norms and gender pay gap scenarios

Scenario K has low starting norms and low fixed GPG; this is compared to Scenario H where only low starting norms are implemented and Scenario E where only low fixed gender pay gap is implemented.

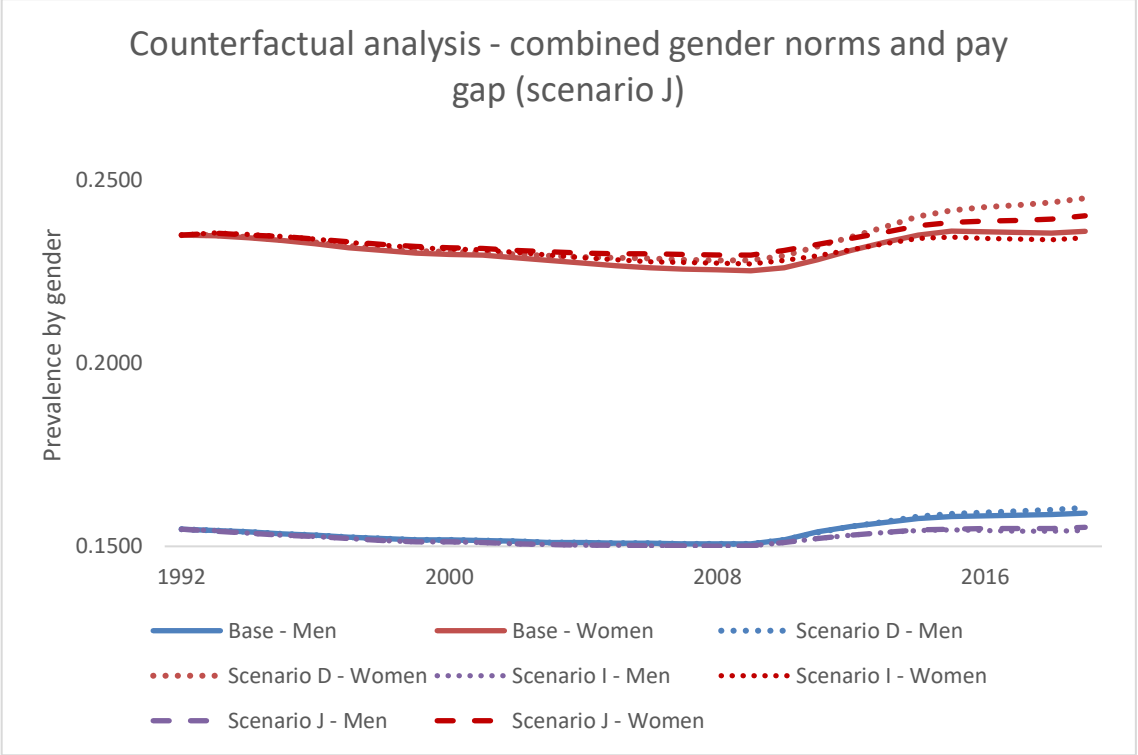


Figure 7-13: Estimated prevalence of poor mental health under different norms and gender pay gap scenarios

Scenario J has high starting norms and fixed gender pay gap at 1992 values; this is compared to Scenario I where only high starting norms are implemented and Scenario D where gender pay gap is fixed at its historical value.

7.5.3 Combining labour market and gender dynamics interventions (Scenarios L and M)

The redistribution in Scenario G raises an obvious question: can the gap in prevalence be closed without shifting the burden onto men? Scenario L combines accelerated norms erosion (G) with stable insecure work (B). If the redistribution in Scenario G is caused by norms equality interacting with a deteriorating labour market, then stabilising the labour market should prevent men's worsening. It does. In Scenario L, prevalence falls among both men and women. The gap closes *less* in L (about 23% versus 40% in scenario G) but achieves a reduction in the gap through universal improvement. The total prevalence falls by ~9% compared to ~2% in Scenario G.

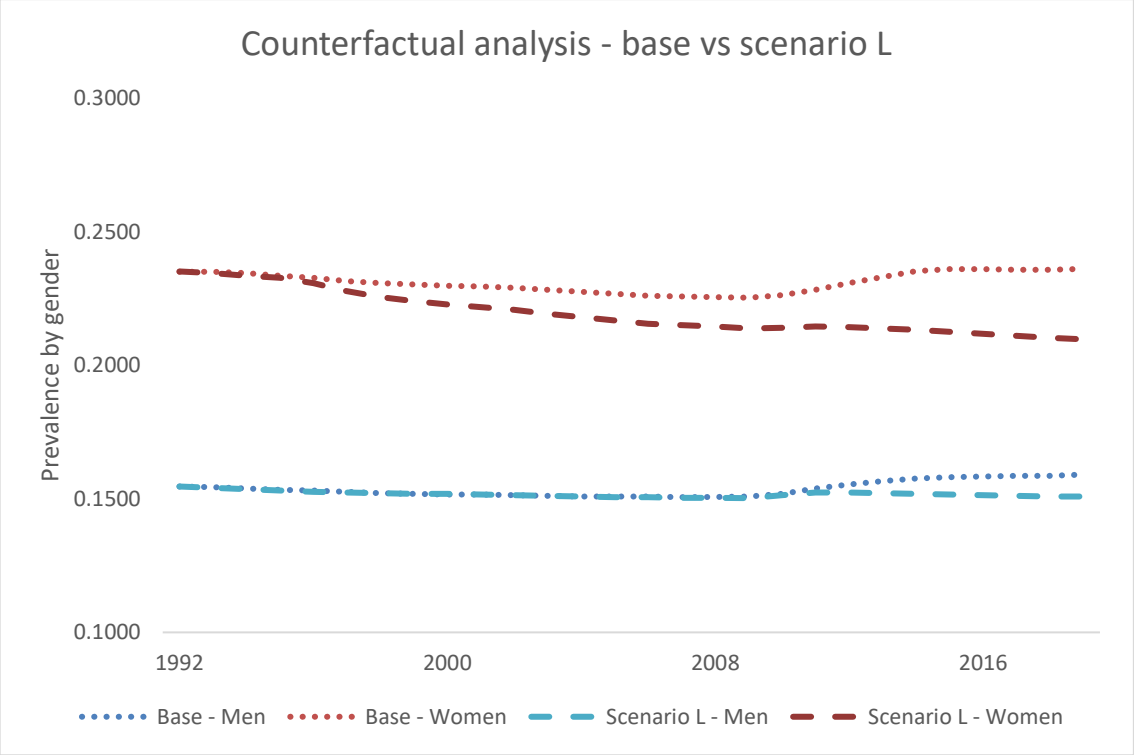


Figure 7-14: Estimated prevalence of poor mental health under more egalitarian norms alongside stable job quality (insecure work)

Displays the results of Scenario L, where accelerated norm erosion with stable insecure work lowers prevalence for both genders, achieving gap reduction through universal improvement.

The underlying mechanics of this redistribution highlight a powerful unintended consequence within the simulated labour market structure. In the base case, traditional gender norms function as an allocative shield for men: they sort men into more secure employment, shielding them from the mental health effects of a deteriorating labour market. Women bear the double burden of higher inequality-driven incidence and higher insecure work exposure. When norms equalise, the shield is removed. Whether this harms men, and by how much, depends on the state of the labour market, but the redistribution is not confined to a deteriorating one: the forward analysis of Section 7.6.1 finds the same direction at realistic lever sizes under benign, stable conditions, with deterioration shaping its magnitude and, under the most adverse conditions, its sign. This suggests that neither gender equality nor labour market quality alone is sufficient for population mental health improvement. Equality without quality redistributes the poor mental health burden from women to men. Quality without equality preserves the gender gap. Within the multi-scenario optimisation presented in Section 7.6, only the combined parameter adjustments yield both a narrower gap and lower prevalence for both genders.

Further, given that gender pay gap and norms operate through overlapping mechanisms, improving gender pay gap in addition to lower norms and stable insecure work, did not result in further gains (Scenario M).

7.5.4 The mental health-labour market feedback (Scenarios O and P)

The loop analysis identified the reinforcing loop R1 as the dominant gendered mechanism. Two scenarios test whether this feedback is load-bearing. Scenario O severs the feedback by setting the sensitivity of all three mental health-to-labour market transitions (job loss, job finding, and exit to inactivity) to zero. Poor mental health no longer affects employment outcomes. Scenario P doubles the sensitivity, strengthening the feedback.

The results are counterintuitive. Severing the feedback (O) makes outcomes worse: men's prevalence rises from 0.159 to 0.167, women's from 0.236 to 0.257, the total increases by ~7% and the gap widens by ~16%. Doubling the feedback (P) makes outcomes better: men fall to 0.156, women to 0.229, and the total reduces by ~3% and the gap narrows by ~5%. Stronger health selection improves population mental health; weaker health selection worsens it.

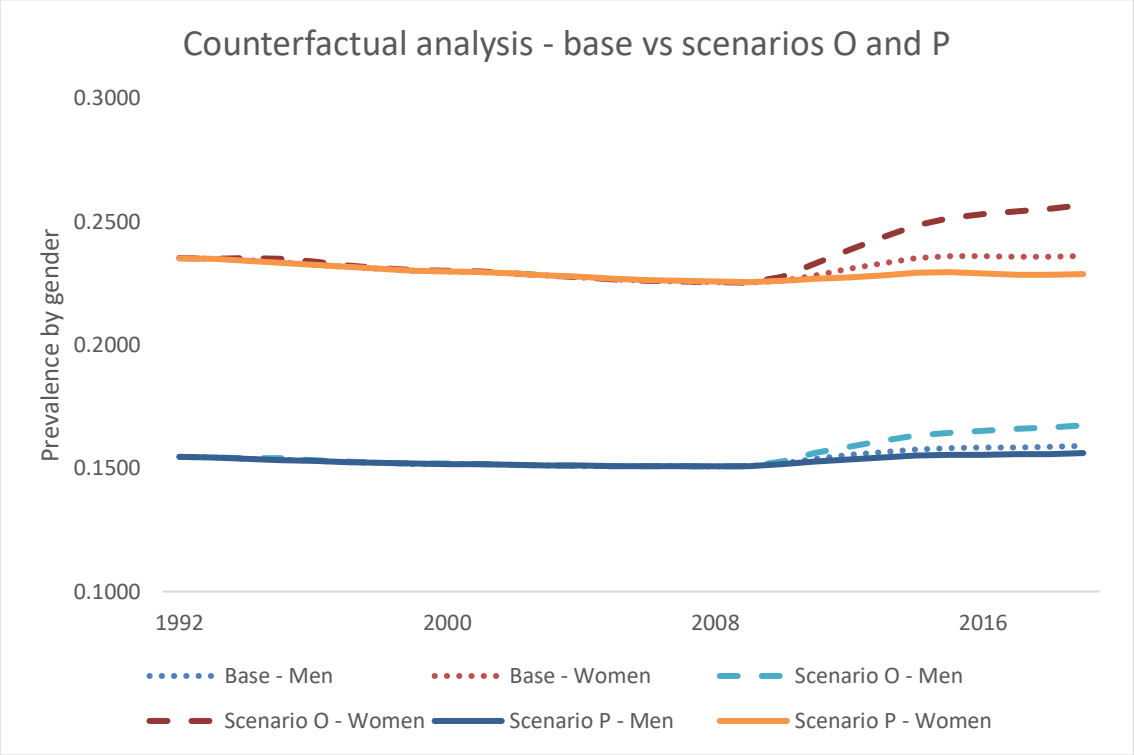


Figure 7-15: Estimated prevalence of poor mental health when severing and doubling the effect of prevalence on labour market outcomes

Shows scenarios modifying the health selection loop (scenarios O and P). It reveals that stronger health selection (labour market exit) paradoxically protects the broader population by limiting job quality deterioration.

Unpacking this requires examining the three mechanisms through which mental health impacts labour market:

1. Poor mental health increases job loss (reinforcing: unemployment worsens mental health)

2. Poor mental health reduces job finding (reinforcing: prolonged unemployment)

3. Poor mental health increases exit to inactivity (removes individuals from the labour market)

While the first two act as reinforcing mechanisms for the individual by keeping them in the stressful state of unemployment, the third removes them from the labour force, removing them from the stressor environment. This also has a system-level consequence: fewer people in the labour market reduces insecure work through labour market tightness.

When all three channels are severed (O), the safety valve closes. More people remain employed, generating higher insecure work given the same labour demand/labour market tightness. The higher insecure work pushes recovery degradation far below base case levels and prevalence rises high enough to activate the prevalence-recovery feedback, causing a spiral.

When the feedback is doubled (P), the opposite occurs. More people exit the labour force when they develop poor mental health, reducing labour force participation and producing lower insecure work. The individual cost (more labour market exit) produces a population benefit.

The implication is that the health selection mechanism has qualitatively different consequences at the individual and population levels. At the individual level, it is harmful: a person who loses their job because of poor mental health faces worse economic circumstances and is likely to deteriorate further. At the system level, it is partially protective: labour market exit removes people from

the stressor environment, preventing the worsening of labour market conditions and protecting the recovery capacity of the remaining population. This explains why R1 reinforcing loop appeared dominant during the 2008–12 crisis, yet the system eventually stabilised. Part of that stabilisation comes from health selection itself: as prevalence rises, more people exit the labour market, reducing stressor load and allowing the drain loop to reassert dominance.

7.5.5 Summary

Table 7-7 synthesises the counterfactual analysis which results in four key findings, discussed below.

Table 7-7: Summary of counterfactual analysis
Synthesises the key findings of the counterfactual scenarios

Intervention	Scenarios	Key finding
Recession removal	A, N	Modest effect; doesn't add much if rise in insecure work was prevented
Insecure work stabilised	B, N	Larger effect than recession removal
Gender pay gap	C, D, E, F	Historical narrowing was protective; further reduction yields modest gains.
Gender norms	G, H, I	Accelerated equality closes the gender gap but redistributes burden to men.
Equality + quality	J, K, L, M	Results in narrower gap *and* lower prevalence for both.
Feedback severance	O, P	Severing the link from mental health to labour market worsens prevalence of poor mental health for both genders; the feedback is protective at system level.

First, within the model, the 2008 recession functions as an activating event for pre-existing structural vulnerabilities rather than as an independent cause of the mental health trend. Removing the demand shock reduces the gap by only 7% and combining it with stable insecure work adds almost nothing. The recession's mental health effect operates almost entirely through the insecure work channel. Insecure work harms mental health primarily by degrading recovery rather than by increasing incidence.

Second, egalitarian gender norms is the most powerful lever for closing the gender gap, but it achieves this partly by redistributing the mental health burden from women to men. This exposes a redistributive trap: equality without quality redistributes the health burden, whereas quality without equality preserves the gender gap. Only simultaneous interventions lower prevalence for both. This suggests that policies addressing gender inequality and labour market quality need to be pursued jointly, not as alternatives.

Third, the mental health–labour market feedback is paradoxically protective at the system level. Exiting the labour market due to poor mental health, although harmful for the individual, and a concern for the labour market and welfare systems, can prevent the deterioration of job quality, mitigating the conditions contributing to the prevalence of poor mental health.

Fourth, considering interventions jointly rather than in isolation surfaces structural insights that single-intervention experiments cannot. First, equality and quality must move together (Scenarios G, B, L). Second, the gender pay gap and norms feed the women's stress index, so addressing one neutralises the marginal benefit of the other (K versus H).

7.6 Addressing the rising and gendered prevalence of poor mental health

So far the chapter has explained why the prevalence of poor mental health rose, why the gender gap persists and what might have happened in counterfactual scenarios. This section asks what might help to close the gap and lower total prevalence going forward, within the model’s structure. Experiments and optimisation, explained in Text Box 3, are used to test the model's sensitivity to structural changes in four causal mechanisms and to identify which levers the model is most responsive to and whether the model's structure implies trade-offs between reducing the gender gap and reducing total prevalence of poor mental health.

Four structural levers are tested — the gender pay gap (reduction), childcare provision (which increases the gender parity in working hours, and thus reduces unpaid work burden and sorting of women into insecure work), gender norms

(reduced inertia), and insecure work (reduction) — against three objective functions (the gap, total prevalence, and both jointly) across nine macroeconomic futures spanning 2024–2040. Methodological detail is in Text Box 3.

The four levers were identified through the convergence of the preceding analysis, the structure of the model and current policy. The structural decomposition attributed most of the gender gap to the incidence pathway (Section 7.3), which in the model runs through the women's inequality composite of the gender pay gap and the unpaid work burden. The pay gap combines an exogenous baseline with an endogenous penalty driven by gender parity in working hours; parity is determined largely by gender norms and can be raised by childcare availability; and the unpaid work burden is driven by norms and reduced by parity. The gender pay gap, childcare and norms are therefore the levers through which the incidence pathway can be modified. Accelerating norm erosion produced the largest single reduction in the gap in the counterfactual analysis (Section 7.5). Through parity in working hours, norms and childcare also shape women's sorting into insecure work. Insecure work, an exogenous trend modified by labour market tightness, carried more of the historical rise than the demand shock (Section 7.5). Three of the levers also correspond directly to current policy: childcare through funded Early Learning and Childcare, the gender pay gap, and job quality under the Fair Work agenda; norms have no single policy instrument and are the least tractable of the four.

The optimisation deploys all four levers together. It cannot show what each lever does in isolation or, to each gender. A falling total does not mean that no one is made relatively worse off, and this distributional question is separate from the aggregate one, and not answered by the optimisation. This section therefore opens with a complementary analysis (per-lever experiments) that isolates each lever, and the combinations that matter, against the baseline, and examines the direction of change for each gender separately.

This per-lever view yields three structural insights that the optimisation then refines: the gender gap closes partly by redistribution of prevalence from women to men, a subset of levers carries almost the whole effect, and the pattern is robust across macroeconomic futures but bounded by the structural floor alluded

to in Section 7.3 and discussed in Section 7.6.5. The gender pay gap is one lever with no cost to men, and the optimiser deploys it to its bound in every macroeconomic future, for both gap and total objectives. Gender norms redistributes poor mental health prevalence onto men, but the optimiser confirms it carries near-zero cost to total prevalence. Finally, the optimiser reiterates the macroeconomic conditioning: reducing insecure work lowers total prevalence in every future, and narrows the gap under adverse conditions but marginally widens it under benign ones. This is because in benign conditions, men's insecure work sits above the threshold at which insecure work degrades the beneficial effect of employment on recovery, and thus they benefit more when insecure work reduces.

An important caveat applies specifically to the forward projections. The model is calibrated on 1992–2019 and treats the 2020–2022 COVID-19 period as an exogenous structural shock that lies outside its boundary. The optimisation therefore runs forward from a pre-COVID model and a 2019 endpoint. Because the model carries no 2019–2024 data, the 2024 state from which the levers are applied is itself a pre-COVID extrapolation rather than the observed 2024 position; the projected levels are conditional on that counterfactual baseline and bound the direction of mechanisms rather than predicting present-day levels. Because the pandemic plausibly disturbed some of those relationships, such as the link between labour demand and employment, the gender division of unpaid work, and the capacity of mental health services, the 2024–2040 results should be read as the behaviour the model's structure implies, not as projections that incorporate the post-2020 trajectory. The same caution extends to other potential structural breaks after the calibration window such as the cost-of-living crisis and changes in employment policy, which could alter the calibrated relationships themselves and not merely the model's inputs; the forward runs assume those relationships persist. The broader implications of this boundary are discussed in Chapter 8.

Text Box 3: Optimisation methodology

The optimisation searches across three objectives, four parameters, and nine macroeconomic futures. The per-lever experiments are aligned with these.

Objectives: The analyses use Stella Architect's in-built optimiser to search for parameter combinations that minimise three alternative objectives: (1) the

gender gap in prevalence at 2040, (2) total prevalence at 2040, and (3) both simultaneously via Pareto optimisation.

Set-up: All interventions activate in 2024, and the model runs forward to 2040.

Parameters: Table 7-8 summarises the parameters. Four parameters are optimised, each representing a potential structural lever that operates through a distinct causal channel identified in the preceding sections. Three relate to additional burden faced by women (gender pay gap (GPG), childcare availability, norms), whereas the fourth, insecure work, impacts both. Norms have an indirect impact on men, as seen in Section 7.5. The ranges are deliberately conservative — a 25% reduction in insecure work, a 50% increase in childcare provision, a halving of the GPG and a halving of norms reinforcement — and are chosen to represent plausible policy effort, not the model's theoretical limit. They therefore define a reachable floor rather than the model's structural minimum; driving the levers to their structural extremes lowers the floor only marginally and leave the best lever combination unchanged. Because the effects of GPG, childcare and norms overlap, the analysis reports both the effect of each lever alone and the combination.

Macroeconomic futures: To test whether the results are contingent on assumptions about the macroeconomic future, each optimisation is run across a 3 x 3 sensitivity grid combining three levels of post-2024 labour demand change (0.90, 1.00, 1.05) and three post-2019 insecure work trend projections (constant, half historical trend, full historical trend). Sensitivity analysis was conducted on labour demand projections from 0.8 to 1.2. Mental health prevalence did not show any significant change above 1 for women, and for men above 1.05. So, a range of 0.9 to 1.05 was selected. 1.2 scaling would mean demand exceeds WAP. Scaling of 0.8 gives a labour demand of 0.682, almost 5% lower than the minimum seen during 1992-2019 (0.704) and was thus excluded.

The grid crosses two labour market forces: labour market tightness and structural casualisation. Because a tighter market lowers insecure work, both forces act on the level of insecure work, and the nine futures collapse onto a single precarity spectrum: from the tight, low-precarity S7, through an intermediate region (S4, S8, and rising through S5, S9), to the high-precarity futures (S6, S1, S2, S3). The base case is S4 (stable demand, stable job quality).

Because labour demand enters as a fraction of working-age population, the demand grid also spans plausible demographic and working-age-population trajectories acting through labour market tightness; demographic change is therefore implicit within the scenario range rather than excluded.

Algorithm and weighting: The optimiser uses differential evolution (population size 40 with 200 generations, crossover rate 0.9, mutation factor 0.6, tolerance 0.00005, seed 123). To appropriately optimise for the combined objective, the two payoffs are weighted by the inverse of their base case 2040 values (gap weight = 16.47, total weight = 5.35), so that a 1% improvement in either objective contributes equally to the payoff function.

Runs: Eighteen single-objective optimisation runs (i.e. nine scenarios each for gap minimisation and total prevalence minimisation) were run followed by nine combined objective runs.

The childcare lever operates through an assumed functional form linking childcare availability to gender parity in working hours, so its result was re-run under alternative specifications of that link. Across these, the optimal solution is unchanged. Full results are in Appendix T.

Table 7-8: Optimisation parameters

Outlines the parameter ranges and rationale for the four causal levers (gender pay gap, childcare, norms, insecure work) used in the model's optimiser to test sensitivity

Parameter	Range	Pathway	Rationale
gender pay gap_scale	0.50–1.00	Gender pay gap reduction → women's stress index → incidence	Exogenous baseline of the pay gap, one of the two components of the incidence pathway (61% incidence pathway)
childcare_scale	1.00–1.50	Childcare expansion → gender parity in working hours	Raises gender parity in working hours, which reduces the gender pay gap, unpaid work (both of which impact incidence) and women's sorting into insecure work; devolved policy
norms_scale	0.50–1.00	Reduces inertia in norms reinforcement → accelerates norms erosion	Drives unpaid work (incidence) and parity in working hours; largest single-lever reduction in gender gap in the counterfactual analysis; persistently active loop
insecure work_scale	0.75–1.00	Reduction in insecure work → increased recovery for both genders (and reduced incidence)	Carried the historical rise; focus of the Fair Work agenda

7.6.1 What is the effect of different levers for men and women?

This section isolates each lever against a no-intervention baseline to show its direction for each gender.

Table 7-9: Effect of each lever at 2040 relative to no intervention

Reports effects under the base case (stable demand, stable insecure work; S4). Shows women's prevalence falls across interventions even though the gap widens in some.

	Prevalence of poor mental health				
Lever	**Men**	**Women**	**Gender gap**	**Total**	**Who benefits**
Gender pay gap	-	↓	narrows	reduces negligibly	Women

	Prevalence of poor mental health				
Lever	**Men**	**Women**	**Gender gap**	**Total**	**Who benefits**
Childcare	↑	↓	narrows	reduces negligibly	Women, small cost to men
Norms	↑	↓	narrows	increases negligibly	Women, small cost to men
Insecure work reduction	↓	↓	widens (most)	reduces	Both, men more
All equality levers combined	↑	↓	narrows (most)	reduces negligibly	Women; redistributes to men
All levers combined	↓	↓	widens	reduces	Both; men more
Insecure work + childcare	↓	↓	widens	reduces	Both; men more

The equality levers redistribute, not only reduce. Some levers move prevalence between the genders. Closing the gap and lowering the total are achieved through different gendered routes, so the choice of lever is, in part, a choice of who is helped. The clearest sign of this is that reducing insecure work is the lever that most helps men: in every one of the nine futures it lowers men's prevalence by as much as, or more than, the full four-lever intervention. Adding the gender-equality levers (GPG, childcare and norms) on top never improves men further; it slightly erodes the gain. The equality levers narrow the gap by lowering women’s prevalence, and through childcare and norms, they do so partly at men's expense. This redistributive direction holds in every future (Table 7-10).

Table 7-10: Each lever's 2040 effect across the nine macroeconomic futures

Three levels of labour demand × three rates of insecure work deterioration, relative to each future's no-intervention case. Total prevalence falls under every lever (except norms where it is unchanged) in every future

Lever	Total prevalence	Gap	Men's prevalence
Gender pay gap	reduces	narrows	unchanged or reduces
Childcare	reduces	narrows	increases slightly in baseline conditions (S4), stable demand

Lever	Total prevalence	Gap	Men's prevalence
			with gradual casualisation (S5) and in high demand with gradual (S8) or rapid casualisation (S9)
Norms	almost unchanged	unchanged or slightly narrows	increases slightly in most futures
Insecure work reduction	reduces	widens at the base case (S4) and high-demand with gradual casualisation (S8); narrows in every other future	reduces
Equality bundle (pay gap + childcare + norms)	reduces; negligible in all high demand scenarios (S7-S9) and base case (S4)	narrows	increases under high demand scenarios (S7-S9), base case (S4) and stable demand with gradual casualisation (S5)
All four levers	reduces	widens at S4 and S8; narrows elsewhere	reduces
Insecure work + childcare *(efficient core)*	reduces	widens at S4 and S8; narrows elsewhere	reduces

Women gain almost unconditionally; men's outcome is conditional. Women's prevalence falls under every lever in every future with a single exception: norms pursued on its own under weak labour demand, where women edge upward only at the very end of the horizon, as the model's endogenous prevalence-recovery mechanism activates. Even in intervention scenarios where the gender gap does not fall, women’s absolute prevalence reduces. On the other hand, reducing insecure work always helps men, but the equality levers raise men’s prevalence in low-precarity futures. This is due to the reduced differential sorting of women into insecure work when certain levers are deployed. Childcare and norms raise women's working hours parity, which reallocates a fixed pool of insecure work

toward men, whereas the gender pay gap does not; the model contains no direct feedback from the pay gap to men's pay or mental health.

Not all levers are needed to get a similar level of reduction. Because there is overlap in the equality levers, it is possible to get equivalent results employing only one. Childcare has the most direct effect on gender parity in working hours, which has an immediate effect on insecure work and the gender pay gap, as well as a lagged effect on future norms. Therefore, childcare and insecure work reduction together reproduce the four-lever reduction closely on both objectives in every future. Adding GPG and norms levers changes the total negligibly and the gap only a little. Norms appears the weakest lever, but this is an artefact of the design: the intervention reduces the resistance to norm change, and within the sixteen-year window it accelerates the erosion of the norms stock only marginally (from roughly 0.32 to 0.30 by 2040), so its downstream effect on prevalence is small.

How much the levers achieve depends on how far the system starts from the structural floor. The directions above hold across all nine macroeconomic futures; what changes is their size. The levers reduce prevalence most where the no-intervention state sits farthest from its structural floor i.e. poor labour markets of weak demand (S1–S3) or high insecure work deterioration (S6), and least where the no-intervention state is already near the floor (the base case and the most comfortable futures). Near the floor there is little aggregate prevalence left to remove, so the levers can only redistribute it, and the gap narrows by shifting burden onto men. Far from the floor, the levers buffer the whole system downward and the reduction reaches both genders, although the redistributive direction persists beneath it.

Specifically, in low-precarity scenarios (S7, S4, S8): women's recovery is near-intact and thus cutting insecure work can widen the gap; this happens in S4 and S8, where men's recovery is degraded but women's is not. This widening is conditional on how the effect of insecure work on recovery is specified within the model: formulated as relative to each gender's 1992 exposure, with no effect at or below that level. In scenarios of rapidly rising or high precarity (S1-3, S5-6 and S9) cutting insecure work narrows the gap as recovery is degraded for both.

7.6.2 What closes the gender gap?

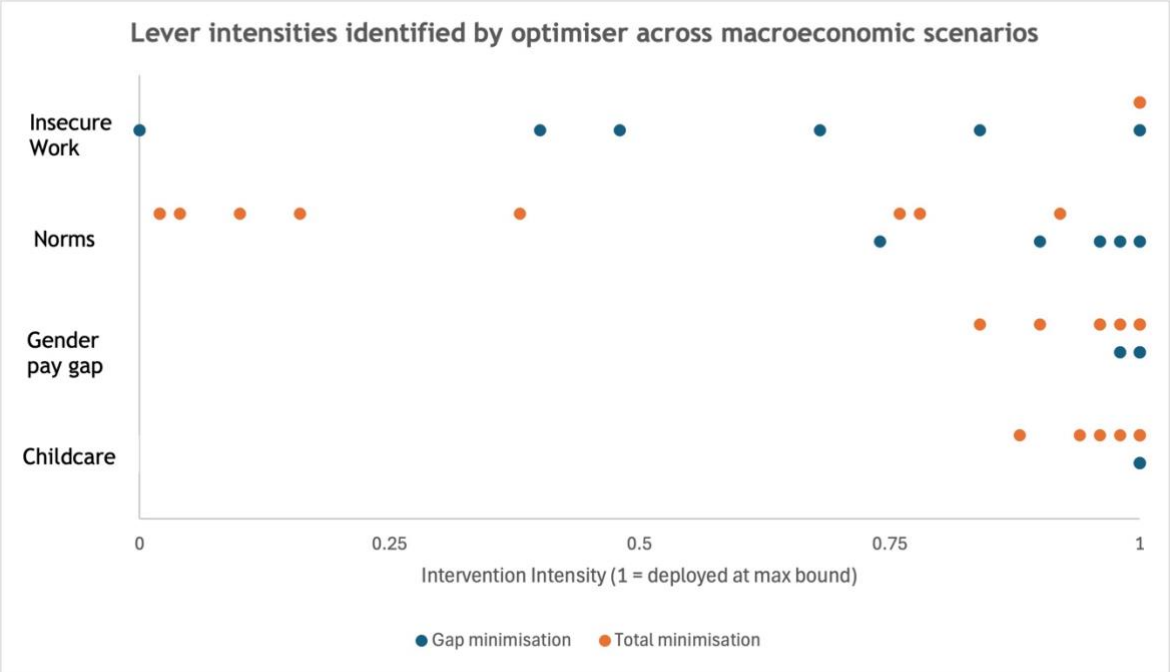


Figure 7-16: Results of single objective optimisations

Summarises where the four structural parameters converge across nine scenarios when solving independently for either a minimised gap or minimised total prevalence. It reveals that the gender pay gap and childcare parameters converge to their maximum across macroeconomic scenarios and objectives; norms converges for gap minimisation and insecure work for total minimisation.

Across the scenarios, two parameters consistently converge to their bounds. The gender pay gap scale reaches its minimum (0.50) and the childcare scale reaches its maximum (1.50) irrespective of macroeconomic assumption for the objective of gap minimisation, and close to the bounds for total minimisation. This convergence aligns with the findings of Section 7.3 that 61% of the gender gap is attributable to the effect of gender inequality on mental health incidence. The gender pay gap and childcare, both feed women's stress index that drives the incidence channel, so the optimiser is identifying the channel the structural decomposition isolates.

The norms lever converges close to its minimum in most scenarios (0.50–0.55) but with more variability than gender pay gap or childcare. Reducing norms reinforcement accelerates norms erosion, which reduces the cultural resistance to gender equality. This reduces the women's stress index by lowering the unpaid work burden but it does not directly reduce men's prevalence. In fact, by raising gender parity in working hours it shifts the sorting of insecure work towards men, slightly increasing prevalence of poor mental health among men.

The insecure work parameter behaves heterogeneously across scenarios for gap minimisation — 0.75 in three scenarios, 1.00 in two, and intermediate values in the rest. The reason emerges from interaction with the gender equality levers: When gender pay gap and childcare levers are at their bounds, which they are in all scenarios, the resulting improvements in gender parity shift women's exposure to insecure work downward. The ratio of women's 2040 insecure work exposure to their 1992 level falls below 1.0 regardless of whether insecure work_scale is set to 1.0 or 0.75. Because the effect of insecure work on recovery is a function that activates only when current exposure exceeds initial exposure, women's recovery is unaffected by insecure work in either case. Men's exposure, by contrast, remains above 1992 levels under most scenarios, so insecure work reduction translates directly into improved recovery for men. This suggests that the model's feedback structure contains a hierarchy: the independence of the incidence and recovery mechanisms identified in Section 7.3 is not symmetric. When the incidence channel is addressed through gender equality improvements, the resulting shift in women's labour market position changes how the recovery channel bears on the gap. The incidence intervention does not merely reduce incidence; it reshapes the gendered exposure structure, pushing women's exposure below the degradation threshold while men's remains above it. Reducing insecure work therefore improves men's recovery but not women's; so, rather than becoming neutral, the recovery channel now acts on the gap asymmetrically.

7.6.3 What lowers total prevalence?

For total prevalence minimisation, three of the four parameters converge close to their bounds across all scenarios: GPG at 0.50–0.58, childcare at 1.44–1.50, and insecure work at 0.75 (9 of 9).

The norms parameter, by contrast, is highly variable (0.54–0.99); total prevalence is essentially indifferent to norms within the 16-year window; the same total prevalence can be reached at substantially different norms levels. Although Section 7.5 first surfaced this in an extreme scenario, the per-lever experiment confirms the redistribution at realistic lever sizes: norms reduction is structurally valuable for equality and carries near-zero cost to total prevalence,

but it nudges men's prevalence. Norms is thus an equality lever whose small cost is borne by men.

This finding also highlights a structural property of the model: norms are a bottleneck for gender equality but not for overall prevalence of poor mental health. The norms entrenchment loop sustains the conditions that produce the gender gap, but its direct impact on prevalence levels is too weak and slow to act as a meaningful lever for reducing the total burden of poor mental health within a 16-year intervention window (2024–2040). Norms operate on a 45-year timescale; a longer simulation horizon might reveal effects on total prevalence that are invisible within this window. Within the simulation, norms do reduce — from 0.54 in 1992 to 0.40 in 2024 and 0.32 in 2040 — so the parameter is not inert. The point is that this substantial change in the norms stock propagates to mental health prevalence only indirectly (through hours parity, sorting into insecure work, and the pay gap) and on a slower timescale than the optimisation window resolves. The small impact of norms on prevalence is aligned with the findings of earlier sections and refines them. Section 7.4 identified the norms entrenchment loop (R3) as the most persistently active informational loop but its pathway to mental health prevalence is indirect and weak contributing only 2% to the (poor-MH-specific) loop dominance analysis in Section 7.2.

7.6.4 Are the two objectives in tension?

The trade-off between the gender gap objective and the total prevalence objective is small. Across the base-case scenario, the entire Pareto front spans 0.005 on each objective; a gap-optimal solution achieves a gap of 0.057 and a total of 0.187 whereas a total-optimal solution achieves a gap of 0.062 and a total of 0.182 (Table 7-11).

This narrow span reflects only how little is given up in moving between the optimiser's two best solutions — the one that minimises the gap and the one that minimises the total. The two objectives are reached through mechanisms with different gender profiles, and that is what makes the choice of lever distributional rather than neutral: the equality levers narrow the gap by lowering prevalence among women, at some cost to men, while reducing insecure work lowers the total but widens the gap, because it lifts men's

recovery. The question is therefore not whether the two objectives compete, but in which macroeconomic context.

Table 7-11: Difference in gender gap and total prevalence in base case under different optimised objectives

Illustrates the marginal trade-off costs between the Pareto-optimal extremes for total prevalence versus gap minimisation. It highlights that the structural trade-off between the two objectives is small.

	Gap	Total	insecure work_scale
Gap-optimal extreme	0.057	0.187	1.00
Total-optimal extreme	0.062	0.182	0.75
Trade-off cost	0.005	0.005	

Appendix T shows the results of the combined optimisation for each of the nine macroeconomic future. Childcare and the pay gap converge to their full-intervention bounds (1.50 and 0.50) in every future. The total prevalence optimiser pushes insecure work to its minimum, while the gap optimiser does the opposite (keep at its maximum) in two key scenarios: under base case conditions (S4) and high demand with gradual casualisation (S8). This finding was confirmed by examining the effect of insecure work on recovery at 2040 under both configurations (insecure work_scale = 0.75 or 1, both with GPG = 0.5, CC = 1.5). The same result held with norms_scale set to 1.0 (no norms intervention) or 0.5 (maximum intervention), confirming that the effect is driven by the gender equality levers of GPG and childcare, not by norms reduction. Similarly, the gap optimiser pushes norms close to its bound (0.5) across scenarios whereas the total optimiser keeps it at its higher end.

The Pareto front is not uniformly present across all macroeconomic scenarios. Of the nine scenarios, four exhibit a divergent front i.e. conflict between the two objectives (insecure work_scale ranging across 0.15 or more of the parameter space), two show a convergent front i.e. general alignment between the two objectives, and three are unified fronts where both objectives agree on the same parameter values (i.e. no trade-off exists).

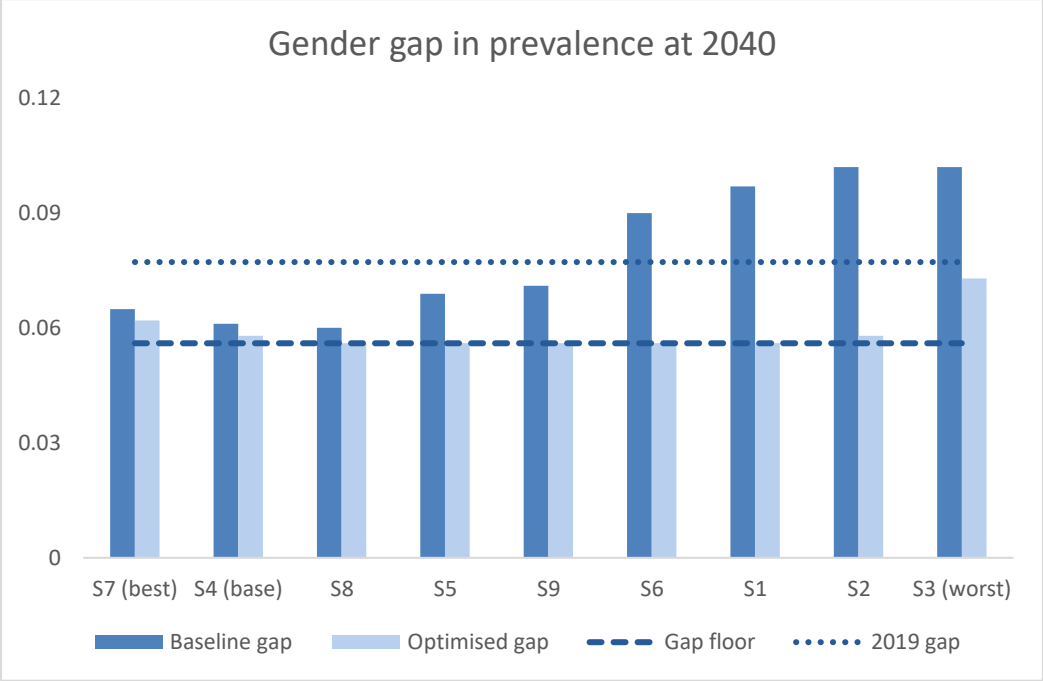
The trade-off emerges under low and medium precarity futures (S4, S5, S8, S9): high enough that reducing it meaningfully helps total prevalence, but not so high that the gap optimiser also benefits from its reduction. Under the high-precarity

futures (S2, S3), the system is under enough stress that both objectives benefit from maximum reduction, and the trade-off disappears. Similarly, when the system is operating comfortably in lowest precarity (S7), the objectives align.

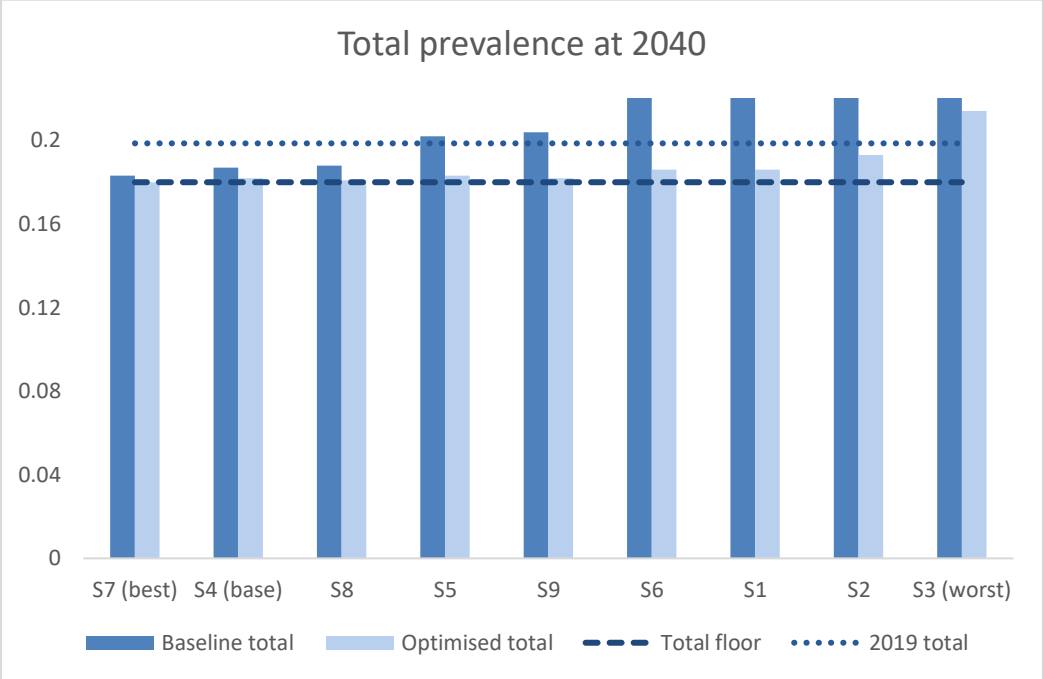
As a further robustness check, the combined-objective optimisation was repeated using the cross-scenario optimisation, which maximises the average payoff across all nine macroeconomic scenarios simultaneously rather than optimising within each scenario individually. Childcare and gender pay gap levers were held at their bound while insecure work and norms were optimised. The scenario-specific trade-off identified in the per-scenario analysis, where insecure work_scale varied between 0.75 and 1.00 depending on macroeconomic conditions, disappears entirely when the optimiser must find solutions that perform well on average across all futures. When policy must be robust to macroeconomic uncertainty, maximum deployment of all four levers is optimal for both objectives simultaneously.

7.6.5 How much improvement is possible?

The no-intervention baselines vary widely across the macroeconomic grid. The worst scenario (S3: weak demand with rapid casualisation) produces a gap of 0.102 (67% wider than in the base case) and a total prevalence of 0.304 (62% higher than the base case). The best assumed future of lowest precarity (S7: high demand with stable job quality) produces a gap of 0.065 and a total of 0.183. The 1.7-fold variation across baselines is driven almost entirely through the labour market channel: the insecure work level varies from 0.277 to 0.504 with little variation in gender dynamics variables (such as norms and gender pay gap). The detailed scenario-by-scenario baselines and optimised outcomes are in Appendix T; the narrative below highlights the convergence pattern that emerges.



Panel A: Gender gap in prevalence



Panel B: Total prevalence

Figure 7-17: The structural floor across macroeconomic futures

Two panels compare no-intervention baselines (dark blue bars) with optimised outcomes (light blue bars) across the nine macroeconomic scenarios (S1–S9; three levels of labour demand (0.90, 1.00, 1.05) crossed with three rates of insecure work deterioration (none, moderate, high); S4 is the base case). Panel A shows the gender gap in prevalence of poor mental health at 2040; Panel B shows total prevalence at 2040. The dashed lines mark the structural floor for each metric (gap: 0.056; total: 0.180). The figure highlights two findings. First, the no-intervention baselines vary by a factor of ~1.7 in both metrics, driven almost entirely by the labour market channel. Second, the optimised outcomes converge to a narrow band (~0.056–0.073 for the gap; ~0.180–0.214 for total) regardless of how far the no-intervention baselines diverge. Within the model, the structural interventions function less as universal prevalence reducers and more as buffers against macroeconomic deterioration: they bring the system toward a floor that is reachable with the levers at their policy bounds and that mirrors the 30% residual identified in Section 7.3.

Across eight of the nine scenarios, the optimised gap converges to a narrow band of 0.056–0.062 (only under reduced labour demand and rapidly increasing insecure work (S3), it reaches 0.073) and the optimised total prevalence to 0.180–0.214, regardless of how far the no-intervention baselines diverge. The reductions are therefore largest in the worst scenarios. The gap optimisation achieves its largest proportional reductions in the scenarios where the baseline gap is widest (S1, S2, S6: 37–43%), and its smallest reductions where the baseline gap is already near the model's minimum (S4, S7, S8: 5–7%). The worse scenarios are driven primarily by labour market deterioration and the optimiser can partially compensate by pushing incidence levers harder.

Similarly, the total optimisation achieves substantial reductions under adverse macroeconomic conditions (S1–S3, S6; 21-30%), whereas in the low-precarity futures (S4, S7, S8), the reductions are modest at 2–4%. Labour market deterioration drives both men's and women's prevalence well above the structural floor, creating substantial room for the optimisation to work; the combined effect of gender equality levers (reducing incidence) and insecure work reduction (improving recovery) can offset a large share of the macroeconomic damage.

The optimised gap converges to a narrow band, which points to a structural floor. Two quantities must be distinguished. The reachable floor is the best outcome attainable with the four levers at their policy-plausible bounds; a gap of ≈ 0.056 and a total prevalence of ≈ 0.180. The model-theoretic floor is the outcome when the same constructs are driven to their structural extremes (insecure work, the pay gap and norm reinforcement eliminated entirely, childcare expanded well beyond observed levels pushing the effect to its maximum); it is only marginally lower: and, tellingly, almost identical in every macroeconomic future: a total of 0.171–0.173 and a gap of 0.051–0.052 across all nine futures, less than 0.01 below the reachable floor on each objective (Appendix T). That the floor barely moves even under the total elimination of every modelled stressor confirms it is structural, not an artefact of where the lever bounds were set: the residual is carried by path dependence, the implied recovery rate difference and unmodelled factors (the 30% residual of Section 7.3), none of which the levers reach. The floor is a property of the model and its calibration, not of the real-world system. The optimisation can address the 70%

of the gap that is driven by identifiable structural mechanisms (incidence inequality, recovery sensitivity); it cannot address the residual that is embedded in the calibration.

Because the floor operates partly through the endogenous prevalence–recovery loop, its sensitivity to that threshold was tested directly. Re-optimising across the macroeconomic grid at halving points of 1.25, 1.40, 1.50 and 1.75 (Appendix T) shows the floor is stable from 1.40 upward: the base case and low-precarity scenarios are unchanged, and the worst-case optimised outcome moves by less than 0.01 in total prevalence. Only at a markedly lower threshold of 1.25 does the loop become self-amplifying, prevalence roughly doubles and the structural levers can no longer compensate. The reported floor is therefore robust across halving points of 1.40–1.75; the conclusion that the levers reduce prevalence toward an irreducible floor, rather than eliminating it, does not depend on the precise threshold within the plausible range.

Within the model, the structural interventions function less as universal prevalence reducers and more as buffers against macroeconomic deterioration: they bring the system toward the model's irreducible minimum regardless of starting conditions. The implication is that the value of intervention is greatest precisely where macroeconomic conditions are worst and the value diminishes as the system approaches the floor.

7.6.6 Summary

The optimisation refines the analysis in five ways.

First, it confirms the priority of the incidence channel: the gender pay gap and childcare levers converge to their bounds in every macroeconomic future for both objectives, identifying the channel that Section 7.3 isolated as the 61% contribution to the gap.

Second, it reveals an asymmetric hierarchy among the mechanisms: once the incidence levers are deployed at full strength, the recovery channel (insecure work) has a small and adverse effect on the gap: because women's exposure falls below the recovery-degradation threshold while men’s does not, reducing

insecure work slightly widens rather than closes the gap. This qualifies rather than contradicts the channel-independence result of Section 7.3, where over the historical period the incidence and recovery operated largely independently: once the incidence levers are held at their bounds, that independence becomes asymmetric and the recovery channel acts on the gap in the opposite direction rather than being neutral to it.

Third, it surfaces a structural floor in the model, an irreducible gap of ~0.056 and total prevalence of ~0.180, reachable with the levers at their policy bounds, mirroring the 30% calibration residual identified in Section 7.3.

Fourth, the relationship between the two objectives is conditional and distributional. The optimiser's Pareto front is narrow because both solutions deploy the same incidence levers; the per-lever experiments show that the gap and the total are reached through different gendered routes; the equality levers lower women's prevalence at some cost to men, while reducing insecure work lowers the total by improving men's recovery more than women's and so widens the gap. This tension is concentrated under moderate macroeconomic conditions as precarity rises. This places the redistribution onto men identified in Section 7.5 at realistic lever sizes rather than in an extreme historical scenario.

Fifth, few levers are sufficient: childcare paired with insecure work reduction reproduces the effects of all four levers closely on both objectives in every future.

7.7 Summary of findings

This chapter answers the research questions posed by this thesis, providing a structural explanation for both the persistence of gender gap and the population-wide deterioration in mental health; the subsections below answer the chapter's research questions in turn.

7.7.1 Gender gap in prevalence of poor mental health

The gender gap in poor mental health persists primarily because structural gender inequality raises the rate at which women develop poor mental health. The structural decomposition attributed 61% of the gap to this incidence

pathway, 8% to women's greater recovery sensitivity, and 3% to differential exposure to insecure work. When combined, the incidence and recovery pathways are mildly sub-additive totalling 67% rather than 69%. The remaining 30% is embedded in the calibration as a difference in the implied recovery rate, reflecting the accumulated consequences of historical disadvantage that predates the simulation period. The robustness analysis confirmed that these proportions are stable across alternative specifications of the recovery function.

The 30% residual points to a recovery environment that is structurally less forgiving for women: women's recovery capacity is degraded at a lower threshold of insecure work exposure. One interpretation relates to the double burden: for women, insecure employment coexists with unpaid work that erode the benefits employment provides for recovery: social participation, structured time, and financial security. The result is a system where the door into poor mental health is wider for women and the door out is narrower, and both are shaped by the same structural forces.

A related implication concerns the gendered threshold at which insecure work neutralises the recovery benefit of employment. Because the model anchors this effect relative to each gender's 1992 baseline, and women began the period with structurally higher exposure to insecure work than men (0.37 versus 0.24), women's recovery benefit is neutralised after a smaller relative increase in insecure work — roughly 20%, at an absolute level of 0.44 — whereas men's requires roughly a 40% increase (absolute level 0.34). In absolute terms women's recovery tolerates more insecure work before neutralisation. This asymmetry is a form of path dependency: it reflects women's higher historical starting exposure embedded in the initial conditions, not a differential intrinsic vulnerability to a given level of insecure work. This mechanism explains why the net recovery benefit of employment can be weaker for women: not because employment intrinsically benefits them less, but because their higher baseline exposure leaves less headroom before insecure work offsets that benefit.

The apparent paradox that equality indicators improved while mental health worsened is explained within the model by the interaction of timescales and by how the model operationalises the perceived burden of inequality. The gender pay gap narrowed over the simulation period (a slow improvement in incidence),

but insecure work rose simultaneously (a medium-speed deterioration in recovery), and norms eroded only partially (a long-delay structural constraint). Critically, the model represents the burden of unpaid work not as the objective division of labour alone, but as its departure from parity weighted by the inverse of gender norms. As norms erode toward equality, the same actual division of unpaid work is experienced as a greater stressor because it increasingly conflicts with egalitarian expectations. This means that partial norms erosion can simultaneously improve some indicators of gender equality (such as the gender pay gap) while intensifying the perceived burden of remaining inequality. The net effect within the model was that the incidence improvement from a narrowing pay gap was offset by rising insecure work degrading recovery and by the intensification of perceived inequality.

Within the simulation, narrowing the gender gap without correspondingly increasing prevalence among men necessitates simultaneous adjustments to parameters governing both gender equality and labour market quality. Accelerating equality narrows the gap by lowering women's prevalence, but the childcare and norms levers that act through parity in working hours shift insecure work onto men and so raise men's prevalence, most clearly under low- and medium- precarity futures, easing as precarity rises. Conversely, focusing solely on reducing insecure work lowers the total but, in the low-precarity futures, widens the gap, because it restores men's recovery more than women's. Only the two together narrow the gap while lowering prevalence for both.

7.7.2 Rise in prevalence of poor mental health

The historical rise in poor mental health for both genders was not primarily driven by the 2008 recession. Within the model, the recession functioned as an activating event for pre-existing structural vulnerabilities rather than as an independent cause of the mental health trend. The counterfactual analysis showed that removing the demand shock reduces the total by only 4%, whereas stabilising insecure work reduces it by 7% and combining the two added nothing beyond insecure work alone. The structural vulnerability of rising insecure work, persistent gender inequality, slow norms erosion was already building before 2008. The recession activated dormant feedback loops, but those loops were dormant because the system was not yet sufficiently stressed, not because the

stress was absent. Within the model, the recession's mental health cost is carried almost entirely by the rise in insecure work it accelerated, not by the demand shock itself.

Within the model, economic shocks do not immediately translate into mental health crises; the reinforcing loop requires both the health causation and the health selection pathways to propagate, the full R1 reinforcing cycle therefore takes 2–3 years to reach peak dominance, as observed between the 2008 onset and the 2011 peak. This temporal architecture is a structural feature of the simulation with implications for the timing of policy interventions: interventions deployed during or immediately after an economic shock could prevent the feedback loop from reaching the dominance it achieved in 2011. Waiting for mental health indicators to signal a crisis means intervening after the vicious cycle has already become the dominant force in the system.

7.7.3 Structural levers

The optimisation and the counterfactual analysis together identify a hierarchy of structural levers within the model. The optimiser deploys childcare and gender pay gap levers to their bounds across every macroeconomic scenario, for both gap and total objectives, but the two differ in their distributional cost: the pay gap reaches only women, the model carrying no direct pathway from the pay gap to men, whereas childcare lever narrows the gap partly at men's expense, by shifting insecure work onto men through working-hours parity. Gender norms reduction adds an equality benefit at near-zero cost to the total, with its small cost again falling on men. Insecure work reduction is the population lever: it lowers total prevalence under every future, but in low-precarity futures it widens the gap, because it improves men's recovery more than women's. The mental health to labour market link (i.e. health selection pathway) is partially system-protective: severing it worsens prevalence for both genders because labour market exit increases labour market tightness, reduces insecure work and shields the recovery capacity of those who remain. A structural floor of ~0.056 (gap) and ~0.180 (total) bounds what any policy-combination of these levers can achieve in this model; it is a property of the calibration, not of the world.

7.7.4 Synthesis

The strength of the findings lies in the internal consistency demonstrated by the convergence of findings across multiple analytical approaches. These methods all point toward the same structural interpretation: women are pushed into poor mental health at higher rates by gendered structural pressures; they are held there for longer by a recovery environment that degrades sooner; the two channels are temporally sequenced, with incidence dominating the first two decades and recovery dominating the last, and once the incidence channel is addressed, the recovery channel does not become irrelevant to the gap – in the low precarity futures, reducing insecure work benefits both men and women, lowers the total yet widens the gap because men benefit more. While this convergence does not constitute independent validation as all methods utilise the same model, it establishes that findings are consistent with every analytical lens this chapter applies. None of these claims requires the simulation to forecast the future of Scottish mental health; what they require is the model's ability to hold incidence, recovery, exposure, and path dependence simultaneously and to test them against the historical record.

The implications of these findings for the broader literature on gender, the labour market, and population mental health, for Scottish policy and the contribution of a system dynamics approach to the public-health evidence base are developed in Chapter 8.

7.8 Strengths and limitations

The central findings converge across four complementary analytical lenses — loop dominance, structural decomposition, counterfactual analysis and forward optimisation — rather than resting on any one. The structure they interrogate has already passed the behavioural-sufficiency tests of Chapter 6. The results are robust to the alternative functional forms tested (the prevalence-on-recovery curve, the childcare form, and the recovery halving point), so they do not hinge on a single modelling choice. Despite these strengths, several limitations specific to the analysis qualify the interpretation of the findings.

First, it is important to note that LTM scores are relative i.e. the share of explained variance in stock rates of change at each timestep; they are not measures of absolute mechanism strength. This means that LTM is better understood as when each mechanism becomes detectable. Additionally, LTM dominance is sensitive to rates of change, not levels: a structurally important variable that is high but stable contributes little, which is partly why step change years like 2017 produce large changes in loop dominance. This means that one-year dominance should not be read as regime shifts.

Second, the 30% residual in the structural decomposition is a direct consequence of the calibration procedure. The implied baseline recovery rate absorbs every structural pressure operating in 1992 that the model does not attribute to a specific mechanism. Different starting assumptions, alternative data sources, or a longer historical calibration period could produce a different residual. This is a boundary of the model, not a finding about the world. The residual, together with the portion of the post-2016 rise the model does not reproduce (as discussed in Chapter 6), is also where unmodelled drivers reside: welfare conditionality, the rollout of Universal Credit, and the benefits freeze, alongside the degradation of mental health services, are leading candidates. These lie outside the model boundary, and the model is accordingly not used to evaluate welfare policy directly.

Third, all gender-specific incidence pressures are channelled through a single inequality variable. The 61% contribution attributed to this variable reflects both the importance of gender inequality as an incidence driver and the structural parsimony of the model. This also contextualises the LTM findings, where economic shocks (such as the unemployment-MH loop) exhibit much higher relative dominance for women, as these shocks compound multiplicatively with the chronic stress of gender inequality. A more disaggregated model with separate pathways for the gender pay gap, unpaid work, occupational segregation, and other dimensions of inequality could reveal that some components matter substantially more than others. The 61% should therefore be read as the collective contribution of the incidence channel rather than a precise measure of any single dimension of inequality. Moreover, this captures aspects of both – the double burden (perceived unpaid work) and gender inequality (gender pay gap), which are both also inherently absorbed into the

gendered differences in starting stocks and the implied baseline recovery rates. Consequently, the model cannot definitively partition the total standalone effect of the double burden.

Fourth, the optimisation analysis projects the model to 2040, twenty-one years beyond the calibration period. The model's structural relationships i.e. the feedback loops, the delay structures, and the sensitivity functions, were estimated from 1992-2019 data. Whether these relationships hold under future conditions is unknown. At the same time, 16 years from intervention activation in 2024 is short relative to the timescale of gender norms that erode through cohort replacement. Effects of this on total prevalence that would only manifest beyond the projection horizon are not captured. The optimisation results should be interpreted as revealing the model's structural sensitivity to parameter changes, not as forecasts of what specific interventions would achieve.

Fifth, there is an implicit assumption of independence in scenario analyses. The decomposition and counterfactual scenarios modify one or two model parameters at a time and report the resulting change in the gender gap or total prevalence. This implicitly assumes that the modified parameters operate independently of the parameters held fixed. The near-additivity of decompositions D1 and D2 is empirical evidence that this assumption holds for the incidence and recovery channels within the model, but it cannot be tested for parameters or pathways that are excluded from the model; notably, there is no modelled relationship from gender inequality variables (gender pay gap and unpaid work burden) to men’s outcomes. It is possible that scenarios in which there is more equality, there are direct consequences for men, as seen through the reduced sorting mechanism resulting in higher insecure work for men.

Sixth, the quantitative magnitudes throughout this chapter are conservative. The integrated model under-simulates the post-2016 rise in women's prevalence by approximately 25% by 2019, and the linear specification of the labour market and gender dynamics modules shields the system from the sharper cascades that can follow once mental health crosses a threshold. Reported effects, particularly for women and for the most recent years, should therefore be read as lower bounds on the strength of the mechanisms rather than as central estimates.

Seventh, insecure work enters the model as an exogenous trend. There is no modelled feedback from labour demand to the level of insecure work, except indirectly through labour market tightness. The model therefore cannot capture the possibility that the 2008 recession itself raised the prevalence of insecure work; the insecure work trajectory is taken as given. This qualifies the finding that accumulated insecure work contributed more than the demand shock: the two are not fully independent, and some of the insecure work accumulation may itself be a downstream consequence of the recession that the model treats as separate. The effect of insecure work on recovery is also specified relative to each gender's 1992 exposure, with no effect at or below that level. The widening of the gap under insecure work reduction in S4 and S8 depends on this specification; a function based on absolute exposure, or one allowing a benefit below the 1992 level, could change this.

Eighth, the levers are bounded by the model. They were identified from the model structure and the preceding analysis rather than with policy makers, and their ranges were set by judgement. Although most of the levers sit well within the current policy context, norms are not necessarily tractable for policy makers, and engaging them could assess the feasibility of intervening, for example by changing the structure of parental leave. The levers are aligned with principles of intervening on a stock (through inflow and outflow) and finding leverage in a system (Meadows, 1999) by intervening on dominant loops (insecure work), changing the underlying paradigm (through norms). However, other levers such as changing the penalty associated with part-time work, which would change the rules of the system, were not tested. Given the undershoot of the rise in prevalence and limited impact of the loop, breaking the reinforcing prevalence trap was not tested as a lever. Labour demand, the only exogenous variable not modelled as a lever, could itself be treated as a lever but is modelled only as a source of uncertainty. The impact of increasing labour demand is therefore an indirect inference: increasing labour demand alone slightly reduces the total but increases the gender gap (S7 vs S4 no intervention; Appendix S). Once the other levers are in place, there is little gain in total and gap reduction when labour demand is higher. Additionally, levers outside the boundary, such as the welfare system and mental health services, cannot be tested. Finally, the levers are represented as changes to structural parameters

rather than as specific policies, and the model contains no costs: it can indicate where intervention acts and in which combination, but not which policy instrument would achieve a given change, or at what cost.

8 Discussion and Conclusion

8.1 Introduction and summary of findings

This thesis set out to develop and test a dynamic causal explanation of the structural mechanisms producing the rising prevalence of poor mental health, and the persistent gender gap in that prevalence, among Scotland's working-age population between 1992 and 2019.

Chapter 1 framed this as the gender paradox: despite measurable progress on gender equality indicators such as the narrowing of the employment and the gender pay gaps, the gender gap in the prevalence of poor mental health in Scotland has not closed.

Chapter 2 argued that this trajectory cannot be adequately explained by the linear or static methods that have dominated epidemiology, because it is generated by circular causality, by historical accumulation in slow-moving stocks, and by the multiplicative interaction of distal and proximal determinants operating at different timescales. Three research questions followed: how reciprocal feedback between labour market conditions and population mental health has contributed to the rise in poor mental health; why the gender gap has persisted; and which structural mechanisms are dominant within the system and can act as leverage points for intervention.

Chapters 3 and 4 explained the methodology adopted to answer these questions, highlighting that the thesis is grounded in complex realism and pursues explanatory verisimilitude rather than prediction.

Chapter 5 synthesised the literature into a causal loop diagram which was refined through expert consultation. It identified a causal hypothesis for the observed trends that included two reinforcing loops between the labour market and poor mental health, operating for both men and women (the unemployment-mental health trap and the job quality decay). It also identified a reinforcing loop that impacted only women’s mental health, acting through gender norms and unpaid work burden. Further, mechanisms in play within the labour market and gender norms (demand saturation and norm reinforcement) interact with

the identified loops, creating balancing mechanisms such as the health-regulated job quality and gender power rebalance.

These loops together capture one explanation of why the literature often reports lower adverse effects of unemployment on mental health for women: the benefits of employment are lower for women within the context of higher insecure work, gender pay gap and unpaid work burden. More importantly, this can lead to the persistence of the gender gap in prevalence of poor mental health despite increased economic participation. Finally, the causal loop diagram identified labour market dynamics that reinforce insecure work and have likely contributed to the rising prevalence of poor mental health even when unemployment is low.

Chapter 6 translated that qualitative causal theory into a system dynamics simulation model, and demonstrates that it offers one plausible explanation that can endogenously replicate the observed trends. Along with evidencing the model's behavioural sufficiency, the chapter also identified its limits, most notably the inability to reproduce the magnitude of the rise in prevalence of poor mental health among women from 2016; these are likely caused by forces outside the model boundary: the model represents the labour market and gender norm pathways to mental health, which are only part of the wider system responsible for the prevalence of poor mental health. This shortfall is a limitation, but it is also reassuring: a model able to reproduce the entire trend from the included factors alone would be implausible, and would risk over-claiming for the labour market–gender norms structure at the expense of forces that lie outside the model boundary.

Chapter 7 reports the analyses conducted on the model which show that:

- the gradual rise of insecure work, which began before the 2008 recession, contributed more to the rising prevalence of poor mental health than the demand shock of the recession itself. The model cannot, however, account for the impact of the recession on the level of insecure work in the labour market, outside of the effect of labour market tightness. The levels of insecure work might have been lower had the 2008 recession not occurred but the model cannot fully capture this.

- the additional pressures on women, operationalised through the gender pay gap and unpaid work burden, are the largest contributors to the higher prevalence among women, within the model. Other structural differences between men and women that act on incidence are subsumed within this factor in the model. Once women are channelled into prevalence of poor mental health at higher rates, they face slower recovery, hypothesised to be partly due to greater exposure to insecure work, the double burden, and the greater impact of reduced support services.

- progress on gender pay gap, in this model, could reduce both the gender gap and total prevalence of poor mental health. The model contains no direct feedback between gender pay gap and men's outcomes. It is possible that a reduction in gender pay gap has indirect consequences for men's wages and their wellbeing that fall outside the model scope. Moving towards egalitarian norms is beneficial for women and reduces the gender gap in prevalence but may increase prevalence among men. Similarly, childcare narrows the gender gap at a cost to men (i.e. it redistributes rather than reduces the prevalence burden). However, the analysis shows that the reduction in gender gap, although increases men's prevalence at times, does not come at the cost of higher total prevalence. Reducing insecure work reduces prevalence for both men and women, and thus the total, but it widens the gender gap under low-precarity futures (where men's recovery is degraded but women's is not), a result that depends on how the model specifies the effect of insecure work on recovery, whereas it narrows the gap as the labour market deteriorates (where women's recovery is also degraded).

The remainder of this chapter situates these findings in the literature (Section 8.2), articulates the contribution of this thesis (Section 8.3), including what the framework offers for understanding population mental health and for decision-making, draws out the implications for policy (Section 8.4), reflects on the key limitations (Section 8.5), identifies directions for future research (Section 8.6), and closes with concluding remarks (Section 8.7).

8.2 Situating the findings in the literature

This section situates the three findings on mental health against the empirical and modelling literature: the role of the 2008 recession in the rise, the persistence of the gender gap, and the structural levers.

8.2.1 The rise in prevalence of poor mental health in Scotland and the 2008 recession

Within the model, the 2008 recession is a driver of incidence — it enters through the rise in unemployment that feeds the economic stressors effect on onset — yet the analysis finds it acted not as an independent cause of the rise in poor mental health so much as an activator of slower-moving structural vulnerability that had been accumulating beforehand. Barr, Kinderman and Whitehead (2015) find that trends in unemployment and wages explained only about a third of the post-2008 increase in reported mental health problems in England, and could not account for its continuation between 2010 and 2013. Among the explanations they advance for the unexplained majority, two are directly relevant here. First, for the rise among those in work, they point to the growth of more precarious forms of work; this account closely supports this model's finding of insecure work as the principal driver. Second, for the larger rise among those out of work, they implicate post-2010 welfare reform, a force this model deliberately holds outside its boundary, and whose omission is consistent with its conservative reproduction of the most recent period.

What a repeated cross-sectional time-trend cannot capture, and what this model adds, is the reciprocal feedback through which these pressures operate. This analysis shows that the reinforcing pressure built up over the crisis years was led first, in 2010, by the rise in insecure work rather than by the demand shock itself, while the unemployment cycle reached its own peak only in 2011. Consistent with their observation that poor mental health began to rise slightly after unemployment did – a lag the model imposes as an assumption rather than derives – the modelled rise continues until 2013; recovery (drain) loops reassert themselves from 2012, as the stock stabilises after a delay.

Two further studies corroborate that the 2008 recession was not the main driver of the rise. Dykxhoorn et al. (2024), using UK primary-care records, find recorded incidence of common mental disorders rising mainly to about 2006 and then broadly stable, with no clear 2007–08 break, and in the devolved nations including Scotland stable to 2012 before declining. Zhang et al. (2023) find the prevalence of psychological distress in Great Britain broadly flat until about 2015, with the clearest deterioration emerging only thereafter. The two studies use different measures but they converge on the same conclusion: the 2008 recession did not produce the sharp, sustained break a simple shock-as-cause account would predict, and the trend is better explained by the gradual structural processes this model makes explicit.

The model reproduces the trend but falls short of its magnitude in the late 2010s, under-reproducing the rise in women's distress. In principle such a shortfall could signal a mis-specified structure rather than a missing one, but two features of the residual make it instead the strongest indirect evidence that forces outside the model boundary did substantive work in this period. It is the magnitude, not the pattern, that is missed: the model reproduces the shape, timing and gendered divergence of the trend — the standard appropriate to a claim of behavioural sufficiency rather than point prediction. And the residual is patterned, not random: the model under-shoots for women in the late 2010s, coinciding with the retrenchment of social security, which the causal literature independently identifies as an important driver but which this model excludes. Simpson et al. (2021) in their systematic review conclude that contractionary reforms (those reducing eligibility or generosity) worsen mental health and widen inequalities, while expansionary reforms do the reverse. The evidence on Universal Credit, using different measures, converges: Wickham et al. (2020) find clinically significant distress (GHQ-12) rising among unemployed claimants relative to a non-eligible comparison group; Marimpi et al. (2025) find a sustained population decline in mental functioning (SF-12) extending beyond the unemployed to employed claimants; and Baxter et al. (2026) find adverse effects across subjective wellbeing measures that were comparatively greater in Scotland and Wales than in England. Additionally, the shortfall is gendered; Thomson, Niedzwiedz and Katikireddi (2018) show that gender inequalities in poor mental health narrowed after the recession but widened under austerity, as

women's prevalence rose after 2012 while men's brief recession-era spike receded. This does not invalidate the model but helps qualify its findings. Because the endogenous structure excludes the welfare system, its recent-year magnitudes should be read as conservative lower bounds, and the size and gendering of the residual point beyond the labour market and gender norm architecture to forces the model deliberately holds outside its frame.

8.2.2 The persistence of the gender gap

This thesis explains the persistence of the gender gap as a difference in women's social position. This view is shared by a body of work that sets out to explain the gap by decomposing it into structural causes. Within the model, the gap is dominated by incidence: the door into poor mental health is wider for women due to gender-specific stressors. The model's explicit differential exposure term is small, but the labour market inequalities these external studies isolate are captured elsewhere in its structure: the gender pay gap acting through incidence, the path-dependent residual absorbing accumulated historical difference, and women's greater sensitivity to insecure work carrying an interaction between labour market conditions and gender roles. This section situates the finding against studies that, by a similar logic, apportion the gender gap to such structural factors.

The most direct of these are two causal decompositions that quantify how much of the gap the labour market accounts for. Gueltzow et al. (2023), simulating a synthetic cohort from US panel data, estimate that equalising women's labour market position (employment, occupation and income) from age 50, would close about a third of the depression gap among older adults. Orchard et al. (2024) apply the same logic to multidimensional mental health, using Monte Carlo simulation on Canadian data to ask how the gap would change if women held men's employment status and job quality: equalising the employment profile narrowed the gap in the most severe cluster of clinical mood disorder, but widened it in the best mental health cluster, leaving even fewer women in “flourishing, no life stress”. Employment thus accounts for part of the gap; gendered employment differences, the authors argue, were masking rather than explaining the wellbeing gap, the remainder reflecting gendered expectations of roles outside work, caregiving and household responsibilities. Platt et al. (2020)

show that the US depression gap narrowed across birth cohorts as women's relative social position improved, supporting the view that the gap is structurally produced. The strongest mediator, however, is found to be women's educational catch-up, while changes in employment and housework played no significant role. The authors attribute this to how these channels were captured: expanded employment brought little benefit because women's jobs remained part-time, temporary and lower-paid, and population-level housework hours miss the perceived equality of the division.

The point is reinforced when the exposure set is widened to consider the social context of employment. In their gendered exposure model that uses multilevel path analysis to examine the effect of work and family stressors on psychological distress among workers in Canada, Bilodeau, Marchand and Demers (2020) find that these factors fully explained the gender difference in psychological distress. They identify gender differences in work-family conflict, decision authority, single-parenthood and self-esteem as significant mediators of psychological distress. These analyses converge on the thesis's finding of a gap rooted in women's structural position. Finally, Kuehner (2017) and Patten and Lee (2004) trace the gender gap in prevalence to incidence (rather than recovery). As Kuehner's (2017) review makes clear, however, women's higher prevalence is multifactorial and includes developmental, biological and interpersonal processes that emerge around puberty.

8.2.3 The structural levers

The thesis finds that coordinated levers achieve what isolated ones cannot, that levers can redistribute poor mental health as much as reduce it, and that the deepest lever is the slowest to move. The first two find limited but convergent support in the systems modelling literature. The finding that levers act better together than singly, and that certain combinations may produce most of the benefit, has its closest support in similar systems modelling (Occhipinti et al., 2021). The distributional caution is not new: Cerdá et al.'s (2014) agent-based simulation of violent victimisation finds that universal interventions reduced population-level rates without reducing the inequalities between groups. The evidence base, however, is thin: Shah et al.'s (2021) umbrella review finds that

the population-level evidence on the structural determinants of mental health is small and largely of low quality.

Intervening directly on norms has a greater effect but it is also the slowest and least tractable lever. In Meadows' (1999) hierarchy of leverage points, the four levers sit at different depths: gender norms at the level of paradigm (level 2); the sorting into insecure work targeted indirectly through norms and childcare at the level of rules (level 5); childcare at the strength of the reinforcing loop it feeds, the working-hours parity that erodes norms (level 7); and the gender pay gap as a parameter that also drives the reinforcing mechanism of norms (level 12). Using this typology of intervention, norms offer the highest leverage. The counterfactual analysis of Chapter 7 bears this out: a shift toward more egalitarian norms produces the largest single-lever reduction in the gender gap. However, the deepest leverage point is also the hardest to shift. Within the feasible policy range tested and the simulation window, intervening on norms only marginally accelerates the erosion of the slow norms stock, so its realised effect is the weakest of the four. Structural leverage and tractability therefore pull in opposite directions: the achievable gains fall to the faster, lower levers.

Current debates around the future of work have increasingly turned from the number of jobs that artificial intelligence (AI) might displace toward its effect on the quality of the work that remains (Institute for the Future of Work, 2025). This emphasis on the transformation of work resonates with the present model's finding that insecure work is the dominant lever on the total burden of poor mental health, and has contributed more to the rise than the demand shock of the 2008 recession. However, Chapter 7 notes that reducing insecure work reduces the total prevalence of poor mental health but could preserve, and in some modelled futures widen, the gender gap in prevalence; an AI effect that falls unevenly across genders could move the gap either way. The model cannot settle this, because it represents neither mechanism at issue: it sorts by gender parity in working hours rather than by occupation, so it misses which jobs AI would expose. The modelled labour market clears through aggregate tightness alone, so it misses the reskilling that shapes who reaches good work. The gendered sign of the effect is itself unsettled: women's greater exposure to jobs exposed to AI carries both greater risk and, through a higher potential for complementarity, greater opportunity (Cazzaniga et al., 2024).

The model does point to one route the debate's focus on occupational exposure overlooks. Because it carries unpaid work as a departure from parity in weekly housework hours, anything reducing the hours women spend on it (without a proportionate reduction for men) would ease their incidence and the role strain with it, even with the underlying norms unchanged; a faster channel than the slow erosion of norms. AI tools that cut the time cost of running a household would, on the model's logic, act on women's mental health.

Across these three threads the literature converges. What these studies establish piecemeal, this thesis assembles into one dynamic structure in which the mechanisms interact over time. That structure is the contribution articulated in the next section.

8.3 Contributions of the thesis

The contributions of this thesis rest on a reorientation: where epidemiology has largely studied poor mental health as a set of individual-level risk factors, this thesis treats it as a population-level system of interacting feedbacks. This leads to the four contributions that follow: the dynamic causal structure built from fragmented evidence, the treatment of gender as a structural determinant, the model as an explanatory thinking tool, and the methodological turn to feedback-driven enquiry. This section discusses each of these in turn, closing with what each offers for understanding population mental health and for decision-making.

8.3.1 A dynamic, population-level causal structure

The first contribution of this thesis is a dynamic, population-level hypothesis built by critically integrating fragmented evidence drawn from epidemiology, economics and sociology. The mutually reinforcing relationship between population mental health and the macroeconomy has long been theorised (Bloom & Canning, 2008; Naik et al., 2019; Thomas et al., 2023). Those literatures, however, stop short of specifying the pathways by which mental health feeds back onto its own macroeconomic determinants, and they do so without operationalising the temporal accumulation that makes the feedback dominant only over particular windows. The model developed here closes that gap by representing mental health as a dynamic stock that is modified by the labour

market, and by specifying the causal relationships through which the reverse channel of health selection runs.

The hypothesis is dynamic in four specific senses: it decouples incidence from recovery, it assigns determinants to different temporal scales, it couples health selection with health causation in a single feedback structure, and it relocates the explanation from the individual to the population level. The first part is, in part, a product of the modelling approach. The epidemiological literature reviewed does not systematically distinguish which social determinants act on the onset of poor mental health and which on its duration. By representing prevalence as a stock governed by distinct inflow (incidence) and outflow (recovery) processes, and by permitting structural factors to act differentially on each, the model makes this distinction operationally explicit. Where other system dynamics models of depression, and of disease more generally, typically represent the condition through multiple states; for instance separating severity, recurrence, or diagnosed from undiagnosed cases (Darabi & Hosseinichimeh, 2020; Graham, Gariepy, & Orpana, 2023), the present model uses only two prevalence stocks per gender, which is what allows the social determinants to be traced through the two channels. The finding that gender inequality acts primarily through the incidence channel while insecure work acts through both is therefore best read as a hypothesis generated by the model, grounded in the chronic precarity literature, and available for empirical testing using longitudinal data.

This grounds a broader methodological proposal. The dominant frameworks for the social determinants of mental health classify determinants into structural and intermediary (Solar & Irwin, 2010). As Chapter 2 argued, following Krieger (2008), such hierarchies capture distance from the outcome but conflate it with the speed at which a determinant acts, leaving them without a way to separate slow-moving structural conditions from fast-moving shocks. This thesis builds on that critique in two ways. It operationalises the missing timescale dimension explicitly — representing determinants that move over generations (gender norms) alongside those that move faster (labour market shocks) — and it adds a distinction the literature has not systematically drawn: whether a determinant acts on the onset (incidence) or on its duration (recovery). The gender pay gap and the prevalence of insecure work occupy the same level of the hierarchy, yet

they call for different interventions once it is recognised that the first acts primarily on incidence and over a slow timescale, whereas the second acts on both incidence and recovery and over a faster one.

These two distinctions also bear on how population mental health is understood. Prevalence, the usual measure of population mental health, conflates the two flows that produce it. Within the model, women's incidence continues to decline across the later period, yet their prevalence does not fall: it is sustained by a failure of recovery (Section 7.2). A prevalence series alone would not distinguish these. For decision-making, the framework therefore suggests asking of any policy which flow it acts on, and when its effect would appear. Improving job quality, for instance, is not only preventive: because insecure work also degrades recovery, it also acts on how long poor mental health lasts. Interventions targeting gender-specific incidence pressures would produce earlier benefits, whereas those addressing recovery capacity would become important only after prevalence has already risen (Section 7.3). Because the reinforcing cycle between unemployment and poor mental health took two to three years to reach dominance after 2008 (Section 7.4), intervention during or immediately after an economic shock could prevent it from doing so. It follows that progress is better tracked on the flows into and out of poor mental health than on prevalence alone.

8.3.2 Gender as a dynamic structural determinant

The second contribution adds to a small body of system dynamics work that locates women's disadvantage in system structure rather than in individual attributes: most directly Palmer’s (2017) account of how gender disparity is built into the Norwegian pension system. This thesis extends this by operationalising gender as a structural determinant of mental health. Gender norms are represented in the model as an intangible stock that erodes through cohort replacement and is reinforced endogenously by the economic signals (the gender pay gap, working-hours parity, the economic inactivity gap) that the norms themselves produce. This representation accomplishes three things that the standard epidemiological treatment of gender as a static covariate cannot.

First, it produces a coherent structural explanation for the persistence of the gender gap despite measurable equality progress. The conceptual model in Chapter 5 and the simulation in Chapter 6 together show that the labour market system actively reinforces traditional norms through unequal unpaid work allocation, and that this reinforcement opposes the gradual erosion of the underlying norm stock. Progress on economic indicators may be faster and more visible, but can coexist with norms that have not shifted to the same degree. More importantly, the model hypothesises that the difference in these speeds can itself increase the felt burden of unpaid work. This unevenness extends within the domain of norms themselves: women's labour market participation has changed faster than the household division of unpaid labour, despite both deriving from the same underlying norms (Gershuny & Kan, 2012).

Second, it offers an explanation of a common empirical finding of the literature that the mental health impact of unemployment is larger for men than for women (Paul & Moser, 2009). Because the overall psychological and financial rewards of employment are structurally lower for women — eroded through the unpaid work burden and the gendered wage penalty — the relative loss upon becoming unemployed is also smaller. Employment is, for many women, a less unambiguously positive state to begin with. This is why the gender difference is largely not seen in societies that have a dual-earner/dual-carer norm (Strandh et al., 2013).

Third, norms configure men’s exposure to labour market risk. As revealed by the redistributive trap, norms on gender roles protect men from the poor labour market conditions by disproportionately sorting more women into undesirable jobs. Chapter 7's finding that norms are persistently active but individually weak, operating indirectly through economic channels, refines the theoretical claim: norms are a structural determinant indirectly transmitting their influence on women’s outcomes.

Treating gender in this way also changes how the gender gap in population mental health is understood, and what can be done about it. First, the gap can be acted on, but largely through the economic channels by which norms transmit their influence (working-hours parity, the gender pay gap and the unpaid work burden) rather than through norms directly. Second, the mental health benefits

of gender equality policy can be expected to lag behind its economic ones (Section 7.6). Employment levels and the gender pay gap are already tracked routinely. However, the drivers the model identifies as crucial but slow-moving — the division of unpaid work, and the gender norms that sustain it — are not. This limits how well the effects of policy can be detected.

8.3.3 The model as an explanatory thinking tool

The third contribution is the model as an explanatory thinking tool where its value lies less in its outputs than in the structural stories it makes thinkable.

- The first story is the quantity–quality trade-off: the interaction of the employment loop and the job quality loop, in which pursuit of employment maximisation without robust regulation of job security can accelerate the population-level rise in poor mental health. This can also be seen in the baseline no-intervention scenarios: a future with lower labour demand but stable insecure work (S4) leads to lower prevalence of poor mental health than one with higher demand but a rising insecure work trend (S9) because the deterioration in job quality outweighs the benefits of stronger demand (Figure 7-17).

- The second is the redistributive trap: equality without job-quality improvement shifts the public health burden rather than reducing it (Sections 7.5 and 7.6). Further inequality can increase the burden but equality by itself cannot remove the harms of poor labour market conditions.

- The third is the buffering role of labour-force exit: at the population level, withdrawal from the labour market limits the rise of insecure work (Section 7.5), which is why policies that compel sustained attachment without repairing the conditions of work can be counter-productive.

- The fourth is the masking effect seen before the recession: through 1992–2008, tight labour markets concealed a gradual decay of institutional protections; when the post-2008 period removed that buffer, the gradual

rise of insecure work it had been masking appears to have crossed a critical threshold (Section 7.4).

The model is, in this respect, a different kind of tool from those most often built to forecast the effect of specific policies. Mahamoud, Roche and Homer (2013), for instance, design their system dynamics model of population health as a "learning tool": one that stakeholders can interrogate through "what if" questions to see how far and how quickly each intervention would move population health, and so to weigh policy directions against one another. Other efforts, notably using dynamic microsimulations, estimate how a specific policy would shift mental health; for example the pandemic-era fiscal response in Kopasker et al. (2024), a universal basic income in Thomson et al. (2024). The contribution here is explanatory rather than predictive: its worth lies in the structural mechanisms it elucidates, the trade-offs and traps set out above, not in forecasting the effect of any single intervention.

For decision-making, the four stories together show that what a policy achieves depends on the conditions under which it is pursued. The first is the combination of levers. Pursued in isolation, a lever can relocate the burden of poor mental health rather than reduce it, as the redistributive trap shows, or reduce it overall while preserving, and in some modelled futures widening, the gender gap. A practical corollary follows from comparing the levers in combination: expanding childcare and reducing insecure work reproduces almost the whole effect on overall prevalence of deploying all four levers, the gender pay gap and norm change adding little once those two are in place (Section 7.6). The second is the macroeconomic state. The levers' effect is largest where the labour market is worst and smallest in low-precarity futures (Section 7.6). The case for structural intervention is therefore strongest precisely under adversity.

8.3.4 Methodological contributions

Finally, this thesis makes two methodological contributions. First, it answers the growing international mandate to apply systems science to highly complex public health challenges (Diez Roux, 2011; Hohn et al., 2023; Rutter et al., 2017; Savigny & Adam, 2009). While system dynamics is well-established in infectious disease modelling and healthcare operations, its application to the social

determinants of mental health remains rare: recent reviews find only a handful of population-level models, most concerned with service delivery or treatment rather than the upstream social determinants (Atkinson et al., 2015; Davahli, Karwowski, & Taiar, 2020; Graham, Gariepy, & Orpana, 2023; Langellier et al., 2019; Long & Meadows, 2017). This thesis demonstrates how fragmented, individual-level epidemiological evidence (such as the discrete odds ratios of unemployment leading to depression) can be successfully integrated with macroeconomic theory (such as labour market tightness and vacancy saturation) to create a cohesive, population-level simulation. Beyond demonstrating feasibility, this shows the approach enables a class of questions regarding feedback, accumulation and timescale in the social determinants that individual-level designs cannot pose. Second, the feedback-driven literature review described in Chapter 4 is a helpful starting point for future researchers. Standard systematic review protocols prioritise the reproducibility of a pre-defined search string; identifying feedback loops requires the heuristic of "following the thread", where the identification of one variable triggers a new search for its reverse impact which might include intermediate variables. Chapter 4 argues, and the resulting model demonstrates, that systems mapping requires a search architecture genuinely different from systematic review, and offers feedback-driven enquiry as one such architecture. The contribution is not the rejection of rigour; it is a re-definition of what rigour means when the object of the review is a structure rather than a relationship.

This research makes the gaps in the evidence explicit, which has implications for public health research. Formalising the causal hypothesis as a simulation required every link to be given a functional form, a delay and a parameter, exposing where these rest on assumption rather than estimation: the shape of most relationships, the effect of rising prevalence on recovery, the degree to which insecure work erodes the recovery benefits of employment, and the timescale over which gender norms change. These assumptions, together with the hypotheses the model generates, such as the decoupling of incidence from recovery, propose an agenda for empirical research, to which Section 8.6 returns.

8.4 Implications for policy

This section draws out what the model's account of how the system operates implies for current policy debates in Scotland. These implications are based on the model's logic: they indicate the direction of an effect and the conditions under which it holds, not a forecast of its size. The policy implications of the thesis must be read against the Scottish context set out in Chapter 1: a Wellbeing Economy ambition operationalised through the Fair Work agenda but bounded by the powers reserved to Westminster (Mackley, 2020; Scottish Government, 2022a, 2024b). The section begins with what prevention implies for mental health spending, turns to the two areas through which the model's levers act, gender equality and insecure work, and closes with the case for pursuing them together.

8.4.1 Mental health spending

Treating poor mental health as a population-level system of interacting feedbacks, the reorientation on which this thesis's contributions rest, reframes the public health challenge: because the rate at which people are pushed into needing help and the pace at which they recover are shaped upstream, downstream services, necessary and valuable as they are, cannot by themselves reverse the trend, and the levers that matter most act on the structural drivers. This emphasis is consonant with the direction of current Scottish public health strategy described in Chapter 1: Scotland's Population Health Framework 2025–2035 announces a deliberate shift toward upstream action, and the Chief Medical Officer for Scotland's most recent annual report places prevention and system design, rather than clinical capacity alone, at the centre (Scottish Government, 2025a, 2026c). The 2026 Mental Health and Wellbeing Delivery Plan carries this emphasis into mental health policy, placing prevention at the centre and committing to track preventative spend (Scottish Government, 2026b).

The Delivery Plan's specific actions on the social determinants of mental health, however, are on housing and employability (Scottish Government, 2026b). The model points to a wider definition. It highlights the contribution of insecure work, gender inequality in pay and unpaid work to the prevalence of poor mental health, and the role of childcare in reducing that inequality (Sections

7.3-7.6). This thesis therefore argues that prevention cannot mean attaching a small programme to a predominantly treatment-focused system: labour market and childcare policy are core to preventing poor mental health, and arguably part of preventative mental health spending.

8.4.2 Equality levers

The model identifies childcare and the gender pay gap as the two primary gap-narrowing levers. They differ, though, in who bears the cost: the gender pay gap reduces the gap at no cost to men, whereas childcare (like norms) narrows it partly at men's expense, by shifting insecure work towards men (Section 7.6). The Scottish gender pay gap has narrowed faster than the UK gap over the simulation period (Scottish Government, 2024a), which Chapter 1 linked in part to devolved measures such as the expansion of childcare and the public sector wage floor. The model suggests that continuing this trajectory is among the higher-priority structural interventions available: the optimiser takes it to or near its bound in every macroeconomic future (Section 7.6). These equality levers sit within a wider menu of policies identified by the Scottish Government's international evidence review on advancing women's equality: childcare and parental leave, pay transparency, and gender-responsive budgeting and procurement among the areas where other governments have effectively narrowed gender inequalities (Craigforth, 2026).

Childcare is fully devolved, and Chapter 1 described the 2021 expansion of funded Early Learning and Childcare to 1,140 hours per year for three- and four-year-olds (Scottish Government, 2019a). The distinctiveness of the Scottish offer now lies less in its scale than in its universality: England's 2025 expansion matches the Scottish hours and reaches younger children, but conditioning entitlement on parental employment in a way the universal Scottish provision does not (Department for Education, 2025). The further expansion planned in Scotland would increase the hours and reach, with the aim of growing the economy and increasing incomes (Scottish Government, 2026d). The model's implication is also that this expansion is structurally well-targeted and offers a different rationale for the same policy, in terms of population mental health and the gender gap. In the model, childcare has the most direct effect on gender parity in working hours, acting immediately on insecure work and the gender pay

gap and, with a lag, on norms. Through these it impacts women's mental health, and through norms potentially generates returns beyond its direct effect by also weakening the gender norms entrenchment channel. On its own, however, it reduces total prevalence only negligibly in the base case (Table 7-9). The thesis therefore highlights that its contribution to population mental health depends on pairing it with action on insecure work.

As noted in Chapter 1, flexible working is promoted as a route to wider participation (Scottish Government, 2022b). The model points to a tension. Women's greater exposure to insecure work arises in the model from their sorting into a segment of the labour market that offers part-time flexibility at the cost of job quality. A recent evidence review by Scottish Government (2025c) concludes that unpaid childcare commonly leads mothers to reduce hours and move into flexible but lower-opportunity jobs. Flexibility, within the model, counts as good work only where it comes with security; otherwise it may preserve the unequal division of unpaid work. Because the model does not distinguish flexibility that workers control from insecurity that employers impose, this is an inference from its structure rather than a tested result.

Norms offer the deepest leverage but are the slowest and least tractable lever, as discussed in Section 8.2, so childcare and the gender pay gap must be pulled in parallel with the slower work on norms.

8.4.3 Insecure work

As set out in Chapter 1, poor health explains a larger share of economic inactivity in Scotland than in the rest of the UK and reducing inactivity is an aim of both the Scottish and UK governments (Department for Work and Pensions, 2024; Randolph, 2024; Scottish Government, 2026d). As the buffering story earlier discussed in 8.3 highlighted, health-related labour market exit is protective at the population level: it slows the rise of insecure work and shields the recovery of those who remain. It follows that expanding employment to reduce economic inactivity without repairing the quality of that work runs against the mechanism: it compels attachment while removing the buffer, and accelerates the rise in insecure work that drives the burden upward. The mechanism operates at the population level, through the additional insecure

work that greater attachment generates for a given level of labour demand; it does not imply that work fails to benefit those who enter it. The recent young-people review presses the same point, that what reduces health-related inactivity is good work rather than any work (Department for Work and Pensions, 2026). The Get Britain Working White Paper accordingly puts the focus on good work (Department for Work and Pensions, 2024).

The model's support for good work does not rest on this mechanism alone. Within the model, insecure work also erodes the benefit of employment for recovery. Scottish public health strategy already recognises that insecure work, like unemployment, has negative effects on mental health (Scottish Government, 2025a). Employment status is therefore an incomplete guide to the mental health effect of employment policy. Yet while the language of policy has moved towards good work, the measures by which its progress is judged, the rates of inactivity and employment, remain measures of quantity. On the model's logic, these measures cannot show whether such support reduces the burden of poor mental health, since a programme can succeed on both while placing people in the insecure work that erodes recovery and, at scale, strengthen rather than break the reinforcing loop between insecure work and poor mental health. Policy-makers would benefit from measuring insecure work as well as participants' mental health over time.

Without any forecast of what AI will do, a shock that raises insecure work, whether through weaker demand or faster casualisation, would move the system into the region where structural intervention matters more (Section 7.6), and there it is the stabilisation of job quality that would carry the reduction. However, a genuine collapse in demand lies beyond the tested range; the model also does not represent the skills mismatch through which such a transition would in part operate; and even fully optimised, the most adverse futures cannot reduce prevalence to levels the benign ones do.

The model's identification of insecure work stabilisation as the necessary partner to gender equality intervention is a policy proposition that cannot be fully delivered by Holyrood acting alone. Insecure work could nonetheless be partly addressed through devolved action, particularly the Community Wealth Building (Scotland) Act 2026, which places duties on local councils and NHS boards to use

procurement and employment for local economic development. The reach of Fair Work First is limited: it attaches conditions to public sector grants, but beyond them relies largely on voluntary employer behaviour. However, its recent evaluation records a stakeholder's suggestion that it be made more enforceable (Scottish Government, 2026a) and, as Chapter 1 noted, the Programme for Government commits to calling on Westminster to fully devolve employment law (Scottish Government, 2026d).

8.4.4 Synchronising job quality and equality

The implication is that the policy framing must be synchronised. A policy pursued in one domain can, within the model, shift the burden of poor mental health into another. An agenda that pursues gender equality without simultaneously addressing insecure work could shift the burden of poor mental health onto men. A Fair Work agenda that pursues secure work standards without simultaneously addressing the gendered allocation of unpaid work could preserve, and in some modelled futures widen, the gender gap, even if it succeeds in lowering overall prevalence.

Pursued together, expanding childcare and reducing insecure work lower prevalence for both men and women in every future tested and narrow the gap in most (Table 7-10). The model therefore gives a reason to judge Fair Work not only by its effect on labour market outcomes but also by its effect on poor mental health and the gender gap. More broadly, employment, Fair Work, childcare and gender equality policy are better pursued, and judged, together than as separate objectives, with population mental health among their shared outcomes. The Scottish Government's Wellbeing Economy framing set out in Chapter 1 emphasises cross-sectoral policy and could enable the synchronisation of these. However, a caveat is necessary: the most plausible drivers of the residual rise in women's distress in the late 2010s — Universal Credit rollout, the broader benefit cuts, and the degradation of mental health services — are all either fully reserved to Westminster or constrained by UK-level funding.

8.5 Limitations

Each stage of the research carries its own limitations: the breadth of the evidence synthesis, the quality of the underlying evidence and data, the translation of the available evidence into a simulation, and the boundary drawn around the system. This section draws together the limitations developed in detail across Chapters 3 to 7.

8.5.1 Gaps in the underlying evidence

As with all such models, this model is no stronger than the evidence synthesised into it, and that evidence is uneven and, in places, absent. Much of it derives from studies of variables in isolation and the strength of support varies across the causal links the model formalises: some, such as the link between unemployment and poor mental health, rest on robust bidirectional longitudinal evidence; others on thinner or single-context findings, a gradient Chapter 5 sets out link by link. Two gaps are consequential. The literature seldom specifies the functional form of a relationship, so the shape of most links was assumed rather than estimated; and the decoupling of incidence from recovery on which the analysis rests is a hypothesis the model generates rather than one the epidemiological literature directly evidences, since that literature does not systematically separate determinants of onset from determinants of duration. The parameters inherit these constraints. Further, a minority are unobservable and were fixed through bounded calibration, and many transition rates are drawn from UK-wide rather than Scottish data.

8.5.2 Researcher discretion in building the model

The model was built in two stages, each carrying discretionary choices. The conceptual model — the causal loop diagram — was assembled through a feedback-driven rather than systematic search (a deliberate response to the demands of mapping structure, Chapter 4), synthesised by a single researcher, and tested against a panel of experts; that protocol can disconfirm individual links and identify omissions within the adopted frame, but it cannot surface assumptions shared across the panel, so the conceptual structure remains a construction shaped by one analyst's reading of an interdisciplinary literature. The simulation model added a second layer of choices: operationalising

intangible constructs such as gender norms as quantified stocks and adopting linear functional forms, for example. Each choice is documented and defensible, but each could have been made otherwise. A different analyst, starting from a different conceptual frame, might formalise the same phenomenon differently.

8.5.3 Boundary exclusions

The model represents a deliberately bounded portion of the system in two senses. Its boundary excludes determinants the evidence and the expert panel judged important: the welfare system and conditionality, poverty, and the capacity of NHS mental health services; and it leaves some feedback loops open, treating insecure work and labour demand as exogenous inputs and representing no pathway from gender inequality to men's outcomes.

The model also treats labour as homogeneous: labour market tightness is based on unemployment and vacancies alone, without considering the workforce's skills composition. Skills mismatch, retraining and skill-biased shifts in demand are the channels through which both the AI transition and the push to move people into employment would in large part operate.

The boundary is temporal as well as structural: the calibration window ends in 2019, so the COVID-19 pandemic and the cost-of-living — which would act on the model's economic stressor composite — fall outside the represented system, and sit within the 2019–2024 interval the optimisation extrapolates across. Much of its effect on the modelled system could, in principle, be represented as an exogenous shock to the model's inputs; but the pandemic may also have altered some of the causal relationships the model holds fixed (for instance, the link between labour demand and employment, or the gendered division of unpaid labour) with potentially material consequences for the post-2019 trajectory.

The model operates at the national level and adopts a binary representation of gender, so it cannot resolve sub-national, intersectional, age-cohort or non-binary variation. Although the boundary was not drawn arbitrarily and the resulting parsimony was justified, these choices mean that its conclusions are population-average and bounded by the loops left open. Because the determinants held outside the boundary would, on the available evidence, have

raised rather than lowered the modelled prevalence, the model's under-reproduction of the post-2016 rise points to conservative estimates.

8.5.4 Limited generalisability

The mechanisms are calibrated to a particular setting in three respects: institutionally, to a broadly neoliberal political economy in which paid work is the primary route to income security; culturally, to Scotland and the UK which follows a modified male breadwinner model; and temporally, to the period 1992–2019. Some features are more context-bound than others. The dependence of the findings on the prevailing welfare regime and the gender norms are the elements least likely to transfer unchanged to another setting. The structural plausibility of the mechanisms is likely to travel to comparable settings while their estimated strengths may not, and any projection beyond the calibration window is an exercise in structural sensitivity rather than a forecast.

8.5.5 Equifinality: sufficiency, not uniqueness

Each of the limitations above could, in principle, be narrowed — by stronger evidence, a second researcher, a wider boundary, or better data. The exception is equifinality. The model demonstrates that the hypothesised causal mechanisms are behaviourally sufficient to generate the observed Scottish trajectory; it does not establish them as the unique generative structure. Different causal structures can in principle produce the same observed behaviour, the property of equifinality. Passing the behaviour reproduction tests reported in Chapter 6 is therefore necessary but not sufficient to establish that the hypothesised structure is the operating structure. Grounding parameterisation in empirical evidence, limiting the use of constrained calibration, and disconfirmatory expert review of the conceptual structure narrow the space of plausible alternatives but they do not eliminate it.

8.6 Directions for future research

The thesis has offered one plausible structural explanation for the Scottish paradox; its claim is behavioural sufficiency, not causal identification, and its contribution is theoretical–structural, supported by careful empirical

parameterisation and a foundation to build on. Four directions for future work follow.

8.6.1 Model validation

The conceptual structure was validated through a disconfirmatory expert protocol, but its operationalisation (the choice of functional forms and effect parameters) was not independently stress-tested; future work should subject it to expert appraisal and refinement.

Targeted data collection on critical but weakly identified parameters would also be valuable: the effect of rising prevalence on recovery (the threshold governing endogenous recovery), which is not a uniquely identified parameter; the degree to which insecure work erodes the recovery benefits of employment, for which direct empirical evidence is presently lacking; and the timescale over which gender norms change, currently set on theoretical grounds over a cohort-replacement window rather than measured. More generally, measuring the flows into and out of poor mental health, rather than prevalence alone, would let the incidence and recovery processes be calibrated directly.

8.6.2 Model refinement

One avenue for refinement concerns the internal structure of the gender dynamics. The behaviours that gender norms govern have moved at different paces: women's labour market participation has risen considerably faster than the household division of unpaid labour has shifted, even though the model represents both as expressions of the same underlying norm. Future work could ask what accounts for this divergence; for instance, examine whether the single norm stock should be disaggregated into components governing paid and unpaid work, and whether the barriers and facilitators to change differ between them.

A further direction is to differentiate the pathways that act on incidence from those that act on recovery. This is most pressing on the recovery side, where the model represents women's slower recovery only as an aggregate difference: unpacking whether it is driven principally by the double burden, by reduced access to support, or by greater exposure to insecure work would add granularity the single recovery rate difference cannot.

8.6.3 Model extension

Future research could benefit from deepening the model's dynamic structure: its feedback loops, and the way its relationships change over time rather than simply including more factors. The most valuable would be to allow causal relationships to evolve endogenously as institutional and economic conditions change. This would allow the effects of welfare reform on the unexplained post-2016 rise to be examined; for example, by letting policy shifts alter the strength of the relationship between unemployment and mental health. Future work would also benefit from closing the loops the present model leaves open — most notably the reverse channel from population mental health to labour demand (through productivity) which is currently omitted from the simulation model although captured in the conceptual model.

Given the importance of the reinforcing loop between prevalence and recovery, a related priority is the explicit representation of the availability and quality of mental health services, which act directly on the recovery side of that loop and are presently held outside the boundary.

A further set of extensions, each adding heterogeneity the present labour market structure omits, would let the model engage with current debates. The most consequential is skill: the model cannot represent skills mismatch, which is central both to the AI transformation of jobs and to moving those in economic inactivity into good work. Disaggregating insecure work by occupation, and by the aspects most exposed to AI, would let that gendered exposure be represented explicitly. AI-enabled flexibility might decouple part-time work from the penalty that currently sorts women into insecure work, or it might deepen a care-accommodation trap in which it leads women to absorb still more unpaid labour; the model yields a direction only once the sign of that effect is fixed.

A different extension concerns who enters the system. Because the model is not age-structured and has no demographic inflow, it cannot represent a changing profile of new entrants; giving it one would let a deteriorating young-adult cohort raise the prevalence stocks directly and set off the prevalence–recovery loop, which the present structure captures only in aggregate.

Extending the model beyond 2019 would also provide a natural test of its framing of economic shocks as activators of pre-existing structural vulnerability: whether the accumulated vulnerability and activator dynamics it identifies also account for the mental health trajectory through the COVID-19 period, including its gendered pattern, or whether that shock operated through channels the present boundary excludes; in which case the relationships themselves would need to be allowed to evolve.

8.6.4 Model application

The model could also be used for more specific policy analysis in collaboration with the Scottish Government, translating its structural insights into politically and operationally feasible intervention specifications. Specifically, it could serve as a platform for policy experimentation: building additional model structure and designing experiments to test policy options in consideration, as well as testing how the levers identified interact when combined and sequenced, and at what point sustained intervention begins to shift the system.

More broadly, the model can be used to generate hypotheses that direct future empirical work. Its proposal, for instance, that the gender pay gap and the unpaid work burden act principally on the incidence of poor mental health, whereas insecure work also degrades recovery, could be tested directly by longitudinal data on transitions into and out of poor mental health. It further implies a slow, reciprocal feedback in which a narrowing pay gap and greater parity in working hours erode traditional gender norms, which panel data spanning a sufficient period could examine. In each case the model's contribution is less the magnitude it estimates than the empirically addressable question it makes precise.

8.7 Concluding remarks

The thesis began with a public health paradox: Scotland has made measurable progress on the visible indicators of gender equality, yet this progress has not translated into a closing of the gender gap in poor mental health. This thesis has offered a structural explanation. The labour market, the household division of unpaid labour, the slow movement of generational norms, and population mental

health together form a single system. Within it, progress on the visible indicators of equality can coexist with the gender gap in prevalence of poor mental health. Through the interaction of timescales and the dynamics of perceived inequality, such progress can even intensify it. The model is one plausible generative structure for the observed trajectory: its claim is behavioural sufficiency, not causal identification. The contribution is theoretical and structural: a dynamic, population-level causal structure that decouples incidence from recovery and treats gender as a structural determinant.

The relevance of that structure to policy lies in framing. Because prevalence conflates the flows that produce it, the framework suggests asking of any policy which flow it acts on, and when its effect would appear; and because, within the model, what a policy achieves depends on the conditions under which it is pursued, it suggests asking also in what combination of levers, and in what macroeconomic state. Because the levers that matter most act on the structural drivers, it also suggests assessing employment, childcare and gender equality policy for their impact on mental health and not just on economic outcomes.

Taken together, the analyses suggest that closing the gender gap in the prevalence of poor mental health will not be accomplished by single levers deployed in isolation. Equality without job quality improvement redistributes poor mental health; job quality improvement without structural equality preserves the gap. Only the two pursued together are likely to lower the burden rather than shift it.

Appendices

Appendix A Search strategy

EBSCOhost

Tuesday, April 09, 2024 8:50:27 AM

#	Query	Limiters/Expanders	Last Run Via	Results
S21	S12 OR S18	Limiters - Publication Date: 20000101-20241231 Narrow by Language: - english Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	1,264
S20	S12 OR S18	Narrow by Language: - english Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	1,532
S19	S12 OR S18	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	1,586
S18	S9 AND S17	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	572
S17	S13 OR S14 OR S15 OR S16	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	3,737,703
S16	TI (UK or "United Kingdom" or Britain or British or England or English or Scotland or Scottish)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	359,288
S15	TI (gender* or women*)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	572,609
S14	TI (review or synthes#s or meta*)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	2,504,821
S13	TI (time-series or trend* or macro* or ecological)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	383,310
S12	S10 AND S11	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	1,023
S11	TI ("unpaid work" or "unpaid labo#r" or "unpaid care*" or "domestic work" or "domestic labo#r" or housework or "household work" or childcare or "gender* role*" or "gender* position*" or "gender* norm*" or "gender* expectation*" or "sex?role*" or "family? role*" or "double burden")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	13,565
S10	S5 OR S8	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	657,369
S9	S5 AND S8	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	4,270
S8	S6 OR S7	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search	559,512

			Database - APA PsycInfo;MEDLINE;EconLit with Full Text	
S7	TI (health OR well#being) AND AB (mental OR depress* OR "psychological distress")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	208,554
S6	TI (depression OR depressive OR depressed OR "psychological distress" OR "mental illness")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	365,545
S5	S1 OR S2 OR S3 OR S4	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	102,127
S4	TI ("precarious job" OR "precarious work" OR "job precari*" OR "job security" OR "job insecurity" OR "insecure work" OR "work instability" OR "unstable labo#r" OR "unstable job" OR "temporary job" OR "temporary work" OR "temporary contract" OR "*permanent job" OR "*permanent work" OR "*permanent contract" OR “non#standard job" OR "non#standard work" OR "non#standard contract" OR "fixed?term work" OR "fixed?term contract" OR "fixed?term job" OR "atypical work" OR "casual work" OR "agency work" OR "zero?hour contract" OR "gig?work*" OR "gig?economy" OR "flexible labo#r" OR "variable hours" OR "variable pay" OR "variable hourly pay" OR “contingent work*" OR "dual labo#r" OR "secondary labo#r")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	2,737
S3	TI (“economic cris?s*" OR “economic contraction*" OR “economic downturn*" OR recession*)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	13,018
S2	TI ("early retirement" OR "disability pension" OR "sick leave" OR "sickness absence" OR "medical leave" OR absenteeism OR presenteeism)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	9,376
S1	TI (employed OR unemployed OR employment OR unemployment OR "paid work" OR under#employed OR under#employment OR jobless* OR "labo#r force participation" OR "labo#r force status" OR "economic* inactiv*" OR "work status" OR "job status" OR "out of work" OR "job loss" OR "laid off" OR "lay#off" OR "part#time work" OR "part#time job")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - APA PsycInfo;MEDLINE;EconLit with Full Text	78,423

Appendix B Expert engagement documents

B1 Invitation email

Subject: Invitation to Participate in Research on Gender, Labour Market and Mental Health

Dear [Recipient's Name],

I hope this email finds you well. My name is Shreya Sonthalia, and I am a PhD student within the School of Health and Wellbeing at the University of Glasgow. I am reaching out to you to invite you to participate in my research based on your expertise in [relevant field].

My PhD research is centred around investigating the relationship between gender, labour market and population-level mental health, with a specific emphasis on understanding gender differences. As part of my study, I will be developing a system dynamics model to explore the causal mechanisms involved. I will build the model by integrating existing research, expert insights, and available data (from Scotland). Your expertise is crucial in refining and enhancing the quality of my model.

Your participation would entail a **45-60 minute online interview** conducted by me between **23rd June and 31st July 2025**, at a time convenient for you. During the interview, I will present my emerging model, seek your feedback, and discuss specific areas where I have identified gaps in the current research. If you are willing to participate, please **respond to this email by 16th June 2025** and suggest a few one-hour slots that you are available between the 23rd of June and 31st of July 2025.

Attached to this email, you will find the participant information sheet which contains further information about the research. Please do not hesitate to contact me if you have any further questions.

Looking forward to hearing from you,

Regards,

Shreya Sonthalia
PhD Student
MRC/CSO Social and Public Health Sciences Unit
University of Glasgow

B2 Participant information sheet

University of Glasgow | College of Medical, Veterinary & Life Sciences

PARTICIPANT INFORMATION SHEET for Gender, Employment and Mental Health in Scotland: A system dynamics study

You are invited to take part in a research study. Before you decide to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Please ask the researchers if there is anything that is not clear or if you would like more information. If you decide to take part in this study, you will be given a copy of this Participant Information Sheet and the signed consent form to keep.

1. What is the purpose of the study?

This research aims to understand the complex relationship between the labour market and population-level mental health in Scotland, with a particular focus on examining how this differs for men and women due to factors such as unpaid work, caregiving role and gendered employment patterns. It will consider how the labour market conditions such as unemployment, job insecurity, working hours and low pay impact the mental health of the Scottish population, as well as how poor mental health impacts these labour market conditions.

The study will develop a system dynamics model, which is a simulation model that focuses on causal mechanisms. It will integrate existing research, expert input and available data. Such a model helps consider the bidirectional nature of the relationship between the labour market and mental health and study this at the population level. The model can help test our causal claims and examine differences in these for men and women.

The research is being conducted as part of a PhD degree, within the School of Health and Wellbeing at the University of Glasgow.

2. Why have I been invited to participate?

You have been invited to take part in this study because of your professional background and expertise that will be valuable to understand how the labour market and mental health are connected at the population level in Scotland, and how this differs for men and women. As part of the study, 8-12 experts will be interviewed.

3. Do I have to take part?

No, you can choose not to take part. Your participation is voluntary. If you do decide to take part, you will be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason.

4. What will happen to me if I take part?

If you agree to participate, you will be asked to sign a consent form and return via email.

Based on your availability, you will be invited to an online interview with the researcher lasting 45-60 minutes. The interview will be conducted via video call. If you prefer, the interview can be conducted as two smaller sessions. During the interview, the researcher will present the preliminary model on the relationship between labour market and mental health in Scotland to you and ask you for feedback. This will include (1) aggregation of causal links from the individual to the population level – for example, rates of unemployment in Scotland having a causal influence on the prevalence of mental ill-health in Scotland, (2) missing variables and links, (3) context relevant factors that are important to consider in Scotland, (4) differences between men and women.

The causal model and supporting material will be shared with you. Further comments following the interview via email or a subsequent interview are welcome. You may be contacted via email for clarification questions.

The interview will take place between June and July 2025. The interview will be (video) recorded and stored securely. For backup, an audio recorder will also be used. Recordings from this will also be stored securely on the University of Glasgow drives. Word to word transcription is not required for this research but the researcher will make use of auto-transcription when needed and may cite your responses as anonymised quotes when presenting the research.

Results will be shared with you via email once the research is completed. You can opt to not receive these updates if you wish.

5. What are the possible disadvantages and risks of taking part?

The risks of participation are low. Interviews will not be covering sensitive topics nor require you to share any personal information. You are also free to decline answering any question.

6. What are the possible benefits of taking part?

There is no intended direct benefit for you. However, it is likely that you will enhance your understanding on the relationship between labour market and mental health in Scotland by actively engaging with the literature-informed model. You will be introduced to a systems thinking tool – causal loop diagram.

7. Will my taking part in this study be kept confidential?

All information which is collected about you, or responses that you provide, during the course of the research will be kept strictly confidential. You will be identified by an ID number, which will be used to link your responses to each other. Once the research is completed, your contact details will be securely deleted. Your responses will be

summarised in a document and/or table, noting where you agreed and disagreed with the presented model. Where relevant, quotes may be included. For context, your expertise/professional background or role of the expert will be incorporated in these summaries. It will not include your name or specific job title. These summaries will be de-identified and securely held in a research data repository for future research and learning. Some of this will be published as supplementary material.

You may be identifiable based on your professional experiences. Summaries will be de-identified as far as possible, but your responses will be recorded and transcribed. These audio files and transcripts will be securely held within the care of the University of Glasgow.

Your responses may be quoted in study reports but these will not be attributed to you, and you will remain anonymous.

8. What will happen to my data?

As part of obtaining consent, we may be collecting and storing identifiable information from you in order to undertake this study. This means that the University is responsible for looking after your information and using it properly. We may keep identifiable information about you until the research has completed and will not pass this information to a third party without your express permission.

Your rights to access, change or move the information we store may be limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible

Researchers from the University of Glasgow collect, store and process all personal information in accordance with the General Data Protection Regulation (2018).

All study data will be held in accordance with The General Data Protection Regulation (2018).

Your identifiable information might be shared with people who check that the study is done properly. Consent forms will be retained in secure university servers for a period of at least 10 years. Audio files and transcripts will be securely held within the care of University of Glasgow. Although audio files and transcripts will not include your name, you may be identifiable from your voice or the information you provide (for example, projects you have been involved with).

Your data will form part of the study result that will be published in expert journals, presentations, student dissertations/theses (if applicable) and on the internet for other researchers to use. Your name will not appear in any publication.

9. What will happen to the results of the research study?

Your input will be used to revise and edit the model. The model will be a key part of the PhD research and thesis. Alongside the thesis, journal articles based on the model are likely to be published. These research outputs will mention the number of experts engaged with and

their backgrounds (e.g. epidemiologist with a focus on employment or research from Public Health Scotland), but not mention any names. Your responses may be quoted in study reports but these will not be attributed to you, and you will remain anonymous.

In addition to these, as a participant, you will receive a report describing how the model has been revised following expert input.

10. Who is organising and funding the research?

The research is conducted as part of a PhD programme that has been funded by the Medical Research Council.

11. Who has reviewed the study?

The research project has been reviewed by the College of Medical, Veterinary & Life Sciences Ethics Committee.

12. Contact for Further Information

If you have any further questions regarding the study, please contact Shreya Sonthalia at or Corinna Elsenbroich (supervisor) at corinna.elsenbroich@glasgow.ac.uk.

Thank you for reading the Information Sheet

B3 Privacy notice

University of Glasgow | College of Medical, Veterinary & Life Sciences

Privacy Notice for Gender, Employment and Mental health in Scotland: A system dynamics study

Your Personal Data
The University of Glasgow *will be what's known as the 'Data Controller' of your personal data processed in relation to* consent. *This privacy notice will explain how The University of Glasgow will process your personal data.*

Why we need it
We are collecting your basic personal data such as name, role, email address *in order to obtain consent. We will only collect data that we need in order to provide and oversee this service to you.*

Legal basis for processing your data
We must have a legal basis for processing all personal data. In this instance, the legal basis is

- Consent – a consent clause needs to be included at the end of the privacy notice
- Public task/Official authority – this covers activities undertaken as part of the core functions of the University

What we do with it and who we share it with

- *All the personal data you submit is processed by staff at the University of Glasgow in the United Kingdom.*

How long do we keep it for
Your data will be retained by the University for the during of the research (i.e. completion of the PhD). *After this time, data will be securely deleted.*

What are your rights?*
You can request access to the information we process about you at any time. If at any point you believe that the information we process relating to you is incorrect, you can request to see this information and may in some instances request to have it restricted, corrected or, erased. You may also have the right to object to the processing of data and the right to data portability.
Where we have relied upon your consent to process your data, you also have the right to withdraw your consent at any time.
If you wish to exercise any of these rights, please submit your request via the webform or contact dp@gla.ac.uk.
*Please note that the ability to exercise these rights will vary and depend on the legal basis on which the processing is being carried out.

Complaints
If you wish to raise a complaint on how we have handled your personal data, you can contact the University Data Protection Officer who will investigate the matter.
Our Data Protection Officer can be contacted at dataprotectionofficer@glasgow.ac.uk
If you are not satisfied with our response or believe we are not processing your personal data in accordance with the law, you can complain to the Information Commissioner's Office (ICO) https://ico.org.uk/

I consent to the University processing my personal data for the purposes detailed above. ☐

I have read and understand how my personal data will be used. ☐

Name of participant

Date

Signature

B4 Consent form

University of Glasgow | College of Medical, Veterinary & Life Sciences

Project Number: 200230153
Title of Project: Gender, Employment and Mental health in Scotland: A system dynamics study
Name of Researcher(s): Shreya Sonthalia

CONSENT FORM

I confirm that I have read and understood the Participant Information Sheet version 3.0 dated 2/6/2025. ☐

I confirm that I have read and understood the Privacy Notice version 3.0 dated 2/6/2025 ☐

I have had the opportunity to think about the information and ask questions, and understand the answers I have been given. ☐

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected. ☐

I confirm that I agree to the way my data will be collected and processed in accordance with relevant Data Protection policies and regulations. ☐

I understand that all data and information I provide will be kept confidential and will be seen only by study researchers and regulators whose job it is to check the work of researchers. ☐

I agree that my name, contact details and data described in the information sheet will be kept for the purposes of this research project. ☐

I understand that if I withdraw from the study, my data collected up to that point will be retained and used for the remainder of the study. ☐

I agree to my interview being audio-/video-recorded. ☐

I understand that my information and things that I say in an interview may be quoted in reports and articles that are published about the study, but my name or anything else that could tell people who I am will not be revealed. ☐

I agree to be contacted by the research team for clarification, sharing results or about taking part in further research. ☐

I give permission for the interview recordings and transcripts and de-identified relevant email correspondence to be securely held within the care of University of Glasgow to evidence data that underpin by conclusions and thesis. Although my name will be removed from these, I am aware that it may be possible for me to be identified because of the information I provide about my professional background and experiences. ☐

I give permission for this consent form to be retained for a period of at least 10 years within a secure drive at the University of Glasgow for the purpose of research integrity. ☐

I give permission for the summaries from interviews be securely deposited for future research and learning and/or published as supplementary material. ☐

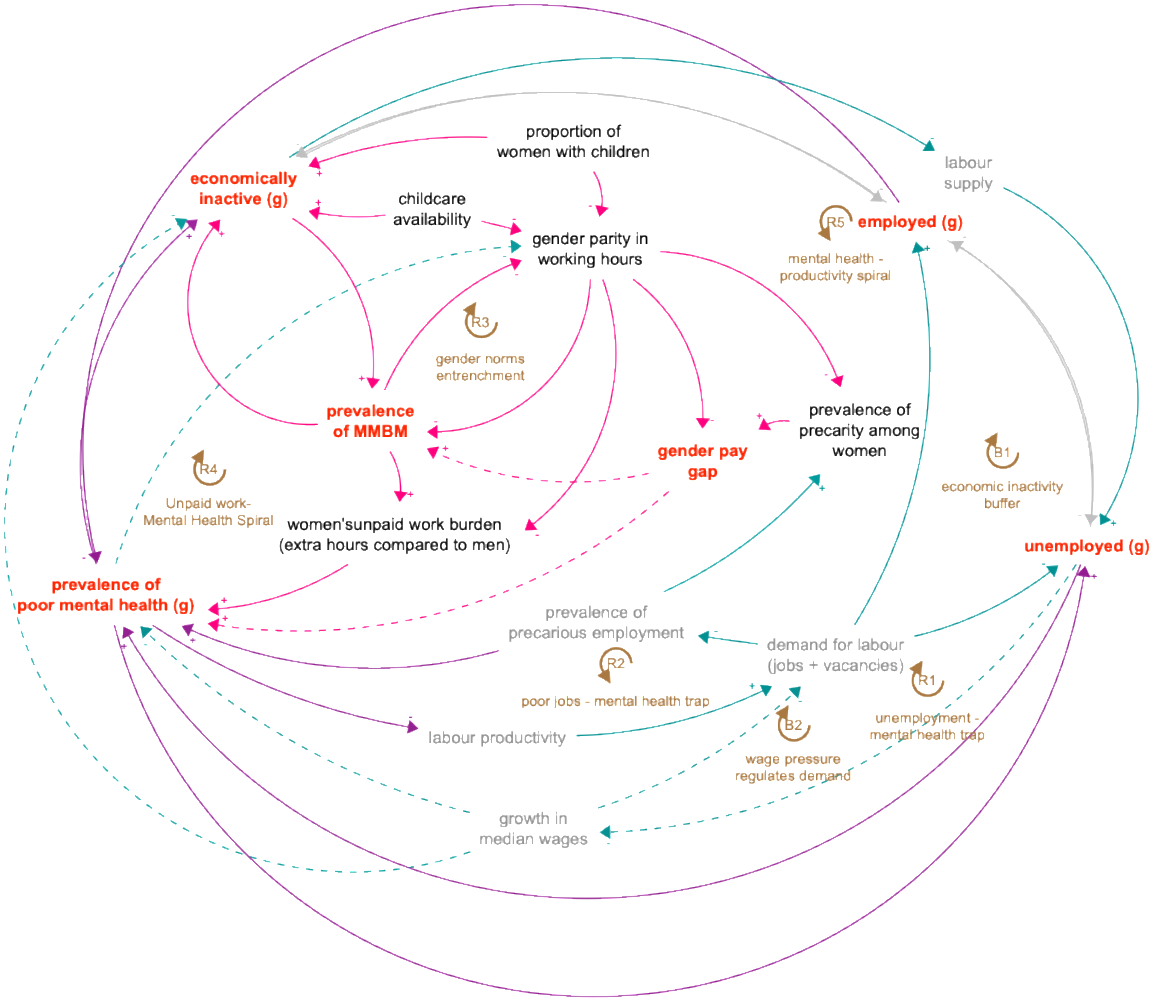
I agree to take part in the study. ☐

Name of participant	Date	Signature

Participant Identification Number:

Name of Researcher	Date	Signature

B5 CLD presented to experts



Appendix C Outputs from expert engagement

C1 Interview summaries

These summaries record the substantive feedback from each expert interview; how it was thematically synthesised and fed back into the CLD is reported in Section 5.5.

Interview 1 - Gender and labour market researcher

- **Conceptualising unpaid work:** While agreeing that housework is a reasonable proxy for the unpaid work burden, the expert emphasised that the big divergence in labour is driven by the birth of the first child and subsequent caregiving responsibilities. They also noted that when women increase their paid work, men don't necessarily do more unpaid work; rather, women do less (either by outsourcing or lowering standards).
- **The "double-edged sword" of part-time work**: A critical nuance was that part-time work can have a positive mental health effect for women, as it allows them to manage the double burden and fulfil socially conditioned gender roles. Women will voluntarily pull back on work hours to prevent burnout, creating a complex tension.
- **The gender pay gap**: The expert endorsed the link from the gender pay gap to the prevalence of the gender roles, noting it often makes economic sense for the lower-earning partner (usually the woman) to reduce hours.
- **Nuancing employment for women**: The expert framed employment as more of a dynamic scale for women than a binary state. Women are "strung all the way along that scale" from underemployment to precarious work, whereas it's more binary for men (employed or unemployed).
- **Life-course perspective:** The expert stressed that caring responsibilities are persistent, moving from children to elderly parents ("sandwich care"), spouses, and grandchildren, perpetuating women's over-representation in care roles. They also highlighted that the financial penalty for taking time out for children is permanent, even if women return to full-time work later.

Interview 2 - labour market economist

- **Precarious employment vs. job quality**: The expert suggested that relying solely on "precarious employment" as a measure of job quality is insufficient. They recommended incorporating other dimensions such as autonomy and flexibility, which are available in datasets like the UK Household Longitudinal Study. In-work poverty was also proposed as a key measure of poor job quality.
- **Mental health and labour market exits:** The expert noted that, compared to other countries, evidence for health shocks leading to labour market exits in the UK is weaker and that a significant mental health episode would be required to cause an exit.
- **Wages and poverty:** The discussion on wages highlighted that "growth in median wages" might not capture the full picture. The expert emphasised the importance of considering the population on very low incomes and in-work poverty, suggesting they are related but distinct concepts.

- **Regional heterogeneity:** A major point of feedback was the challenge of modelling the UK/Scotland as a single labour market due to significant regional variations, such as the disparity between the southeast and the northeast in the UK.
- **Policy levers:** The expert stressed the importance of framing the model around actionable policy levers, particularly considering the devolved powers of the Scottish Government.
- **Data Sources**: A significant concern was raised about the reliability of the Labour Force Survey (LFS) and recommended using UKHLS instead.

Interview 3 - public health, Scotland focus

- **Regional variation within Scotland**: The expert questioned whether treating Scotland as a single economy was appropriate, pointing to weaker labour markets in certain areas which could influence job availability and quality. However, they acknowledged the data challenges of a sub-regional analysis.
- **Missing populations:** The expert highlighted the absence of lone parents, who are likely to be disproportionately affected. They also raised the issue of caring responsibilities for older adults, not just children, impacting women's labour market participation.
- **Education and skills:** Education was identified as a critical missing variable. The expert noted its protective effect against both unemployment and poor mental health. The concept of over-qualification or skills-based underemployment was suggested as a more nuanced and potentially more important factor for gender differences than just educational attainment levels.
- **Data sources:** The expert confirmed that using the Labour Force Survey (LFS) for historical calibration seemed reasonable, given the concerns are more recent. They suggested starting the working-age analysis at 25 to exclude those in full-time education.

Interview 4 - public health researcher focusing on health inequalities and economic policy

- **Mental health measurement:** The expert raised concerns about using the GHQ-12 to measure mental health, noting that it yields a higher rate of false positives that are correlated with socioeconomic position. They recommended utilising the SF-12 mental health measure (available in the UKHLS) as a more robust alternative for the model.
- **Socioeconomic status and intersectionality:** The expert emphasised the deep intersection between class and gender, suggesting that modelling at a purely population-wide average level might fail to capture how socioeconomic status modifies many of the relationships being modelled.
- **Geographical clustering and skills mismatch:** The expert highlighted that poor mental health and unemployment are often distinctly geographically clustered in areas suffering from low education, low skills, and low demand for labour. Referencing interventions in places like Preston, they pointed out a skills mismatch reality: creating low-skilled, low-wage jobs can actually have a better impact on mental health for historically deprived populations than creating high-skilled jobs they cannot realistically access.
- **Wages:** The discussion on pay emphasised that median wages are insufficient for examining these issues. The expert noted a much stronger

correlation between poor mental health and wages specifically at the bottom of the wage distribution (e.g., the 20th percentile) rather than the median.
- **Education:** Finally, education and human capital were identified as crucial missing variables in the model. The expert noted that these factors are foundational drivers that relate heavily to the wages side of the model and fundamentally modify many of the included labour market relationships.

Interview 5 - labour market advisor
- **Job quality vs. precarity:** The expert suggested reframing "precarious employment" as job quality. They discussed the challenges of measurement, noting that contractual measures (e.g., temporary or zero-hours contracts) are flawed, as many people on such contracts are satisfied. They suggested using a composite measure from UKHLS including perceived insecurity and autonomy.
- **Longitudinal/cumulative effects:** A key conceptual point was the importance of cumulative experience, such as the history of unemployment spells impacting later health outcomes or career interruptions affecting the gender pay gap.
- **Data issues:** The expert confirmed the significant problems with UK labour market data in recent years, particularly the LFS and APS, which affects inactivity statistics.
- **The benefit system:** They mentioned the relevance of evidence on welfare conditionality's impact on mental health.
- **Childcare:** The link between childcare availability and mothers' working hours was reinforced with a reference to a survey of Universal Credit claimants. The lack of "wraparound" childcare in more deprived areas was also noted as a barrier.

Interview 6 - qualitative researcher, poverty and welfare focus
Did not consent for interview summary to be published.

Interview 7 - public health, work and employability in Scotland focus
- **"Fair Work" in Scotland:** The expert mentioned that the Scottish Fair Work Framework provides a strong alternative to precarious employment. They noted that measurement is difficult but pay (real living wage) is one tracked metric.
- **Drivers of precarious employment:** The expert suggested that poor employers can persist where workers are easily replaceable (low-skilled, low-paid roles), providing little incentive to invest in good work.
- **Preventing labour market exits:** A range of factors to prevent health-related exits from work were identified: flexible working, workplace adjustments, supportive line managers, the Access to Work scheme, a better Fit Note system, and enhanced sick pay.
- **Occupational segregation and health:** the expert noted that sectors with high numbers of workers with health conditions (care, sales, retail, education) are also highly feminised, suggesting the nature of the work itself may increase health risks.
- **Youth inactivity:** The risk of young people becoming economically inactive early was highlighted as a major concern, as it predisposes them to future unemployment and poor quality work.

Interview 8 - work and gender researcher

- **Occupational segregation as a key driver:** The expert argued that the demanding nature of "interactive work" (e.g., social care, education), which is predominantly done by women, is a key driver of stress and poor mental health.
- **The "modified male breadwinner model" as a misnomer:** The expert questioned the terminology, noting that the model is often a forced outcome due to economic necessity (one wage is no longer sufficient) and structural gender norms, rather than a simple choice.
- **Data limitations:** The expert noted the difficulty of capturing complex social phenomena like gender norms in a quantitative model, stating that feminist economics has long highlighted the importance of the social reproductive sphere, which is often un- or under-quantified.
- **Social desirability in survey responses:** A warning was given about taking survey responses at face value, particularly regarding the rewards of caring. Social desirability bias means people may report taking pride in a role they find immensely stressful.
- **Flexibility:** The expert noted from qualitative work that while flexibility is appreciated, poor management (e.g., unpredictable shift patterns) can create significant stress, complicating the link between flexible hours and well-being.

Interview 9 - Health policy analyst

- **Macroeconomic context:** The expert highlighted that the relationships between labour supply, demand, wages, and unemployment are highly conditional on the broader macroeconomic environment and business cycle. For instance, high labour demand might fuel wage growth instead of reducing unemployment if the economy is already at capacity.
- **Nuances of employment states:** A key insight was the need to consider people who are employed but off sick. The ability to take sick leave while remaining employed is an indicator of job quality, and this group is often overlooked in statistics focused on unemployment or inactivity.
- **Heterogeneity of inactivity:** The expert noted that the link between poor mental health and economic inactivity is conditional on factors like education level. They hypothesised that the "male breadwinner model" might be more prevalent in affluent families who can choose that lifestyle, whereas inactivity due to poor mental health might be more common in lower socioeconomic groups.
- **The link between inactivity and unemployment:** The causal link where increased inactivity leads to lower unemployment (the "economic inactivity buffer") was described as "counterintuitive" and overly simplistic, as its effect is highly conditional.

Interview 10 - health economist/analyst

- **Heterogeneity of inactivity:** A key critique was that the economically inactive category is too broad. The link between inactivity and the "male breadwinner model" was questioned, as the inactive group includes students and others for whom the link is less relevant. The expert pointed out that the drivers for women caring are likely different from those for young men becoming inactive, which may not reinforce the same gender norms.

- **Precarious employment:** The expert suggested that low pay might function differently from other forms of precarity (e.g., zero-hours contracts), with the latter being more about stress and worry from unpredictability.
- **Gap vs. absolute Measures:** The expert questioned the use of "gap" measures (gender pay gap, parity in hours) versus absolute levels. They suggested an intermediate variable might be needed; for example, the gender pay gap could increase the proportion of women with low earnings, and it is the low earnings (an absolute measure) that directly impacts mental health.
- **Education:** The expert identified skills and education as an important missing component. They also noted the importance of being a carer for children and the associated mental health struggles that can persist even after children start school.

C2 Refinements to the CLD following expert consultation

Table C-1: Refinements to the CLD following expert consultation. Lists each candidate causal link considered and whether it was retained in, removed from, or added to the CLD, with the reason

CLD link	Action
Unemployment → Prevalence of Poor Mental Health	Removed; replaced with single link from Employed to Prevalence of Poor Mental Health
Insecure Work → Prevalence of Poor Mental Health	Retained
Prevalence of Poor Mental Health → Labour Productivity	Retained
Prevalence of Poor Mental Health → Economic Inactivity	Removed; replaced with single link from Prevalence of Poor Mental Health to Employed
Prevalence of Poor Mental Health → Unemployment	Removed; replaced with single link from Prevalence of Poor Mental Health to Employed
Employment → Prevalence of Poor Mental Health	Retained
Norms on Gender Roles → Gender Parity in Working Hours	Retained
Norms on Gender Roles → Economic Inactivity (Women)	Retained
Norms on Gender Roles → Unpaid Work Burden	Retained
Unpaid Work Burden → Prevalence of Poor Mental Health (Women)	Retained
Gender Parity in Working Hours → Unpaid Work Burden	Retained
Gender Parity in Working Hours → Gender Pay Gap	Retained
Gender Parity in Working Hours → Insecure Work among Women	Retained

Gender Parity in Working Hours → Norms on Gender Roles	Retained
Demand → Unemployed	Removed; not a direct link
Gender Pay Gap → Prevalence of Poor Mental Health (Women)	Retained
Gender Pay Gap → Norms on Gender Roles	Retained
Economic Inactivity → Norms on Gender Roles	Retained
Childcare Availability	Removed; removed from CLD for parsimony but retained in model as exogenous factor
Insecure Work among Women → Gender Pay Gap	Retained
Labour Market Tightness → Insecure Work	Retained
Insecure Work → Productivity	Retained
Productivity → Demand	Removed; demand removed
Demand → Labour Market Tightness	Removed: demand removed for parsimony from CLD; simulation model will use labour demand as exogenous input
Demand → Vacancies	Removed; retained in simulation model
Vacancies → Labour Market Tightness	Removed; retained in simulation model
Demand → Employed	Removed; not a direct link; goes through Labour Market Tightness
Employed → Vacancies	Removed; retained in simulation model
Economic Inactivity → Labour Supply	Removed; no variable of labour supply in model (labour market tightness is based on both employed and unemployed separately)
Labour Supply → Unemployed	Removed; no variable of labour supply in model (labour market tightness is based on both employed and unemployed separately)
Labour Market Tightness → Employed	Added
Labour Market Tightness → Economic Inactivity	Removed; CLD does not contain inactivity; labour market tightness also impacts job gain and job loss, so impact is best modelled on employed.
Prevalence of Poor Mental Health → Employed	Added
Employed → Labour Market Tightness	Added

Appendix D Simplified causal hypothesis for the simulation

This appendix shows the causal loop diagram aspects that translated into the simulation, and maps each loop and link to its simulated counterpart. It supports Section 6.2, which explains how the Chapter 5 CLD was distilled for simulation.

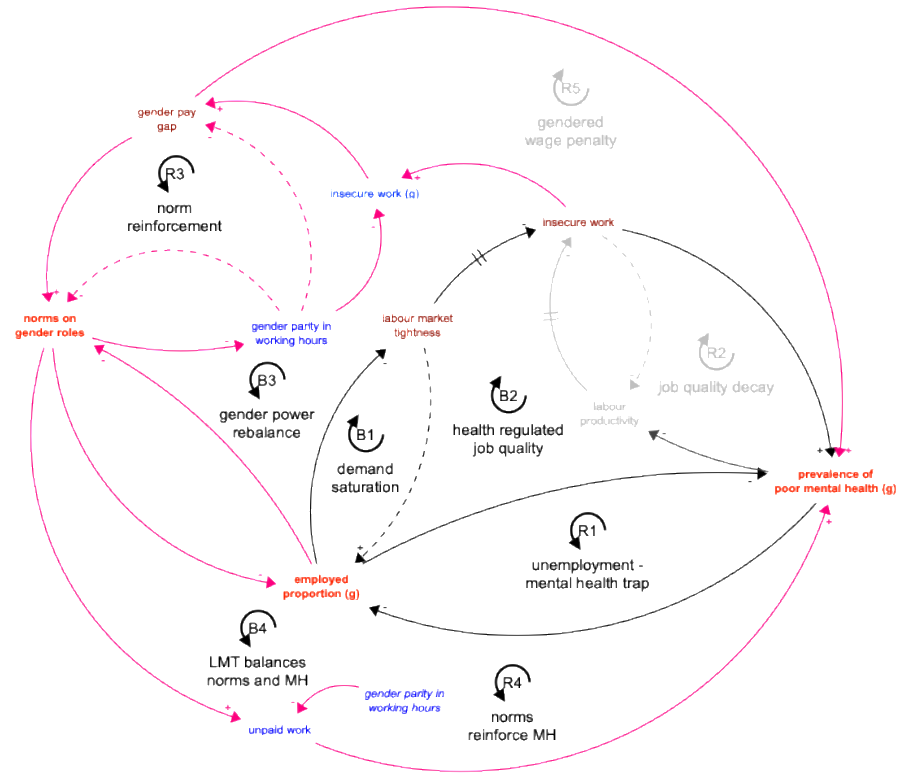


Figure D-1: CLD that was modelled

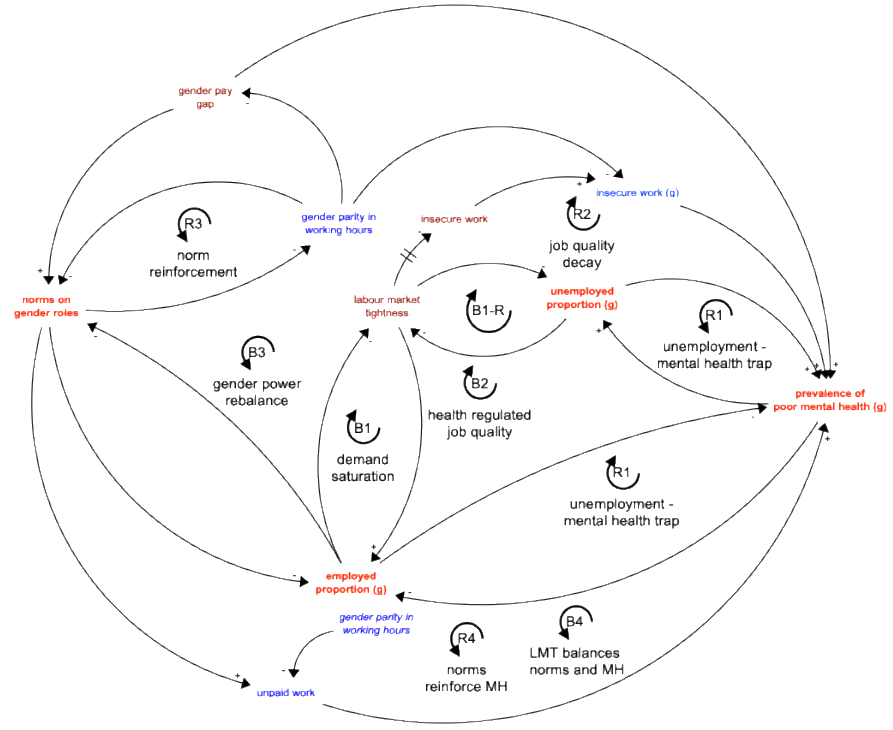


Figure D-2: Loops and Links in the Simulated Model

Appendix E Parameter tables

This appendix gives the full parameterisation of every model parameter — description, derivation, value, method, confidence, and sensitivity — in three per-module tables. It is the detailed companion to the summary tables in Sections 6.3–6.5. The next appendix provides details on sensitivity analysis for relevant parameters. Parameters that provided exogenous inputs to the model are reported separately.

E1 Labour market

Parameter	Description	Derivation	Value (unit)	Method	Confidence
baseline job gain rate	Annual rate at which unemployed individuals transition to employment, before adjustments for poor mental health and LMT.	Quarterly hazard rate estimated from ONS dataset X02 (Labour Market Flows; UK, aged 16–64; Oct–Dec 2001 to Oct–Dec 2019), converted to an annual rate by $1 - (1 - q)^4$. Regression analysis using Excel's LINEST function isolated the rate after accounting for the effect of labour market tightness, giving a raw rate of 0.734. This was then adjusted upward for the effect of poor mental health using DWP analysis of UKHLS (Salis et al., 2021), yielding the adjusted baseline rate of 0.768. (A subset of analyses uses 0.764 due to an early arithmetic error; the difference is immaterial.)	0.768 (1/year)	Data-driven (ONS)	High; UK-level rates applied to Scotland
baseline job loss rate	Annual rate at which employed individuals transition to unemployment, before adjustments.	Same data source and approach as the job gain rate — quarterly hazards from ONS X02 (UK, aged 16–64, Oct–Dec 2001 to Oct–Dec 2019), annualised, and regression-isolated using LINEST after accounting for LMT. Raw rate: 0.042. Adjusted downward for the effect of poor mental health using DWP/UKHLS analysis (Salis et al., 2021): adjusted rate 0.035.	0.035 (1/year)	Data-driven (ONS)	High; UK-level rates applied to Scotland
baseline rate of becoming	Annual rate of direct transition from	Rate estimated from ONS X02 (UK, aged 16–64, from Oct–Dec 2001) via LINEST regression isolating	0.226 (1/year)	Data-driven (ONS)	High; UK-level

Parameter	Description	Derivation	Value (unit)	Method	Confidence
employed from inactivity	inactivity to employment — captures re-entry to the labour market without a measurable unemployment spell at the model's annual time step (e.g. graduates, returning carers).	the rate after accounting for labour market tightness.			
baseline rate of becoming inactive from employment	Annual rate at which employed individuals exit to economic inactivity (e.g. for care responsibilities, sickness or retirement).	Average rate estimated from ONS X02 (UK, aged 16–64, Oct–Dec 2001 to Oct–Dec 2019). Raw rate: 0.072. Adjusted for the effect of poor mental health and effect of state pension change and gender norms; the average of the multipliers for women was calculated and weighted by the employed proportion, and the average rate was back-adjusted by this weighted multiplier to derive the implied baseline rate	0.054 (1/year)	Data-driven (ONS)	Medium; UK-level; three sequential adjustments compound uncertainty
baseline entry rate (inactive → unemployed)	Annual rate at which economically inactive individuals re-enter the labour force as unemployed.	Average rate from ONS X02 (UK, aged 16–64, Oct–Dec 2001 to Oct–Dec 2019). Raw rate: 0.176. Adjusted for the effect of gender norms — the average effect multipliers for men and women were calculated, and the weighted average (by share of inactive) was used to derive the implied baseline rate of 0.221.	0.221 (1/year)	Data-driven (ONS X02); adjusted for endogenous gender-norms multiplier	Medium; UK-level
baseline exit rate (unemployed	The only gender-disaggregated baseline rate, set by calibration to correct	Average UK rate from ONS X02 (16–64, Oct–Dec 2001 to Oct–Dec 2019): 0.552. A symmetric search around this mean (±50%) failed to identify a viable value for women. A secondary empirically-guided	Men: 0.69, Women: 1.1 (1/year)	Constrained calibration (asymmetric	Medium (men); Low-Medium (women) —

Parameter	Description	Derivation	Value (unit)	Method	Confidence
→ inactive), by gender	a structural misalignment in women's inactivity.	search revealed a gendered divergence — men's rates clustered near the UK mean, women's required materially higher values. A third calibration stage adopted asymmetric bounds (men capped at +25%, women at +100%), yielding final values of 0.69 (men) and 1.1 (women). Full calibration logs are in Appendix H2.		bounds; UK mean anchor)	significant departure from UK mean
LMT elasticity elasticity of job gain to labour market tightness elasticity of job loss to labour market tightness elasticity of becoming employed from inactivity to labour market tightness	Cobb-Douglas-aligned elasticity by which Labour Market Tightness modifies the job gain rate, the job loss rate, the rate of becoming employed from inactivity, and the level of insecure work.	Regression analysis using LINEST on ONS quarterly data (Oct–Dec 2001 to Oct–Dec 2019), with the quarterly rate as dependent variable and labour market tightness as independent. Estimated elasticities: job gain 0.3099; job loss −0.3236; becoming employed from inactive 0.2692. For job gain there was a structural break in 2008 (elasticity 0.5 → 0.3); this was model-tested but not retained for parsimony. Given the similar elasticities a single value of 0.3 was selected (applied with the relevant polarity), consistent with Bank of England analysis (Pizzinelli & Speigner, 2017) reporting elasticities between 0.277 and 0.333.	0.3/-0.3 (dimension less)	Data-driven; literature-corroborated	High; UK level data

Parameter	Description	Derivation	Value (unit)	Method	Confidence
normal labour market tightness	Normalisation constant — the effect multipliers equal 1 when LMT equals this value, ensuring the endogenous impact matches the implicit impact already present in the exogenous baseline.	Set to the historical average of labour market tightness over the calibration period (1992–2019).	1.44	Data-driven	High
adjustment time for insecure work	First-order delay for the endogenous effect of LMT on insecure work, reflecting the lag between macroeconomic conditions and renegotiation of employment terms.	3 years based on theory and best data match	3 (years)	Theoretical	Low; not empirically estimated
effect of statutory pension-age change (on women)	Multiplier on women's flow from employment to inactivity, reflecting the gradual increase in state pension age from 60 to 65 between 2009 and 2018.	Two pre-policy datasets (2002–2007) were compared: (1) ONS flow rate data for men and women aged 16–64; (2) Jenkins and Chandler (2010) for men aged 16–64 and women aged 16–59. The resulting ratio of 0.89 was interpreted as the relative decrease attributable to the higher SPA. The effect was formulated as a non-linear table function reducing from 1 to 0.89 between 2009 and 2018, gradually decreasing at first. Sensitivity analysis confirmed no discernible	Linear table function over TIME Anchors: (2009, 1.00); (2018, 0.90) (dimensionless)	Data-driven anchors; theoretical shape	Medium; UK-level; short comparison window

Parameter	Description	Derivation	Value (unit)	Method	Confidence
		difference across alternative shapes; a linear form was selected.			
effect of gender norms on women's inactivity	Multiplicative effect of the prevalence of traditional gender norms on the flow of women from employment to inactivity — captures the structural pull of norms toward female labour market exit.	Originally a sigmoid lookup function chosen to capture threshold effects, with theoretical anchor points: (1) Egalitarian norms (Norms = 0): multiplier 1.0 (women's flow equals men's, no normative pull); (2) Maximum traditional norms (Norms = 1): multiplier 2.0, doubling the baseline rate to reflect strong traditional pressure for women to exit. Sensitivity analysis showed sigmoid/concave/convex/linear shapes gave functionally identical trajectories; a linear form was adopted for parsimony.	Linear table function; Anchors: (0, 1), (1, 2). (dimension less).	Theoretically-driven shape + theoretical anchor points	Low (anchors theoretical); direction of effect supported by literature
effect of poor mental health on the job-gain rate	Multiplier by which population-level prevalence of poor mental health reduces the unemployed → employed (job gain) flow.	Anchored on two literature sources: Butterworth et al. (2012) report that those with poor mental health take 1.36 times the duration to find a job; Germinario et al. (2022) find that depression reduces employment by up to 18%. The curve was anchored at a maximum reduction of 10% (multiplier = 0.9) at 0.5 prevalence. Sensitivity analysis confirmed the shape does not alter the behavioural modes of the employed-proportion stock; a linear form was adopted. Analysis showed that (0.5,0.9) was too weak an effect, almost not capturing the causal link, so this was changed to (0.5, 0.8)	Linear table function; Anchors: (0, 1.0), (0.5,0.8) (dimension less).	Theoretically-driven shape + literature-driven anchor points (Butterworth et al. 2012; Germinario et al. 2022)	High; not Scotland-specific
effect of poor mental health on the job-loss rate	Multiplier by which population-level prevalence of poor mental health	Anchored on van Rijn et al. (2014) meta-analysis OR of 1.61, used to estimate the effect at 0.5 prevalence (1.3). At 0 prevalence the multiplier is 1.0 (no effect). The shape was originally specified	Linear table function; Anchors:	Theoretically-driven shape + literature-driven anchor	High; not Scotland-specific

Parameter	Description	Derivation	Value (unit)	Method	Confidence
	increases the employed → unemployed (job loss) flow.	as non-linear with diminishing returns — as poor mental health becomes more widespread, job loss attributable to it reduces (lack of employer alternatives, reduced stigma): linear up to 0.3 prevalence, slope reducing 0.3–0.4, tapering beyond. Sensitivity analysis confirmed the shape does not alter behavioural modes; a linear form was adopted.	(0, 1.0), (0.5, 1.3). (dimension less).	point (van Rijn 2014)	
effect of poor mental health on labour market exit	Multiplier by which population-level prevalence of poor mental health increases the employed → inactive flow.	Anchored on the Amiri and Behnezhad (2019) meta-analysis pooled risk ratios for the effect of depression on disability pension — 1.82 (95% CI 1.45–2.28) for men and 1.62 (1.31–2.02) for women — broadly consistent with Campbell et al. (2022) (OR 1.16 for non-employment, OR 1.56 for sickness/disability). These were used to estimate the effect at 0.5 prevalence (1.41 men, 1.31 women). At 0 prevalence the multiplier is 1. The shape was originally sigmoidal from 0 to 0.5 prevalence: at low prevalence employers may provide support preventing exit; at high prevalence they cannot afford further staff losses; productivity effects further dampen exits. Sensitivity analysis confirmed the shape does not alter behavioural modes; a linear form was adopted.	Linear table function; Anchors: (0, 1.00), (0.5, 1.36) (dimension less)	Theoretically-driven shape + literature-driven anchor points (Amiri 2019; Campbell et al. 2022)	High; not Scotland-specific

E2 Gender dynamics

Parameter	Description	Derivation	Value (unit)	Method	Confidence
time for norms to change	Characteristic time over which the norms stock responds to a sustained change in inflow or outflow; governs both reinforcement and erosion.	Represents the average time for a population-level shift in gender ideology, assumed to occur via cohort replacement (Perales, Lersch, & Baxter, 2019) — the stickiness of the norm is structurally tied to demographic turnover of the working-age population. The search range (35–50 years) was set from the length of a typical career (16–64 years) and the duration of post-childbirth labour market participation. Within these theoretically plausible bounds, constrained calibration settled at 45 years, aligned with a typical career length.	45 (years)	Constrained calibration (theoretical bounds 35–50 years)	Medium; defensible but not uniquely identified by data
economic-signal composite weights	Weights combining the gender pay gap, parity in working hours (as equality gap), and economic inactivity gap into the composite signal driving norm reinforcement.	Theoretically motivated. Hours and GPG receive equal weight (0.4 each) as no obvious theoretical reason favours one over the other. Inactivity receives a lower weight (0.2) because a reduction in the inactivity gap is a weaker signal of egalitarianism when women are moving into part-time work that itself reinforces traditional roles. Given each variable could only range from 0 to 1, the composite is also bounded 0 to 1. Sensitivity analysis across all theoretically valid combinations (with inactivity weight below GPG/hours weights, and GPG weight ≤ hours weight) showed negligible difference in the prevalence of norms.	GPG 0.4 / Hours 0.4 / Inactivity 0.2 (dimensionless)	Theoretical (sensitivity-tested)	Medium
resistance factor	Multiplier on the reinforcement inflow capturing the cultural stickiness of gender norms — even	Originally formulated as a sigmoid with range (0.1, 0.8) — a floor of 0.1 at low prevalence of traditional norms and a plateau at 0.8 at high prevalence, reflecting cultural lock-in under traditional norms with the system effectively	Linear; resistance = prevalence of norms	Theoretical (simplified after SA)	Low–Medium; substantive necessity confirmed;

Parameter	Description	Derivation	Value (unit)	Method	Confidence
	when economic conditions favour change, norms shift slowly.	decoupled from economic counter-signals. Sensitivity analysis confirmed resistance is structurally necessary — without it, norms erode faster than observed even at the 50-year maximum time constant. A linear formulation with the range simplified to (0, 1) — resistance = prevalence of norms itself — reproduces the historical trend as well as the sigmoid while removing one calibrated parameter. The linear form was therefore adopted; the time constant settles at 45 years.	(dimensionless)		precise form not identified
effect of economic signal on norm reinforcement	Multiplier translating the economic-signal composite into reinforcement pressure on traditional norms.	Originally a concave function to capture hypersensitivity to even marginal inequality — minor disparities serve as sufficient proof to validate traditional roles, with the steepest decline in norm reinforcement occurring only as the system approaches near-perfect parity. Sensitivity analysis showed concave gave the best historical fit at time-for-norms = 40 yrs, linear/convex at 50 yrs, and sigmoid was rejected (best fit > 50 yrs, outside the theoretical bound). For consistency with the 45-year time constant and for parsimony, a linear form was adopted — the effect equals the economic-signal composite itself. This simplification removed one calibrated parameter.	Linear; multiplier = economic-signal composite. anchors: (0, 0), (1, 1). (dimensionless)	Theoretical (simplified after SA)	Medium
gender parity in working hours based on norms	Table function determining the ratio of women's to men's average weekly working	Three anchor points: (1) Traditional anchor (Norms = 1.0, Parity = 0.5) — strict male-breadwinner/female part-time split, aligned with ONS historical averages of part-time vs full-time hours;	Sigmoid table function; Anchors: (0, 1),	Theoretically-driven shape + one literature-driven anchor	Low; mid-range anchor from a single study

Parameter	Description	Derivation	Value (unit)	Method	Confidence
	hours as a consequence of the prevailing gender norms.	(2) Egalitarian anchor (Norms = 0.0, Parity = 1.0) (3) Mid-range anchor (Norms ≈ 0.5, Parity = 0.7), grounded in Andrew et al. (2020), reflecting a post-childbirth average of 28 hours. Shape unidentifiable at the edges; linear in the historical range of norms 0.4 to 0.6.	(0.5, 0.7), (1, 0.5). (dimensionless)	point and two theoretical (Andrew et al.; ONS hours data)	
unpaid work burden based on norms	Table function determining the proportion of unpaid (domestic) work performed by women as a function of the prevailing gender norms.	A concave shape captures the greatest behavioural sensitivity at the egalitarian end of the spectrum. Three anchor points: (1) Traditional (1.0, 0.95): historical peak of gendered specialisation (1960s) per Kan, Sullivan and Gershuny (2011) for liberal and Nordic countries; (2) Egalitarian (0.0, 0.5): theoretical perfect parity; (3) Mid-range (0.55, 0.80): connects 1991 norm data (55% traditional) with observed UK domestic shares of ~78–80% per Kan, Sullivan and Gershuny (2011); Zhou and Kan (2019). The mid-range anchor is set higher than 0.8 because this is before reduction through gender parity in hours. Sensitivity analysis confirmed only a concave shape reproduces magnitude — convex and linear forms fall short even when extended to the theoretical maximum range.	Concave table function. Anchors: (0, 0.5), (0.55, 0.8), (1, 0.95). (dimensionless)	Theoretically-driven shape + literature-driven anchor points (Kan et al. 2011; Zhou & Kan 2019)	Medium–High; alternative shapes cannot reproduce magnitude
effect of gender parity in working hours on unpaid work	Modulator reducing (but not eliminating) the norm-driven unpaid-work disparity as women's hours approach parity.	Originally a piecewise function ranging from (0.5, 1) to (1, 0.5): at maximum disparity in working hours (parity = 0.5) there is no reduction in unpaid-work burden — it is fully determined by norms ($0.5 + (X-0.5)\cdot 1 = X$); at full parity (1) the gap halves (e.g. norm-driven UWB of 0.6 reduces to 0.55). Following sensitivity analysis the	Linear table function; anchors (0.5,1) to (1,0.5)	Theoretical (simplified after SA)	Medium

Parameter	Description	Derivation	Value (unit)	Method	Confidence
		function was reformulated to a simple linear form: (1.5 − parity). Sensitivity analysis showed that the presence of the multiplier matters (without it the unpaid-work burden exceeds the observed level) but the piecewise-vs-linear shape choice does not; linear was adopted for parsimony.	(dimensionless)		
effect of gender parity in working hours on insecure work	Sorting mechanism — distributes the macroeconomic level of insecure work between men and women according to how women's hours fall below parity.	A convex shape was hypothesised: small deviations from equal hours (parity = 1.0) have minimal impact, but as women's hours fall significantly (e.g. below 0.75 parity) the likelihood of being "sorted" into insecure work accelerates disproportionately. The function is anchored on Florisson and Gable (2022) data showing a 1.3 ratio of insecure work prevalence for women relative to men at the current parity value (~0.6). Sensitivity analysis confirmed only a steep concave shape produces the observed gender disparity; other shapes (convex, linear, sigmoid) reverse the gender ordering, giving men higher insecure work prevalence than women.	Concave table function; Range (0.5, 1.30) to (1.0, 1.00); anchor: (0.6, 1.3) (dimensionless)	Theoretically-driven shape + one literature-driven anchor point (Work Foundation 2022)	Medium
effect of gender parity in working hours on gender pay gap	Additive structural penalty on the gender pay gap reflecting the disadvantage of part-time work	Following Manning and Petrongolo (2006), the raw pay penalty for part-time work of 25 log points reduces to 12 log points after controlling for personal characteristics. This 12-log-point structural penalty was converted to a percentage change using e^(log points) − 1 = 0.113, used to set the multiplier value at parity ~0.6 (the study-period level). The function is modelled as concave so that even small deviations from full-time work yield immediate penalties (career-track	Concave table function; (0.5, 0.12) → (1.0, 0) (dimensionless)	Theoretically-driven shape + one literature-driven anchor point (Manning & Petrongolo 2006)	Medium; anchored on a single high-quality source

Parameter	Description	Derivation	Value (unit)	Method	Confidence
		exclusion). Sensitivity analysis confirmed the magnitude is sensitive to shape: a steep concave function gives the best fit, slightly overestimating the gap in early years but matching the long-run trajectory.			
fraction of gap closed by childcare	Modulates the extent to which government-funded childcare reduces the parity gap dictated by norms. Ensures childcare can alleviate time constraints but the ceiling for parity is still governed by prevailing norms.	Modelled as convex to represent increasing marginal utility i.e. low levels of provision (a few hours per week) are insufficient to change labour market status from part-time to full-time. Anchored at two extremes: (1) Zero availability (0, 0): no mitigating effect on the gender parity dictated by norms; (2) Maximum availability (1, 0.5): the parity gap (1.0 − baseline parity) is reduced by 50% (e.g. a baseline parity of 0.8 increases to 0.9). The 50% closure cap acknowledges childcare facilitates higher participation but full parity (1.0) requires the long-term erosion of norms. Sensitivity analysis showed the shape produced no material difference in historical behaviour given low historical availability; concave and linear forms overestimate gains; convex retained on theoretical grounds.	Convex table function; Anchors: (0, 0), (1, 0.5) (dimension less)	Theoretically -driven shape + theoretical anchor points	Low; no material difference in historical trajectory given low availability

E3 Mental health

Parameter	Description	Derivation	Value / form	Method	Confidence
baseline onset rate	Annual probability that an individual in good mental	An initial UK estimate was drawn from DWP analysis (Salis et al. 2021, Table 2.2): across the seven UKHLS waves analysed, on average 14.1% of	0.08 (1/year)	Calculated — back-adjusted	High; multiple data paths

Parameter	Description	Derivation	Value / form	Method	Confidence
	health transitions to poor mental health, before the effects of endogenous model variables; represents the combined influence of all unmodelled risk factors.	the at-risk population experienced an onset. The continuous rate was obtained from the survival relationship S(T) = e^(−rT): setting (1 − P) = e^(−r) for an annual period yields r = −ln(1 − 0.141) = 0.151. This was back-adjusted by the weighted multiplier to remove the structural burden of economic and gender stressors present in 1992, yielding a UK-based baseline onset rate of 0.121. This was subsequently superseded with Scottish data: primary analysis of the Scottish BHPS/UKHLS sub-sample found observed 1992 onset rates of 0.086 (Men) and 0.119 (Women), and back-adjustment by the weighted 1992 multiplier yielded an implied baseline rate of 0.08 used in the final model.		from observed Scottish 1992 onset by removing the weighted 1992 effect multipliers	converge on a consistent value
baseline recovery rate	Annual probability of transition from poor mental health to good mental health; captures the combined effect of all real-world factors not explicitly modelled.	An initial UK estimate was drawn from Table 2.2 of Salis et al. (2021): on average 49.3% of those in poor mental health transitioned to good mental health, giving r = −ln(1 − 0.493) = 0.679. Using primary Scottish BHPS/UKHLS data, the implied baseline rates were calculated from the 1992 equilibrium condition (onset = recovery at t = 0): 0.409 for men and 0.346 for women. Following integration of the three sub-modules, these were recalculated to 0.4096 (men) and 0.348 (women) to ensure the system initialises in a state of partial equilibrium, because the resting state of the integrated model differs slightly from the isolated historical proxies used during module development.	Men: 0.4096; Women: 0.348 (1/year)	Calculated from the 1992 equilibrium constraint; recalculated post-integration	Medium; depends on the equilibrium assumption and on the values of the multipliers used in back-adjustment
effect of economic	Multiplier on the incidence rate	Anchored in odds ratios from meta-analyses. Thomson et al. (2023) provides UK gender-	Concave table function;	Theoretically-driven shape	High; odds ratios from

Parameter	Description	Derivation	Value / form	Method	Confidence
stressors on mental health	from the composite of unemployment, change in unemployment, and insecure work.	disaggregated ORs for unemployment (M 2.02 / W 1.48) and unexposed prevalence (15.02% M / 20.75% W) — suggesting female employment is already subject to unmodelled stressors. The overall OR of 1.66 is adopted as a universal baseline for economic stress to avoid "double-taxing" women's vulnerability — the gender gap then emerges from interaction with the gender-specific GPG and unpaid-work multipliers. Rönnblad et al. (2019) provides OR 1.52 (95% CI 1.35–1.70) for insecure work. Curtis et al. (2019) provides ORs comparing low/stable to higher/stable employment; the reciprocal of 0.833 → OR 1.20, RR 1.165 for change in unemployment. ORs converted to RRs using RR = OR / ((1 − P) + (P × OR)) with the single empirical anchor of 18.24% unexposed prevalence from Thomson et al. (2023): unemployment 1.48, insecure work 1.39, change in unemployment 1.16. Weighted average = (0.3 × 1.48) + (0.4 × 1.39) + (0.3 × 1.16) = 1.348. A concave shape captures diminishing marginal impact at high stressor levels (stigma reduction). Sensitivity analysis found little difference but concave produces highest MH response.	Anchors: (0.0, 1.0), (1.0, 1.35); (dimensionless)	+ literature-driven anchor points (meta-analyses: Thomson 2023; Rönnblad 2019; Curtis 2019)	meta-analysis
economic-stressor composite weights	Composite aggregating unemployment (30%), insecure work (40%) and change in	Theoretically grounded — effect sizes for unemployment are generally larger than insecure work, so the two unemployment pathways (unemployment, change in unemployment) jointly receive 0.6 to ensure insecure work has sufficient impact. Both unemployment pathways are	Insecure 0.4 / Unemployed 0.3 / Change in unemployment 0.3	Theoretical (sensitivity-tested)	Medium

Parameter	Description	Derivation	Value / form	Method	Confidence
	unemployment (30%).	weighted equally given the focus on population-level effect. The first two components are gender-disaggregated; the change-in-unemployment component is at aggregate level. Sensitivity analysis across weight combinations showed minimal trajectory difference.	(dimensionless)		
effect of gender inequality on mental health	Multiplier on women's incidence rate from the women's-inequality composite.	Anchored in two studies with ORs converted to RRs using historical average prevalence in men (P = 0.15) as the unexposed baseline. Platt et al. (2016) provides an OR of 2.43 for the effect of wage disparities on women's likelihood of depression, converted to RR 2.01. Piovani and Aydiner-Avsar (2021) report an OR of 1.021 per additional hour of weekly unpaid work performed by women; with total per-week hours estimated at 20 (Wheatley, Lawton, & Hardill, 2018), a 10-hour increase from parity yields OR 1.021^10 = 1.22, RR 1.18. Weighted maximum RR = (1.18 × 0.6) + (2.01 × 0.4) = 1.512 ≈ 1.5. A sigmoid shape scaled to a "threshold of perception": at low levels disparities perceived as statistical noise (small impact); as disparity widens it becomes an undeniable signal of injustice (steep increase); beyond a threshold the marginal psychological impact diminishes as the stressor saturates as a chronic feature of the social environment. Sensitivity analysis found some difference but sigmoid and concave similar, retained over linear/convex for theoretical reasons.	Sigmoid table function; Anchors: (0, 1.0), (0.143, 1.011), (0.286, 1.051), (0.429, 1.164), (0.571, 1.336), (0.714, 1.449), (1.0, 1.500). (dimensionless)	Theoretically-driven shape + literature-driven anchor points (Platt 2016; Piovani & Aydiner-Avsar 2021; Wheatley et al. 2018)	Medium; based on individual studies rather than meta-analyses
women's-inequality	Weights combining the	Evidence-based: expert engagement and the literature review confirmed the double-burden	Unpaid work 0.6 / GPG 0.4	Theoretical (evidence-	Medium

Parameter	Description	Derivation	Value / form	Method	Confidence
composite weights	unpaid-work burden (60%) and gender pay gap exposure (40%) into the women's-inequality composite.	link between unpaid domestic work and distress is well established. The 60/40 split favours unpaid work, reflecting the stronger evidence base for its mental health effect. The GPG component is multiplied by the female employed proportion (population-level exposure rises as women's employment rises, even when the gap shrinks). Sensitivity analysis found that magnitude differs across weight combinations but trajectory is stable.	(dimensionless)	based weighting; sensitivity-tested)	
normative sensitivity multiplier (within women's-inequality composite)	Multiplier amplifying the felt intensity of unpaid work as norms erode — the persistence of the double burden is realised as a sharper stressor when it conflicts with modern egalitarian expectations.	Formulated as 1 + (1/norms − 1). At full traditional norms (norms = 1) the multiplier equals 1 (no amplification). As norms erode toward 0, 1/norms increases sharply, amplifying the felt unpaid-work stressor. Sensitivity analysis showed varying the normative sensitivity altered the composite magnitude but the stressor trajectory remained stable across scenarios, except when set to zero (no amplification).	1 + (1/norms − 1) (dimensionless)	Theoretical	Low; formulation theoretical
perceptual delay (SMTH3 for change in unemployment)	Smoothing parameter on the change-in-unemployment input to the economic-stressors composite, capturing the lag	Third-order smoothing (SMTH3) with an estimated two-quarter perceptual delay; starting value 0.11, based on UK unemployment levels in 1991–1992 (unemployment was rising in 1992). The smoothing function represents the physiological toll of volatility as accumulative rather than instantaneous — when the external threat subsides, the internalised damage dissipates slowly.	SMTH3 with adjustment time 0.5 yr (≈2 quarters); initial value 0.11 (years)	Theoretical	Low

Parameter	Description	Derivation	Value / form	Method	Confidence
	between economic volatility and its psychological toll.				
effect of employment on recovery	Multiplier capturing the positive impact of employment on the recovery rate, before adjustment for insecure work.	Derived from DWP analysis of UKHLS Table 4.12 (Salis et al., 2021), reporting a regression coefficient of 0.156 from a log-hazard model; the corresponding discrete-time Hazard Ratio is e^0.156 ≈ 1.169. A linear relationship is modelled because the pathway is considered to operate at the individual level only. The anchor point at employed proportion = 1 is set slightly higher (1.2) than the empirical hazard ratio (1.169) to represent the maximum recovery dividend available in the labour market.	Linear table function; Anchors: (0, 1.000), (1.0, 1.200). (dimensionless)	Linear; literature-driven anchor point (Salis et al. 2021)	Medium; UK-level discrete-time hazard
effect of insecure work on the recovery benefit of employment	Modifier that erodes the recovery benefit of employment as insecure work rises above its 1992 level; it captures insecure work as a structural barrier that keeps people in poor mental health, rather than just a stressor that causes it.	Adjustment was formulated as relative, because direct empirical data for this effect were unavailable. An increase in insecure work from 1992 reduces the positive impact of employment; at levels equal to or below 1992 the multiplier equals 1 (no erosion). At a calibrated threshold (see next row) the positive effect is reduced to zero (multiplier = 1; only baseline recovery applies). Reaching this threshold does not imply secure employment ceases to provide individual recovery benefits; it signifies a point of net structural neutralisation at the population level. The full revised formulation is: 1 + (effect_of_employment_on_recovery − 1) × table function(change in insecure work / INIT(insecure work)).	Linear table function applied to change in insecure work (relative to 1992). Men: (1.00, 1) → (1.40, 0); Women: (1.00, 1) → (1.20, 0) (dimensionless)	Theoretical (relative adjustment, no direct empirical anchor; threshold calibrated separately)	Low

Parameter	Description	Derivation	Value / form	Method	Confidence
Insecure work threshold (effect on recovery)	Gender-specific change in insecure work (relative to 1992) at which the recovery benefit of employment is neutralised.	Estimated separately for men and women using a phased four-stage constrained-calibration strategy. Stage 1 (gender-neutral) confirmed gender-neutral assumptions fail to replicate the gendered pattern (9-fold increase in error for the female sub-system). Stage 2 (unconstrained) identified the "basins" for gendered resilience. Stage 3 (gender-refined within sociologically defensible bounds, e.g. 20–25% capacity margins) showed the model's fit is robust within those bounds. Stage 4 (conditional freeze) fixed the labour market thresholds and re-estimated the prevalence-clogging effect, confirming the stability of the prevalence thresholds across phases. The calibrated thresholds diverge by gender, illustrating path dependency anchored in 1992 baselines: women's higher 1992 insecure work (0.37 vs men's 0.24) implies that their recovery capacity is neutralised after a smaller relative increase albeit at a higher absolute level. Calibration yielded neutralisation at absolute insecure work = 0.44 for women (~20% relative increase) versus 0.34 for men (~40% relative increase).	Men: 1.4 (≈40% rise from 1992); Women: 1.2 (≈20% rise from 1992) (dimensionless)	Constrained calibration (4-phase strategy)	Low
effect of prevalence on recovery rate	Endogenous reduction in the recovery rate as community prevalence of poor mental health rises above its 1992 value —	Originally specified as a sigmoid with a floor at half the baseline recovery rate. Loop-dominance analysis showed this contributed essentially nothing to model dynamics during 1992–2019. A sigmoid was also theoretically inappropriate, assuming a buffer zone where small demand increases have no effect on service capacity — implausible for systems operating near capacity.	Halving point = 1.5 (recovery halves when prevalence reaches 150% of 1992 baseline). Men: concave;	Constrained calibration (shape + halving point; theoretical retention of 1.5 after bounds hit)	Low; gender asymmetry, calibrated

Parameter	Description	Derivation	Value / form	Method	Confidence
	captures the reinforcing dynamic where rising prevalence strains services, slows recovery, and allows prevalence to rise further (Clark et al., 2018; Siciliani, 2025)	Following a four-stage constrained calibration, the relationship was simplified to a linear table function for women and a concave function for men, with an implied halving point of 1.5 retained on theoretical grounds because calibration hit the search bounds without uniquely identifying the value. Loop dominance analysis confirmed the new formulation activates the previously dormant prevalence-recovery reinforcing loop.	Women: linear; Anchors (1.00, 1.000), (1.50, 0.500) (dimensionless)		

Appendix F Sensitivity-tested parameters

This appendix provides details on sensitivity analysis for parameters across the three modules.

F1 Labour market module

Parameter	Estimation method (parameter type)	Sensitivity	Sensitivity-analysis finding
baseline exit rate (unemployed → inactive), by gender	Constrained calibration (calibrated constant)	Medium	Empirically-guided search revealed gendered divergence. Revised rates (compared to data-driven estimates) significantly shift the system's resting state (lower female unemployment, higher female inactivity) and correct a structural misalignment in women's inactivity; analysis confirms the *direction* of results is unchanged, only magnitude.
effect of gender norms on women's inactivity	Theoretical (table function)	Low	Sigmoid, concave, convex and linear shapes produce trajectories identical in direction and timing, differing only marginally in magnitude; a linear shape was adopted for parsimony.
effect of poor mental health on the job-gain rate effect of poor mental health on the job-loss rate effect of poor mental health on labour market exit	Literature-based + theoretical shape (table function)	Low	The shape does not alter the behavioural modes of the employed-proportion stock; a linear shape was adopted for parsimony.
effect of statutory pension-age change	Data-driven + theoretical shape (table function)	Low	No discernible difference across alternative shapes; a linear shape was selected.

F2 Gender Dynamics Module

Parameter	Estimation method (parameter type)	Sensitivity	Sensitivity-analysis finding
effect of gender parity in working hours on insecure work	Literature-based + theoretical shape (table function)	Critical	Only a steep concave shape produces the observed gender disparity; other shapes reverse the gender ordering, giving men a higher prevalence of insecure work than women.
resistance factor	Theoretical (table function)	High	Without the resistance factor, norms erode faster than observed; the factor is structurally necessary to reproduce the historical trend. No significant difference between linear and sigmoid in the simulation period; concave gives slower deterioration, convex faster. Linear with range simplified to (0, 1) removed one unknown parameter.
unpaid work burden based on norms	Literature-based + theoretical shape (table function)	High	Only a concave shape reproduces the observed magnitude of the unpaid-work burden; convex and linear shapes fall short even when extended to their theoretical maximum range.
time for norms to change	Constrained calibration (calibrated constant)	Medium	Within the 35–50-year theoretical range, the prevalence-of-norms trajectory shifts in magnitude and interacts with the shape of the economic-signal effect; 45 years is suitable across shapes but is not uniquely identified by the data.
effect of economic signal on norm reinforcement	Theoretical (table function)	Medium	concave, linear, convex and sigmoid forms give materially different rates; a linear form was adopted after analysis for parsimony.
gender parity in working	Literature-based + theoretical	Medium	Magnitude shifts across shapes but the qualitative trajectory does not.

Parameter	Estimation method (parameter type)	Sensitivity	Sensitivity-analysis finding
hours based on norms	shape (table function)		
economic signal composite weights	Theoretical (composite weights)	Low	Negligible difference across all theoretically valid weight combinations.
effect of gender parity in working hours on unpaid work	Theoretical (table function)	Low	The presence of the multiplier matters (without it the unpaid-work burden exceeds the observed level) but the shape choice does not; a linear form was adopted for parsimony.
effect of gender parity on gender pay gap	Literature-based + theoretical shape (table function)	Medium	The magnitude of the simulated gender pay gap depends on this shape
fraction of gap closed by childcare	Theoretical (table function)	Low	The shape produced no material difference in historical behaviour given low historical availability; a convex shape was retained on theoretical grounds.

F3 Mental Health Module

Parameter	Estimation method (parameter type)	Sensitivity	Sensitivity-analysis finding
prevalence threshold	Constrained calibration (calibrated constant)	Medium	Multivariate analysis shows that it is robust across the calibration-supported range.
Insecure work threshold (effect on the recovery benefit of employment)	Constrained calibration (calibrated constant)	Medium	The gender-specific neutralisation important for calibration; women's trajectory is largely insensitive to the precise value within the explored bounds (1.20–1.40)

Appendix G Parameters providing exogenous input

This appendix documents every historical time series fed into the model as an input, with sources and estimation methods. It supports Section 6.2.4 (parameterisation philosophy) and the module parameterisation in Sections 6.3.2, 6.4.7 and 6.5.5; the labour-demand and insecure work series in particular required composite estimation, detailed here.

G1 Demographic variables

Working-age population (WAP) and **proportion of WAP that is male** was directly extracted from NOMIS Population estimates: local authority based by single year of age (Source: National Records of Scotland (NRS)).

	WAP	**proportion of WAP that is male**
1992	3289639	0.4929
1993	3289390	0.4927
1994	3292415	0.4924
1995	3293104	0.4921
1996	3288190	0.4915
1997	3284299	0.4912
1998	3281902	0.4907
1999	3281482	0.4902
2000	3278692	0.4897
2001	3286645	0.4898
2002	3295351	0.4899
2003	3303864	0.4900
2004	3321295	0.4898
2005	3345841	0.4899
2006	3368993	0.4903
2007	3400324	0.4907
2008	3423498	0.4909
2009	3441098	0.4910
2010	3462158	0.4910
2011	3487433	0.4916
2012	3470459	0.4911
2013	3461727	0.4911
2014	3454426	0.4906
2015	3458694	0.4902
2016	3463312	0.4901
2017	3463600	0.4897
2018	3455588	0.4892
2019	3457100	0.4891
2020	3448390	0.4888
2021	3439686	0.4884
2022	3457768	0.4888
2023	3479044	0.4892

G2 Labour demand

Labour demand is the sum of workforce jobs and vacancies. Jobs data is extracted from NOMIS "workforce jobs by industry (SIC 2007) - seasonally adjusted" (June values). Workforce jobs before 1996 are derived by scaling UK jobs to Scotland. ONS vacancy data for UK is available from 2001. Five distinct methods for estimating Scottish vacancies from 1992 were considered (1) direct scaling of UK data (using either workforce jobs or (2) unemployment ratios), (3) a model-based back-casting (using a UK Beveridge Curve (ln(Vacancies)=ß +ß x ln(Unemployment Rate)), scaled by either workforce jobs or (4) unemployment ratios) and (5)a hybrid approach (using direct scaling on official data when available (post-2001) and a Scotland-specific model for back-casting). The back-casting methods rely on the core assumption that the inverse relationship between unemployment and vacancies remained structurally stable between the pre-2001 and post-2001 periods. Analyses indicated that the series from the direct scaling approach reflects the volatility of the source ONS data, while the model-based approaches filter out the short-term noise, representing a more stable, underlying trend. The methods scaled by unemployment consistently produced higher mean vacancy estimates, suggesting a structural difference in Scotland's share of UK jobs versus its share of UK unemployment. The hybrid series was selected as it incorporates measured volatility of the labour market.

1992	2355275
1993	2316906
1994	2330142
1995	2364363
1996	2514503
1997	2514776
1998	2523519
1999	2508610
2000	2518856
2001	2653639
2002	2640546
2003	2660565
2004	2678692
2005	2736678
2006	2785144
2007	2798798
2008	2838197
2009	2693915
2010	2643075
2011	2662961
2012	2680818
2013	2669920
2014	2775563
2015	2786497
2016	2803072
2017	2830814
2018	2808633
2019	2836470

2020	2786775
2021	2791025
2022	2911799
2023	2875129
2024	2968018

G3 Baseline gender pay gap

It is the median hourly earnings (excluding overtime) for full-time employees, which isolates factors such as direct discrimination and historical systemic inequities from the endogenous impacts of part-time work. Data post-1997 is anchored in the ONS Annual Survey of Hours and Earnings (ASHE). For the pre-1997 period, this was estimated using data from New Earnings Survey. Trend extrapolation was selected over proportional scaling and linear regression. While scaling and regression methods introduced discontinuities (+2.6 and +1.1 percentage points) by relying on mismatched GB-mean proxies, the extrapolation method provided a near-perfect join at the 1996/1997 transition (18.22% vs. 18.40%). This prioritises simulation stability and avoids introducing shocks to the model during initialisation. The primary limitation is that this masks any non-linear economic dynamics of the early 1990s (e.g. recovery from recession) which could have altered GPG's trajectory.

1992	0.2062
1993	0.2002
1994	0.1942
1995	0.1882
1996	0.1822
1997	0.184
1998	0.191
1999	0.167
2000	0.181
2001	0.171
2002	0.157
2003	0.132
2004	0.117
2005	0.09
2006	0.106
2007	0.119
2008	0.108
2009	0.085
2010	0.074
2011	0.066
2012	0.084
2013	0.077
2014	0.091
2015	0.077
2016	0.064
2017	0.066
2018	0.056
2019	0.072

2020	0.03
2021	0.03
2022	0.03
2023	0.014
2024	0.022

G4 Insecure work

A composite was constructed to measure cumulative risk in the population. (1) Low Pay: Derived from Scottish RLW data (2012–2024) and spliced with GB historical trends (pre-2012). (2) Involuntary Temporary Employment (3) Underemployment, time-based. (4) Zero-Hours Contracts (ZHC): assigned a double weighting. Data came from ONS; Low-pay figures for 1991 to 98 came from "FIGURE 1: The minimum wage has almost eradicated ‘low’ pay among employees" (Low Pay Britain 2025, Resolution Foundation). ONS data was used for the remaining years. A linear trend is derived from the 1992 and 2019 values, after partialling out the contemporaneous effect of labour market tightness.

	Total	Effect of LMT	Baseline trend
1992	0.2708	1.4664	0.1847
1993	0.2807	1.4563	0.1927
1994	0.2764	1.4240	0.1941
1995	0.2863	1.3710	0.2089
1996	0.2867	1.2837	0.2234
1997	0.2807	1.2000	0.2339
1998	0.2754	1.1503	0.2394
1999	0.2747	1.1258	0.2440
2000	0.2735	1.1145	0.2454
2001	0.2665	1.0773	0.2473
2002	0.2641	1.0370	0.2547
2003	0.2506	1.0153	0.2468
2004	0.2537	1.0009	0.2534
2005	0.2568	0.9848	0.2608
2006	0.2533	0.9646	0.2626
2007	0.2451	0.9520	0.2575
2008	0.2477	0.9463	0.2617
2009	0.2825	0.9774	0.2890
2010	0.2991	1.0631	0.2813
2011	0.3041	1.1245	0.2704
2012	0.3161	1.1414	0.2770
2013	0.3455	1.1423	0.3025
2014	0.3498	1.1189	0.3127
2015	0.3436	1.0846	0.3168
2016	0.3341	1.0643	0.3139
2017	0.3326	1.0476	0.3175
2018	0.3315	1.0373	0.3196
2019	0.3120	1.0307	0.3027

G5 Childcare availability index

This is a theoretically constructed exogenous index representing the availability and generosity of funded childcare in Scotland. It is constructed from policy milestones using two components: (1) Intensity — government-funded hours per year as a proportion of a standard full-time working year (35 hours/week × 47 weeks); (2) Coverage — proportion of 1–4 year olds covered by the policy. The two components are multiplied.

Year	Policy	Hours Funded	Ages Covered	Intensity	Coverage	Generosity index
1996	412.5 hours per year for 4-year olds	412.5	1	0.2508	0.2500	0.0627
2002	412.5 hours per year for 3 and 4-year olds	412.5	2	0.2508	0.5000	0.1254
2007	475 hours per year for 3 and 4-year olds	475	2	0.2888	0.5000	0.1444
2014	600 hours per year for 3 and 4-year olds and eligible 2-year olds	600	2.13	0.3647	0.5325	0.1942
2019	600 hours per year for 3 and 4-year olds and eligible 2-year olds	600	2.13	0.3647	0.5325	0.1942
2021	1140 hours per year for 3 and 4-year olds and eligible 2-year olds	1140	2.13	0.6930	0.5325	0.3690

Working hours per week	35
Weeks per year	47
Total employment hours per year	1645

Intensity is funded hours/total employment hours

Coverage is ages covered/4

Appendix H Parameterisation via calculation and constrained calibration

H1 Calculation of baseline onset and recovery rates

This appendix shows the back-calculation of the baseline onset rate and the equilibrium-constraint calculation of the baseline recovery rate, supporting Section 6.5.5. These calculations ensure the module initialises in equilibrium so that subsequent dynamics are feedback-driven.

	Men	Women
Observed Onset Rate in 1992	0.086	0.119
1992 multipliers on Incidence	1.075	1.508
Implied Baseline Onset Rate	0.080	0.079
Prevalence of Poor Mental Health in 1992	0.155	0.235
1992 multipliers on Recovery	1.150	1.120
Implied Baseline Recovery Rate	0.409	0.346
Recalculated post integrated	0.4096	0.348

H2 Constrained calibration for baseline exit rates

This appendix documents the three-stage calibration of the unemployment-to-inactivity exit rate, referenced in Section 6.3.2. It records the optimiser settings and shows why UK-wide benchmarks were insufficient for women, motivating the asymmetric calibration bounds.

Key Configuration Settings: Payoff elements were weighted using the inverse of their historical means (e.g., Men's Employment [0.752] weight = 1.3; Men's Unemployment [0.097] weight = 10.3). This ensures that the optimizer assigns equal significance to percentage deviations across all stocks, regardless of their absolute magnitude. This prevents large stock bias, where the model might otherwise ignore fit errors in the smaller, theoretically critical unemployment stocks.

Element	Labour market. Unemployed proportion [Men]	Labour market. Unemployed proportion [Women]	Labour market. Inactive proportion [Men]	Labour market. Inactive proportion [Women]	Labour market. Employed proportion [Men]	Labour market. Employed proportion [Women]
Weight	15.6	24.7	5.9	3.4	1.3	1.5

Maximum Simulations (= -1): This setting allows the optimiser to run indefinitely until the specified tolerance is met, ensuring the algorithm converges fully to the best mathematical fit without premature termination.

Tolerance (0.00001): A high precision threshold was set to ensure the final parameters are stable and robust.

Additional Starts (10): To guard against the risk of "local optima"—where the optimiser gets stuck in a sub-optimal solution—ten independent searches were initiated from randomized starting points across the parameter space.
Initial Step (Init_step = 0.01): In the final gender-disaggregated calibration, the initial step was reduced to 0.01 to ensure a fine-grained search within the constrained theoretical bounds.
Optimisation 3 was repeated after table functions were replaced with linear formulations, and the same values were returned.

Optimisation 1: Literature-based

Parameter:	Labour market.baseline exit rate[Men]	Labour market.baseline exit rate[Women]
min_value	0.27	0.27
max_value	0.83	0.83
scaling	1	1

Payoff	Labour_market.Unemployed_proportion [Men]	Labour_market.Unemployed_proportion [Women]	Labour_market.Inactive_proportion [Men]	Labour_market.Inactive_proportion [Women]	Labour_market.Employed_proportion [Men]	Labour_market.Employed_proportion [Women]
622.796	172.738	299.19	52.8118	71.4841	11.0809	15.492

	Labour market.baseline exit rate[Men]	Labour market.baseline exit rate[Women]
Lower	0.785568898717	*
Value	0.806706552195	*
Upper	0.82755146823	*

Optimisation 2: Empirically guided

Parameter:	Labour market.baseline exit rate[Men]	Labour market.baseline exit rate[Women]
min_value	0.2	0.2
max_value	2.2	2.2
scaling	1	1

Payoff	Labour_market.Unemployed_proportion [Men]	Labour_market.Unemployed_proportion [Women]	Labour_market.Inactive_proportion [Men]	Labour_market.Inactive_proportion [Women]	Labour_market.Employed_proportion [Men]	Labour_market.Employed_proportion[Women]
443.69	161.773	156.19	62.2048	35.633	12.6127	15.2767

	Labour market.baseline exit rate[Men]	Labour market.baseline exit rate[Women]
Lower	0.865295131183	1.65385077729
Value	0.94223226719	1.8274537997
Upper	1.0116570045	1.96432502113

Optimisation 3: Literature-guided and gender disaggregated based on empirical results

Parameter:			Labour market.baseline exit rate[Men]			Labour market.baseline exit rate[Women]
min_value			0.41			0.41
max_value			0.69			1.1
scaling			0.1			0.1
Payoff	Labour_market.Unemployed_proportion[Men]	Labour_market.Unemployed_proportion[Women]	Labour_market.Inactive_proportion[Men]	Labour_market.Inactive_proportion[Women]	Labour_market.Employed_proportion[Men]	Labour_market.Employed_proportion[Women]
508.634	175.204	201.014	63.6272	47.3217	9.35703	12.1098

	Labour market.baseline exit rate[Men]	Labour market.baseline exit rate[Women]
Lower	*	*
Value	*	*
Upper	*	*

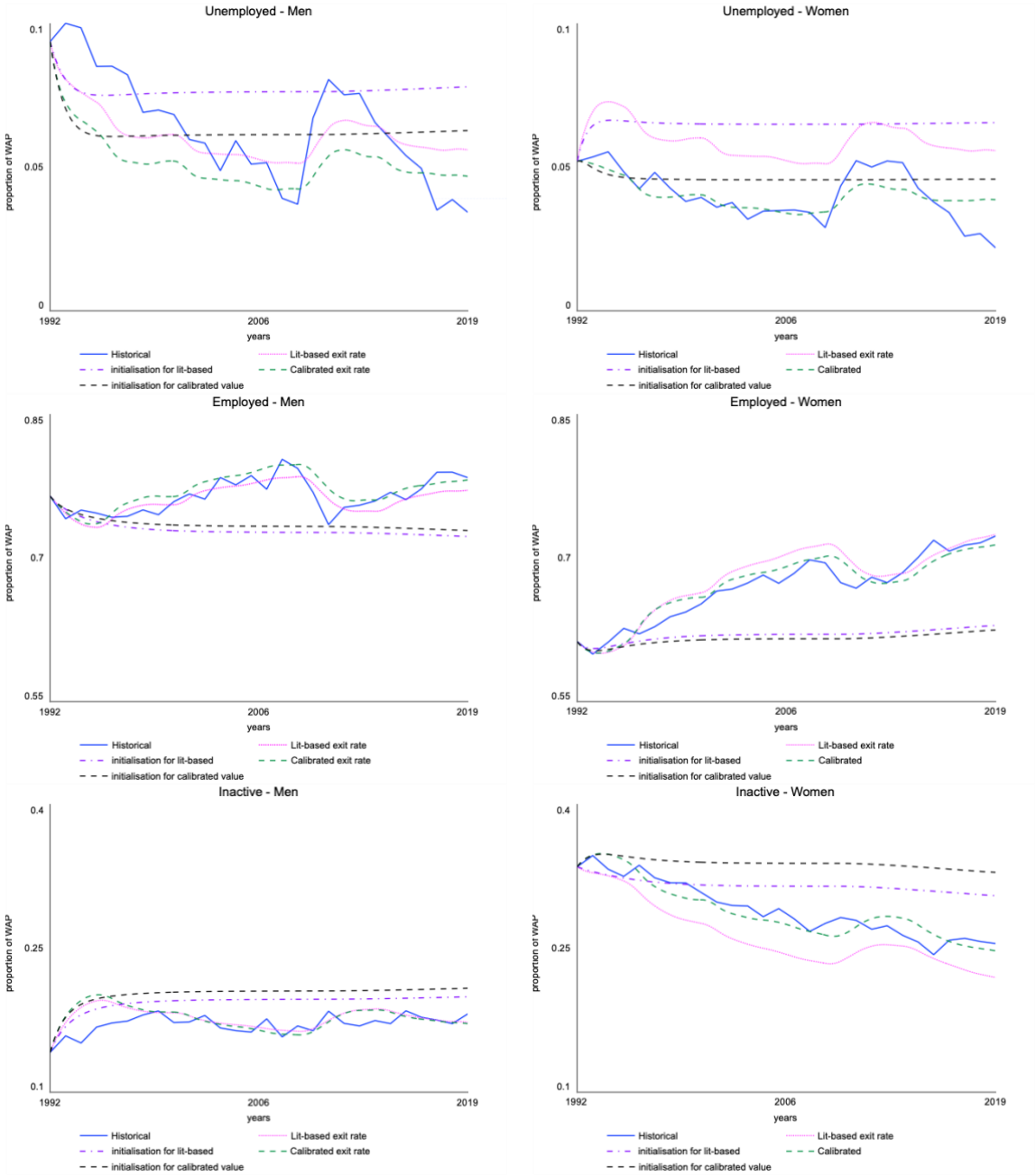


Figure H-1: Resting state of the labour market under calibrated and literature-based baseline exit rates

Demonstrates that the calibrated rates correct a structural misalignment in women's inactivity

H3 Constrained calibration for time for norms to change

Method	stoponerr	tolerance	maxiter	init_step
Powell		0.00001	1000	1

Payoff:	**Fit to Norms**
Action	minimize
Kind	Calibration
Element	Gender norms.Prevalence of Norms on Gender Roles
Weight	1
Comparison Variable	Historical Data.Historical Gender Norms
Comparison Run	-2
Comparison Type	Absolute Error
Comparison Tolerance	0

Parameter:	Gender norms.time for norms to change
min_value	30
max_value	55
scaling	1

	Gender norms.time for norms to change	**Fit to Norms**
Starting at	45	5.86976
After 56 runs	47.4334	5.70576

Computing confidence bounds(+/-) for payoff change of 2

	Gender norms.time for norms to change
Lower	40.1377328629
Value	47.4334000251
Upper	*

Rounded to the near 5 to 45.

H4 Constrained calibration for insecure work effect on recovery

This appendix documents the four-stage calibration of the insecure work thresholds that neutralise the recovery benefit of employment, referenced in Section 6.5.3.

Phased calibration strategy

To identify the structural parameters governing mental health recovery, a four-stage calibration process was employed. This strategy was designed to test the hypothesis of gender-based resilience and to ensure that the identified thresholds were sociologically plausible rather than artifacts of mathematical over-fitting.

Phase 1 (gender-neutral control): An optimisation forcing identical thresholds for both sexes. This stage confirmed that gender-neutral assumptions fail to replicate the "Scottish Paradox," resulting in a nine-fold increase in error for the female population.

Phase 2 (unconstrained discovery): Parameters were allowed to converge freely across wide boundaries. This identified the "basins for gendered resilience: high sensitivity for women and high resilience for men.

Phase 3 (gender refinement): Optimisation was restricted to theoretically defensible "Conceptual Benchmarks." This phase proved that the model's fit remains robust when thresholds are constrained to plausible sociological buffers (e.g., 20-25% capacity margins).

Phase 4 (conditional freeze): To avoid parameter confounding, the labour market thresholds were fixed, and the systemic "prevalence clogging" effects were isolated and re-estimated. The stability of the prevalence thresholds across these phases confirms they represent an independent generative mechanism.

Translation to Structural Functions

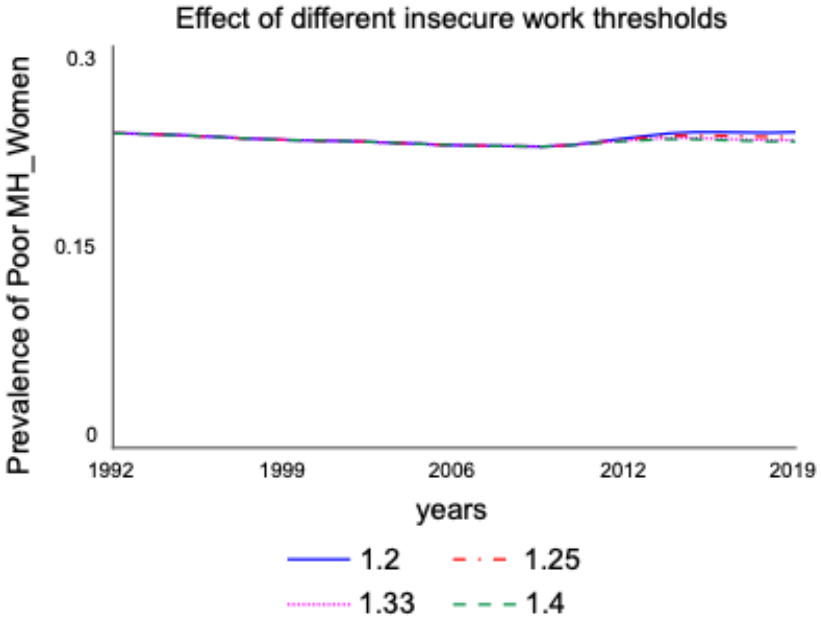
While table functions are superior for extreme condition robustness, such functions lack the continuous mathematical gradient required for automated optimisation algorithms. Consequently, the recovery multipliers were temporarily translated into a continuous algebraic form to provide a smooth numerical surface for Phase 1-4 of the calibration. Once the optimal thresholds were identified, they were re-implemented as the midpoints of the final structural table functions - utilising a linear decay for insecure work to ensure long-term model stability and theoretical alignment.

Employment Multiplier = MAX(1, (1+((effect_of_employment_on_recovery-1)*((insecure_work_threshold-change_in_insecure_work)/(insecure_work_threshold-1)))))

Prevalence Multiplier =1 - (0.5 x (Ratio of change in prevalence – 1)/(threshold – 1))

Insecure work threshold

Subsequent sensitivity analysis on the fully integrated model revealed that the overall trajectory of women's poor mental health is largely insensitive to the precise value of the female insecure work threshold within the explored bounds (1.20 to 1.4). Stage 3 of the constrained calibration pressed against its lower search bound of 1.25, indicating the unconstrained optimum lies at or below that bound; a value of 1.20 was therefore adopted, at the lower end of the insensitive range.



Prevalence threshold

While early isolated module calibration suggested a highly sensitive prevalence threshold for women, loop dominance analysis of the integrated model revealed that the endogenous prevalence-recovery reinforcing loop remained largely dormant during the historical calibration period in the fully integrated model. Because this loop's macroscopic impact was minimal when fully coupled with the labour market and gender dynamics modules, maintaining a distinct, highly sensitive threshold for women risked structural over-parameterisation without yielding additional explanatory power. In the original formulation the prevalence threshold was set at 2.0 for both genders (a sigmoid form in which prevalence must double to halve the baseline recovery rate). This was the more conservative assumption, but loop-dominance analysis of the integrated model showed it left the prevalence–recovery loop dormant throughout 1992–2019, and it implied a theoretically implausible behaviour: because recovery is stock-proportional, an undiminished rate meant recovery effectively strengthened as prevalence rose. The function was therefore respecified — linear for women, concave for men, with a halving point of 1.5 — so that recovery capacity degrades, rather than strengthens, as population prevalence rises. This 1.5 formulation is used throughout Chapters 6 and 7; the 2.0 sigmoid is retained here only as the superseded original.

Overview of optimisation settings

All runs utilized the Powell (Conjugate Direction) Method with 10 additional random restarts to ensure the identification of the global minimum payoff. The payoff was defined as the Sum of Absolute Errors between the model's prevalence output and historical data for men and women (1992–2019).

Stage 1

Parameter:	Mental health.insecure work threshold	Mental health.prev threshold
min_value	1.1	1.1
max_value	5	5
scaling	1	1

Payoff	Mental_health.Prevalence_of_Poor_Mental_Health[Men]	Mental_health.Prevalence_of_Poor_Mental_Health[Women]
34.8088	3.53853	31.2703
	Mental health.insecure work threshold	Mental health.prev threshold
Lower	1.31757049268	*
Value	1.5211123948	1.75902443792
Upper	1.72004867481	*

Stage 2

Parameter:	Mental health.insecure work threshold[Men]	Mental health.insecure work threshold[Women]	Mental health.prev threshold[Men]	Mental health.prev threshold[Women]
min_value	1.1	1.1	1.1	1.1
max_value	5	5	5	5
scaling	1	1	1	1

Payoff	Mental_health.Prevalence_of_Poor_Mental_Health[Men]	Mental_health.Prevalence_of_Poor_Mental_Health[Women]
12.3231	3.37145	8.95166

	Mental health.insecure work threshold[Men]	Mental health.insecure work threshold[Women]	Mental health.prev threshold[Men]	Mental health.prev threshold[Women]
Lower	1.86621293252	1.15749506819	4.52072082286	1.2860910036
Value	1.88678006387	1.15948851739	4.9635242323	1.28730545556
Upper	1.91366525994	1.1623169199	*	1.28903949091

Stage 3

Parameter:	Mental health.insecure work threshold [Men]	Mental health.insecure work threshold [Women]	Mental health.prev threshold [Men]	Mental health.prev threshold [Women]
min_value	1.25	1.25	2	1.25
max_value	2	1.75	5	1.75
scaling	1	1	1	1

Payoff	Mental_health.Prevalence_of_Poor_Mental_Health[Men]	Mental_health.Prevalence_of_Poor_Mental_Health[Women]
14.4902	6.78708	7.70313

	Mental health.insecure work threshold [Men]	Mental health.insecure work threshold [Women]	Mental health.prev threshold [Men]	Mental health.prev threshold [Women]
Lower	1.40023935288	*	4.54751886129	1.30277181817
Value	1.40589376474	1.25	4.99339766379	1.30423145198
Upper	1.41371875551	1.25572791702	*	1.30621077476

Stage 4

Parameter:	Mental health.prev threshold[Men]	Mental health.prev threshold[Women]
min_value	2	1.25
max_value	5	1.75
scaling	1	1
Payoff	Mental_health.Prevalence_of_Poor_Mental_Health[Men]	Mental_health.Prevalence_of_Poor_Mental_Health[Women]
16.1781	10.3951	5.78304
	Mental health.prev threshold[Men]	Mental health.prev threshold[Women]
Lower	4.5464155179	1.31048226474
Value	4.99369805447	1.31193357724
Upper	*	1.31549451205

H5 Constrained calibration for endogenous recovery rate

This appendix documents the calibration of the endogenous recovery mechanism — the profile-grid search, the structural constraints applied, and the rejection of a post-2010 structural break — supporting Section 6.5.4.

Unconstrained optimisation of all three parameters simultaneously produced equifinal solutions — mathematically optimal but theoretically implausible parameter combinations in which the men's shape parameter was driven to its upper search bound. To address this, a theory-constrained profile calibration was conducted in three stages.

The endogenous recovery multiplier was formulated as an equation using constant parameters that can be used during calibration. The equation defines a bounded, non-linear degradation of the recovery rate as community distress rises, and was given by:

MAX(0.5, (1 - (0.5 / (capacity multiplier^shape of curve[Gender])) * (MAX(0, (change_in_prev-1))^shape of curve[Gender])))

The function imposes a strict structural floor via the initial MAX(0.5, ...) condition, ensuring the recovery multiplier never falls below 50% of its baseline value, which represents the system's minimum intrinsic recovery capacity.

The Capacity Multiplier mathematically defines the "halving point" of the system: it specifies the exact relative increase in prevalence (change_in_prev - 1) required to trigger the maximum 50% penalty to the recovery rate.

The shape parameter determines the trajectory of systemic failure, capturing whether the erosion of recovery capacity occurs linearly, accelerates (convex curve), or decelerates (concave curve) as the population distress approaches the defined halving point threshold.

Stage 1: Profile grid

The capacity multiplier was fixed at nine values from 0.4 to 1.8. At each value, Powell optimisation in Stella Architect was used to find the best-fitting shape parameters for men and women independently (between 0.4 and 1.6).

Implied Halving Point if linear shape	Capacity multiplier	Men's shape	Women's shape (95% CI)	Payoff
1.4	0.4	*	1.12 (0.81, *)	10.0
1.6	0.6	*	1.05 (*, *)	10.1
1.8	0.8	*	1.03 (*, *)	10.1
2	1.0	*	1.05 (*, *)	10.1
2.2	1.2	*	1.01 (*, *)	10.1
2.4	1.4	*	1.01 (*, *)	10.1
2.6	1.6	*	1.01 (*, *)	10.1
2.8	1.8	*	1.02 (*, *)	10.1

Asterisks (*) denote values at the search bound, indicating the parameter could not be identified. Men's shape hit the bound at every capacity value tested.

Women's shape stayed very similar to linear but also hit the bounds at every value tested, suggesting that the data cannot distinguish the shape. Thus a linear shape for women was selected, for parsimony.
Four constraints were applied to eliminate configurations that, while mathematically viable, were inconsistent with the model's structural requirements:
Activation: The mechanism must reduce recovery by more than 1% at the historical operating point (input = 1.06). This eliminates configurations where the mechanism is present in the equations but has no practical effect. Thus, men's shape was fixed at 1.5
Stability: The recovery multiplier must remain above 0.40 at input = 1.15 and above 0.25 at input = 1.30. This prevents the function from producing runaway dynamics under moderate demand increases, consistent with the observed bounded rise in prevalence.
Structural stability: The reinforcing effect of the prevalence-recovery mechanism must be weaker than the balancing effect of the stock-dependent recovery outflow at the operating point. If the reinforcing effect were stronger, small perturbations would amplify rather than decay, and the model would be structurally unstable.
The relative strength of these two loops can be quantified as a *loop gain ratio*. When this is less than 1, a small increase in prevalence is self-correcting: the increased outflow (balancing) exceeds the reduced recovery rate (reinforcing), and the system returns toward equilibrium. When it is greater than 1, the system is structurally unstable: perturbations amplify rather than decay.

Stage 2: Calibrating halving point
The capacity multiplier was optimised between 0.4 to 1.8, fixing men's shape to 1.5 and women's shape to 1. Optimisation found 0.48 (*,*) with a payoff of 10.1. The bounds were hit, suggesting an inability to identify the true value. The implied halving point of 1.5 was retained grounded in theory. This significantly improved the Theil's statistics. Theil's inequality statistics decompose model error into three components: systematic bias (the model consistently over- or under-shoots), variance error (the model mismatches the size of fluctuations), and unsystematic error (random noise the model should not be expected to capture). A well-specified model should concentrate its error in the unsystematic component.

Table H-1: Theil's inequality statistics compared to original formulation, calculated for the isolated module.

Formulation	R2	MAE/Mean	RMSE	Bias (UM)	Variance (US)	Covariance (UC)
Original formulation						
Men	0.1762	0.0667	0.0140	0.0967	0.3967	0.5066
Women	0.2082	0.0734	0.0206	0.1811	0.4137	0.4052
Revised						
Men	0.1751	0.0689	0.0143	0.1310	0.2630	0.6059
Women	0.3992	0.0700	0.0200	0.3471	0.1834	0.4695

Revised - Simplified						
Men	0.1770	0.0700	0.0145	0.1472	0.2271	0.6257
Women	0.4054	0.0700	0.0201	0.3565	0.1721	0.4714
Revised – Simplified with Structural Break						
Men	0.1758	0.0702	0.0145	0.1499	0.2163	0.6338
Women	0.4245	0.0691	0.0201	0.3819	0.1378	0.4803

Stage 3: Simplification and integration
The relationship was simplified to a linear table function for women, and a concave function with a curvature of 1.25 (in the in-built Stella tool) for men, to closely approximate the equation values.

	R-square	MAE/Mean	RMSE	Bias	Variance	Covariance
Men - OLD integrated	0.153	0.062	0.014	0.002	0.750	0.248
Women - OLD integrated	0.081	0.072	0.021	0.084	0.687	0.229
Men - Revised	0.159	0.061	0.014	0.003	0.732	0.265
Women - Revised	0.145	0.071	0.021	0.094	0.682	0.224

Stage 4: Exploration of a structural break
The Theil's statistics for the revised integrated model compared to the original formulation integrated model showed an improvement. However, the rise in prevalence post-2016 was still not fully captured, and the R-square remained low. The idea of a regime shift was explored where there is a structural break in women's shape post 2010 and 2012. Given the almost linear estimation and small gains in Theil's statistics in the isolated model, this structural break was not implemented.

Structural break implementation year	Women's Shape pre-break (95% CI)	Women's Shape post-break(95% CI)	Optimizer Payoff
2010	Fixed at 1	0.94 (*, *)	10.1
2010	*	0.94 (*, *)	10.0
2010	Fixed at 1.5	0.94 (*, *)	10.0
2010	Simplified table function for both men and women	0.96 (*, *)	10.1
2012	Fixed at 1	0.94 (0.59, *)	10.1
2012	*	0.93 (0.58, *)	10.0
2012	Fixed at 1.5	0.94 (0.58, *)	10.0
2012	Simplified table function for both men and women	0.94 (0.59, *)	10.1

In all cases, men's shape was fixed at 1.5 and halving point at 1.5. Loop Dominance Analysis confirms that the new formulation activates the previously dormant prevalence-recovery reinforcing loop, reaching 3% dominance for women and 1% for men in 2013.

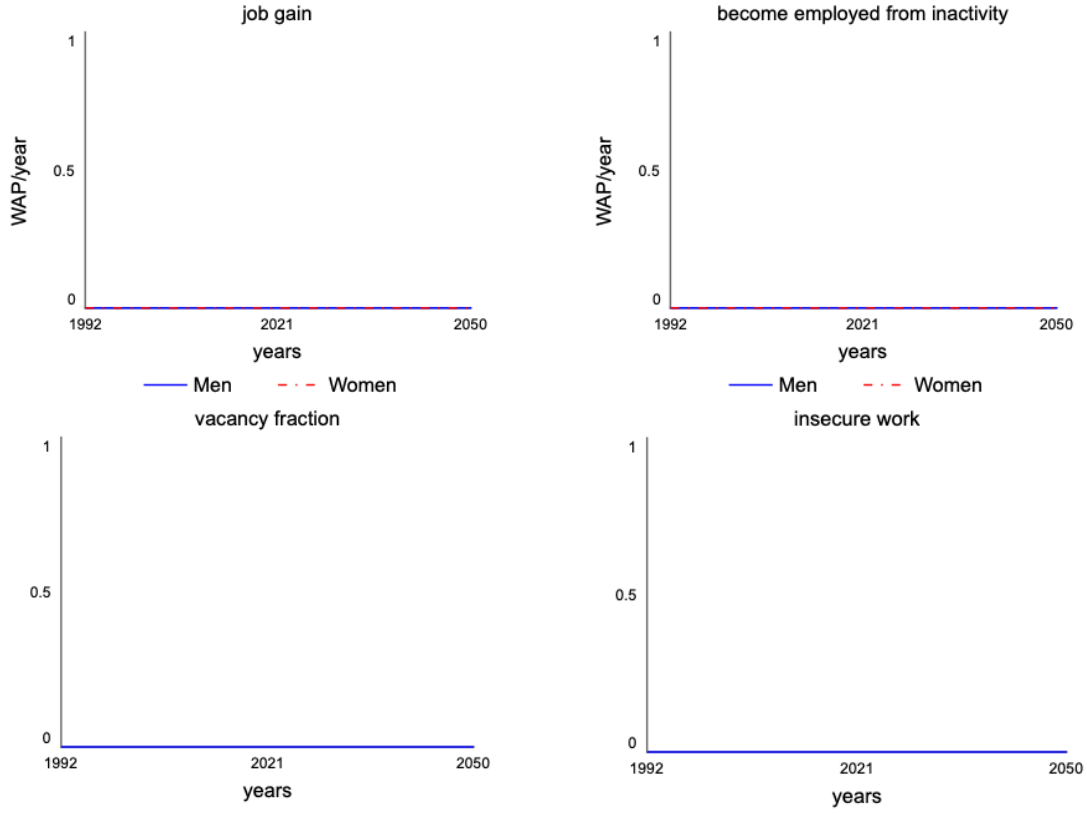
Appendix I Structural verification tests for the labour market module

This appendix reports the full extreme-conditions, boundary-adequacy, behavioural-integrity and sensitivity tests for the labour market module, summarised in Section 6.3.3. Each test states its condition, expected outcome, and the structural insight gained.

I1 Extreme conditions

This module was subjected to extreme condition testing prior to integration into the full system. These tests force the model inputs to theoretical extremes (e.g., zero population, zero demand) to ensure the equations remain mathematically robust and logically consistent without generating 'phantom' values (such as negative stocks or flows occurring without causality). The aim was to confirm that the model behaves logically under extreme, but plausible, conditions. If a model's structure is logical, it should behave in a qualitatively predictable way. A failure to do so would point to a significant flaw in an equation's formulation or a missing structural link.

Baseline labour demand was set to zero across time. The *job gain* flow fell to zero, but the *become employed* did not change as there was no input from the labour demand on the flow initially. An effect multiplier from *labour market tightness* to *become employed* was added to correct this. In the repeated test (Figure I-1), the *Unemployed Proportion* and *Inactive Proportion* stocks increased and the *Employed Proportion* approached zero. It did not reach zero because the current *Employed Proportion* is not directly linked to the *labour demand*. Over time, as the inflow to employed is turned off, and the outflows continue, this stock approaches zero. The difference in the unemployed between men and women, reflects the higher level of inactive among women.



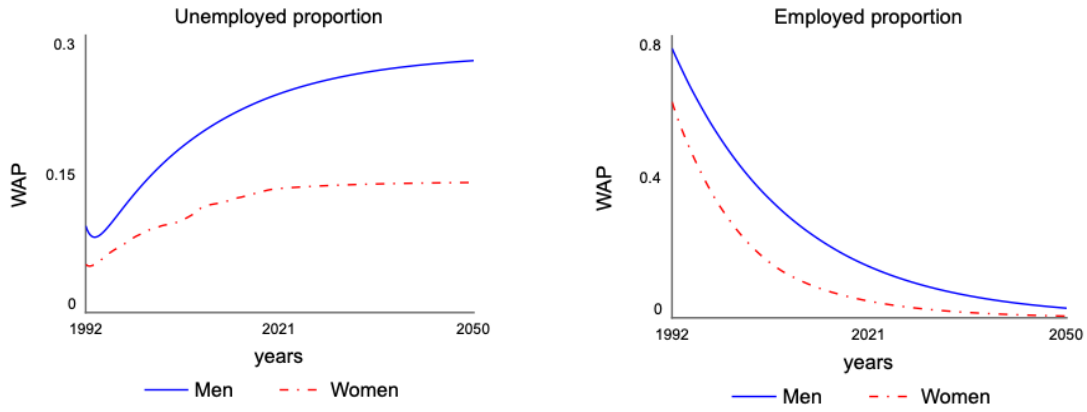


Figure I-1: Extreme conditions test - zero labour demand

Similarly, when **baseline labour demand was set to a very high value** (twice the *working-age population* (WAP)), the *Unemployed Proportion* reduced to a low value and labour market tightness increased to a high value as expected (Figure I-2).

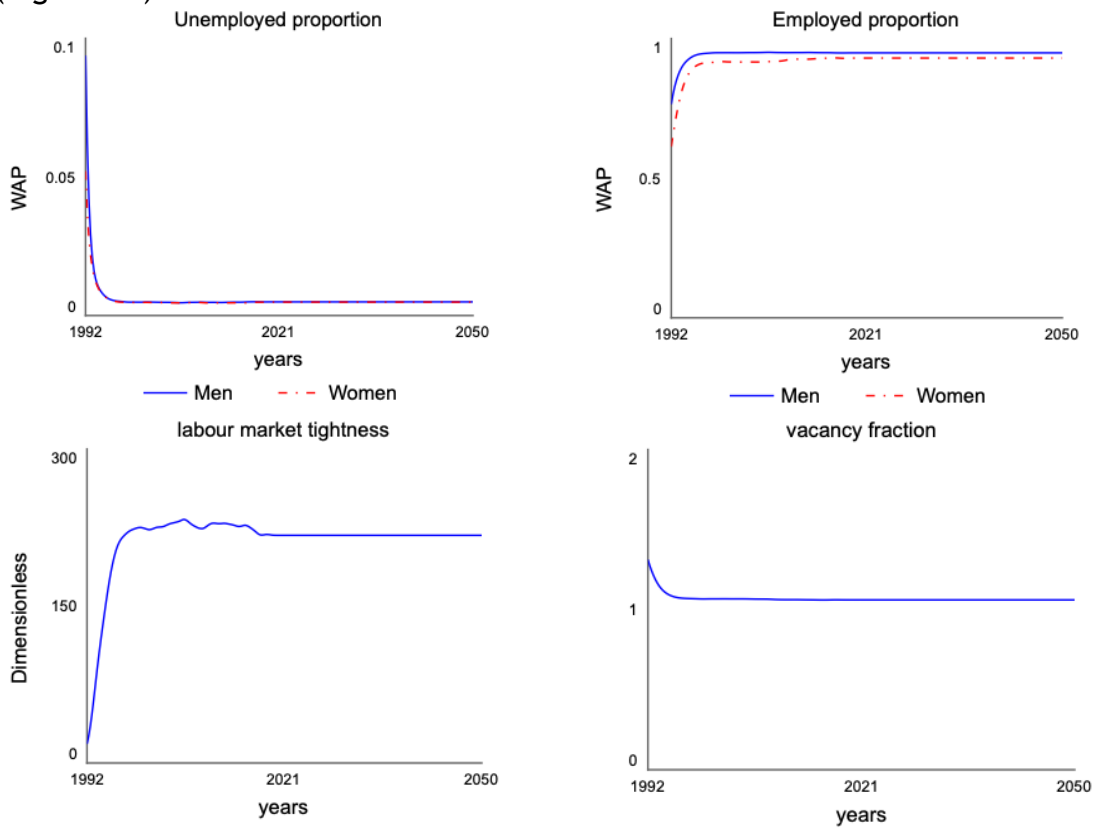


Figure I-2: Extreme conditions test - high labour demand

The model was also subjected to a **full employment extreme conditions test**, where initial unemployment and inactivity were set to zero for both genders. As the simulation progressed, the employed proportion for both genders immediately began to decay, proving that the model correctly identifies the outflow pressures (job loss and movement to inactivity) as being independent of the initial stocks (Figure I-3). The transient behaviour observed in men—a slight dip and subsequent rise in employment—reflects the system's internal re-balancing through the *Labour Market Tightness (LMT)* signal. Initially, the absence of any unemployed causes LMT to collapse toward zero, which in turn triggers an intensified job loss rate until a "frictional" level of unemployment is

endogenously generated to satisfy the system's structural requirements. The finding that inactivity stabilises at a higher level for women, while unemployment is higher for men, confirms that the model's gender-specific multipliers (such as gender norms and exit rates) remain dominant even under extreme mathematical shocks. Crucially, the LMT's rapid fluctuation from zero to a peak of 0.9 before settling validates the smooth asymptotic frictional floor formulated in the vacancy equations, ensuring the model can handle deep economic volatility without generating negative values or numerical instability.

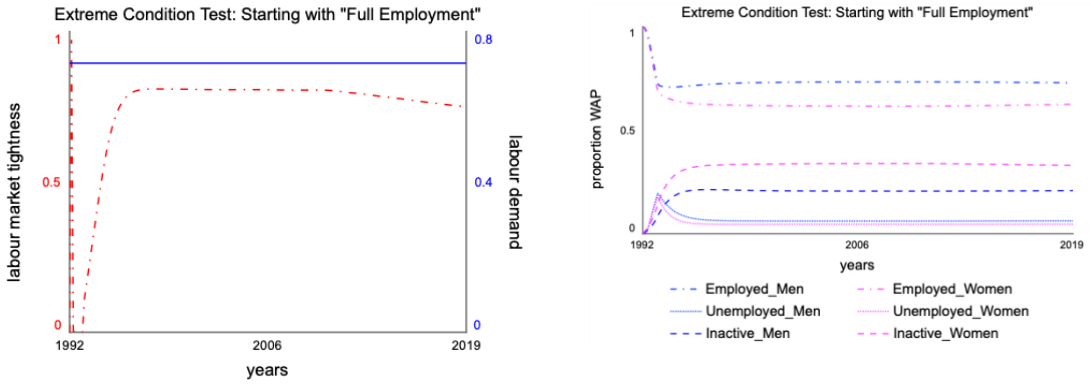


Figure I-3: Extreme conditions test - full employment

I2 Robustness to boundary conditions

Sensitivity to exogenous inputs was tested by varying the exogenous inputs by +-20%. As seen in Figure I-4, the **magnitude of insecure work** had little effect on the overall trajectory and magnitude of prevalence of poor mental health. This is because if the estimate of insecure work is historically changed, the baseline onset and recovery rates are re-calculated to ensure the model matches historical data and initialises in equilibrium.

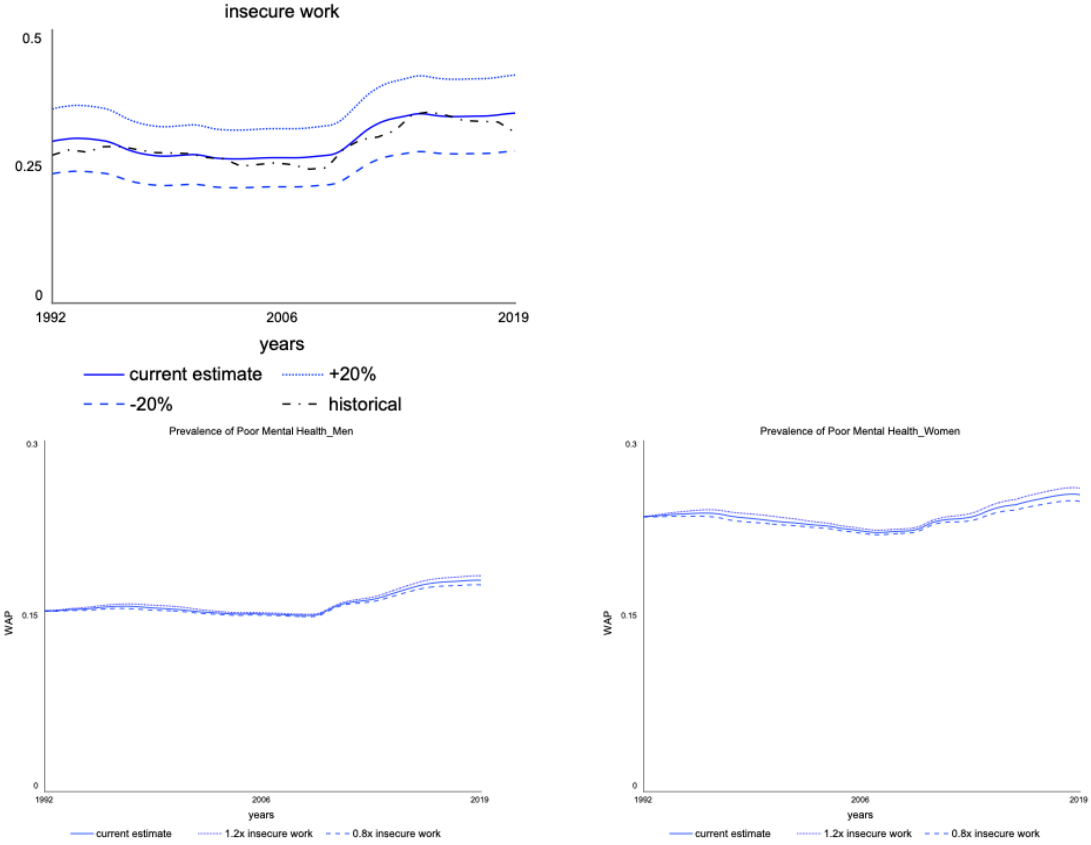


Figure I-4: Boundary adequacy test - insecure work

Similarly, **changes in the estimate of demand** had little impact on the trajectory although a large difference to magnitude (Figure I-5). However, the behaviour was counter intuitive – the prevalence of poor mental health was higher when demand was higher; this was because when demand increased, employed increased and thus the exposure to gender pay gap among women increased, increasing the effect of gender inequality on women's onset. Higher prevalence among women reduces recovery rate for both men and women as discussed in the mental health module later.
The model responds smoothly to changes in exogenous input values, confirming that the internal feedback structure is robust enough to handle uncertainty in the boundary data.

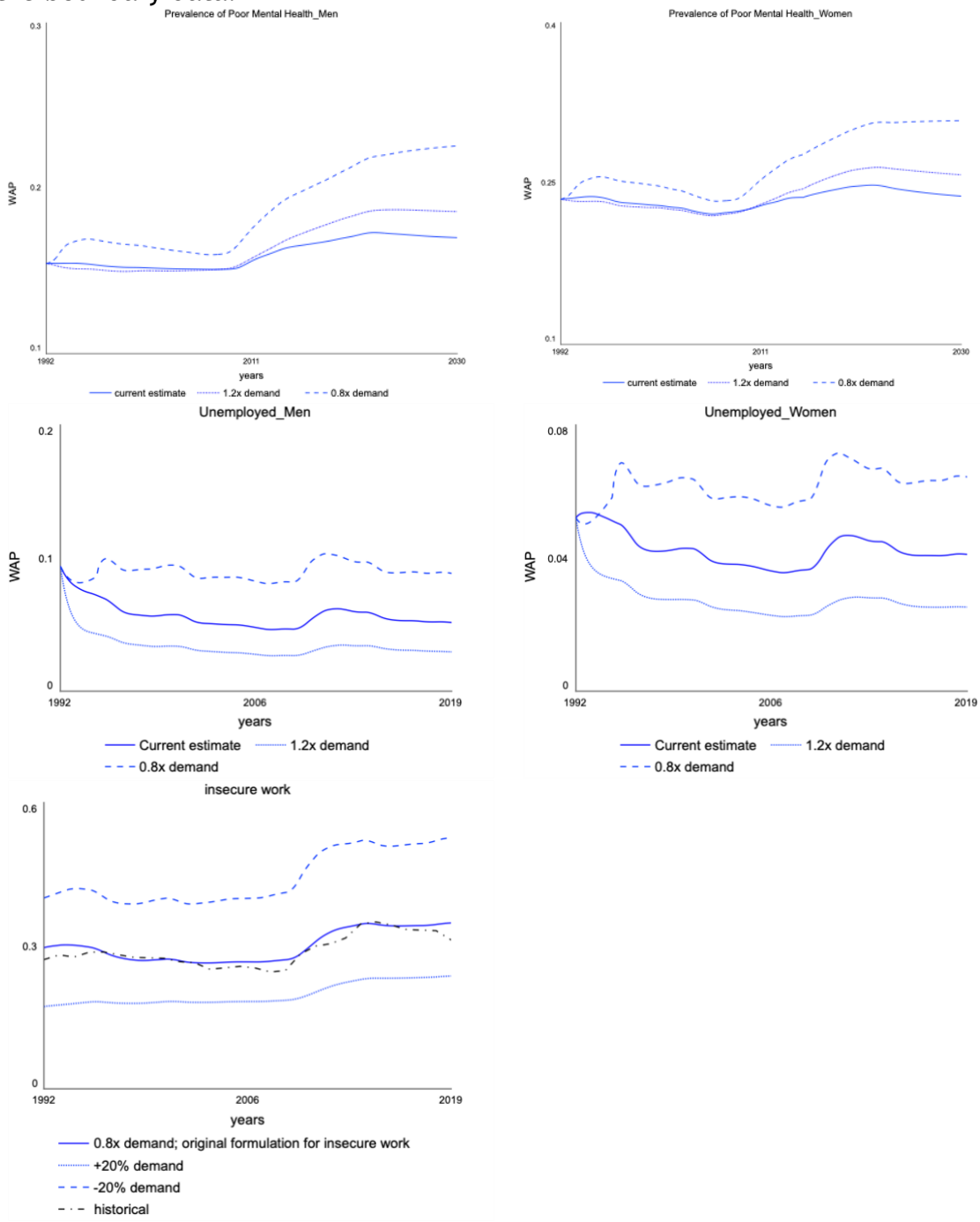


Figure I-5: Boundary adequacy test - labour demand

I3 Behavioural tests of structural integrity

To evaluate the structural stability and internal momentum of the Labour Market module, a **temporary increase in labour demand** was applied between 1999 and 2000, creating a short-lived economic "pulse" (Figure I-6). As expected, the model responded with an immediate decline in the Unemployed Proportion for both men and women. However, once the demand stimulus was removed in the year 2000, the system exhibited a compensatory "overshoot," where unemployment rose slightly above the original baseline between 2000 and 2001 before decaying back to equilibrium by 2005. This non-linear recovery shows that the model captures the "transient response" of the labour market, demonstrating that the system possesses a structural "memory" governed by the internal time-constants of the transition flows rather than simply tracking exogenous inputs.

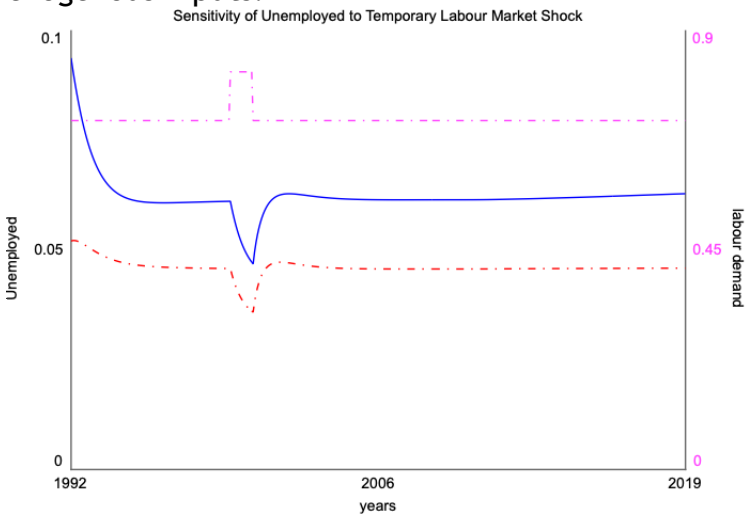


Figure I-6: Stability test - temporary increase in labour demand

The **structural necessity of Balancing Loop B1 (Demand Saturation)** was verified through a three-scenario "knock-out" analysis, comparing a base case (A) against a step increase in demand with (B) and without (C) the elasticity link between *Labour Market Tightness (LMT)* and labour flows (Figure I-7). In Scenario B, the active balancing loop regulated the system, allowing LMT to stabilise at a value of 1.8 by 2019 as the increased employment levels exerted a satiation effect on the market. Conversely, in Scenario C, where the LMT elasticity was set to zero, the system became effectively decoupled. In this "knock-out" state, the LMT rose to a higher, unregulated steady state of 2.0, while the unemployment stocks exhibited zero response to the demand shock, remaining flat at their 1995 levels. This failure of the unemployment stocks to react in the absence of the LMT link provides rigorous evidence that Loop B1 is the primary structural engine driving macroeconomic adjustment in the model.

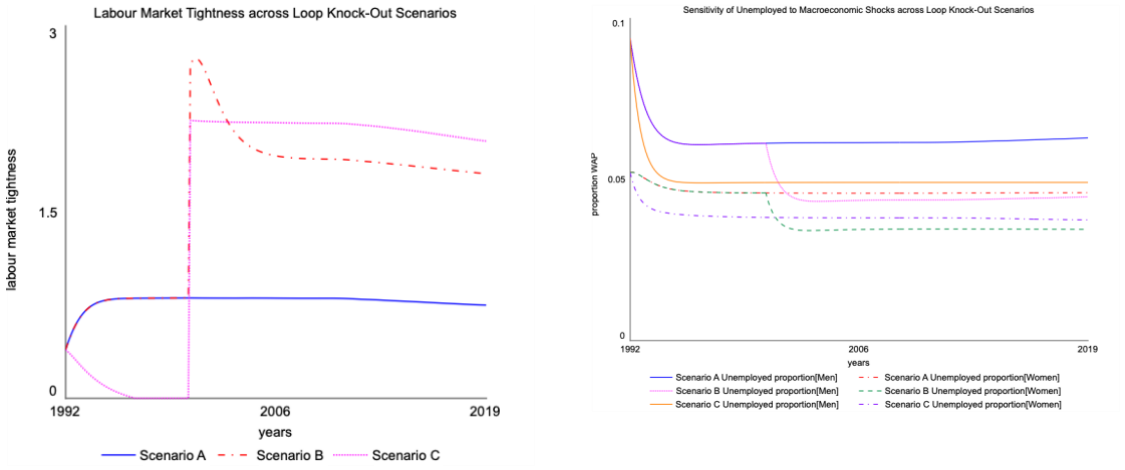


Figure I-7: B1 Loop knockout

I4 Sensitivity analysis

The purpose of these tests is to move from verifying the model's logic to understanding its behaviour and systematically identify which parameters, assumptions, and boundary conditions exert the most influence on the model's outcomes. This stage explores the model's behaviour in the face of uncertainty. To isolate the model's endogenous feedback structure, all key exogenous drivers were held constant at their 1992 initial values for the entire 1992-2045 simulation period. By running the sensitivity tests under these "constant" conditions, any observed divergence in behaviour between runs can be unambiguously attributed to the single internal parameter being varied, revealing the true leverage points of the model's core theoretical structure.

The key unknown parameters are the table functions determining the effect of gender norms and poor mental health on labour market flows. Sensitivity analysis was conducted to examine the impact of these (Figure I-8). All four test shapes (sigmoid, concave, convex, and linear) produce trajectories that are functionally identical in terms of direction and timing. They all follow the historical trajectory of women's inactivity with only minor differences in magnitude. Given the high structural robustness and preference for parsimony, this was replaced by a linear formulation.

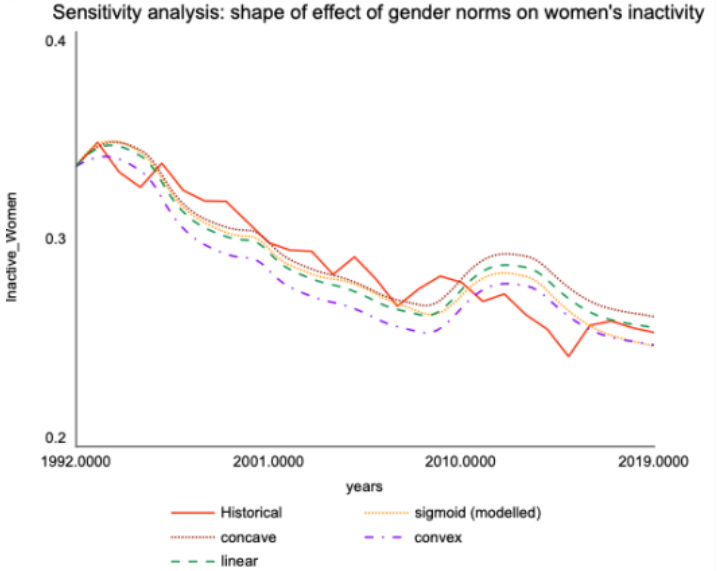


Figure I-8: Sensitivity analysis: relationship between norms and women's inactivity

Similarly, sensitivity analysis testing the shape of the effect of prevalence of poor mental health had little impact on the primary stock it impacts, employed proportion (Figure I-9). The shape of these relationships did not alter the fundamental behavioural modes of the system. Consequently, a linear formulation was adopted to prioritise structural parsimony.

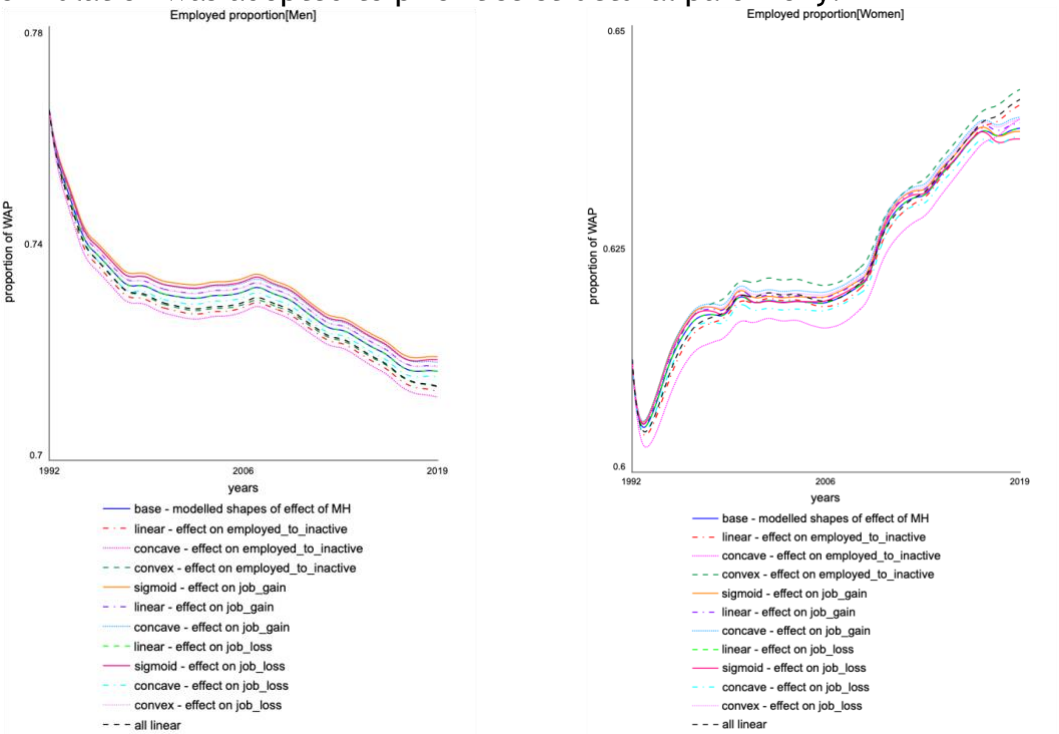


Figure I-9: Sensitivity analysis - shape of effect of poor mental health

The sensitivity to the specific table function for the effect of statutory pension age change was examined. There was no discernible difference and thus a linear formulation was selected.



Figure I-10: Sensitivity of effect of change in statutory pension age

I5 Sensitivity analysis for baseline exit rates

This appendix reports the multivariate (Latin Hypercube, n=100) sensitivity analysis over the explored exit-rate ranges, referenced in Section 6.3.2. It confirms that the rising-distress trajectory and the persistent gender gap are robust to exit-rate uncertainty.

To evaluate the impact of parameter uncertainty on the model's core findings, a multivariate sensitivity analysis was conducted using Latin Hypercube sampling (n=100) across the full explored range of baseline exit rates (Men: 0.55–0.69; Women: 0.55–1.10). The results, detailed in Appendix H2, demonstrate that while the magnitude of labour market stocks varies slightly, the longitudinal trajectory of rising psychological distress and the persistence of the gender mental health gap remain structurally robust. The lack of significant divergence across these ranges confirms that the model's outcomes are driven by the internal feedback architecture—specifically the interplay between gender norms and economic stressors—rather than being contingent upon precise parameter tuning of the transition flow.

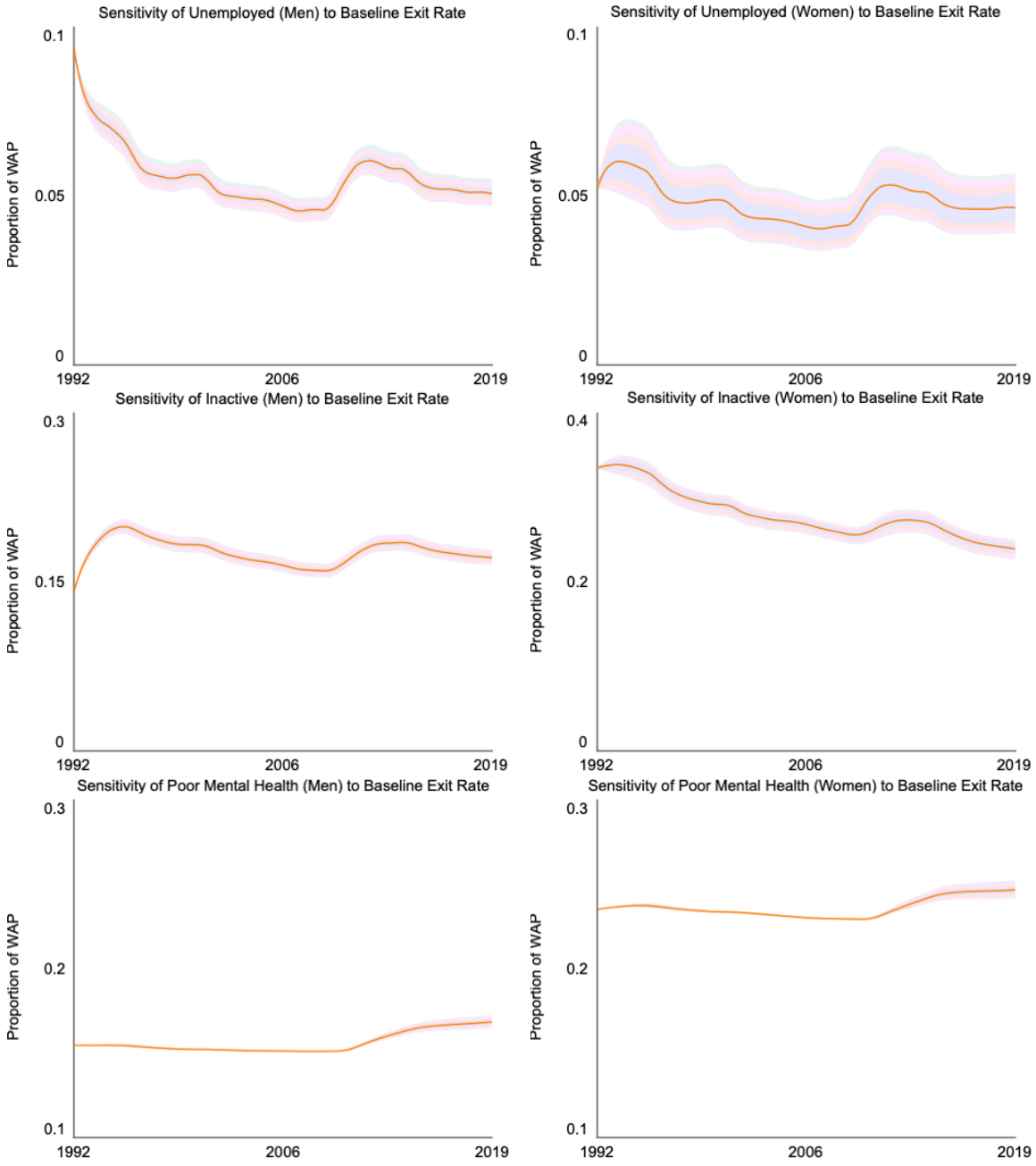


Figure I-11: Sensitivity analysis - baseline exit rate

I6 Behaviour reproduction for labour market

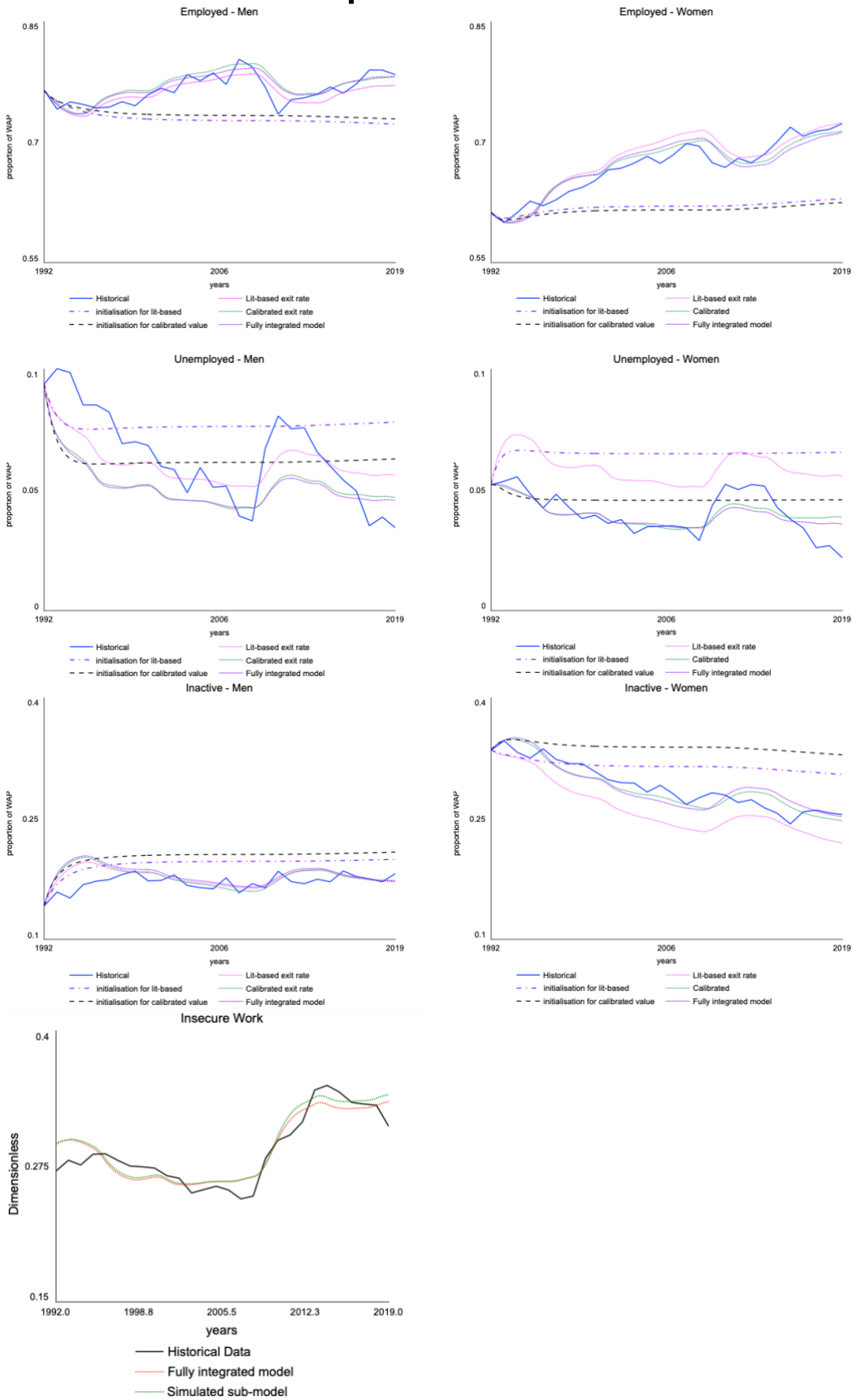


Figure I-12: Behaviour reproduction of labour market variables Compares the labour market module's fit to historical data when run in isolation versus fully integrated. It shows that integration changes magnitude but not the structural trajectory.

Appendix J Structural verification of the gender dynamics module

This appendix reports the extreme-conditions and boundary tests for the gender dynamics module, summarised in Section 6.4.8, including the unpaid-work-burden equation error identified and corrected during testing.

J1 Extreme conditions tests

During extreme conditions testing, an important structural error was identified and corrected in the formulation of the unpaid work burden. The initial multiplicative formulation produced illogical results under **egalitarian conditions (norms = 0)**, causing the unpaid work burden to fall below its theoretical minimum of 0.5. This revealed that the model was incorrectly applying the time availability multiplier to the baseline of perfect equality, rather than just to the inequality component driven by norms. To correct this, the equation was reformulated to first isolate the inequality created by norms (norm_driven_inequality = effect_of_gender_norms - 0.5) and then apply the time availability multiplier only to this component. The final, structurally correct equation (unpaid_work_burden = 0.5 + (norm_driven_inequality * effect_of_time_availability)) represents the causal theory that time availability mitigates the degree of inequality, but does not alter the definition of equality itself, ensuring the variable behaves logically across all possible conditions.

The logic of the childcare availability index as a mitigating factor was also validated. A controlled test was conducted where the *Prevalence of Norms on Gender Roles* was held constant at 0.5. When the **childcare availability index was set to zero**, the gender parity in working hours was determined solely by the baseline effect of norms. When the index was set to its maximum, the *fraction of gap closed by childcare availability* multiplier became fully active, successfully closing a significant portion of the "parity gap" and resulting in a much higher level of gender parity in working hours. The expected downstream effects of a lower gender pay gap and unpaid work burden were also observed. This test confirmed the correct formulation of the childcare mechanism as a dynamic modulator that counteracts, but does not eliminate, the influence of prevailing gender norms.

To verify the stickiness of norms, an isolation test was performed. The economic signal composite was frozen at its 1992 baseline value (0.24) throughout the simulation. Norms did not remain static but exhibited a slow, endogenous decline. This confirms that the stock formulation decays independent of current economic conditions. When the progressive **economic signal was explicitly switched off (effect =0)**, the stock deteriorated significantly faster than historical data. When economic signals were held constant at 1, the norms stayed constant.

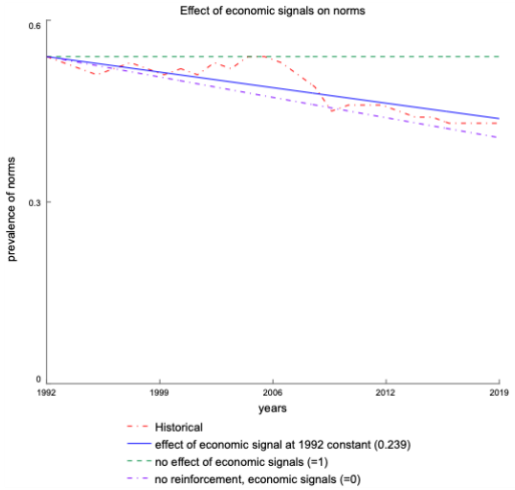


Figure J-1: Extreme conditions test - no effect of economic signals

To verify the structural inertia of the norm stock, a parity test was conducted by forcing the **Economic Signal Composite to zero** in the year 2000. Despite this absolute removal of economic reinforcement, the prevalence of traditional norms did not collapse instantly, reaching only 0.244 by 2050; this extreme-condition test was extended ten years beyond the projection horizon to give the slow-moving norm stock sufficient time to manifest its inertial behaviour. This slow, endogenous decline confirms that the module correctly distinguishes between transient economic signals and the deep-seated structural "memory" of the norm stock, which is governed by the physical delay of cohort replacement.

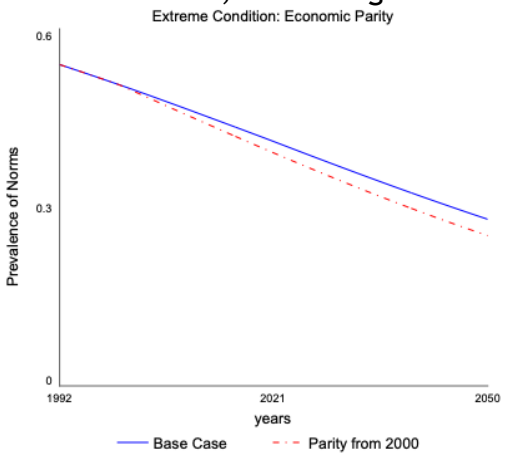


Figure J-2: Extreme conditions test - economic parity

J2 Robustness to boundary conditions

Sensitivity to exogenous inputs was tested by varying the baseline gender pay gap by +- 20%. As seen in Figure J-3, there is almost no difference in the prevalence of norms.

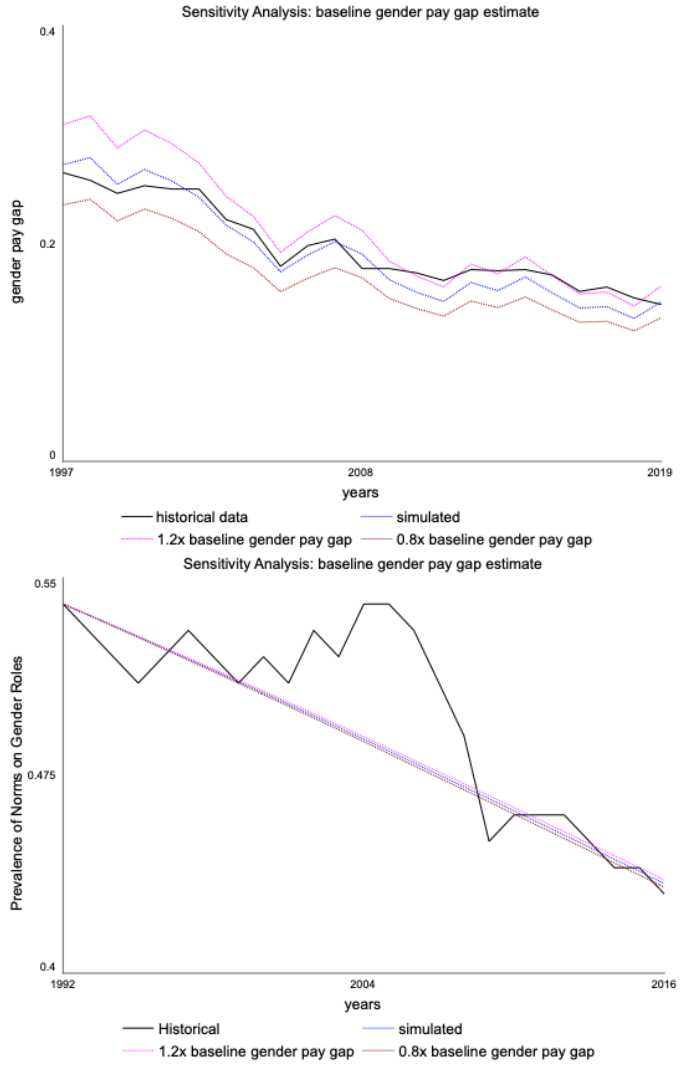


Figure J-3: Sensitivity analysis for baseline gender pay gap

This test examined the effect of disconnecting the feedback loop.
In the first set of cases, this was done by modifying the relationship between gender pay gap and gender parity in working hours. When the effect of gender parity in working hours on gender pay gap was removed, the gender pay gap was simulated to be significantly lower and gender parity in working hours was slightly higher. When the effect of gender parity in working hours was doubled, the gender pay gap increased significantly and gender parity in working hours slightly reduced.

In the second set of cases, this was done by modifying the relationship between gender pay gap and norm reinforcement. When the effect of gender pay gap on norm reinforcement was held constant at 1, gender pay gap was slightly higher, and gender parity in working hours was significantly lower. When the effect of gender pay gap on norm reinforcement was simulated to be zero, the gender pay gap was slightly lower and gender parity in working hours was higher.

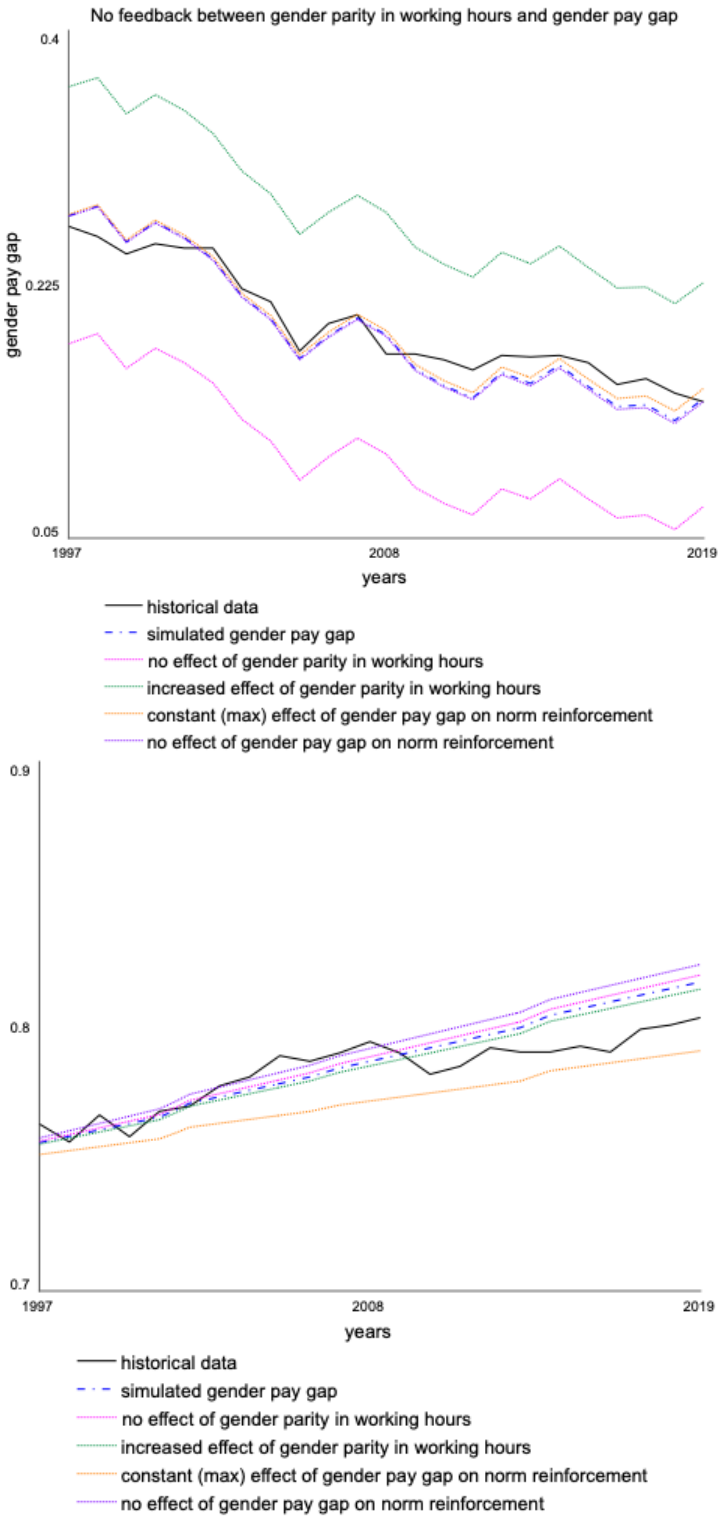


Figure J-4: Loop R3 knock out analysis

J3 Sensitivity analysis for gender norms module

This appendix reports the sensitivity analyses behind the gender dynamics formulation choices — the economic-signal composite weights, the resistance factor, and the effect of working-hours parity on unpaid work — summarised in Sections 6.4.3–6.4.5. It documents the simplifications that removed two table-function parameters.

Economic signal composite

Hours and GPG are assigned 0.4 weight, while inactivity is 0.2; sensitivity analysis was conducted on alternative weighting. Figure J-5 shows that the trajectory is robust to the weighting choice, only the slopes differ.

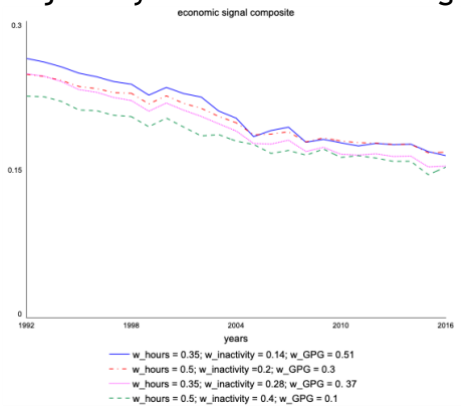


Figure J-5: Sensitivity analysis for economic signal composite

The composite signal is translated into a multiplier on norm reinforcement using a concave function anchored at (0,0) and (1,1); indicating that there is no reinforcement due to economic signals at full parity and full reinforcement at maximum inequality. Concavity captures hypersensitivity to even marginal levels of inequality: minor disparities serve as sufficient proof to validate traditional roles and the steepest decline in norm reinforcement occurs only when the system approaches near-perfect parity.

Sensitivity analysis (
Figure J-6) shows that the prevalence of norms is sensitive to the non-linearity of this multiplier: concave function shows a slower decline in norms compared to the linear, convex or sigmoid alternatives. The concave shape gives the best historical fit at a 40-year time constant, against 50 years for linear and convex, and a poor fit at any value for sigmoid (which is therefore rejected).

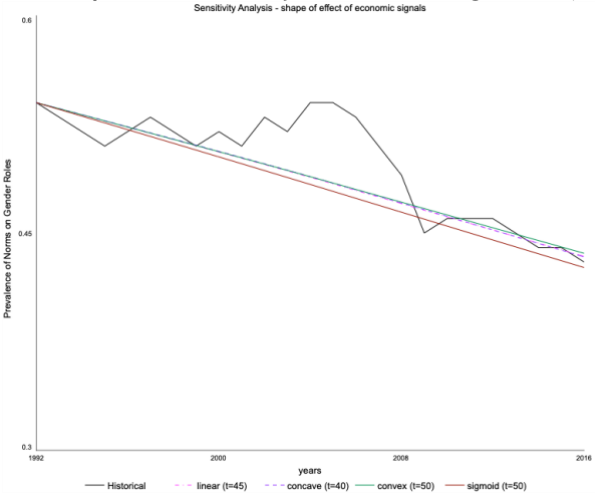


Figure J-6: Sensitivity analysis for effect of economic signals on norms

Resistance to economic signals

Analysis on the sensitivity to the formulation of the resistance factor (Figure J-7) found that there was no significant difference between a linear and sigmoid formulation, the concave formulation resulted in slower (with 35 years giving a better fit) and convex in faster deterioration of norms. This is possibly due to the simulation period not covering the extreme values of norms where the sigmoidal shape is most different to a linear shape. Given the linear formulation had a good fit and is the simplest formulation, it was selected and subject to further analysis. Different ranges of the effect were tested. The theorised range of (0.1, 0.8) was expanded to (0, 1), essentially assuming that resistance = prevalence of norms. This resulted in a time of 50 years instead of 40 years. 50 years is within the theoretical bounds for this parameter. However, this was preferred for parsimony, as one unknown parameter was reduced.

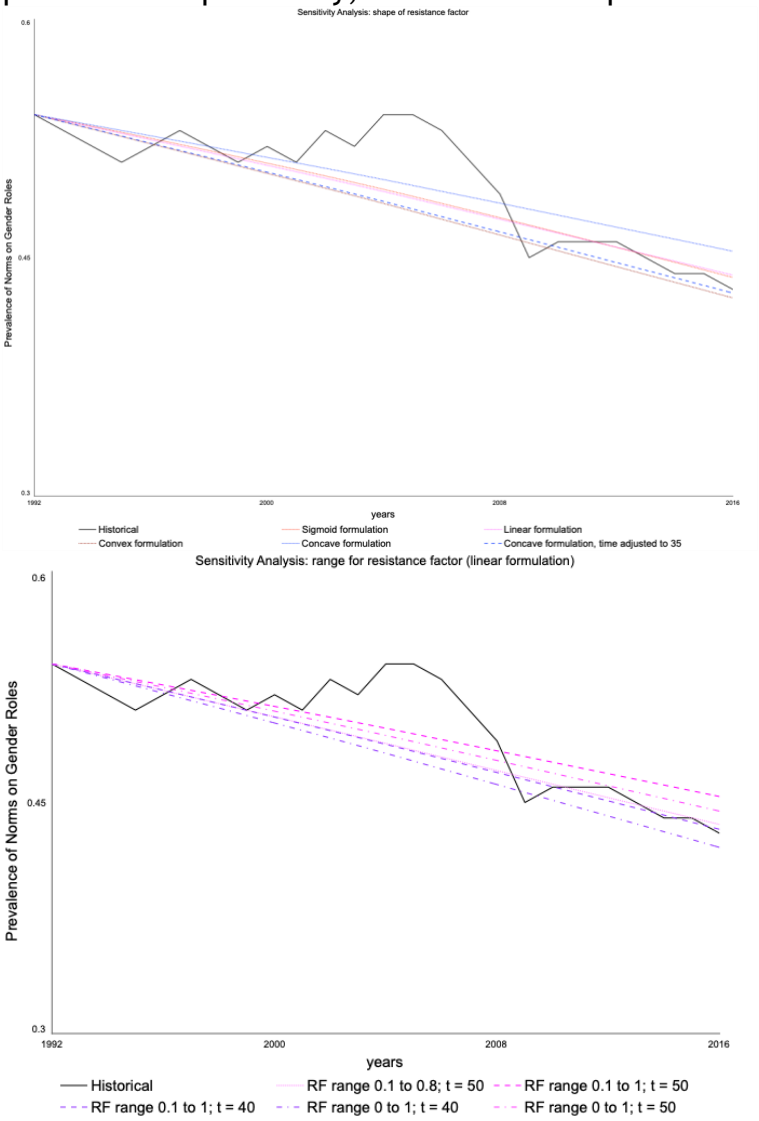


Figure J-7: Effect of resistance formulation on norm reinforcement

The final sensitivity analysis focused on the combination of the shape of the effect of economic signals and time for norms to change (Figure J-8). Time was bounded within the theory-based range of 35 to 50 years. Zooming in on these, it can be seen that each shape suggests a different rate at which norms change. The sigmoid curve is rejected because the best fit will be when time is greater than 50 years.

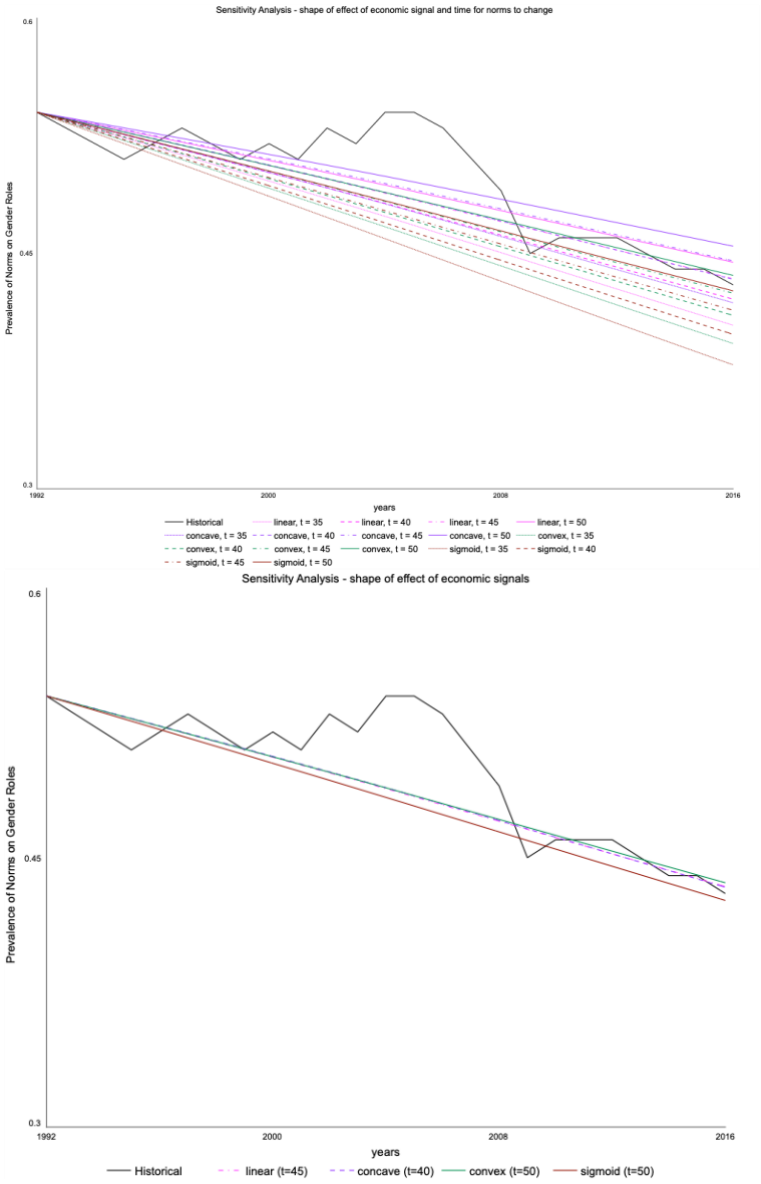


Figure J-8: Sensitivity analysis: effect of economic signal on norm reinforcement

Given the minor differences, alignment of 45 years with a typical career for most of the simulation duration, and the preference for parsimony, the linear formulation was selected; the effect of the economic signals equals the economic signal composite itself.

Effect of gender parity in working hours on unpaid work

Sensitivity analysis found that without the multiplier the unpaid work burden would be higher than observed. The magnitude could possibly be matched if the range for unpaid work burden (resulting from norms) was artificially reduced to 0.5 to 0.84 rather than 0.5 to 0.95. However, the decline in unpaid work burden is not as well captured compared to the presence of this effect. This effect was retained, but following the sensitivity analysis it was reformulated to a linear function i.e. the effect of gender parity in working hours on unpaid work was reformulated as (1.5 – gender parity in working hours).

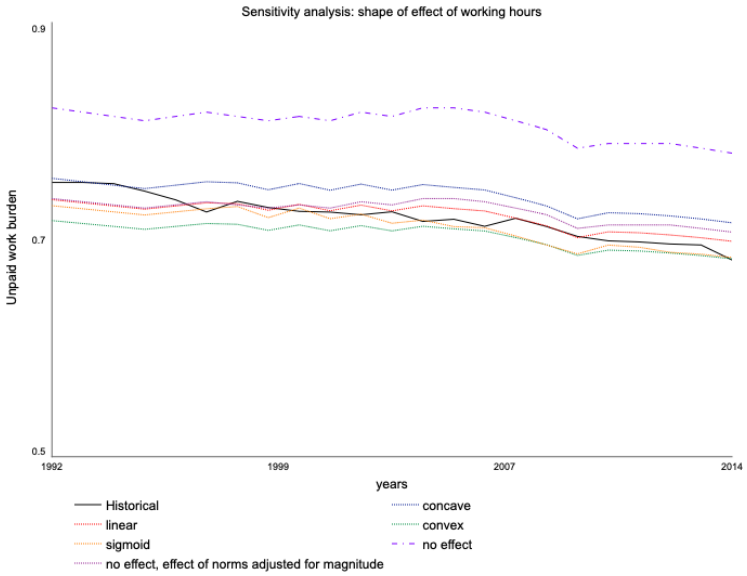


Figure J-9: Sensitivity analysis - relationship between gender parity in working hours and unpaid work burden

Appendix K Structural verification of the mental health module

This appendix reports the equilibrium-initialisation, extreme-conditions, initialisation-sensitivity and parameter-sensitivity tests for the mental health module, summarised in Section 6.5.6.

K1 Basic simulation results

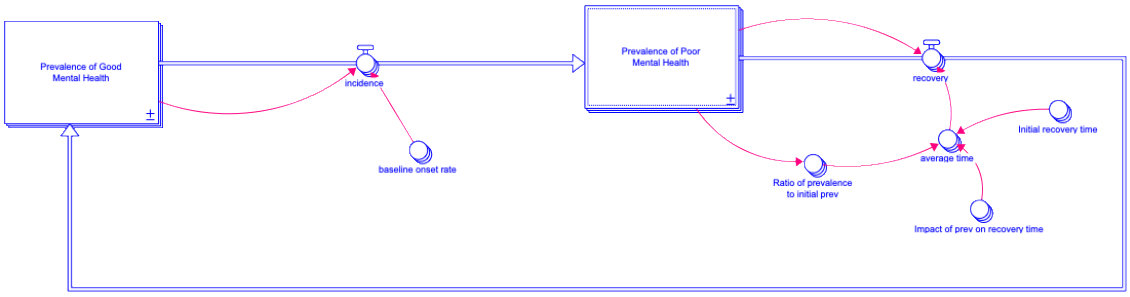


Figure K-1: Stock and flow model

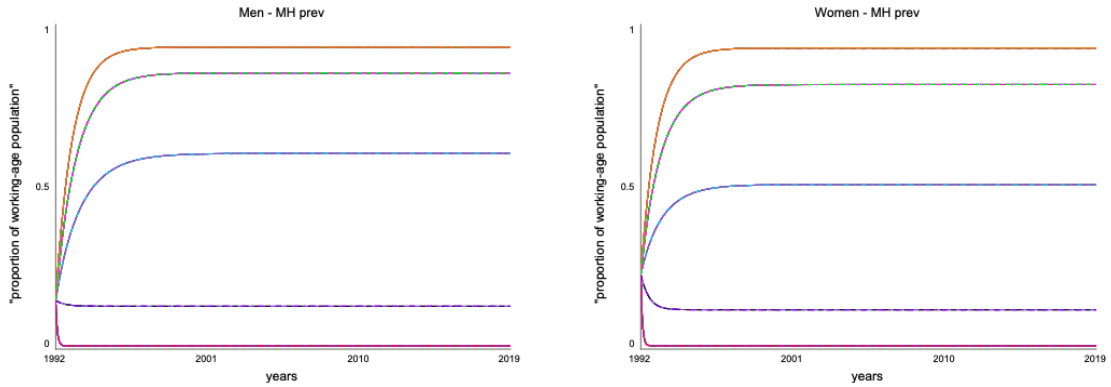


Figure K-2: Graphs produced by different parameters for the above model structure

K2 Equilibrium initialisation check

When *stressors* are held constant at 1992 levels, the stocks of *Mental Health* remain flat. This shows that the *implied baseline* is mathematically sound. It signals that any movement in the main results is due to dynamics.

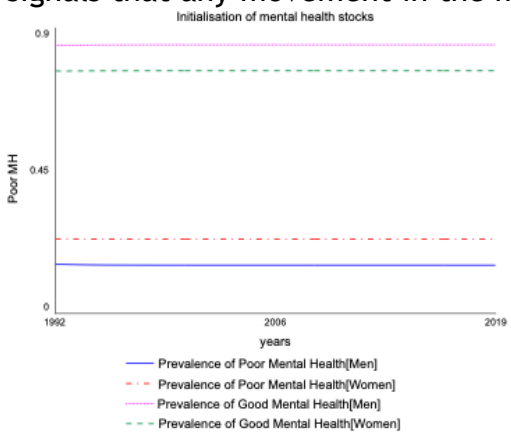


Figure K-3: Initialisation

A test was also conducted to test the effect of not initialising in equilibrium but using the raw baseline recovery rates (Figure K-4). Three scenarios were considered (a) no changes apart from baseline recovery rates set to be those derived from 1992 BHPS Scottish data, (b) repeating scenario a holding onset and recovery drivers at 1992 constants, and (c) setting the baseline recovery rate to the average of 1992-94. In all three conditions, the prevalence of poor mental health settles at a lower level. This is expected given the higher prevalence of poor mental health in 1992 compared to 1993, indicating a trend toward recovery. The data is noisy, making the raw recovery rate uncertain. However, in all cases (apart from 1992 constant values), there is a rise in prevalence at the same point as the base case, reflecting the changes in the onset and recovery drivers. This shows that the behaviour is robust to the specific parameters.

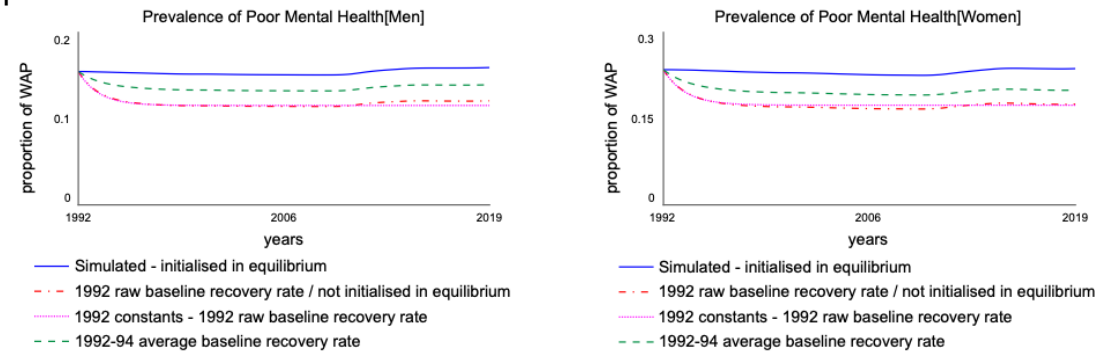


Figure K-4: Starting the module in equilibrium vs non-equilibrium

K3 Extreme conditions

A conservation of flow test was performed by **initialising both stocks to zero** while maintaining positive inflow rates. The objective was to verify that the model does not generate population through structural errors in the flow equations. As shown in Figure K-5, the *prevalence of mental health* states remains at exactly zero throughout the simulation. This result confirms that all inflows in the model are correctly formulated as "dependent" on existing stocks or valid source populations i.e. the 'draining' processes are strictly coupled to the available population.

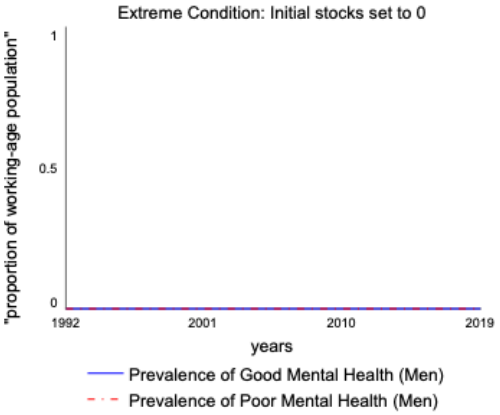


Figure K-5: Extreme conditions test - starting stocks zero

A further extreme conditions test involved setting the **baseline recovery rate to 0**. As expected, the prevalence of poor mental health stabilised at 1 (Figure K-6).

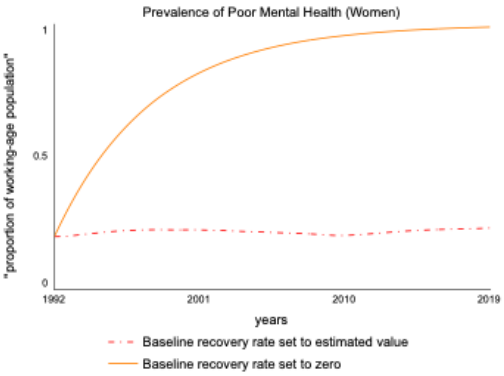


Figure K-6: Extreme conditions test – zero recovery rate

Next, the **Economic Stressors and Women's Inequality composites were forced to their theoretical maximum (1.0)** using a STEP function starting in the year 2000. Under these extreme conditions, the prevalence of poor mental health reached a saturated state of approximately 0.5 for women and 0.2 for men, but critically, neither stock exceeded the total population boundary of 1.0. This result confirms that the flow equations are correctly formulated as stock-dependent—the "draining" of the good mental health stock is strictly coupled to the available population. The observed gender gap under saturated stress further validates that the gendered divergence in the model is an emergent property of the internal multipliers rather than a result of arbitrary parameter tuning.

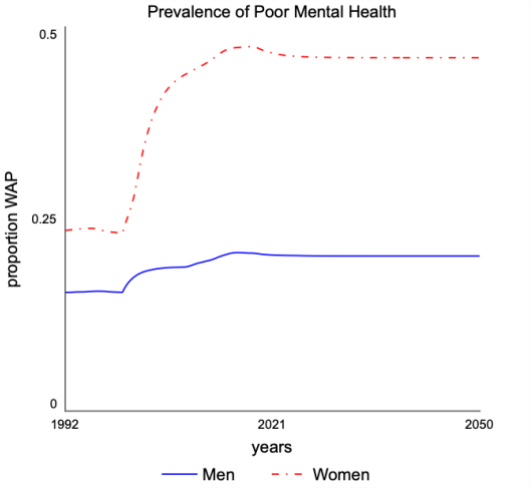


Figure K-7: Extreme conditions – maximum stress

This test was repeated under the revised formulation for endogenous recovery rate. Instead of a STEP function, the composites took their maximum value of 1 after 1992 among both men and women. The stock rose sharply and displayed asymptotic behaviour as expected. The stock did not exceed 1.

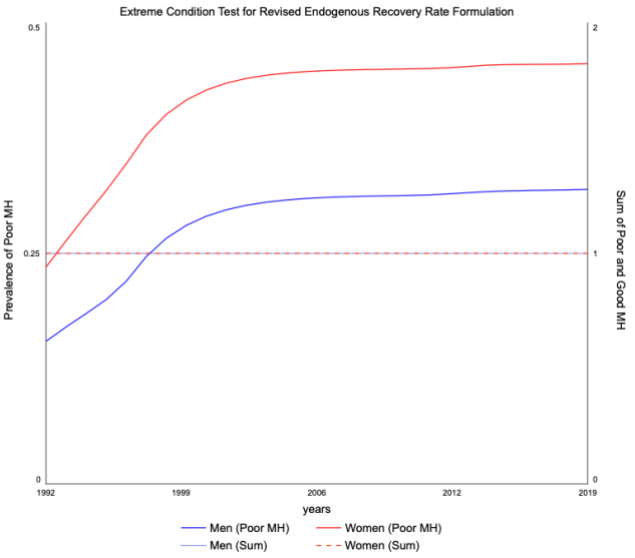


Figure K-8: Extreme conditions test - endogenous recovery

K4 Initialisation sensitivity and equilibrium recovery

To prove that the model's long-term results are driven by its internal feedback structure rather than its starting values, an initialisation sensitivity test was conducted Figure K-10. The **initial 1992 prevalence was artificially inflated** to 0.25 for men and 0.3 for women, significantly above the empirical baseline. Despite this initial perturbation, the prevalence for both genders decayed toward the internalised "resting state" of the system, reaching its lowest point in 2007 (0.15 for men; 0.23 for women). When repeated with all exogenous stressors held constant at 1992 values, the system sought a clear equilibrium at 0.152 and 0.234. This "return-to-base" behaviour provides rigorous evidence that the Implied Baseline calculations are mathematically sound. It demonstrates that the longitudinal rise in psychological distress observed is not an artefact of initial state momentum.

K5 Absolute vs relative risk formulation

While a relative deviation formulation was explored to enhance model sensitivity to historical volatility, it was ultimately rejected in favour of the absolute risk composite to maintain greater structural defensibility and initialisation stability. This acknowledges that the Scottish labour market in 1992 was already characterised by a high degree of economic stress. Comparative analysis of the resulting onset multipliers revealed that while the relative formulation produced higher amplitudes during acute shocks like the 2008 recession, both versions generated remarkably similar longitudinal trajectories. This consistency suggests that the model’s core behaviour is an emergent property of the fundamental feedback structure rather than a consequence of specific mathematical parameterisation or formulation. The absolute formulation is more aligned with the research aims given a multiplier of 1 suggests the absence of economic stress.

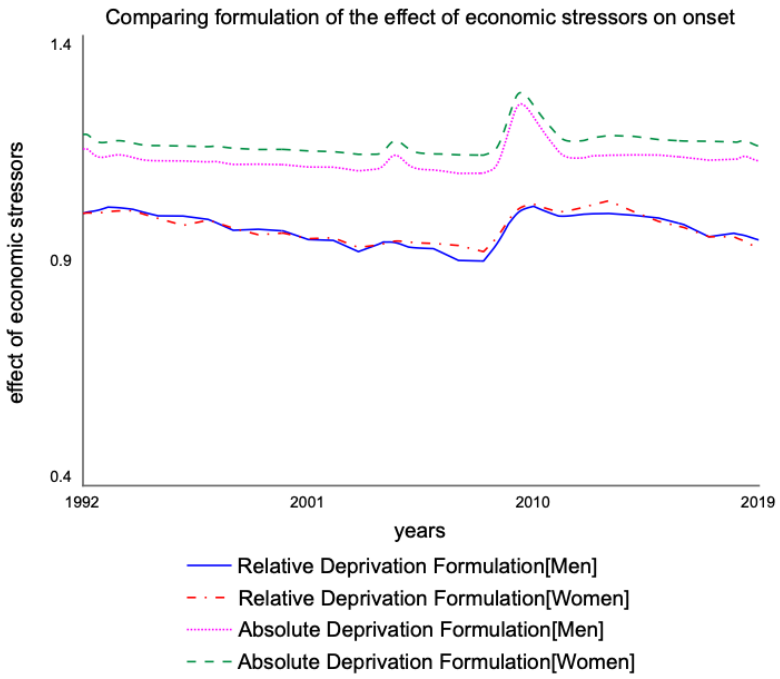


Figure K-9: Comparative analysis of different formulations for economic stressor

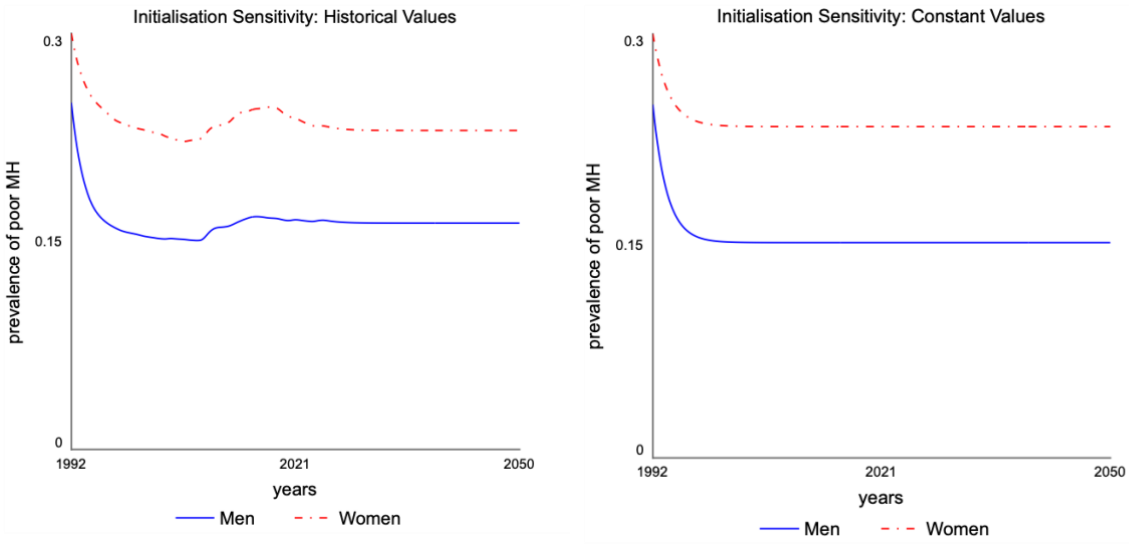


Figure K-10: Sensitivity to initial conditions

K6 Sensitivity analysis

Sensitivity analyses were conducted to examine the effect of the choice of table functions. There was some difference, and the simulated prevalence with selected parameters (solid black line) sits consistently above the other formulations. The selected shapes (concave for economic stress and sigmoid for gender inequality) were retained due to the discussed theoretical reasons. Although the sensitivity analysis shows a difference, this difference is not very large, indicating the robustness of the model.

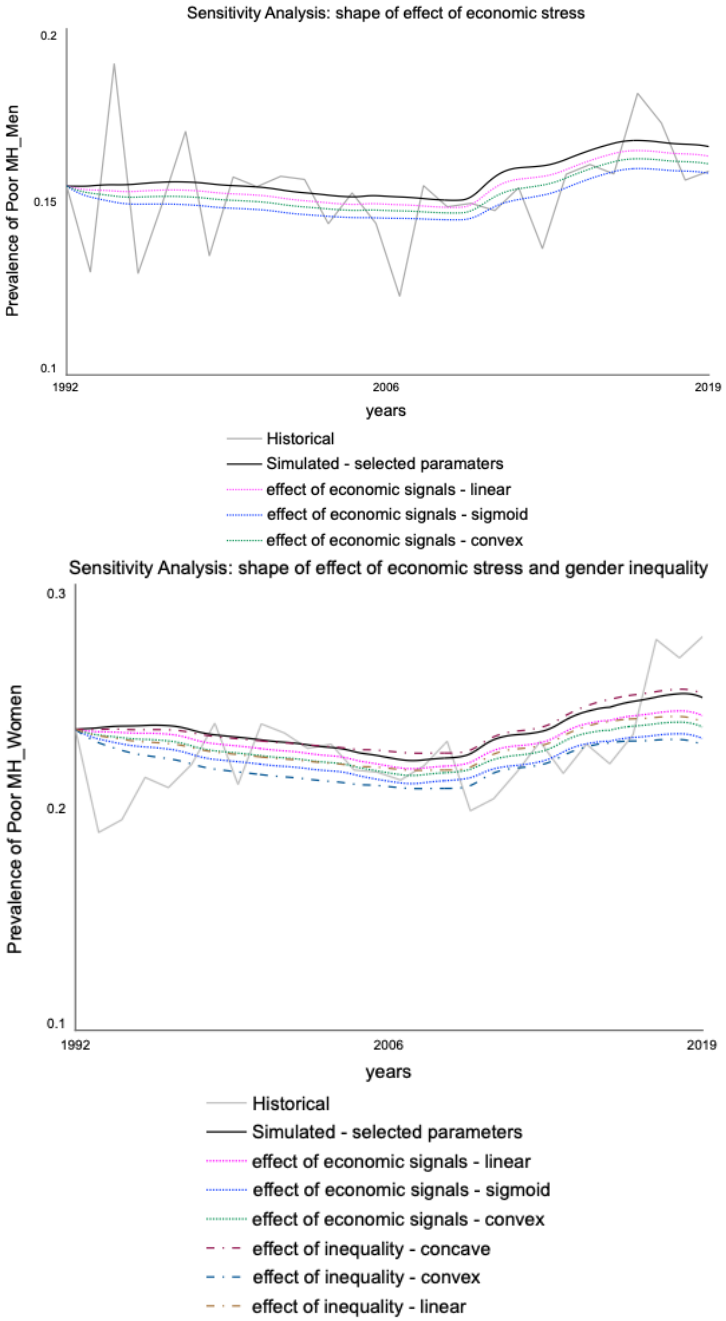


Figure K-11: Sensitivity analysis to effect multipliers on onset

Similarly, the effect of the weights in the two composites was minimal. The magnitude is different for effect of gender inequality on onset, but this would be reflected in the calculation of the implied baseline rate.

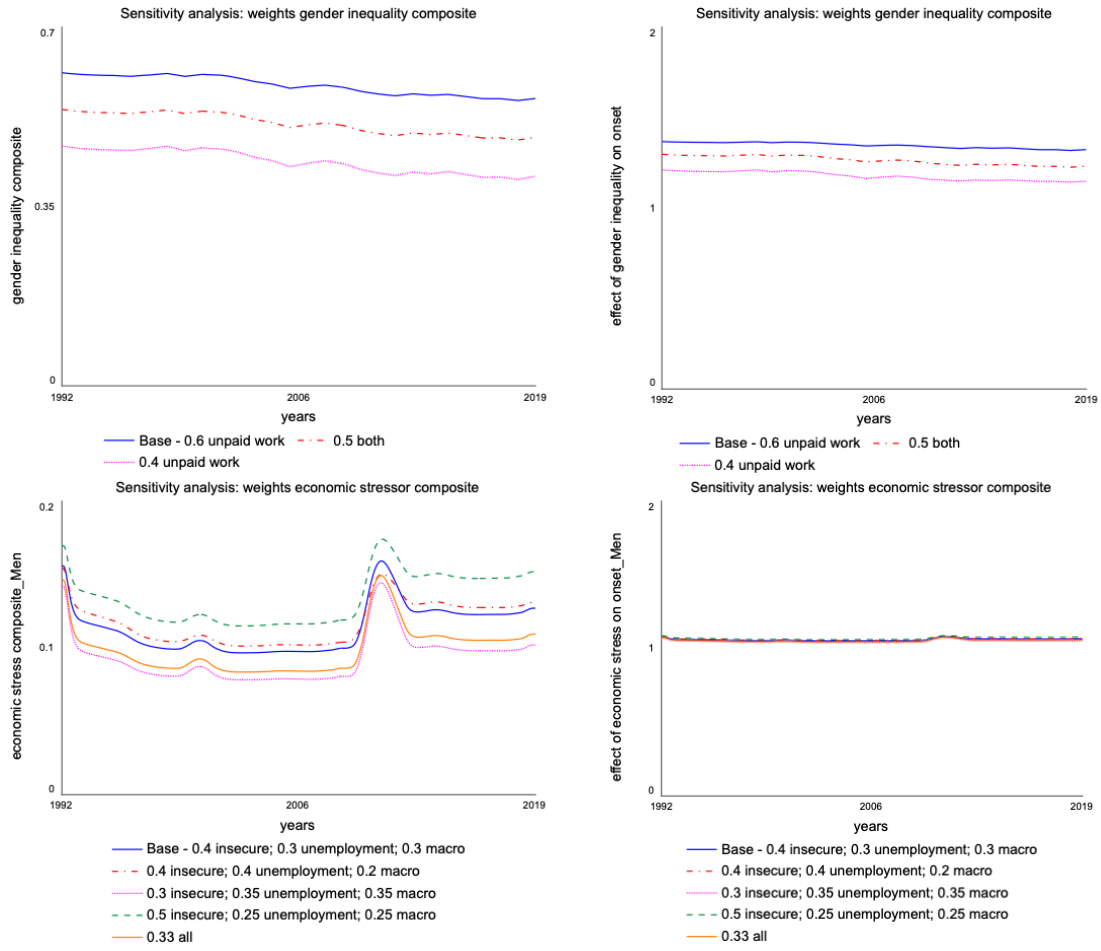
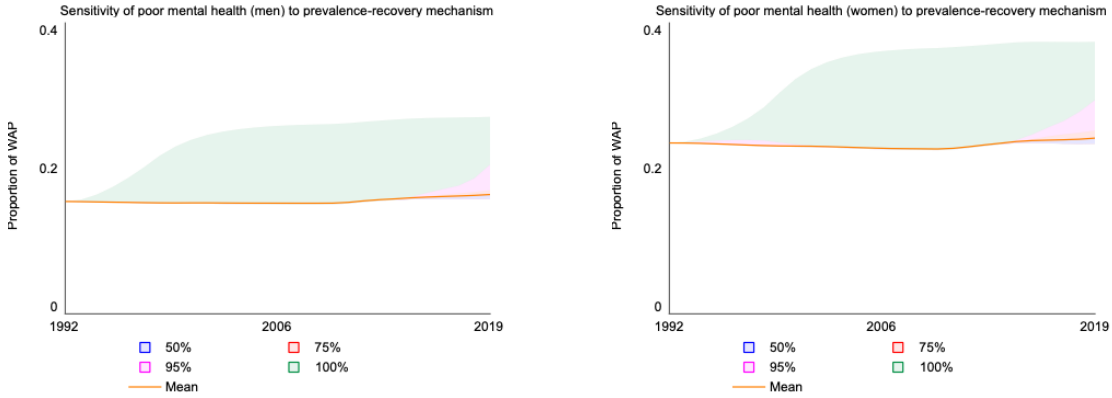


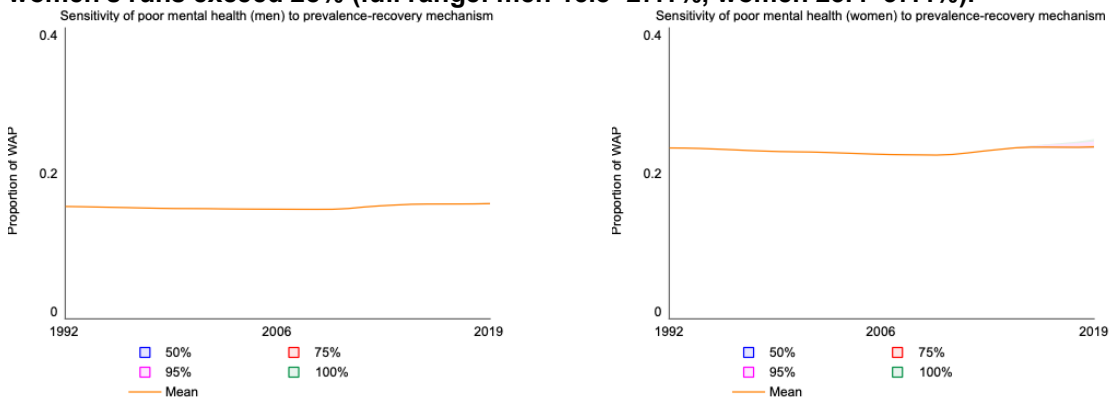
Figure K-12: Sensitivity analysis - composite weights

Sensitivity analyses were conducted on the prevalence-recovery mechanism — the halving point and the gender-specific degradation shapes — using one-way sweeps and a four-parameter, 200-run Latin-hypercube analysis on the fully integrated model (halving points 1.2–2.0; shapes 0.5–2.0, both genders). With all four parameters varied together the model is robust: median 2019 prevalence stays at the retained base (men 15.9%, women 23.5%) and the central 95% of runs lie within 15.8–20.1% (men) and 23.4–28.3% (women). The distribution is right-skewed — the full range reaches 27.1% (men) and 37.4% (women), but only 6 of 200 women's runs exceed 28% and 3 exceed 32%. The response is shape-driven and one-sided: as the one-way panels show, 2019 prevalence responds to the degradation-shape parameters far more than to the halving points, rising only when the shape is more convex than the retained value.

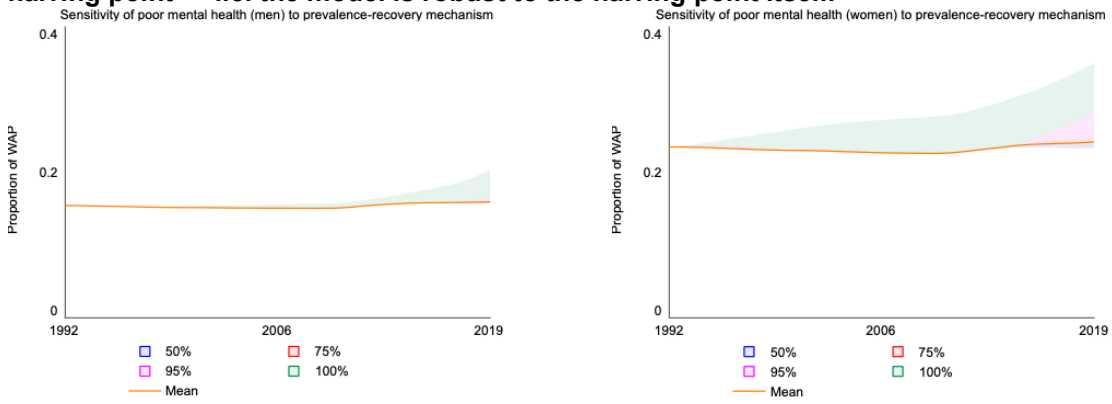
The high tail is one-sided and driven by a strongly convex women's degradation shape — recovery collapsing at the first rise in prevalence. Men remain bounded chiefly by construction: the calibration holds men's shape concave (1.5); the asymmetric panel confirms men stay ≤ ~20% under that shape, whereas allowing men's shape convex in the full four-parameter space lifts men's prevalence too (to 27.1%). This sensitivity therefore does not identify the mechanism as a leverage point; consistent with the loop-dominance analysis (largely dormant in 1992–2019, ~3% for women / ~1% for men by 2013), it is a bounded structural uncertainty retained so it can activate under the more extreme conditions.



Panel A: Full multivariate sweep. All four parameters (HP and shape, each gender) sampled simultaneously by Latin-hypercube (n = 200; HP 1.2–2.0, shape 0.5–2.0). The 2019 prevalence is centred on the base (median: men 15.9%, women 23.5%) with a right-skewed tail; the central 95% of runs lie within 15.8–20.1% (men) and 23.4–28.3% (women), and only 6 of 200 women's runs exceed 28% (full range: men 15.8–27.1%, women 23.4–37.4%).



Panel B: Halving-point-only sweep. The halving point alone is sampled 1.2–2.0 (n = 100), with the shapes held at their retained values (men 1.5, women 1.0). The fan is narrow for both genders, showing that the spread in Panel A is driven by the degradation shape, not the halving point — i.e. the model is robust to the halving point itself.



Panel C: Asymmetric sweep isolating the men–women contrast. Men's parameters are fixed at their retained values (HP 1.5, shape 1.5); women's halving point (1.2–2.0) and shape (0.5–1.5) are varied (n = 100).

Figure K-13: Sensitivity of poor mental health prevalence to the prevalence-recovery mechanism, by gender. The mechanism scales the baseline recovery rate; its two free parameters are the halving point (HP) and the gender-specific degradation shape (s). Panels sweep these over HP 1.2–2.0 and s 0.5–2.0 (retained values: HP 1.5 both genders; s = 1.0 women [linear], 1.5 men [concave]). Across the panels the model is robust to the halving point and to any linear-or-concave shape: median 2019 prevalence stays at ≈ the retained base (base: men 15.9%, women 23.6%), and a materially higher prevalence arises only under a strongly convex women's shape (s < 1).

The mechanism scales the baseline recovery rate by the multiplier defined; its two free parameters are the halving point (HP) and the gender-specific degradation shape (s). The mechanism modifies the baseline recovery rate by effect of change in prevalence:
recovery rate = baseline recovery rate × m, where
m = MAX(0.5, 1 − 0.5 × [MAX(0, r − 1) / (HP − 1)]^(shape)),
r = P(t) ⁄ P(1992) is the overall prevalence of poor mental health relative to its 1992 value, HP is the halving point (retained 1.5), and s is the degradation shape (retained: women 1.0 = linear, men 1.5 = concave). At r = 1 the multiplier is 1 (no effect); at r = HP it is 0.5 (recovery halved); the floor of 0.5 caps the reduction. The inner MAX(0, r − 1) guards against non-real values when prevalence dips below its 1992 level. This continuous form — equivalently written with a capacity multiplier equal to (HP − 1), as in the calibration in — was used for the sensitivity sweeps because table functions are not suitable to run large number of sensitivity runs; at the retained parameters it reproduces the final table-function model implemented in the simulation.

Appendix L Theil's statistics for isolated modules

This appendix reports the Theil's inequality statistics for each module run in isolation, supporting Section 6.6.4. It establishes the independent mechanical accuracy of each module before coupling.

L1 Labour market

Variable	R square	MAE/ Mean	RMSE	Bias	Variation	Covariation
Unemployed Proportion[Men]	0.59	0.23	0.02	0.49	0.22	0.30
Unemployed Proportion[Women]	0.60	0.11	0.01	0.02	0.44	0.54
Inactive Proportion[Men]	0.12	0.07	0.02	0.30	0.03	0.67
Inactive Proportion[Women]	0.80	0.04	0.01	0.00	0.01	0.99
Employed Proportion[Men]	0.63	0.01	0.01	0.10	0.02	0.87
Employed Proportion[Women]	0.87	0.02	0.01	0.01	0.04	0.95
Labour Market Tightness	0.85	0.14	0.25	0.03	0.21	0.77
Insecure Work	0.85	0.04	0.01	0.03	0.04	0.92

L2 Gender dynamics

Variable	R square	MAE/ Mean	RMSE	Bias	Variation	Covariation
Insecure Work[Men]	0.99	0.01	0.00	0.03	0.37	0.60
Insecure Work [Women]	0.99	0.01	0.00	0.06	0.35	0.59
Prevalence of Norms on Gender Roles	0.69	0.03	0.02	0.21	0.04	0.76
Gender Parity in Working Hours	0.82	0.01	0.01	0.12	0.34	0.55
Unpaid Work Burden	0.94	0.01	0.01	0.10	0.47	0.43
Gender Pay Gap	0.97	0.06	0.01	0.16	0.47	0.37

L3 Mental health

Variable	R square	MAE/ Mean	RMSE	Bias	Variation	Covariation
Poor Mental Health [Men]	0.1770	0.0700	0.0145	0.1472	0.2271	0.6257
Poor Mental Health [Women]	0.4054	0.0700	0.0201	0.3565	0.1721	0.4714

L4 Theil's Statistics to examine effect of integration

This appendix decomposes how integration changes behavioural reproduction, supporting Section 6.6.4. It quantifies the cost of coupling — chiefly the labour market module's smoothing of recession volatility.
While the prevalence of poor mental health exhibits a low R^2 due to stochastic year-to-year fluctuations in the historical data, the system dynamics paradigm prioritises endogenous behavioural reproduction over point-predictive accuracy. The low bias confirms the model successfully captures the fundamental structural trajectory. The high bias in unemployed_men and inactive_men suggests that the model consistently simulates these to be different to the historical data (higher for inactive and lower for unemployed) which is not unexpected given the flow rates were derived from UK level data. The low R square inactivity_men reflects that the model does not examine its causes and include them endogenously. It is important to note that the error itself is quite small, making these less problematic.

Theil's statistics for prevalence of poor mental health

	R-square	MAE/Mean	RMSE	Bias	Variance	Covariance
Men - integrated	0.1587	0.0614	0.0137	0.0029	0.7317	0.2654
Men - all historical inputs	0.1774	0.0695	0.0144	0.1426	0.2345	0.6229
Men - only LM historical	0.1781	0.0696	0.0144	0.1407	0.2266	0.6328
Men - only gender historical	0.1531	0.0616	0.0138	0.0030	0.7308	0.2662
Women - integrated	0.1445	0.0712	0.0208	0.0943	0.6818	0.2239
Women - all historical inputs	0.4044	0.0698	0.0201	0.3542	0.1749	0.4710
Women - only LM historical	0.4227	0.0727	0.0202	0.3747	0.2155	0.4099
Women - only gender historical	0.1346	0.0682	0.0204	0.0831	0.5609	0.3560

For women, the fully isolated configuration produces a substantial improvement in behavioural reproduction for women. R-squared increases. However, the Theil decomposition reveals that integration reduces covariance but also reduces bias. This reflects the fact that the isolated module's simulated mean (0.233) exceeds the historical mean (0.225): the module systematically overshoots, particularly during the 1992-2008 period when the historical data oscillates below the simulated trajectory. This overshoot arises because the isolated module receives the full force of historical stressor signals — including the sharp 2009 recession spike in the economic stressor composite — without the smoothing that occurs when these signals pass through the endogenous labour market module. The Gender module's contribution is negligible. The gender-incorporated model produces an R-square virtually identical to the fully isolated configuration, suggesting that the integration cost to mental health behavioural reproduction arises almost entirely from the labour market module.

For men, the reduction in R-square due to integration is small. However, there is a notable reduction in covariance, owing to an increase in variance. This is likely because the simulated version smooths the signals, whereas the historical data captures raw changes which can be volatile.

The integrated model produces a near-flat trajectory throughout the simulation period whereas the isolated module shows a clear inflection around 2009-2010. By 2019, the gap between the isolated and integrated simulations is approximately 0.010 — representing roughly 25% of the total divergence between the historical trajectory (0.277) and the integrated simulation (0.234).

The largest single discrepancy between isolated and integrated runs occurs during the 2008-2010 recession:

Year	Isolated [Women]	Integrated [Women]	Gap	Integrated as % of Isolated
2008	0.1313	0.1395	-0.008	106% (integrated higher)
2009	0.2343	0.1520	+0.082	65% (integrated captures only 65%)
2010	0.2725	0.1997	+0.073	73%
2011	0.1804	0.1899	-0.010	105%

The integrated model captures only 65% of the 2009 recession's stressor impact for women. This is because the integrated LM module smooths the unemployment spike — the endogenous response is less volatile than the historical data. Though transient, this under-response has cumulative consequences: the MH stock receives less "push" during the crisis, starts the recovery period at a lower prevalence, and never catches up.

Post-2012 the picture reverses: By 2019, the integrated composite is actually HIGHER than the isolated (0.172 vs 0.153 for women). This is because the integrated LM module's post-crisis recovery is also less volatile — it settles slightly above the historical values. The stressor composite gap therefore changes direction, and the post-2012 optimism gap is not in the stressor composite itself but in other variables.

The concave function shape attenuates upstream signal differences. The function is steepest near the origin and flattest in the operating range (composite 0.10-0.25). This means moderate upstream optimism produces small onset differences. Only large discrepancies (like the 2009 recession — 55% gap in composite) produce meaningful onset differences, and even those are transient.

Appendix M Summary of testing conducted on isolated modules

This appendix consolidates every verification test run on the three modules in isolation.

Test Name	Test Condition	Expected Outcome	Actual Outcome	Structural Insight
Boundary adequacy — insecure work	Insecure work varied by +/-20%	Smooth model response; no structural breaks	Smooth response; internal feedback structure robust to boundary uncertainty	Model not over-sensitive to boundary data estimation
Boundary adequacy — labour demand	Labour demand varied by +/-20%	Proportional response in employment levels	Counter-intuitive: higher demand increased women's MH distress (increased GPG exposure through higher employment)	Reveals cross-module coupling: more employment exposes women to structural penalty
Transient pulse	Temporary increase in labour demand (1999-2000)	Immediate unemployment decline; return to baseline after removal	Compensatory "overshoot" above baseline (2000-2001); decay to equilibrium by 2005	Model captures structural "memory" via internal time-constants; non-linear transient response
B1 loop knockout	Step increase in demand WITH vs WITHOUT LMT elasticity link	With: LMT stabilises at 1.8 (satiation). Without: LMT rises unregulated to 2.0; unemployment non-responsive	As expected. Without LMT link, unemployment stocks show zero response to demand shock	Loop B1 (Demand Satiation) is the primary structural engine for macroeconomic adjustment
Zero stocks initialisation	Both MH stocks set to zero with positive inflow rates	Prevalence remains at zero (no phantom population generation)	Prevalence stays exactly zero throughout	Flow equations correctly formulated as stock-dependent; no spurious generation
Zero recovery rate	Baseline recovery rate set to 0	Prevalence rises to maximum (1.0)	Prevalence stabilised at 1.0	Conservation confirmed; all individuals eventually transition to poor MH without recovery

Test Name	Test Condition	Expected Outcome	Actual Outcome	Structural Insight
Maximum stress	Economic stressors and women's inequality composites forced to 1.0 (maximum) via STEP from 2000	Prevalence rises but does not exceed 1.0; gender gap emerges from multiplier structure	Women ≈ 0.5, Men ≈ 0.2; neither exceeds population boundary	Stock-dependent flows correctly bounded; gender divergence is emergent property of multipliers, not arbitrary parameterisation
Initialisation sensitivity	1992 prevalence artificially inflated to 0.25 (Men), 0.30 (Women) — well above empirical baseline	System decays toward internal "resting state"	Prevalence decayed to lowest point by 2007 (Men: 0.15, Women: 0.23); with constant stressors, equilibrium at 0.152/0.234	Long-term behaviour driven by feedback structure, not starting values; implied baseline calculations mathematically sound
Equilibrium initialisation check	Stressors held constant at 1992 levels	MH stocks remain flat (no autonomous drift)	Stocks remain flat	Confirms equilibrium initialisation is mathematically correct
Non-equilibrium initialisation	Raw baseline recovery rates used instead of equilibrium-adjusted rates	Prevalence settles at lower level but retains rising trajectory when stressors change	In all conditions (except 1992 constants), rise in prevalence observed; lower settling level as expected	Behaviour (rising trajectory) is robust to specific parameter values; only magnitude differs
Norm stock inertia; frozen signal	Economic signal composite frozen at 1992 baseline (0.24) for entire simulation	Norms should exhibit slow endogenous decline (cohort replacement effect)	Slow decline confirmed; norms decay independent of current economic conditions	Stock formulation correctly captures demographic turnover as autonomous decay mechanism
Norm stock inertia — signal switched off	Effect of economic signals explicitly set to 0 (no reinforcement)	Norms should deteriorate faster than historical	Norms deteriorated significantly faster than observed	Without economic reinforcement, norms erode at full cohort replacement rate; confirms reinforcement is structurally necessary
Norm stock — maximum reinforcement	Economic signals held constant at 1 (maximum inequality signal)	Norms should remain constant (erosion = reinforcement)	Norms stayed constant	Confirms structural balance: at maximum signal,

Test Name	Test Condition	Expected Outcome	Actual Outcome	Structural Insight
				reinforcement exactly offsets erosion
Parity shock	Economic Signal Composite forced to zero from 2000	Norms should decline but slowly (structural memory)	Norms reached only 0.244 by 2050 despite absolute removal of reinforcement	Module correctly distinguishes transient signals from deep-seated structural memory; governed by cohort replacement delay
Boundary adequacy — GPG	Baseline gender pay gap varied by +/-20%	Some response in norms trajectory	Almost no difference in prevalence of norms	Norms trajectory robust to GPG boundary uncertainty

Appendix N Model equations

This appendix presents the full model — exogenous inputs, calibration data, the equations of all three modules, initial values, parameter values, and graphical functions — as a set of structured tables. Calibration and test toggles are documented separately and are not part of the operative model.

N1 Labour market module

Exogenous Input	Unit	Years	Drives
Working age population (WAP)	Working-age population (WAP)	1992–2023	Normalises labour demand to a proportion
proportion_of_WAP_that_is_male	Dimensionless	1992–2023	Gender-weighting of overall employed/unemployed proportions
historical_labour_demand	Jobs	1992–2024	Labour-demand fraction → vacancy fraction → LMT
baseline_insecure_work_-_historical_trend	Dimensionless	1992–2019	Base level of insecure work (then modified endogenously by LMT)
effect_of_increase_in_statutory_pension_age_for_women	Dimensionless	2009–2018	Women's become_inactive_from_employment flow

Stock [Gender] Unit: proportion of WAP	Net change = inflows − outflows
Employed proportion	(job gain + become employed from inactivity) − (job loss + become inactive from employment)
Unemployed proportion	(job loss + join labour force) − (job gain + exit labour force)
Inactive proportion	(exit labour force + become inactive from employment) − (join labour force + become employed from inactivity)

Initial stock values (1992 LFS data)	Men	Women
Employed proportion	0.7647	0.6127
Unemployed proportion	0.0938	0.0521
Inactive proportion	0.1415	0.3352

Flow (Gender) Unit: 1/year	Equation	Gender Difference
job gain	Unemployed proportion × baseline job-gain rate × effect of LMT on job gain × effect of poor MH on job gain	None
job loss	Employed proportion × baseline job-loss rate × effect of LMT on job loss × effect of poor MH on job loss	None
become employed from inactivity	Inactive proportion × baseline rate of becoming employed from inactivity × effect of LMT on job gain	None
become inactive from employment	Employed proportion × baseline rate of becoming inactive from employment × effect of poor MH on labour market exit	Women’s is men’s (left column) × effect of gender norms on women's participation × effect of statutory pension-age increase
join labour force	Inactive proportion × baseline entry rate	Women’s is men’s (left column) ÷ effect of gender norms on women's participation (reciprocal: norms reduce women's entry)
exit labour force	Unemployed proportion × baseline exit rate	baseline exit rate differs by gender

Key auxiliaries	Unit	Equation / definition
labour-demand fraction	proportion of WAP	historical labour demand ÷ WAP
vacancy fraction	proportion of WAP	MAX(labour-demand fraction − overall employed proportion, 0), implemented as a smooth curve with a small non-zero frictional floor: $(\text{gap} + \sqrt{(\text{gap}^2) + 0.0001}) / 2$
labour market tightness (LMT)	dimensionless	vacancy fraction ÷ overall unemployed proportion
normalised LMT	dimensionless	LMT ÷ normal LMT (1.44)

Key auxiliaries	Unit	Equation / definition
effect of LMT on job gain	dimensionless	normalised LMT ^ elasticity
effect of LMT on job loss	dimensionless	normalised LMT ^ (−elasticity)
effect of LMT on insecure work	dimensionless	normalised LMT ^ (−elasticity)
insecure work	dimensionless	baseline insecure work trend × delayed effect of LMT
delayed effect	dimensionless	DELAY1(effect of LMT on insecure work, 3 years)
effect of poor MH on job gain / job loss / LM exit	dimensionless	table functions of poor-MH prevalence
effect of gender norms on women's participation	dimensionless	table function of norms
overall employed / unemployed proportion	proportion of WAP	gender-weighted average using proportion of WAP that is male

N2 Gender dynamics module

Exogenous Input	Unit	Years	Drives
baseline_gender_pay_gap	dimensionless	1992–2024	Gender pay gap
childcare_availability_index_estimated	dimensionless	1992–2024	Fraction of the hours-parity gap closed by childcare

Element	Equation	Units	Initial value
Prevalence of Norms on Gender Roles (stock)	net change = norm reinforcement − norm erosion	dimensionless	0.54
norm erosion (outflow)	Norms ÷ time for norms to change	1/year	
norm reinforcement (inflow)	MIN([Norms × (E + (1 − E) × Norms)] ÷ time for norms to change , (1 − Norms) ÷ DT	1/year	

Element	Equation	Units	Initial value
	) where E = economic-signal composite		

Key auxiliaries	Units	Equation / definition
economic-signal composite	dimensionless	GPG × weight_GPG + (1 − hours parity) × weight_hours + inactivity gap × weight_inactivity
gender parity in working hours	dimensionless	parity-from-norms + (1 − parity-from-norms) × fraction of gap closed by childcare
parity in working hours based on norms	dimensionless	table function of norms
unpaid work burden	dimensionless	0.5 + (unpaid-work burden based on norms − 0.5) × (1.5 − gender parity in working hours)
unpaid-work burden based on norms	dimensionless	table function of norms
gender pay gap	dimensionless	baseline gender pay gap + effect of gender parity on GPG (additive, not multiplicative)
effect of gender parity on GPG	dimensionless	table function of gender parity
gendered exposure to insecure work — women	dimensionless	effect of parity on insecure work × insecure work
gendered exposure to insecure work — men	dimensionless	residual: (insecure work − women's exposure × women's employed share) ÷ men's employed share
effect of parity on insecure work	dimensionless	table function of gender parity
fraction of gap closed by childcare	dimensionless	table function of the childcare availability index

N3 Mental health module

Element	Equation	Gender difference
Prevalence of Good Mental Health (stock) [Gender]	net change = recovery − incidence	None
Prevalence of Poor Mental Health (stock) [Gender]	net change = incidence − recovery	None
incidence [Gender]	Good-MH prevalence × baseline onset rate × effect of economic stressors on MH	Women's is men's (left column) × effect of gender inequality on MH
recovery [Gender]	Poor-MH prevalence × recovery rate × effect of employment on recovery (insecure work adjusted)	[Gender]

Initial stock values (1992 primary analysis of BHPS survey)	Men	Women
Prevalence of Good Mental Health	0.8454	0.7649
Prevalence of Poor Mental Health	0.1546	0.2351

Key auxiliaries	Equation / definition	Gender difference	Unit
recovery rate [Gender]	baseline recovery rate × effect of prevalence on recovery rate	baseline rate differs by gender (	1/year
effect of prevalence on recovery rate [Gender]	table function of change in prevalence	men concave, women linear	dimensionless
change in prevalence	overall poor-MH prevalence ÷ its 1992 initial value	NA	dimensionless
economic-stressor composite [Gender]	unemployed × 0.3 + insecure work × 0.4 + perceived unemployment change × 0.3	None	dimensionless
effect of economic stressors on MH [Gender]	table function of the economic-stressor composite	None	dimensionless

Key auxiliaries	Equation / definition	Gender difference	Unit
perceived unemployment change	SMTH3(MAX(0, year-on-year % rise in unemployment), 0.5 yr, initial 0.11)	NA	dimensionless
women's stress index	GPG exposure × 0.4 + perceived unpaid work × 0.6	Only applies to women	dimensionless
GPG exposure	gender pay gap × women's employed proportion	Only applies to women	dimensionless
perceived unpaid work	((unpaid work − 0.5) ÷ 0.5) × effect of norms on unpaid work	Only applies to women	dimensionless
effect of norms on unpaid work	1 + (1/norms − 1)	Only applies to women	dimensionless
effect of gender inequality on MH	table function of the women's stress index	Only applies to women	dimensionless
effect of employment on recovery (insecure-adjusted) [Gender]	1 + (effect of employment on recovery − 1) × effect of insecure work on recovery	None	dimensionless
effect of employment on recovery [Gender]	table function of employed proportion	None	dimensionless
effect of insecure work on recovery [Gender]	table function of change in insecure work	men and women use different thresholds	dimensionless
change in insecure work [Gender]	insecure work ÷ its 1992 initial value	None	dimensionless

Appendix O Calibration data

O1 Labour market module

Source: ONS Labour Force Survey (LFS) extracted from NOMIS specifically June figures for labour force survey - national and regional - headline indicators (seasonally adjusted) from 1992

Year	Employed [M]	Employed [W]	Unemployed [M]	Unemployed [W]	Inactive [M]	Inactive [W]	Insecure work
1992	0.7647	0.6127	0.0938	0.0521	0.1415	0.3352	0.2708
1993	0.7411	0.5996	0.1002	0.0536	0.1587	0.3468	0.2807
1994	0.7502	0.6121	0.0986	0.0554	0.1512	0.3325	0.2764
1995	0.7470	0.6265	0.0852	0.0484	0.1678	0.3251	0.2863
1996	0.7426	0.6208	0.0852	0.0425	0.1721	0.3367	0.2867
1997	0.7436	0.6282	0.0822	0.0482	0.1741	0.3236	0.2807
1998	0.7504	0.6387	0.0691	0.0427	0.1805	0.3185	0.2754
1999	0.7453	0.6436	0.0699	0.0381	0.1847	0.3183	0.2747
2000	0.7590	0.6521	0.0684	0.0395	0.1726	0.3083	0.2735
2001	0.7671	0.6655	0.0597	0.0361	0.1732	0.2983	0.2665
2002	0.7616	0.6676	0.0584	0.0377	0.1800	0.2947	0.2641
2003	0.7842	0.6739	0.0488	0.0319	0.1670	0.2942	0.2506
2004	0.7765	0.6823	0.0593	0.0347	0.1642	0.2830	0.2537
2005	0.7863	0.6736	0.0511	0.0349	0.1626	0.2915	0.2568
2006	0.7720	0.6841	0.0516	0.0351	0.1764	0.2808	0.2533
2007	0.8033	0.6981	0.0392	0.0343	0.1576	0.2677	0.2451
2008	0.7938	0.6950	0.0371	0.0290	0.1691	0.2760	0.2477
2009	0.7688	0.6743	0.0672	0.0435	0.1640	0.2822	0.2825
2010	0.7349	0.6685	0.0806	0.0523	0.1845	0.2792	0.2991
2011	0.7530	0.6800	0.0753	0.0501	0.1717	0.2700	0.3041
2012	0.7552	0.6743	0.0757	0.0522	0.1690	0.2736	0.3161
2013	0.7597	0.6848	0.0656	0.0516	0.1747	0.2636	0.3455
2014	0.7687	0.7007	0.0598	0.0427	0.1715	0.2566	0.3498
2015	0.7611	0.7187	0.0541	0.0379	0.1849	0.2434	0.3436

Year	Employed [M]	Employed [W]	Unemployed [M]	Unemployed [W]	Inactive [M]	Inactive [W]	Insecure work
2016	0.7723	0.7073	0.0496	0.0342	0.1781	0.2585	0.3341
2017	0.7897	0.7135	0.0351	0.0260	0.1752	0.2605	0.3326
2018	0.7898	0.7159	0.0387	0.0269	0.1715	0.2572	0.3315
2019	0.7841	0.7231	0.0343	0.0219	0.1816	0.2550	0.3120

O2 Gender dynamics module

- Norms: Extracted from Zhou and Kan (2019), figure 2 which reports the "Weighted distribution of the five types of the division of labour for the selected couples, 1991 to 2017". The analysis is based on harmonized longitudinal data from the British Household Panel Survey (BHPS) and the Understanding Society (UKHLS), including a representative sample from Great Britain (England, Scotland, and Wales). The data pertains to heterosexual couples (either married or cohabiting), where the female partner is aged 20–44 and the male partner is aged 20–59. The male breadwinner and modified male breadwinner model are summed.
- GPG: ASHE from Scottish government - "Gender Pay Gap - Median Hourly Earnings [excluding Overtime], All Employees, 1997-2024"; https://www.gov.scot/publications/annual-survey-of-hours-and-earnings-2024/documents/
- Unpaid work: Extracted from Wheatley, Lawton and Hardill (2018), figure 8.7; these were weighted by labour market participation rates to estimate the weekly housework average for men and women and then the ratio of unpaid work burden.
- Gender Parity in working hours: ASHE from NOMIS annual survey of hours and earnings - workplace analysis

Year	Prevalence of Norms (stock)	Gender pay gap (GPG)	Unpaid work burden	Gender parity in working hours
1992	0.54		0.7527	0.6150
1993	0.53		0.7527	0.6100
1994	0.52		0.7516	0.6100
1995	0.51		0.7447	0.6100
1996	0.52		0.7367	0.6100
1997	0.53	0.2650	0.7255	0.6081
1998	0.52	0.2580	0.7353	0.6116
1999	0.51	0.2460	0.7296	0.6198
2000	0.52	0.2530	0.7261	0.5950
2001	0.51	0.2500	0.7254	0.6198

Year	Prevalence of Norms (stock)	Gender pay gap (GPG)	Unpaid work burden	Gender parity in working hours
2002	0.53	0.2500	0.7229	0.6061
2003	0.52	0.2220	0.7257	0.6160
2004	0.54	0.2130	0.7168	0.6222
2005	0.54	0.1790	0.7188	0.6086
2006	0.53	0.1980	0.7123	0.6086
2007	0.51	0.2040	0.7197	0.6343
2008	0.49	0.1770	0.7121	0.6286
2009	0.45	0.1770	0.7030	0.6114
2010	0.46	0.1730	0.6989	0.6000
2011	0.46	0.1660	0.6979	0.6000
2012	0.46	0.1760	0.6960	0.6029
2013	0.45	0.1750	0.6951	0.6000
2014	0.44	0.1760	0.6813	0.6000
2015	0.44	0.1710		0.6000
2016	0.43	0.1560		0.6000
2017		0.1600		0.6000
2018		0.1500		0.6000
2019		0.1440		0.6200

O3 Mental health module

Source: Primary analysis of BHPS data

Year	Poor MH prevalence [M]	Poor MH prevalence [W]
1992	0.1546	0.2351
1993	0.1297	0.1888
1994	0.1900	0.1945
1995	0.1293	0.2135
1996	0.1491	0.2089
1997	0.1704	0.2193

Year	Poor MH prevalence [M]	Poor MH prevalence [W]
1998	0.1344	0.2378
1999	0.1572	0.2103
2000	0.1544	0.2375
2001	0.1574	0.2334
2002	0.1565	0.2265
2003	0.1437	0.2284
2004	0.1526	0.2168
2005	0.1438	0.2158
2006	0.1227	0.2123
2007	0.1547	0.2187
2008	0.1486	0.2297
2009	0.1496	0.1985
2010	0.1475	0.2039
2011	0.1541	0.2159
2012	0.1365	0.2295
2013	0.1580	0.2153
2014	0.1608	0.2284
2015	0.1581	0.2196
2016	0.1814	0.2326
2017	0.1728	0.2754
2018	0.1563	0.2671
2019	0.1589	0.2767

Appendix P Loop dominance analysis

This appendix provides the loop-dominance source material supporting Sections 7.2 and 7.4.

- Table P-1 reports the results of a sensitivity analysis for the Section 7.2 LTM analysis filtered for prevalence of poor mental health, enabling gendered comparison. Equalising women's insecure work threshold changes the single shared run, so both genders' relative shares shift between this run and Table 7-3; the men's split is expected to move and is not an inconsistency. The tested finding — the gender asymmetry in drain dominance — is robust.
- Table P-2 supports Section 7.4 by giving the full-system, year-by-year dominance series.

A more technical description of the model's temporal architecture (the three-timescale structure) and chronology follows, going beyond the main Chapter 7 text.

Table P-1: Relative loop dominance scores for sensitivity LTM analysis holding insecure work threshold same for men and women

Gender	CLD Loop Set	Loop Type	Relative Dominance	Relative Dominance (2008-2013)	Peak Year	Active (2008-10)	Years Active	Years Active (> 1%)
M	Unemployment-MH	balancing	0%	0%	N/A	0	0	0
M	Unemployment-MH	reinforcing	4%	5%	2011	2	5	3
M	Insecure work - MH	balancing	3%	0%	1993	1	12	1
M	Insecure work - MH	reinforcing	3%	0%	2006	3	27	1
M	Norms - MH	balancing	0%	0%	N/A	0	0	0
M	Norms – MH	reinforcing	0%	0%	N/A	0	0	0
M	prevalence - recovery	balancing	0%	0%	N/A	0	0	0
M	prevalence - recovery	reinforcing	0%	0%	N/A	0	0	0
M	churn	reinforcing	51%	81%	2013	3	27	7
M	drain	balancing	40%	14%	2019	3	27	18
W	Unemployment-MH	balancing	4%	6%	2011	3	27	3
W	Unemployment-MH	reinforcing	18%	25%	2011	3	27	3
W	Insecure work - MH	balancing	0%	0%	1993	0	11	1
W	Insecure work - MH	reinforcing	7%	13%	2010	3	27	3
W	Norms - MH	balancing	2%	2%	2010	3	27	2
W	Norms - MH	reinforcing	1%	2%	2010	3	27	2
W	prevalence - recovery	balancing	0%	0%	N/A	0	0	0
W	prevalence - recovery	reinforcing	0%	0%	N/A	0	0	0
W	churn	reinforcing	42%	47%	2012	3	27	13
W	drain	balancing	25%	5%	2002	3	27	19

Table P-2: Full-system year-by-year loop dominance

CLD Loop Set	Loop Type	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Unemployment-MH	balancing	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	4.93%	0.07%	0.00%	0.00%	0.03%
Unemployment-MH	reinforcing	0.00%	0.00%	0.06%	0.01%	0.00%	0.00%	35.48%	0.13%	0.00%	0.00%	0.25%
Insecure work - MH	balancing	8.75%	1.22%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Insecure work - MH	reinforcing	0.72%	0.06%	1.14%	0.08%	0.00%	0.00%	0.00%	0.11%	0.01%	0.00%	0.21%
Norms - MH	balancing	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
Norms - MH	reinforcing	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%
prevalence - recovery	balancing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
prevalence - recovery	reinforcing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
churn	reinforcing	5.14%	5.61%	17.57%	13.25%	8.57%	2.27%	3.83%	7.49%	12.16%	16.66%	36.21%
drain (other)	balancing	37.48%	53.28%	40.71%	61.76%	41.85%	48.89%	39.66%	49.99%	64.56%	33.51%	28.16%
demand saturation	balancing	18.22%	4.82%	5.17%	1.35%	2.41%	1.06%	0.75%	6.82%	1.63%	2.72%	4.85%
demand saturation	reinforcing	6.60%	5.24%	11.28%	18.93%	1.03%	1.96%	4.54%	7.25%	14.72%	0.37%	7.32%
norms	balancing	0.00%	0.00%	0.01%	0.12%	0.00%	0.01%	0.00%	0.00%	0.03%	0.00%	0.01%
norms	reinforcing	13.30%	16.90%	8.99%	2.28%	30.60%	43.42%	5.28%	25.90%	4.83%	16.70%	4.06%
drain - recovery	balancing	9.80%	12.83%	15.04%	2.17%	15.46%	2.32%	5.54%	2.16%	1.98%	30.00%	18.85%

CLD Loop Set	Loop Type	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Unemployment-MH	balancing	0.00%	0.01%	0.00%	0.00%	0.00%	0.05%	1.63%	14.53%	0.00%	0.00%	0.00%
Unemployment-MH	reinforcing	0.04%	0.05%	0.16%	0.00%	0.01%	1.14%	12.24%	71.93%	0.00%	0.01%	0.00%
Insecure work - MH	balancing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	0.04%	0.02%	0.08%	0.03%
Insecure work - MH	reinforcing	0.05%	0.13%	9.68%	0.00%	0.07%	0.00%	31.53%	2.56%	0.01%	0.03%	0.01%
Norms - MH	balancing	0.00%	0.01%	0.43%	0.00%	0.01%	0.00%	10.00%	0.00%	0.00%	0.00%	0.00%
Norms - MH	reinforcing	0.00%	0.01%	0.40%	0.00%	0.01%	0.00%	10.60%	0.00%	0.00%	0.00%	0.00%
prevalence - recovery	balancing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.11%	0.11%
prevalence - recovery	reinforcing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	4.10%	0.85%
churn	reinforcing	17.13%	22.29%	7.02%	1.99%	12.33%	3.74%	20.86%	8.74%	21.08%	13.49%	9.77%
drain (other)	balancing	43.50%	41.99%	3.80%	54.18%	49.97%	61.18%	1.68%	0.71%	57.63%	35.89%	63.38%
demand saturation	balancing	4.57%	3.08%	2.78%	1.74%	4.90%	5.90%	0.48%	0.03%	0.32%	1.95%	1.90%
demand saturation	reinforcing	5.99%	11.56%	49.45%	2.05%	11.29%	6.95%	7.10%	0.01%	6.73%	0.11%	12.08%
norms	balancing	0.00%	0.02%	16.60%	0.00%	0.04%	0.00%	1.10%	0.00%	0.00%	0.00%	0.03%
norms	reinforcing	12.01%	3.97%	9.37%	34.10%	14.25%	19.80%	1.25%	0.26%	3.33%	22.30%	9.75%
drain - recovery	balancing	16.63%	16.92%	0.29%	5.88%	7.18%	1.30%	1.38%	1.18%	10.88%	18.93%	2.12%

CLD Loop Set	Loop Type	2015	2016	2017	2018	2019
Unemployment-MH	balancing	0.00%	0.00%	0.09%	0.00%	0.00%
Unemployment-MH	reinforcing	0.00%	0.00%	0.94%	0.00%	0.00%
Insecure work - MH	balancing	0.05%	0.11%	68.57%	0.00%	0.19%
Insecure work - MH	reinforcing	0.00%	0.03%	2.15%	0.00%	0.02%
Norms - MH	balancing	0.00%	0.00%	1.57%	0.00%	0.00%
Norms - MH	reinforcing	0.00%	0.01%	24.88%	0.00%	0.01%
prevalence - recovery	balancing	0.09%	0.26%	0.55%	0.03%	3.51%
prevalence - recovery	reinforcing	1.39%	1.12%	0.02%	0.61%	4.55%
churn	reinforcing	5.59%	16.03%	0.90%	2.12%	22.12%
drain (other)	balancing	55.48%	46.52%	0.03%	58.43%	33.99%
demand saturation	balancing	5.92%	3.69%	0.00%	1.34%	3.38%
demand saturation	reinforcing	4.62%	20.42%	0.08%	3.07%	7.41%
norms	balancing	0.00%	0.08%	0.08%	0.00%	0.01%
norms	reinforcing	22.40%	7.65%	0.10%	32.20%	7.34%
drain - recovery	balancing	4.44%	4.09%	0.03%	2.27%	17.48%

P1 The model's temporal architecture

The model's feedback loops operate at three timescales, set by the delays built into its design. These three layers are design features, not findings. The findings are discussed in the chronology that follows, which traces how loops at each timescale switch on over the historical period. Setting the layers out first explains why the system reacts differently to a short shock than to slow structural change.

- **Short-delay layer (drain and the unemployment-MH trap (R1)):** Drain dominates in most years, interrupted by sharp R1 spikes during crises. R1 takes 2–3 years to complete a full cycle because the health causation and health selection pathways each require roughly one year to propagate; once it peaks, drain reasserts within a year. This layer governs the system's cyclical behaviour.
- **Medium-delay layer (insecure work-MH pathways and demand saturation):** A three-year delay links labour market tightness to insecure work, so its effect surfaces slowly and reflects cumulative/historical labour market conditions. At the same time, demand saturation fluctuates with the business cycle but peaks when the labour market reaches structural capacity. This layer governs the slow build-up of structural vulnerability.
- **Long-delay layer (norms entrenchment (R3) and norms-MH (R4):** The norms stock erodes through cohort replacement on a 45-year delay, so across the 27-year simulation it is in mid-transition. This is why R3 is the most persistently active informational loop in the system; it retreats during crises when the short-delay loop dominates and surges during periods of stability. This layer governs the structural evolution of gender inequality.

P2 Detailed chronology for key years

- **1999**: R1's activation does not principally come at drain's expense; drain only loses about 6 percentage points between 1998 and 1999. R1 rises *as norms-reinforcing collapses*, from over 43% in 1998 to just over 5% in 1999. This is therefore at least as much a story of norms-reinforcing falling out of dominance as it is a story of R1 rising into it. The absolute link scores confirm that R1 is materially more active than in 1998, and that the direct effect of norms on norm reinforcement reduces whereas the effect of economic signal on norm reinforcement increases. The shift in relative dominance explains why the LTM percentage was as large as it was, but the link scores confirm that the loop came alive in 1999.
- **2006**: The system reaches structural capacity. Demand saturation (B1) surges to 49.5% while the material drains collapse (drain 3.8%), and the recovery of poor mental health is essentially switched off (drain-recovery 0.3%). With labour market tightness at its simulation high, the economy can absorb no further employment growth. The insecure work–mental health reinforcing loop becomes meaningfully active for the first time (9.7%), and the gender-power-rebalance loop (B3) peaks at 16.6% as women's rising participation tightens the market, widens the gender inactivity gap, and feeds back to reinforce traditional norms. No mental health crisis is yet visible, but the structural preconditions are now in place.

- **2009**: R1 first becomes detectable at 1.1%. The recession has raised unemployment, but the health causation pathway (unemployment worsening mental health) has not yet generated enough additional prevalence to activate the health selection pathway (poor mental health reducing employment). As unemployment increases, effect of economic stressors on mental health rises and labour market tightness (LMT) collapses by over 50% in one year — the largest one-year movement in this variable in the simulation). However, prevalence of poor mental health has not changed. The health-causation pathway has fired (unemployment up → stressors up) but the prevalence stock has not yet moved enough for the health-selection direction (prevalence reducing employment) to register. R1 is a vicious cycle that needs both directions to complete one full pass before its loop score becomes meaningful, and at this point it has only completed half a turn.
- **2010**: The crisis enters a compounding phase. Three mental health linked reinforcing loops activate simultaneously: insecure work-mental health (31.5%), unemployment-mental health (12.2%), and norms-mental health (10.6%). Drain collapses to 3.1% (general drain 1.7%; recovery drain-recovery 1.4%). The activation sequence reflects the model's delay structure. The insecure work channel dominates first (31.5% versus R1's 12.2%), followed by the unemployment-mental health channel (R1) and finally the norms-mental health channel activates at 10.6%. This suggests that the long-delay norms layer, normally a background force, produces direct mental health effects when crisis conditions strip away the recovery capacity that normally masks it.
- Insecure work jumps significantly and its effect on recovery reduces recovery for men for the first time. At the same time, LMT falls further as unemployment rises and the effect of economic stressors on mental health jumps further for both genders. R2 dominates ahead of R1 in 2010 for two structural reasons. First, insecure work is in its *fastest rate of change* of the entire simulation in 2010 (the single year step change is unprecedented), so its link scores are very large. Second, the insecure work pathway operates through both *recovery* and incidence. Recovery operates on a faster timescale than the slow accumulation of poor-mental health prevalence that R1 requires, and so the LTM algorithm picks up the R2 signal first.
- **2011**: R1 achieves complete dominance at 71.9%; the single highest loop score in the entire simulation. This R1 dominance is the culmination of a 2–3 year build, not an immediate consequence of the recession. The delay reflects the system's temporal architecture: the feedback loop requires both the health causation direction and the health selection direction to propagate, each requiring approximately one year, before the full reinforcing cycle reaches peak force.
- Unemployment continues rising, now three years into a sustained climb from the 2008 trough. The effect of insecure work on recovery now drops recovery for both genders - the largest one-year movement in the simulation. The poor mental health stocks are now visibly rising. The full vicious cycle is now propagating: unemployment up → stressors up → prevalence up → unemployment up. Both directions of the loop have fired, and because R1 is the short-delay loop in the system (cycle time 2–3 years), it dominates once both arms are active. With R1 absorbing over 70% of the relative dominance, every other loop loses share by

construction. R2 has not become inactive, but its share of the variance is now small in relative terms.

- **2012:** With unemployment falling, the material drains reassert (drain 57.6%) and recovery re-engages (drain-recovery 10.9%, up from ~1% at the peak) and the system begins to settle.
- **2013**: The endogenous prevalence-recovery loop appears for the first time, contributing 4.1% reinforcing and 3.1% balancing. They are entirely dormant before the crisis; their activation signals that prevalence has now risen far enough to enter the range where the prevalence–recovery feedback operates. This is the new, post-crisis vulnerability the system carries forward.
- **2017:** Most variables are stable: insecure work has plateaued since 2012, the recovery-effect multipliers have stabilised since 2014, and prevalence has been near its simulation high for several years. What changes in 2017 to give B2 its 68.6% dominance score is that tightness jumps (+0.107). At the same time, there is an improvement in gender equality seen in more egalitarian norms and parity in working hours. This reflects the ongoing reduction in women's inactivity pushing employment up. The B2 mechanism is firing as labour market tightness is rising. Alongside B2, the norms–mental health reinforcing loop reaches 24.9%: once insecure work has degraded recovery, the otherwise masked norms-MH channel becomes briefly visible.

Appendix Q Structural decomposition

This appendix provides the technical detail behind the structural decomposition in Section 7.3.

- It reports the recovery-rate experiments (Decomposition D4): it equalises the implied baseline recovery rate in each direction — women given men's rate (D4a), men given women's (D4b) — and adds a run that also equalises starting stocks (D4c), isolating the residual gap attributable to path dependence and showing that the rate already subsumes the effect of the starting stocks.
- It then validates those implied rates against recovery rates derived from Scottish Health Survey transitions, confirming that the gender asymmetry (men recovering faster than women) is empirically grounded rather than a calibration artefact.
- It also reports the sensitivity of the 2019 gap to the shape of the prevalence-on-recovery function, tested across four specifications spanning linear to concave for both genders; the gap moves by under 1.5%.
- Finally, it reports the equal-starting-conditions tests, which recalibrate the recovery rate at equalised starting stocks and confirm that the structural mechanisms alone do not generate a gender gap.

Q1 Recovery rate scenarios

To understand what the implied recovery rate represents, it is important to consider the calibration logic. In 1992, the model must start in equilibrium i.e. the incidence and recovery must balance exactly, so that any subsequent change in prevalence is driven by the feedback mechanisms, not by starting momentum. To achieve this balance, the recovery rate was set to whatever value makes the equation work, given all the structural pressures (gender inequality, sensitivity differences, exposure levels) that were also operating in 1992. The implied recovery rate is therefore not an independently measured quantity; it is the equilibrium-preserving residual that absorbs every structural pressure the model cannot attribute to a specific mechanism. For women, this residual rate is lower (0.348) than for men (0.410) because women face more structural pressures that the equilibrium condition must accommodate. The rate difference captures the accumulated consequence of every structural disadvantage operating at the starting point, including those the model represents explicitly and those it does not.

Two experiments were conducted to equalise the baseline recovery rates: (1) from 1995, women were given men’s recovery rate (D4a), and (2) from 1995, men were given women’s recovery rate (D4b).
The direction of equalisation does not materially affect the result: giving women men's higher rate produces a gap of 0.002; giving men women's lower rate produces a gap of 0.003. The direction of equalisation matters, however, for overall population health. When women's recovery rate is raised to match men's, the gap closes through improvement — women's prevalence drops from 0.184 to 0.161 while men remain at 0.159, reducing total population burden. When men's recovery rate is lowered to match women's, the gap closes through deterioration — men's prevalence rises from 0.159 to 0.181 while women remain at 0.184, increasing total population burden.

Table Q-1: Residual estimation decompositions Contrasts scenarios where the baseline recovery rates are equalised either towards men's or women's implied rates. It isolates the residual gap attributable to path dependence and calibration limits. Prevalence of poor mental health among men and women, and the gender gap in prevalence in 2019.

Decomposition	Men	Women	Gap
D3 (reference)	0.1585	0.1838	0.025
D3 + women given men's rate (D4a)	0.1588	0.1607	0.002
D3 + men given women's rate (D4b)	0.1812	0.1836	0.003

In the second scenario (D4b), women's trajectory is identical to D3 throughout the simulation as expected, since no women's parameters were modified. The gap closes entirely through men's deterioration: men's prevalence rises from 0.159 to 0.181. This indicates that the test is internally consistent and that the gap reduction reflects the rate difference.

An additional simulation run (D4c) was conducted to assess whether starting stocks contribute independently of the implied recovery rate. In addition to D3 conditions, women were given men's recovery rate from 1995 and men's starting stocks of good and poor mental health. This resulted in a gap of 0.002 (Men: 0.165; Women: 0.163), which is identical to the gap when only the recovery rate was equalised. This demonstrates that the implied recovery rate subsumes the effect of different starting stocks: the recovery rate was derived from those stocks during calibration, so overriding the rate already neutralises the stocks' influence. This means the 30% residual attributed to the implied recovery rate does not contain an independent starting stocks component; the starting stocks and the recovery rate are two expressions of the same underlying path dependence.
The estimated recovery rate adjusts for net effect of *every* structural pressure operating in 1992: the modelled factors (gender inequality on incidence, recovery sensitivity differences), the different starting stocks, differential exposure, *and* any gender differences in recovery or incidence that the model does not explicitly represent. This means that the 30% residual includes the effect of (1) path dependence: women started with a higher prevalence (because structural mechanisms had been operating for decades and this history is embedded in the starting stocks and, through the calibration, in the recovery rate), and (2) unmodelled factors.

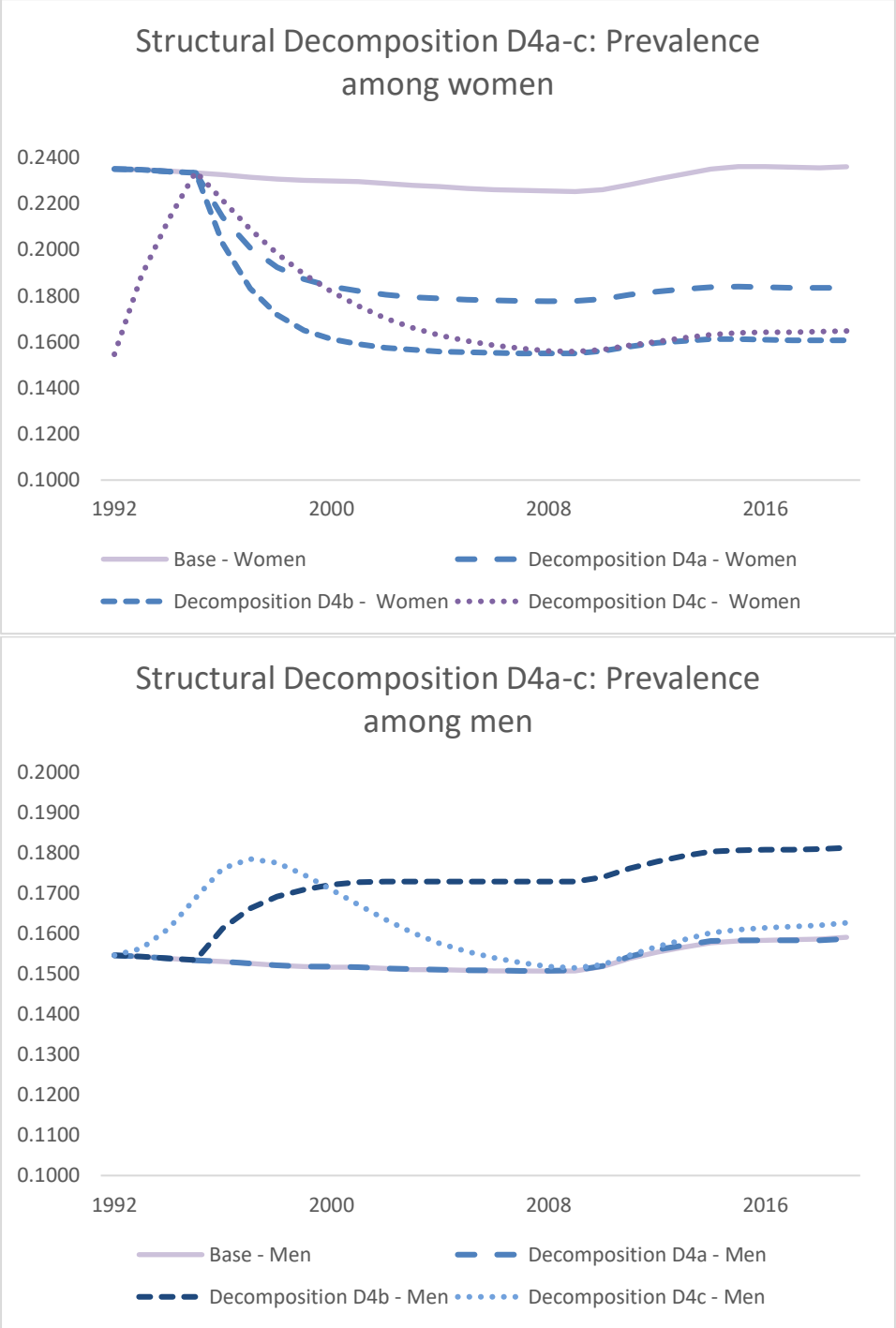


Figure Q-1: Trajectories of poor mental health Visualises the simulated trajectories of poor mental health for both men and women across additional D4 structural decompositions (D4a-D4c)

Q2 Validating the implied recovery rate

The finding that the implied recovery rate carries 30% of the gender gap raises a methodological question: is the gender difference in this rate an artefact of the calibration procedure, or does it reflect an empirically grounded asymmetry? The implied baseline recovery rate is 0.348 for women and 0.410 for men. This can be compared against data-driven recovery rates calculated from Scottish Health Survey transitions. When observed recovery rates are adjusted for the model's effect multipliers to isolate the baseline component, the resulting rates are:

Table Q-2: Comparing recovery rates Contrasts the model's data and equilibrium-adjusted implied recovery rates against only data-adjusted recovery rates. It shows that the empirical direction of gender asymmetry (men recovering faster than women) is supported by data.

	Data-adjusted (1992)	Data-adjusted (average 1992–94)	Data-adjusted for equilibrium
Men	0.564	0.477	0.410
Women	0.505	0.428	0.348
Male/female ratio	1.12	1.11	1.18

The implied rates are lower than the data-adjusted rates for both genders, which is expected because prevalence was declining in the early 1990s. The gender ratio from the implied rates (1.18) is modestly higher than the data-driven ratio (1.11–1.12). This divergence is attributable to the implied rate absorbing structural pressures beyond those captured in the recovery data.

The direction and approximate magnitude of the gender difference in recovery rates are empirically supported. Men recover faster than women in both the data and the calibration. The implied rates amplify this difference slightly due to the equilibrium constraint, but they do not create an asymmetry absent from the data.

Q3 Sensitivity of gender asymmetry in effect of prevalence on recovery

The model specifies that men's recovery capacity follows a concave function of prevalence (resilient until a high threshold, then declining sharply), while women's follows a linear function (each unit increase in prevalence degrades recovery proportionally). This asymmetry was identified through calibration; sensitivity tests are needed to confirm that the findings do not depend on the precise functional specification.

Four specifications are tested, spanning the full range from linear to concave for both genders. The gender gap is effectively unchanged across all specifications. The maximum variation in the gap is 0.0011 i.e. less than 1.5% of the base case gap. This means that the gender difference does not depend on the shape of the effect of prevalence on recovery rate. Whether this function is linear, mildly concave, or strongly concave for either gender, the gender gap persists at essentially the same magnitude.

Table Q-3: Sensitivity analysis specification and results for effect of prevalence on recovery Displays the 2019 prevalence gaps resulting from alternative model specifications of the effect of prevalence on recovery table function. It shows that the gender gap is robust and independent of the specific shape of the prevalence-on-recovery curve. Prevalence of poor mental health among men and women, and the gender gap in prevalence in 2019.

Specification	Men	Women	Gap
Base (women linear, men concave)	0.1591	0.2361	0.0770
Women mild concave (men concave)	0.1589	0.2349	0.0760
Women = men's concave	0.1589	0.2348	0.0759

Specification	Men	Women	Gap
Men changed to linear (women linear)	0.1603	0.2366	0.0763

The difference in the drain loop dominance between genders was large (59% for men vs 22% for women). The sensitivity analysis demonstrates that this difference is driven primarily by the *level* of men's and women's prevalence stocks (which determines where each gender sits on the recovery function) and by the implied recovery rate (which sets the baseline recovery capacity), rather than by the *shape* of the function itself. The concave-vs-linear distinction describes *how* the system responds at the margin, but the system's actual behaviour over the simulation period is governed by the starting position and calibrated baseline, not by the marginal response shape.

The gender inequality effect on incidence is not subjected to the same sensitivity analysis. The recovery threshold sensitivity analysis above tested whether the model's findings depend on the precise functional form of the effect of prevalence on recovery, a parameter whose shape (concave for men, linear for women) was determined through calibration and could plausibly take other forms. The gender inequality effect on incidence is structurally different. It is a composite variable derived from the gender pay gap and the unpaid work burden, translated into an incidence multiplier through a table function. Testing alternative specifications of this table function would require recalculating the baseline onset rate (back-calculated from 1992 data using the table function's multiplier values at that point) and then recalibrating the implied recovery rate to restore equilibrium. This would produce a different model, not a robustness test of the same model. The recovery threshold analysis worked because changing the function shape did not disturb the calibration: the equilibrium absorbed the change, and the finding was that the gap is invariant to the shape (less than 1.5% variation). No equivalent invariance test is possible for the incidence pathway without fundamentally altering the model's calibration. The robust finding is qualitative: gender-specific incidence pressures are the dominant contributor to the gender gap, substantially exceeding the contribution of recovery sensitivity.

Q4 Corroborating path dependence: equal starting conditions

This section provides evidence that path dependence is a substantial component of that residual. If starting stocks are equalised and the recovery rate is recalibrated to maintain equilibrium, does a gender gap emerge from the structural mechanisms alone?

Where D4 held the starting stocks constant and overrode the recovery rate from 1995, these tests change the starting stocks from the outset and let the recovery rate adjust through recalibration. If both approaches show that the structural mechanisms alone do not generate a gender gap from equal starting conditions, then confidence in the path dependence interpretation is strengthened. Two complementary tests were conducted:

- Women's starting stocks are set equal to men's (lower) values. Women's baseline recovery rate is recalculated to maintain equilibrium at the new stock levels, given women's actual effect multipliers and exposure patterns.

- Men's starting stocks are set equal to women's (higher) values. Men's baseline recovery rate is recalculated to maintain equilibrium at the new stock levels, given men's actual effect multipliers and exposure patterns.

In both cases, all other model parameters remain unchanged. Equilibrium was verified by running each decomposition with no exogenous changes and confirming that stocks remained stable.

Table Q-4: Path dependence test results Demonstrates the negligible gaps that remain when men's and women's starting stocks are equalised and recovery rates are recalibrated. It shows that equal starting conditions remove the gender gap, highlighting that path dependence heavily drives the model's outcomes. Prevalence of poor mental health among men and women, and the gender gap in prevalence in 2019.

Decomposition	Men	Women	Gap
Base	0.159	0.236	0.077
Women’s equalised at men's level	0.161	0.159	–0.002
Men’s equalised at women's level	0.239	0.233	–0.006

In both tests, the gender gap effectively disappears. The residual differences (0.002 and 0.006) are negligible. The gap closes in both cases, but it closes through improvement only when women are given men’s (lower) starting values. The direction reverses in both cases: men's prevalence is marginally higher than women's, indicating that the structural mechanisms alone do not produce a female disadvantage from equal starting conditions. The slight reversal reflects the fact that the economic mechanisms in the model are not inherently gendered in direction. They amplify whichever gender starts at a higher baseline.

The equilibrium recalibration forces the recovery rate to compensate for all structural disadvantages at the starting stock level. The structural factors (gender inequality, differential sensitivity) are present, but when the recovery rate is recalibrated to absorb them at a new starting position, they maintain the level of prevalence without driving divergence. The consistency across both methodological approaches strengthens confidence that the 30% residual contains a substantial path dependence component. However, the model cannot distinguish path dependence from unmodelled factors, because the recalibration absorbs both.

Appendix R Counterfactual scenario analysis

This appendix supports the counterfactual analysis in Section 7.5.

- It reports the gender-pay-gap scenarios (C–F): freezing the gap at its 1992 level (C, D) widens the 2019 gender gap by about 10%, while pushing it lower (E, F) closes the gap by a similar margin — confirming that the historical narrowing was protective.
- It then reports the norms scenarios (G–I), showing how accelerated norms erosion (G) redistributes the burden from women to men in a deteriorating labour market, while more traditional starting norms (I) leave almost no gender difference.
- The combination scenarios (J, K) show that norms and the pay gap run through overlapping channels, so addressing one neutralises the marginal benefit of the other.
- Table R-1 supports the labour market scenarios by reporting the insecure work parameterisation sensitivity behind the 9% result, confirming it is unchanged under a heavier weight on insecure work and a steeper stress-onset curve.

R1 What if the gender pay gap were different? (Scenarios C, D, E and F)

Four scenarios test the gender pay gap. Scenarios C and D freeze the GPG at its 1992 level: C holds the gender pay gap at the 1992 constant value (0.3), turning off the endogenous multipliers operating on it; scenario D holds the exogenous baseline gender pay gap at its historical 1992 value, allowing the endogenous multipliers to operate and causing a slow decline to 0.28. Scenarios E and F set the GPG to lower values (E fixes the gender pay gap at 0.10; F fixes the baseline gender pay gap at 0.05, allowing a decline from 0.14 to 0.12).

Freezing the GPG at its 1992 level worsens outcomes. In Scenario C, the gap widens to 0.085 (+10.4%). Women's prevalence rises to 0.246 (+0.010 from base), while men's rises marginally to 0.161. In Scenario D, the results are almost identical (+9.1%). This suggests that the historical narrowing of the GPG, from 0.30 in 1992 to 0.15 by 2019, was protective. Without it, the women's stress index remains elevated at 0.60 rather than declining to 0.56, sustaining higher incidence pressure on women throughout the simulation.

Reducing the GPG further produces improvements, with both Scenarios E and F closing the gap in 2019 by approximately 10%. Men are essentially unaffected in both cases given that the GPG operates exclusively through the women's incidence channel.

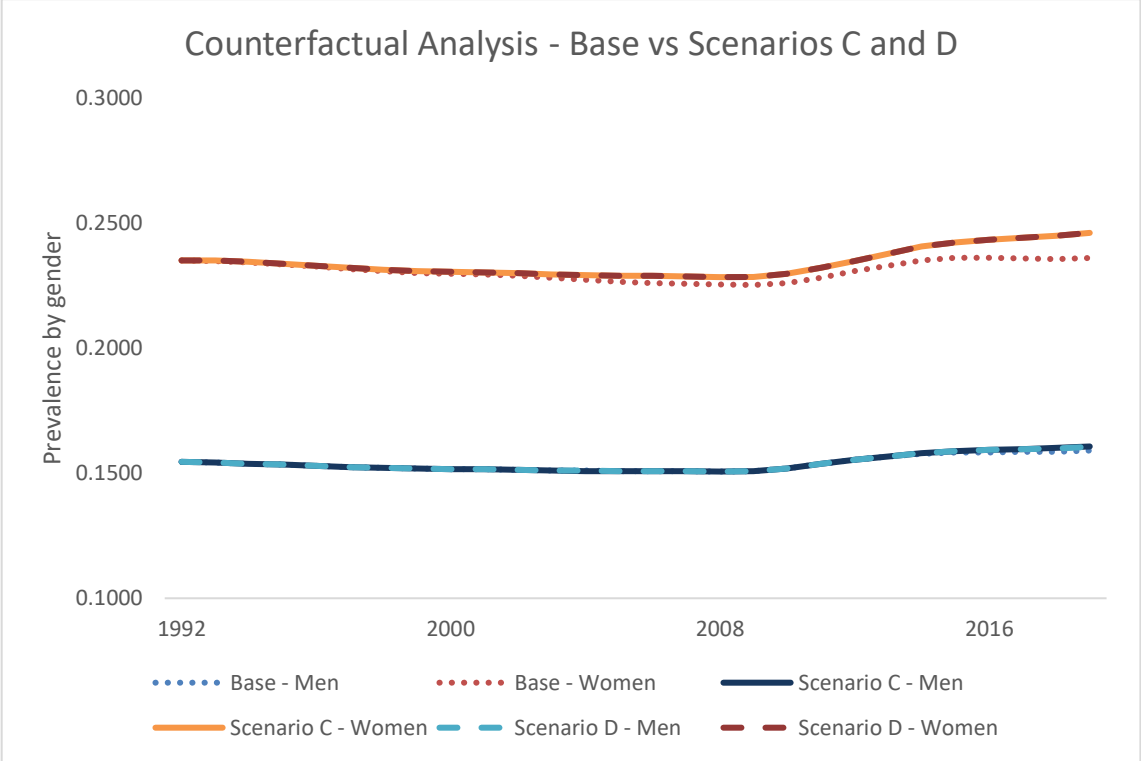


Figure R-1: Estimated prevalence trajectories under unchanging gender pay gap Compares Scenario C (gender pay gap held at its 1992 constant value of 0.30, with endogenous multipliers disabled) and Scenario D (gender pay gap baseline held at 0.30 while endogenous multipliers continue to operate, producing a slow decline to 0.28). Both produce a ~10% widening of the gap; the near-identical result shows the historical narrowing was protective

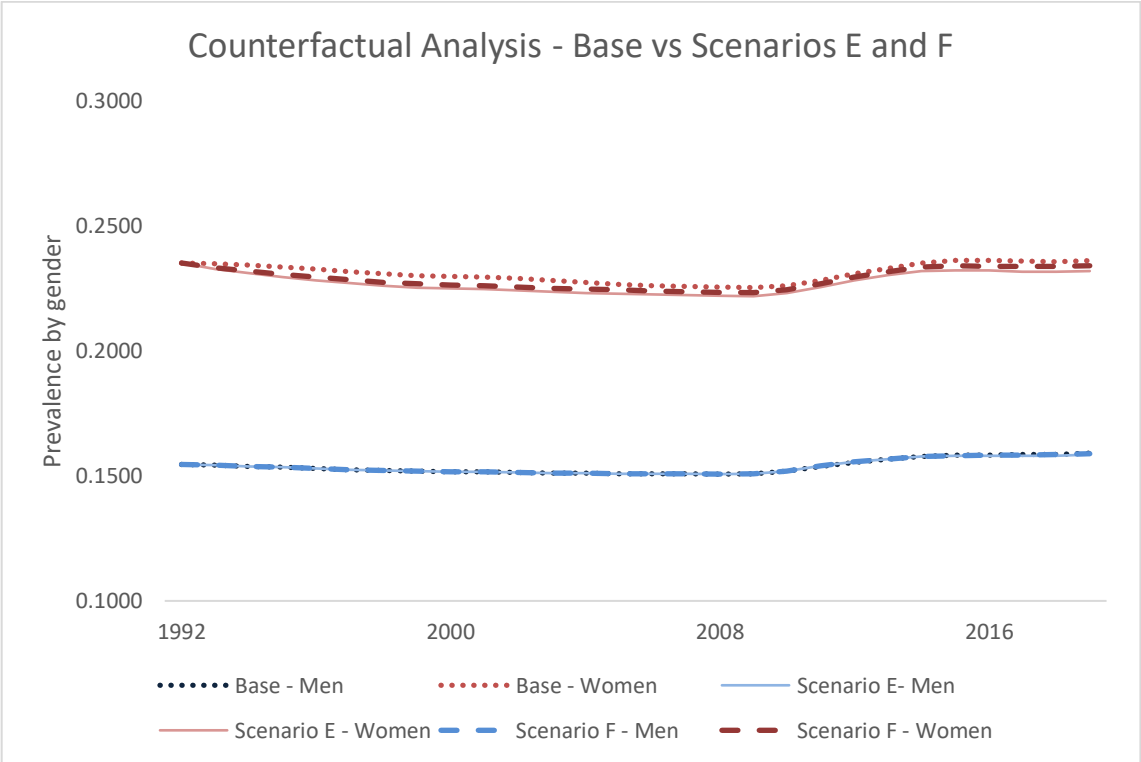


Figure R-2: Estimated prevalence trajectories under lower gender pay gap Compares Scenario E (gender pay gap fixed at 0.10) with Scenario F (baseline gender pay gap fixed at 0.05, allowing a decline from 0.14 to 0.12). Both close the 2019 gap by ~10%; men are essentially unaffected because the GPG operates exclusively through the women's incidence channel

R2 What if norms were different? (Scenarios G, H, and I)

Scenario G imposes a rate of norm change the model cannot generate endogenously: norms erode only through 45-year cohort replacement, and the model represents neither cultural backlash nor social-movement-driven shifts. It is therefore a structural sensitivity test of how the system responds to rapid normative change, not a forecast of a plausible future. The scenario reduces norms from 0.54 to approximately 0.22 by 1996, operationalised through a one-time STEP increase in norm erosion in 1995. This reduces the 2019 gender gap to 0.046, a reduction of about 40%. However, while women improve substantially (0.236 to 0.217), men's prevalence worsens (0.155 to 0.171). This redistributive effect is specific to a scenario where norms change rapidly while labour market quality is simultaneously deteriorating. It does not imply that gender equality harms men's mental health in general; rather, it shows that the benefits of equality depend on the structural quality of the labour market into which both genders are being equalised.

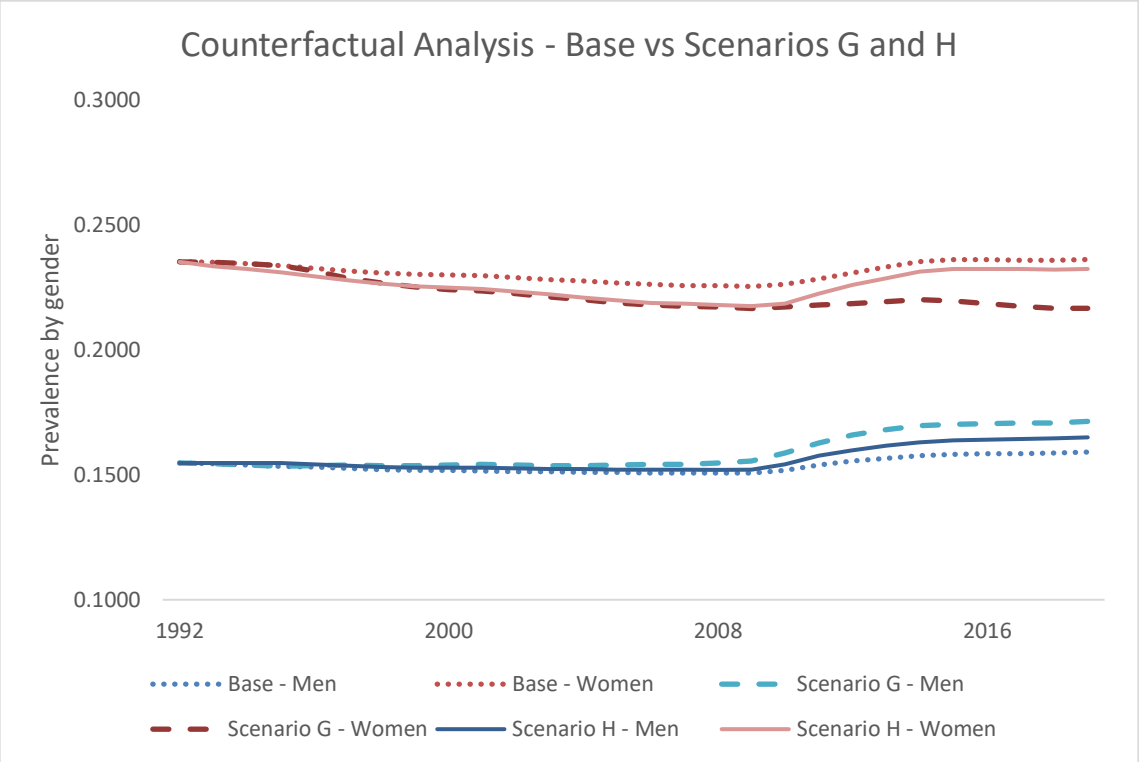


Figure R-3: Estimated prevalence trajectories under more egalitarian norms Models the effect of accelerated norms erosion. It highlights that a redistributive trap where rapid equality in a deteriorating labour market inadvertently pushes men into higher insecure work and worsens their prevalence

When norms equalise rapidly, hours parity rises from 0.74 to 0.95. In the base case, traditional norms sort men into more secure full-time employment: low hours parity means men's effective exposure to insecure work is well below women's. When norms equalise, men's exposure rises to 0.33 by 2019, approaching women's level of 0.38. In the base case, this convergence is gradual and harmless because the aggregate level of insecure work is moderate. In Scenario G, the convergence occurs rapidly *and* coincides with a labour market where aggregate insecure work is rising to 0.35 due to reduced labour market tightness as a result of higher employment among women. Men's exposure to insecure work erodes employment's positive effects on recovery. Women's

recovery effect, by contrast, remains at 0.85 because their exposure was already high in the base case; the equalisation moves men *toward* women's level. On the other hand, women benefit from the norms change because reduced inequality lowers the stress index from 0.61 to 0.50, reducing incidence pressure, which outweighs the recovery cost from continued high exposure.

The mechanism through which norms affect incidence operates through the perceived burden of unpaid work. Unpaid work burden is not operationalised as the objective division of labour but as its departure from parity weighted by the inverse of norms: as norms erode toward equality, the same actual share of unpaid work is experienced as a greater stressor because it increasingly conflicts with egalitarian expectations. In Scenario G, rapid norms erosion reduces the stress index despite this intensification effect because it simultaneously increases women's labour market participation and narrows the pay gap; the net effect on the composite is a reduction.

In Scenario H, norms start at 0.25 but remain higher than in Scenario G because the erosion rate is unchanged from the base case. The gap reduces as the same redistribution dynamic operates, but at lower intensity. Scenario I tests the opposite: norms starting at 0.75 (more traditional). With lower hours parity, men are sorted more strongly into secure work, but women’s exposure to insecure work does not increase as strongly since their employment drops under more traditional norms. This leads to a marginal improvement on the effect of employment on recovery among women. Unemployment also reduces among women leading to almost no difference in the economic stress affecting incidence. Stress due to gender inequality increases due to higher unpaid work and lower hours parity leading to higher gender pay gap. However, the gains in recovery outweigh these, leading to a lower prevalence by 2019. Men’s prevalence is even lower in 2019 (due to higher employment), resulting in almost no gender difference compared to the base case.

R3 What if both norms and gender pay gap were different? (Scenarios J and K)

Norms and the gender pay gap operate through overlapping channels (both feed the stress index), and reducing one makes the marginal effect of reducing the other negligible. When low starting norms and low fixed gender pay gap are simulated simultaneously (Scenario K), there is no difference in the gender gap in prevalence compared to when only norms are changed (H). There is some difference before the crisis, which disappears by 2019.

However, when high norms and unchanging (i.e. reducing) gender pay gap are simulated simultaneously (Scenario J), there is further worsening of women’s prevalence compared to high norms alone (Scenario I). In this case, there is almost no difference until the crisis, but a 0.006 higher prevalence for women and 0.001 higher prevalence for men in 2019. For women, this is due to higher incidence and the negative effect of prevalence on recovery kicking in, whereas for men, it is due to the latter.

R4 Sensitivity analysis to examine stable insecure work under different model specifications

To test whether the modest gap reduction depends on the specific parameterisation of the insecure work incidence pathway, two alternatives were

tested: increasing insecure work's weight in the economic stress composite (from 0.4 to 0.6), and steepening the curve relating economic stress to incidence onset. Neither materially changed the finding.

Table R-1: Sensitivity analysis for insecure work counterfactual analysis Reports results from alternative parameterisations for the effect of insecure work. Prevalence of poor mental health among men and women, and the gender gap in prevalence in 2019.

Scenario	Men	Women	Gap	Gap reduction
Base	0.1591	0.2361	0.077	
Scenario B	0.1496	0.2195	0.0699	9.2%
Base - revised weights	0.1612	0.2397	0.0785	
Scenario B + revised weights	0.1498	0.2189	0.0691	12.0%
Base - revised shape	0.1574	0.2336	0.0762	
Scenario B + revised shape	0.1467	0.2162	0.0695	8.8%

The shape of the curve for effect of economic stress on onset was varied (changed from a default of 1 to 3.5 in Stella Architect). The baseline onset and recovery rates were recalculated to adjust for the effect of multipliers in 1992 and initialise in equilibrium. Scenario B is the halting of rise in insecure work/the exogenous trend is held at the 1992 value.

The effect of insecure work is unchanged across the three parameterisations for both genders. Incidence levels differed, but the difference between the baseline and stable-insecure work scenarios was negligible for women (the effect of economic stress rose from 0.04 to 0.05). For men it was slightly larger (0.02 to 0.04–0.05), which drives the small difference in gap reduction. In the revised-weights parameterisation, the difference in the effect on incidence between the base and stable insecure work scenarios moved from 0.02 to 0.04 for men [versus 0.04 to 0.05 for women] — men benefitted more in relative terms.

Appendix S Single and combined levers

This appendix complements the optimisation presented in Section 7.6 by isolating what each lever does, to each gender, relative to no intervention — the distributional information the Pareto comparison aggregates away. Each lever was run on its own, and in the combinations that matter for the trade-off — the equality bundle (pay gap + childcare + norms), the full four-lever package, and insecure work + childcare — first in the base case (S4) and then across the nine macroeconomic futures, with every included lever set to its Section 7.6 optimiser bound (pay gap halved, childcare at 1.5×, norms reinforcement halved, insecure work at its 0.75 floor). Endpoint (2040) prevalence is reported for men and women, with the gap and the total. The total is the model's population-weighted prevalence.

- Table S-1 reports the base case (S4) outcomes at 2040 — men, women, gap and total, with the change relative to no intervention — for each lever and the three combinations, with the insecure work + childcare pair added as the final lever row.
- Table S-2 reports robustness checks at the base case. Childcare functional form compares the convex (base) and sigmoid specifications. Every single-lever sign is robust to the form (effects are larger under the sigmoid). The one form-sensitive result is the net effect of the combined package on the gap: it widens by a small amount under the convex form and narrows by a similar amount under the sigmoid. Recovery structure gives women men's insecure work threshold (R1) and men's prevalence-based recovery curve (R2). Both are inert at the base case — women already sit below their recovery threshold there, so its exact value is immaterial — and leave the trade-offs unchanged. Intervention intensity (panel C) confirms that the redistribution onto men is not an artefact of deploying the levers at full strength: at the mid-range childcare (1.25) and norms (0.75) settings the effect on men is smaller but still positive while the gap still narrows, so there is no intensity at which the gap narrows without some cost to men.
- Table S-3 does this across the nine macroeconomic futures.
- Table S-4 gives the base case (S4) diagnostics: each gender's insecure work exposure relative to 1992 (and whether it crosses 1.0, the recovery-degradation threshold), the multiplier this places on the recovery benefit of employment, the economic-stress contribution to incidence, and the women's stress index. At no intervention men's insecure work exposure sits above the 1992 level (ratio 1.105, so their recovery benefit is degraded to a multiplier of 0.74) while women's sits below it (0.971, multiplier 1.00). The genders therefore straddle the recovery-degradation threshold, and this straddle drives the base case trade-offs. Reducing insecure work pulls both exposures down and restores recovery — but it restores it for the degraded gender, men (multiplier 0.74 → 1.00), and largely bypasses women, so men gain more and the gap widens. The equality levers do the opposite through a different route: by raising women's working-hours parity they reallocate a fixed pool of insecure work toward men, lifting men's exposure (to 1.130 under the bundle) and pushing their recovery multiplier further down (0.74 → 0.68), so men's prevalence rises. The women's stress index moves only with the equality levers; the recovery multiplier moves only with insecure work; and reducing insecure work also lowers the economic-stress contribution to incidence for both genders — so insecure work acts on both pathways, while the equality levers act on women's incidence and (for childcare and

norms) on men's recovery via the reallocation. The gender pay gap moves men not at all, because the model carries no direct pay-gap-to-men feedback.

- The cross-future behaviour follows from where each gender sits on the recovery curve. Childcare (like the equality bundle) raises men's exposure in every future, so its direct effect on men is always mildly adverse; whether men's prevalence rises or falls then depends on how much recovery headroom men have. Where men sit on the degrading part of the curve, the reallocation degrades recovery further and men worsen — a pure redistribution from women to men. Where men's recovery is already floored, the reallocation can do no further harm. The single case in which an intervention raises women — norms alone under weak demand — is because the prevalence-recovery trap activates.
- Table S-5 gives the each gender's insecure work exposure relative to 1992 across the macroeconomic futures. In S4 and S7, where insecure work is lowest, women's effect of insecure work on recovery sits at 1 — employment's full recovery benefit is retained; for men this holds only in S7. Under the highest levels of insecure work (S1, S2, S3, S6) women's effect falls to 0, removing employment's recovery benefit entirely; for men this occurs only in S2, S3 and S6 (S1 leaves men just short, at 0.07, because their degradation threshold is higher). Near the floor (benign futures) there is little aggregate prevalence to remove, men's recovery has room to be degraded, and the equality levers can only redistribute. Far from the floor (deteriorated futures), the levers buffer the whole system toward the floor and the reduction reaches both genders.

Table S-1: Effect of different levers under baseline macroeconomic assumptions. Prevalence of poor mental health among men and women, and the total and gender gap in prevalence in 2019, and the change in these compared to no intervention.

Intervention	Men	Women	Gap	Total	Δ Men	Δ Women	Δ Gap	Δ Total
No intervention	0.1562	0.2168	0.0606	0.1872				
Gender pay gap (0.5)	0.1562	0.2163	0.0601	0.1869	0.0000	-0.0005	-0.0005	-0.0003
Childcare (1.5)	0.1568	0.2153	0.0585	0.1867	0.0006	-0.0015	-0.0021	-0.0005
Norms (0.5)	0.1568	0.2167	0.0599	0.1874	0.0006	-0.0001	-0.0007	0.0002
Insecure work (0.75)	0.1501	0.2142	0.0641	0.1828	-0.0061	-0.0026	0.0035	-0.0044
Equality bundle (GPG+CC+norms)	0.1574	0.2147	0.0573	0.1867	0.0012	-0.0021	-0.0033	-0.0005
All four levers	0.1502	0.2121	0.0619	0.1818	-0.0060	-0.0047	0.0013	-0.0054
Insecure work + childcare	0.1501	0.2126	0.0625	0.1821	-0.0061	-0.0042	0.0019	-0.0051

Table S-2: Robustness checks Prevalence of poor mental health among men and women, and the total and gender gap in prevalence in 2019, and the change in these compared to no intervention.

Intervention	Men	Women	Gap	Total	Δ Men	Δ Women	Δ Gap	Δ Total
A. Childcare functional form								
No intervention – convex	0.1562	0.2168	0.0606	0.1872				
Childcare 1.5 – convex	0.1568	0.2153	0.0585	0.1867	0.0006	-0.0015	-0.0021	-0.0005
Equality bundle – convex	0.1574	0.2147	0.0573	0.1867	0.0012	-0.0021	-0.0033	-0.0005
All four – convex	0.1502	0.2121	0.0619	0.1818	-0.0060	-0.0047	0.0013	-0.0054
No intervention – sigmoid	0.1564	0.2163	0.0599	0.1870				
Childcare 1.5 – sigmoid	0.1582	0.2120	0.0538	0.1857	0.0018	-0.0043	-0.0061	-0.0013
Equality bundle – sigmoid	0.1589	0.2115	0.0526	0.1857	0.0025	-0.0048	-0.0073	-0.0013
All four – sigmoid	0.1503	0.2090	0.0587	0.1803	-0.0061	-0.0073	-0.0012	-0.0067
B. Recovery structure (R1 = women given men's insecure work threshold; R2 = women given men's recovery curve)								
No intervention	0.1562	0.2168	0.0606	0.1872				
Insecure work	0.1501	0.2142	0.0641	0.1828	-0.0061	-0.0026	0.0035	-0.0044
Equality bundle	0.1574	0.2147	0.0573	0.1867	0.0012	-0.0021	-0.0033	-0.0005
All four levers	0.1502	0.2121	0.0619	0.1818	-0.0060	-0.0047	0.0013	-0.0054
No intervention – R1	0.1562	0.2168	0.0606	0.1872				
Insecure work – R1	0.1501	0.2142	0.0641	0.1828	-0.0061	-0.0026	0.0035	-0.0044

Equality bundle — R1	0.1574	0.2147	0.0573	0.1867	0.0012	-0.0021	-0.0033	-0.0005
All four — R1	0.1502	0.2121	0.0619	0.1818	-0.0060	-0.0047	0.0013	-0.0054
No intervention — R2	0.1562	0.2168	0.0606	0.1872				
Insecure work — R2	0.1501	0.2142	0.0641	0.1828	-0.0061	-0.0026	0.0035	-0.0044
Equality bundle — R2	0.1574	0.2147	0.0573	0.1867	0.0012	-0.0021	-0.0033	-0.0005
All four — R2	0.1502	0.2121	0.0619	0.1818	-0.0060	-0.0047	0.0013	-0.0054
C. Intervention intensities (mid-range; check for redistribution effect on men's prevalence)								
No intervention	0.1562	0.2168	0.0606	0.1872				
Childcare (1.5)	0.1568	0.2153	0.0585	0.1867	0.0006	-0.0015	-0.0021	-0.0005
Childcare (1.25)	0.1565	0.2161	0.0596	0.1870	0.0003	-0.0007	-0.0010	-0.0002
Norms (0.5)	0.1568	0.2167	0.0599	0.1874	0.0006	-0.0001	-0.0007	0.0002
Norms (0.75)	0.1565	0.2167	0.0602	0.1873	0.0003	-0.0001	-0.0004	0.0001

Table S-3: Effect of different levers across the nine macroeconomic futures. Prevalence of poor mental health among men and women, and the total and gender gap in prevalence in 2019, and the change in these compared to no intervention.

Scenario	Labour demand scalar	Insecure work rise	Intervention	Men	Women	Gap	Total	Δ Men	Δ Women	Δ Gap	Δ Total	% gap	%total
S1	0.90	none	No intervention	0.1982	0.2951	0.0969	0.2477						
S1	0.90	none	Gender pay gap	0.1970	0.2926	0.0956	0.2458	-0.0012	-0.0025	-0.0013	-0.0019	-1%	-1%
S1	0.90	none	Childcare	0.1964	0.2885	0.0921	0.2435	-0.0018	-0.0066	-0.0048	-0.0042	-5%	-2%
S1	0.90	none	Norms	0.1998	0.2959	0.0961	0.2488	0.0016	0.0008	-0.0008	0.0011	-1%	0%
S1	0.90	none	Insecure work	0.1543	0.2179	0.0636	0.1868	-0.0439	-0.0772	-0.0333	-0.0609	-34%	-25%
S1	0.90	none	Equality bundle	0.1966	0.2864	0.0898	0.2425	-0.0016	-0.0087	-0.0071	-0.0052	-7%	-2%
S1	0.90	none	All four levers	0.1554	0.2158	0.0604	0.1862	-0.0428	-0.0793	-0.0365	-0.0615	-38%	-25%
S1	0.90	none	Insecure work + childcare	0.1549	0.2163	0.0614	0.1863	-0.0433	-0.0788	-0.0355	-0.0614	-37%	-25%
S2	0.90	moderate	No intervention	0.2232	0.3251	0.1019	0.2753						
S2	0.90	moderate	Gender pay gap	0.2201	0.3210	0.1009	0.2716	-0.0031	-0.0041	-0.0010	-0.0037	-1%	-1%
S2	0.90	moderate	Childcare	0.2140	0.3123	0.0983	0.2642	-0.0092	-0.0128	-0.0036	-0.0111	-4%	-4%
S2	0.90	moderate	Norms	0.2227	0.3243	0.1016	0.2746	-0.0005	-0.0008	-0.0003	-0.0007	0%	0%
S2	0.90	moderate	Insecure work	0.1622	0.2250	0.0628	0.1943	-0.0610	-0.1001	-0.0391	-0.0810	-38%	-29%
S2	0.90	moderate	Equality bundle	0.2117	0.3085	0.0968	0.2612	-0.0115	-0.0166	-0.0051	-0.0141	-5%	-5%
S2	0.90	moderate	All four levers	0.1635	0.2216	0.0581	0.1932	-0.0597	-0.1035	-0.0438	-0.0821	-43%	-30%

S2	0.90	moderate	Insecure work + childcare	0.1629	0.2226	0.0597	0.1934	-0.0603	-0.1025	-0.0422	-0.0819	-41%	-30%
S3	0.90	high	No intervention	0.2527	0.3542	0.1015	0.3046						
S3	0.90	high	Gender pay gap	0.2481	0.3494	0.1013	0.2998	-0.0046	-0.0048	-0.0002	-0.0048	0%	-2%
S3	0.90	high	Childcare	0.2354	0.3361	0.1007	0.2869	-0.0173	-0.0181	-0.0008	-0.0177	-1%	-6%
S3	0.90	high	Norms	0.2520	0.3532	0.1012	0.3037	-0.0007	-0.0010	-0.0003	-0.0009	0%	0%
S3	0.90	high	Insecure work	0.1775	0.2571	0.0796	0.2182	-0.0752	-0.0971	-0.0219	-0.0864	-22%	-28%
S3	0.90	high	Equality bundle	0.2306	0.3306	0.1000	0.2817	-0.0221	-0.0236	-0.0015	-0.0229	-1%	-8%
S3	0.90	high	All four levers	0.1776	0.2505	0.0729	0.2149	-0.0751	-0.1037	-0.0286	-0.0897	-28%	-29%
S3	0.90	high	Insecure work + childcare	0.1772	0.2520	0.0748	0.2154	-0.0755	-0.1022	-0.0267	-0.0892	-26%	-29%
S4	1.00	none	No intervention	0.1562	0.2168	0.0606	0.1872						
S4	1.00	none	Gender pay gap	0.1562	0.2163	0.0601	0.1869	0.0000	-0.0005	-0.0005	-0.0003	-1%	0%
S4	1.00	none	Childcare	0.1568	0.2153	0.0585	0.1867	0.0006	-0.0015	-0.0021	-0.0005	-3%	0%
S4	1.00	none	Norms	0.1568	0.2167	0.0599	0.1874	0.0006	-0.0001	-0.0007	0.0002	-1%	0%
S4	1.00	none	Insecure work	0.1501	0.2142	0.0641	0.1828	-0.0061	-0.0026	0.0035	-0.0044	6%	-2%
S4	1.00	none	Equality bundle	0.1574	0.2147	0.0573	0.1867	0.0012	-0.0021	-0.0033	-0.0005	-5%	0%
S4	1.00	none	All four levers	0.1502	0.2121	0.0619	0.1818	-0.0060	-0.0047	0.0013	-0.0054	2%	-3%
S4	1.00	none	Insecure work + childcare	0.1501	0.2126	0.0625	0.1821	-0.0061	-0.0042	0.0019	-0.0051	3%	-3%
S5	1.00	moderate	No intervention	0.1670	0.2363	0.0693	0.2024						
S5	1.00	moderate	Gender pay gap	0.1668	0.2352	0.0684	0.2018	-0.0002	-0.0011	-0.0009	-0.0006	-1%	0%
S5	1.00	moderate	Childcare	0.1671	0.2319	0.0648	0.2002	0.0001	-0.0044	-0.0045	-0.0022	-6%	-1%
S5	1.00	moderate	Norms	0.1677	0.2355	0.0678	0.2023	0.0007	-0.0008	-0.0015	-0.0001	-2%	0%
S5	1.00	moderate	Insecure work	0.1508	0.2154	0.0646	0.1838	-0.0162	-0.0209	-0.0047	-0.0186	-7%	-9%
S5	1.00	moderate	Equality bundle	0.1677	0.2305	0.0628	0.1998	0.0007	-0.0058	-0.0065	-0.0026	-9%	-1%
S5	1.00	moderate	All four levers	0.1510	0.2133	0.0623	0.1828	-0.0160	-0.0230	-0.0070	-0.0196	-10%	-10%
S5	1.00	moderate	Insecure work + childcare	0.1509	0.2138	0.0629	0.1830	-0.0161	-0.0225	-0.0064	-0.0194	-9%	-10%
S6	1.00	high	No intervention	0.1903	0.2799	0.0896	0.2361						
S6	1.00	high	Gender pay gap	0.1896	0.2780	0.0884	0.2347	-0.0007	-0.0019	-0.0012	-0.0014	-1%	-1%
S6	1.00	high	Childcare	0.1882	0.2737	0.0855	0.2319	-0.0021	-0.0062	-0.0041	-0.0042	-5%	-2%
S6	1.00	high	Norms	0.1905	0.2798	0.0893	0.2361	0.0002	-0.0001	-0.0003	0.0000	0%	0%
S6	1.00	high	Insecure work	0.1550	0.2166	0.0616	0.1865	-0.0353	-0.0633	-0.0280	-0.0496	-31%	-21%
S6	1.00	high	Equality bundle	0.1878	0.2718	0.0840	0.2307	-0.0025	-0.0081	-0.0056	-0.0054	-6%	-2%

S6	1.00	high	All four levers	0.1562	0.2144	0.0582	0.1859	-0.0341	-0.0655	-0.0314	-0.0502	-35%	-21%
S6	1.00	high	Insecure work + childcare	0.1556	0.2150	0.0594	0.1860	-0.0347	-0.0649	-0.0302	-0.0501	-34%	-21%
S7	1.05	none	No intervention	0.1505	0.2150	0.0645	0.1834						
S7	1.05	none	Gender pay gap	0.1505	0.2145	0.0640	0.1832	0.0000	-0.0005	-0.0005	-0.0002	-1%	0%
S7	1.05	none	Childcare	0.1505	0.2135	0.0630	0.1827	0.0000	-0.0015	-0.0015	-0.0007	-2%	0%
S7	1.05	none	Norms	0.1505	0.2149	0.0644	0.1834	0.0000	-0.0001	-0.0001	0.0000	0%	0%
S7	1.05	none	Insecure work	0.1490	0.2127	0.0637	0.1815	-0.0015	-0.0023	-0.0008	-0.0019	-1%	-1%
S7	1.05	none	Equality bundle	0.1509	0.2129	0.0620	0.1826	0.0004	-0.0021	-0.0025	-0.0008	-4%	0%
S7	1.05	none	All four levers	0.1491	0.2105	0.0614	0.1805	-0.0014	-0.0045	-0.0031	-0.0029	-5%	-2%
S7	1.05	none	Insecure work + childcare	0.1490	0.2111	0.0621	0.1808	-0.0015	-0.0039	-0.0024	-0.0026	-4%	-1%
S8	1.05	moderate	No intervention	0.1577	0.2174	0.0597	0.1882						
S8	1.05	moderate	Gender pay gap	0.1577	0.2168	0.0591	0.1879	0.0000	-0.0006	-0.0006	-0.0003	-1%	0%
S8	1.05	moderate	Childcare	0.1584	0.2152	0.0568	0.1874	0.0007	-0.0022	-0.0029	-0.0008	-5%	0%
S8	1.05	moderate	Norms	0.1583	0.2169	0.0586	0.1882	0.0006	-0.0005	-0.0011	0.0000	-2%	0%
S8	1.05	moderate	Insecure work	0.1497	0.2137	0.0640	0.1824	-0.0080	-0.0037	0.0043	-0.0058	7%	-3%
S8	1.05	moderate	Equality bundle	0.1590	0.2144	0.0554	0.1873	0.0013	-0.0030	-0.0043	-0.0009	-7%	0%
S8	1.05	moderate	All four levers	0.1498	0.2116	0.0618	0.1814	-0.0079	-0.0058	0.0021	-0.0068	4%	-4%
S8	1.05	moderate	Insecure work + childcare	0.1497	0.2122	0.0625	0.1816	-0.0080	-0.0052	0.0028	-0.0066	5%	-4%
S9	1.05	high	No intervention	0.1680	0.2387	0.0707	0.2041						
S9	1.05	high	Gender pay gap	0.1678	0.2378	0.0700	0.2035	-0.0002	-0.0009	-0.0007	-0.0006	-1%	0%
S9	1.05	high	Childcare	0.1682	0.2345	0.0663	0.2021	0.0002	-0.0042	-0.0044	-0.0020	-6%	-1%
S9	1.05	high	Norms	0.1687	0.2379	0.0692	0.2041	0.0007	-0.0008	-0.0015	0.0000	-2%	0%
S9	1.05	high	Insecure work	0.1503	0.2148	0.0645	0.1833	-0.0177	-0.0239	-0.0062	-0.0208	-9%	-10%
S9	1.05	high	Equality bundle	0.1689	0.2331	0.0642	0.2017	0.0009	-0.0056	-0.0065	-0.0024	-9%	-1%
S9	1.05	high	All four levers	0.1506	0.2127	0.0621	0.1823	-0.0174	-0.0260	-0.0086	-0.0218	-12%	-11%
S9	1.05	high	Insecure work + childcare	0.1504	0.2133	0.0629	0.1825	-0.0176	-0.0254	-0.0078	-0.0216	-11%	-11%

Table S-4: Base case (S4) mechanism diagnostics at 2040 Exposure is each gender's 2040 insecure work level relative to 1992; recovery degrades once it exceeds 1.0. Recovery multiplier is the effect of insecure work on the recovery benefit of employment (1 = benefit fully retained, 0 = fully removed). Men begin above the threshold (exposure 1.105, multiplier 0.74) and women below it (0.971, multiplier 1.00).

Run	Exposure (M)	Exposure (W)	Recovery mult. (M)	Recovery mult. (W)	Econ. stress (M)	Econ. stress (W)	Women's stress index
No intervention	1.1050	0.9711	0.7375	1.0000	0.1181	0.1524	0.5281
Gender Pay Gap	1.1062	0.9719	0.7345	1.0000	0.1182	0.1525	0.5248
Childcare	1.1176	0.9645	0.7061	1.0000	0.1193	0.1515	0.5191
Norms	1.1124	0.9677	0.7190	1.0000	0.1196	0.1521	0.5279
Insecure work	0.8301	0.7295	1.0000	1.0000	0.0919	0.1171	0.5281
Equality bundle	1.1298	0.9608	0.6755	1.0000	0.1206	0.1511	0.5162
All four levers	0.8481	0.7212	1.0000	1.0000	0.0937	0.1160	0.5162
Insecure work + childcare	0.8389	0.7239	1.0000	1.0000	0.0927	0.1162	0.5191

Table S-5: Insecure work exposure (relative to 1992) and its effect on the recovery benefit of employment, by gender, across the nine futures (no intervention, 2040) Recovery effect = 1 only where insecure work is lowest (women: S4, S7; men: S7); it falls to 0 under the highest insecure work (women: S1–S3, S6; men: S2, S3, S6 — S1 leaves men at 0.07, just short, because their threshold is higher).

Scenario	Labour demand	Insecure work rise	Exposure (W)	Recovery effect (W)	Exposure (M)	Recovery effect (M)
S1	0.90	none	1.201	0.00	1.372	0.07
S2	0.90	moderate	1.395	0.00	1.594	0.00
S3	0.90	high	1.589	0.00	1.815	0.00
S4	1.00	none	0.972	1.00	1.106	0.74
S5	1.00	moderate	1.132	0.34	1.288	0.28

Scenario	Labour demand	Insecure work rise	Exposure (W)	Recovery effect (W)	Exposure (M)	Recovery effect (M)
S6	1.00	high	1.287	0.00	1.466	0.00
S7	1.05	none	0.872	1.00	0.991	1.00
S8	1.05	moderate	1.017	0.92	1.156	0.61
S9	1.05	high	1.161	0.20	1.320	0.20

Table S-6: Level of insecure work and effect of insecure work lever on gender gap in prevalence of poor mental health

Scenario	**Demand**	**Casualisation**	**Insecure work level in 2040**	**Effect of insecure work on recovery**	**Gap in 2040 prevalence when insecure work lever is deployed**
S7	1.05	stable	0.277	No change for men or women	Narrows slightly
S4 (low precarity)	1	stable	0.309	No change for women, degraded for men	widens
S8 (low precarity)	1.05	gradual	0.323	Slight degradation for women, degraded for men	widens
S5	1	gradual	0.359	Degraded for both	narrows
S9	1.05	rapid	0.369	Degraded for both	narrows
S1	0.9	stable	0.381	Degraded for both	narrows
S6	1	rapid	0.409	Degraded for both	narrows
S2	0.9	gradual	0.443	Degraded for both	narrows
S3	0.9	rapid	0.504	Degraded for both	narrows

Table S-7: Per-scenario optimisation (both objectives) across all four levers Prevalence of poor mental health (total and gender gap in 2019).

	Demand	Rise in insecure work	Insecure work scale	Childcare	GPG	Norms	Objective	Insecure work	CC	GPG	Norms	Total	Gap	Payoff
S1	0.9	none	0.75–0.83	1.5	0.5–0.52	0.5–1	combined	0.83	1.5	0.5	0.5	0.1896	0.0560	1.9362
							gap	0.83	1.5	0.5	0.5	0.1896	0.0560	1.9362
							total	0.75	1.5	0.5	1	0.1860	0.0610	1.9995
S2	0.9	moderate	0.75	1.5	0.5	0.5–1	combined	0.75	1.5	0.5	0.5	0.1932	0.0581	1.9905
							gap	0.75	1.5	0.5	0.5	0.1932	0.0581	1.9905
							total	0.75	1.5	0.5	0.99	0.1931	0.0591	2.0070
S3	0.9	high	0.75	1.5	0.5	0.5–1	combined	0.75	1.5	0.5	0.5	0.2155	0.0723	2.3433
							gap	0.75	1.5	0.5	0.5	0.2155	0.0723	2.3433
							total	0.75	1.5	0.5	1	0.2154	0.0734	2.3614
S4	1	none	0.75–1	1.5	0.5–0.52	0.5–0.92	combined	1	1.5	0.5	0.5	0.1867	0.0572	1.9412
							gap	1	1.5	0.5	0.5	0.1867	0.0572	1.9412
							total	0.75	1.5	0.5	0.84	0.1818	0.0619	1.9928
S5	1	moderate	0.75–0.9	1.49–1.5	0.5–0.52	0.5–0.93	combined	0.9	1.5	0.52	0.5	0.1881	0.0557	1.9235
							gap	0.9	1.5	0.52	0.5	0.1881	0.0557	1.9235
							total	0.75	1.5	0.5	0.87	0.1828	0.0624	2.0053
S6	1	high	0.75–0.79	1.48–1.5	0.5–0.54	0.5–0.99	combined	0.79	1.5	0.5	0.5	0.1879	0.0560	1.9274
							gap	0.79	1.5	0.5	0.5	0.1879	0.0560	1.9274
							total	0.75	1.5	0.53	0.98	0.1857	0.0589	1.9632
S7	1.05	none	0.75	1.5	0.5	0.5–0.78	combined	0.75	1.5	0.5	0.5	0.1805	0.0615	1.9778
							gap	0.75	1.5	0.5	0.5	0.1805	0.0615	1.9778
							total	0.75	1.5	0.5	0.76	0.1805	0.0615	1.9785
S8	1.05	moderate	0.75–1	1.49–1.5	0.5–0.51	0.5–0.88	combined	1	1.5	0.5	0.5	0.1873	0.0554	1.9144
							gap	1	1.5	0.5	0.5	0.1873	0.0554	1.9144
							total	0.75	1.5	0.5	0.81	0.1814	0.0619	1.9894
S9	1.05	high	0.75–0.88	1.49–1.5	0.5–0.53	0.5–0.98	combined	0.88	1.5	0.5	0.59	0.1872	0.0559	1.9219
							gap	0.88	1.5	0.5	0.59	0.1872	0.0559	1.9219
							total	0.75	1.5	0.51	0.95	0.1822	0.0623	2.0014

Appendix T Optimisation

This appendix provides the optimisation results behind Section 7.6.

- Table T-1 reports how the insecure work scale converges across the nine macroeconomic futures when minimising the gender gap, showing that it fluctuates between its bounds with macroeconomic conditions while the gender pay gap and childcare hold at their bounds throughout.
- Table T-2 reports the Pareto fronts across the nine futures, classifying each as divergent, convergent or unified and showing that any tension between the gap and total-prevalence objectives emerges only under moderate conditions.
- Table T-3 and Table T-4 do the same for the norms scale.
- Table T-5 reports the no-intervention baselines at 2040, which vary widely across the grid, driven almost entirely by the labour market channel.
- Table T-6 and Table T-7 compare those baselines against the optimised outcomes — for the gap and for total prevalence respectively — showing that interventions make their largest proportional difference where macroeconomic conditions are worst. The gap floor is 0.056 under gap-minimisation and 0.061 under total-minimisation.
- The optimiser configuration and the functional-form robustness checks for the childcare lever are reported in the final section.

Table T-1: Optimisation results across scenarios for insecure work scale for the objective of gap minimisation Outlines how the insecure work parameter fluctuates between its bounds based on underlying macroeconomic assumptions when attempting to minimise the gender gap.

Scenario	Value
S1 (weak demand, stable insecure work)	0.83
S2 (weak demand, moderately increasing in insecure work)	0.75 (min)
S3 (weak demand, continued trend of rising insecure work)	0.75 (min)
S4 (base case)	1.00 (max)
S5 (base demand, moderately increasing in insecure work)	0.90
S6 (base demand, continued trend of rising insecure work)	0.79
S7 (high demand, stable insecure work)	0.75 (min)
S8 (high demand, moderately increasing in insecure work)	1.00 (max)
S9 (high demand, continued trend of rising insecure work)	0.88

Table T-2: Pareto front for insecure work scale across different scenarios Details the distribution of the Pareto front trade-offs across diverse macroeconomic scenarios, classifying fronts as divergent, convergent, or unified. It highlights that the tension between gap and total prevalence objectives emerges largely under moderate macroeconomic conditions

Scenario	Macroeconomic worsening	Pareto front	Insecure work scale	Trade-off cost (gap)
S4 (base case)	Neither	Divergent	0.75–1.00	0.005
S8	Only insecure work	Divergent	0.75–1.00	0.006
S5	Only insecure work	Divergent	0.75–0.90	0.006
S9	Only insecure work	Divergent	0.75–0.88	0.006
S1	Only demand	Convergent	0.75–0.83	0.005
S6	Only insecure work	Convergent	0.75–0.79	0.002
S2	Both	Unified	0.75	0.000
S3	Both	Unified	0.75	0.000
S7	Neither	Unified	0.75	0.000

Insecure work scale shows the range of values across the 40 different Pareto combinations for each scenario and the trade-off cost gives the difference in gender gap in prevalence in 2040 between the optimal Pareto solution for total minimisation and the optimal Pareto solution for gap minimisation

Table T-3: Optimisation results across scenarios for norms scale for the objective of total minimisation Outlines how the norms parameter fluctuates between its bounds based on underlying macroeconomic assumptions when attempting to minimise the total prevalence.

Scenario	Value
S1 (weak demand, stable insecure work)	1 (max)
S2 (weak demand, moderately increasing in insecure work)	0.99
S3 (weak demand, continued trend of rising insecure work)	1 (max)
S4 (base case)	0.84
S5 (base demand, moderately increasing in insecure work)	0.87
S6 (base demand, continued trend of rising insecure work)	0.98
S7 (high demand, stable insecure work)	0.76
S8 (high demand, moderately increasing in insecure work)	0.81
S9 (high demand, continued trend of rising insecure work)	0.95

Table T-4: Pareto front for norms scale across different scenarios Details the distribution of the Pareto front trade-offs across diverse macroeconomic scenarios, classifying fronts as divergent, convergent, or unified. It highlights that the tension between gap and total prevalence objectives emerges largely under moderate macroeconomic conditions. The trade-off is difference between the totals when holding norms at 0.5 compared to the total-optimised value, holding the other three levers at their bounds.

Scenario	Macroeconomic worsening	Pareto front	norms scale	Trade-off cost (total)
S8	Only insecure work	Divergent	0.5-0.92	0.000
S5	Only insecure work	Divergent	0.5-0.88	0.000
S9	Only insecure work	Divergent	0.5-0.93	0.000
S1	Only demand	Divergent	0.5-0.98	0.000
S6	Only insecure work	Divergent	0.5-1	0.000
S2	Both	Divergent	0.5-0.99	0.000
S3	Both	Divergent	0.5-1	0.000
S7	Neither	Divergent	0.5-1	0.000
S8	Only insecure work	Divergent	0.5-0.78	0.000

Table T-5: No-intervention baselines at 2040 Simulates the 2040 prevalence, gaps, and insecure work metrics under nine different baseline macroeconomic futures without any intervention applied. It gives the baseline condition to use as comparison for optimisation results.

Scenario	Men	Women	Gap	Total	Insecure work	Demand
S7 (best)	0.150	0.215	0.065	0.183	0.277	0.896
S4 (base)	0.156	0.217	0.061	0.187	0.309	0.853
S8	0.158	0.217	0.060	0.188	0.323	0.896
S5	0.167	0.236	0.069	0.202	0.359	0.853
S9	0.168	0.239	0.071	0.204	0.369	0.896
S6	0.190	0.280	0.090	0.236	0.409	0.853
S1	0.198	0.295	0.097	0.247	0.381	0.768
S2	0.223	0.324	0.102	0.275	0.443	0.768
S3 (worst)	0.252	0.353	0.102	0.304	0.504	0.768

Table T-6: Difference in gender gap in prevalence under no-intervention compared to optimised interventions Compares the baseline gaps from Table 16 to the optimised gaps. It illustrates that interventions make the largest proportional impact when underlying macroeconomic conditions are the worst.

Scenario	Baseline gap	Optimised gap	Reduction
S1	0.097	0.056	42%
S2	0.102	0.058	43%
S6	0.090	0.056	37%
S3 (worst)	0.102	0.073	29%
S9	0.071	0.056	21%
S5	0.069	0.056	20%
S8	0.060	0.056	7%
S4 (base)	0.061	0.058	7%
S7 (best)	0.065	0.062	5%

Table T-7: Difference in total prevalence under no-intervention compared to optimised interventions Compares the baseline total prevalence from Table 16 to the optimised total prevalence. It illustrates that structural interventions can successfully buffer total prevalence against macroeconomic deterioration.

Scenario	Baseline total	Optimised total	Reduction
S3 (worst)	0.304	0.214	30%
S2	0.275	0.193	30%
S1	0.247	0.186	25%
S6	0.236	0.186	21%
S9	0.204	0.182	11%
S5	0.202	0.183	9%
S8	0.188	0.181	4%
S4 (base)	0.187	0.182	3%
S7 (best)	0.183	0.180	2%

T1 Sensitivity of the structural floor to the prevalence–recovery halving point

The structural floor identified in Section 7.6 operates partly through the endogenous prevalence–recovery loop, whose strength is governed by the halving point — the prevalence level, relative to 1992, at which the baseline recovery rate is halved. This parameter was set to 1.5 through constrained calibration but was not uniquely identified by the historical data: the calibration pressed against its search bounds, and 1.5 was retained on theoretical grounds as the value at which the loop is active without destabilising the system. Because both the floor and the post-2013 recovery vulnerability run through this loop, their robustness to the halving point was tested directly.

The model was re-run at halving points of 1.25, 1.40, 1.50 (calibrated) and 1.75, under two configurations at 2040: no intervention, and all four levers deployed at their optimised bounds (which reproduce the optimiser's converged solution). Rather than repeat the full nine-scenario grid at each threshold, three bracketing futures were used: the most adverse (S3: weak demand, high insecure work deterioration), the base case (S4: fixed demand, stable insecure work), and a benign future (S7: high demand, stable insecure work). These bracket the grid because the base and benign futures carry the lowest no-intervention prevalence and the adverse future the highest; intermediate futures lie between. The logic is confirmed by the results: where the base case is unchanged and the worst case improves (raising the threshold), all intermediate futures are bounded between them; and where even the base case destabilises (lowering the threshold to 1.25), every higher-prevalence future does so at least as strongly.

Table T-8: Total prevalence and gender gap at 2040 by halving point in baseline and optimised scenarios across select macroeconomic futures

Future	Halving Point	Baseline gap	Baseline total	Optimised gap	Optimised total
S3 (worst)	1.25	0.098	0.348	0.088	0.332
	1.50	0.102	0.305	0.073	0.215
	1.75	0.083	0.230	0.070	0.211
S4 (base)	1.25	0.088	0.313	0.089	0.307
	1.50	0.061	0.187	0.062	0.182
	1.75	0.061	0.187	0.062	0.182
S7 (best)	1.25	0.093	0.309	0.089	0.305
	1.50	0.065	0.183	0.061	0.181
	1.75	0.065	0.183	0.061	0.181
Floor	1.25	—	—	0.088	0.305
	1.50	—	—	0.061	0.180

Future	Halving Point	Baseline gap	Baseline total	Optimised gap	Optimised total
	1.75	—	—	0.061	0.180

Two findings follow. First, the floor is stable across halving points of 1.40 to 1.75. The minimum optimised total prevalence is 0.180 at each of 1.40, 1.50 and 1.75 (and the minimum optimised gap 0.061), reproducing the full-grid floor reported in Section 7.6; the base case and benign optimised outcomes are identical across this range, and the adverse-case optimised total moves by less than 0.01 (0.221 → 0.211). The calibrated value of 1.5 therefore sits on a stable plateau with clear margin, not at a critical edge.
Second, the system destabilises below this. At a halving point of 1.25 the prevalence–recovery loop becomes self-amplifying: no-intervention total prevalence rises by roughly two-thirds to ninety per cent across the three futures, and the levers lose their power, reducing the adverse-case total by only about five per cent (0.348 → 0.332) against roughly thirty-six per cent at 1.40 (0.344 → 0.221). The mechanism is a threshold effect: at halving points of 1.40 and above, the interventions can pull prevalence back below the recovery threshold and so escape the loop, whereas at 1.25 the threshold is low enough that even full intervention leaves prevalence above it and the loop continues to fire. The 1.25 regime places base case 2040 prevalence (≈0.31) above any level observed during the calibration period.

Taken together, the floor and the lever effectiveness on which Section 7.6 relies are robust to the halving point across the plausible range (1.40–1.75), with instability confined to an implausibly low threshold of 1.25. This both justifies the retention of 1.5 — the lowest value that activates the trap, as theory requires, while leaving the system well clear of instability — and confirms that the structural-floor finding does not depend on the precise value of a parameter the historical data could not uniquely identify.

T2 The theoretical floor

Section 7.6.5 reports a structural floor reached with the four levers held within their policy-plausible bounds. To establish that this floor is a property of the model's structure rather than of those bounds, the model was re-run with every lever driven to its model-theoretic extreme: insecure work, the gender pay gap and gender-norm reinforcement eliminated entirely, and childcare set to 3 — the point at which its effect on gender parity in working hours saturates, beyond which additional provision produces no further parity gain (
Table T-9). These settings lie well outside any plausible policy range and represent the furthest the model's structure can be pushed through these channels.

Three results follow. First, the theoretical floor is almost invariant to the macroeconomic future: across all nine scenarios, total prevalence settles at 0.171–0.173 and the gender gap at 0.051–0.052; in fact at just three distinct values, one for each level of labour demand: 0.1707 (high-demand S7–S9), 0.1714 (stable-demand S4–S6) and 0.1730 (weak-demand S1–S3). The insecure work trend (casualisation) axis disappears entirely, because eliminating insecure work removes the only channel through which it acts; the sole surviving cross-future variation is exogenous labour demand, which is not one of the four levers.

This band is far narrower than the 1.7-fold spread of the no-intervention baselines (0.183–0.305 total) and narrower even than the all-four-at-bounds deployment (0.180–0.215 total).

Second, the minimum the model can reach falls only marginally below the policy floor: total prevalence moves from 0.181 to 0.171 and the gap from 0.058 to 0.051, under 0.01 on each objective, despite the complete removal of every modelled stressor. (Where a future's policy optimum sits well above the floor — most visibly the worst case, S3, at 0.214 — the larger step to the theoretical floor reflects the bounded levers' limited reach in adverse conditions, consistent with the finding that the value of intervention is greatest where macroeconomic conditions are worst, not the floor itself moving.) The great majority of the floor — roughly 95% of the total prevalence and 92% of the gap — is therefore untouchable by any of the four levers, at any strength.

Third, the floor is a near-equilibrium property: extending the horizon from 2040 to 2060 lowers it by less than 0.001 on both metrics. What holds the floor in place is the residual identified in Section 7.3 — path dependence, the implied recovery-rate difference between women and men, and the accumulated disadvantage embedded in the 2019 starting position — none of which is reachable by interventions acting on insecure work, the pay gap, norms or childcare. The floor is thus a property of the model's calibration, not of the real-world system.

Table T-9: The theoretical (model-limit) floor across the nine macroeconomic futures
Prevalence of poor mental health at 2040 (and 2060) when all four levers are driven to their model extremes (insecure work, the gender pay gap and gender-norm reinforcement eliminated; childcare set to the saturation point of its working-hours-parity effect). Compared against each future's no-intervention baseline and to scenario where all four policy levers are deployed at their bounds.

	Total prevalence of poor mental health (2040 or 2060)				Gender gap in prevalence of poor mental health (2040 or 2060)			
Scenario	**No intervention**	**Levers at bound**	**Theoretical floor**	**Theoretical floor (2060)**	**No intervention**	**Levers at bound**	**Theoretical floor**	**Theoretical floor (2060)**
S7 (best)	0.183	0.181	**0.171**	0.170	0.065	0.061	**0.051**	0.050
S4 (base case)	0.187	0.182	**0.171**	0.171	0.061	0.062	**0.052**	0.050
S8	0.188	0.181	**0.171**	0.170	0.060	0.062	**0.051**	0.051
S5	0.202	0.183	**0.171**	0.171	0.069	0.062	**0.052**	0.051
S9	0.204	0.182	**0.171**	0.170	0.071	0.062	**0.051**	0.050
S6	0.236	0.186	**0.171**	0.171	0.090	0.058	**0.052**	0.051

	Total prevalence of poor mental health (2040 or 2060)				Gender gap in prevalence of poor mental health (2040 or 2060)			
Scenario	**No intervention**	**Levers at bound**	**Theoretical floor**	**Theoretical floor (2060)**	**No intervention**	**Levers at bound**	**Theoretical floor**	**Theoretical floor (2060)**
S1	0.248	0.186	**0.173**	0.173	0.097	0.060	**0.052**	0.051
S2	0.275	0.193	**0.173**	0.173	0.102	0.058	**0.052**	0.051
S3 (worst)	0.305	0.215	**0.173**	0.173	0.102	0.073	**0.052**	0.051
Range	0.183–0.305	0.181–0.215	**0.171–0.173**	0.170–0.173	0.060–0.102	0.058–0.073	**0.051–0.052**	0.050-0.051

T3 Robustness analysis

The childcare lever acts through an assumed functional form linking childcare availability to working-hours parity. Given the uncertainty of the functional form, the optimisation was re-run under four specifications of it — convex, sigmoid, concave and linear. Across all four, the optimal solution is unchanged.

Table T-10: Results for childcare functional-form check under base macroeconomic assumptions (stable demand and insecure work) The optimiser converges to the same parameter combination under all four forms.

Fraction of gap reduced by childcare	Labour market.insecure work scale	Gender norms.childcare scale	Gender norms.GPG scale	Gender norms.norms scale
Base case (convex)	0.75(min) to 1 (max)	1.5 (max)	0.5(min) to 0.52	0.5(min) to 0.92
concave	0.75(min) to 1 (max)	1.5 (max)	0.5(min) to 0.52	0.5(min) to 1 (max)
sigmoid	0.75(min) to 1 (max)	1.5 (max)	0.5(min) to 0.52	0.5(min) to 0.98
linear	0.75(min) to 1 (max)	1.5 (max)	0.5(min) to 0.51	0.5(min) to 0.91

Table T-11: Optimisation Configuration The cross-scenario run ("maximize average") maximises the average payoff across all nine macroeconomic futures at once rather than within each future separately, so a single parameter set must perform well under every assumption. Childcare and gender pay gap are set to their bounds.

Method	seed	cr	f	tolerance	pop_size	generations	crossover_type	k	recombine_type
Differential Evolution	123	0.9	0.6	0.00005	40	50	bin		rand

Payoff:	**Total prevalence**	**Gap in Prevalence**
Action	minimize	minimize
Kind	Outcome Optimization	Outcome Optimization
Element	Analysis.Total prev	Analysis.Gap in prev
Weight	5.35	16.47
Time Range	Time 2040	Time 2040
Parameter:	Labour market.insecure work scale	Gender norms.norms scale
min_value	0.75	0.5
max_value	1	1
increment	0.01	0.01
Solution	0.75 (min)	0.5 (min
Payoff	**Total prevalence**	**Gap in Prevalence**
	1.00438	1.02218

Glossary

Behaviour reproduction test	A validation test assessing whether the model can endogenously generate the historical patterns of observed behaviour.
Constrained calibration	A parameterisation approach in which unobservable parameters, or those for which no data exist, are estimated using calibration tools within bounds set by theory and available evidence.
Elasticity	The percentage change in one variable in response to a one percent change in another; used in this model to describe how labour market flows respond to Labour Market Tightness.
Endogenous variable	A variable determined within model boundary, arising from the system's own feedback structure rather than from external inputs.
Equifinality	The property where the same outcome can be reached in different ways. In calibration, this means that different combinations of parameter values can generate equivalent model behaviour.
Exogenous variable	A variable determined outside the model boundary: it influences model behaviour but is not itself affected by any variable within the model.
Extreme conditions tests	Tests that force the model inputs to theoretical extremes (e.g., zero population, zero demand) to ensure the equations remain mathematically robust.
Feedback loop	A closed chain of cause and effect in which a change in a variable propagates through the system and eventually feeds back to influence that same variable. Reinforcing loops amplify change; balancing loops resist it.
First- and third-order delay	A first-order delay is a simple smoothing structure in which the output converges on the input over a characteristic delay time; used here for the lag between labour market tightness and insecure work. A third-order delay is a more gradual structure, producing an S-shaped delay; used here for perceptual response to economic volatility.
Flow	A rate of change feeding into or out of a stock (e.g. the onset rate, the recovery rate). Flows have units of quantity per unit time.
Loop dominance	The analytical identification of which feedback loop is most powerfully driving the system's behaviour at a given time.

Path dependence	The property whereby the sequence of past events constraints the outcomes available later, so that early structural conditions become entrenched.
Sensitivity analysis	Tests examining how model behaviour changes across plausible range of uncertain parameter values.
Stock	An accumulating quantity, a level that can rise or fall over time but cannot change instantaneously (e.g. the prevalence of poor mental health, the prevalence of traditional gender norms).
Table function	A graphical (lookup) function specifying how one variable depends on another; used where the relationship is non-linear and cannot conveniently be expressed in closed form.
Theil's inequality statistics	A decomposition of simulation data error into bias, variance, and covariance components, used to assess whether residual error is systematic or unsystematic.

List of References

Allan, B. A., Autin, K. L., & Wilkins-Yel, K. G. (2021). Precarious work in the 21st century: A psychological perspective. *Journal of Vocational Behavior*, *126*, 103491. https://doi.org/10.1016/j.jvb.2020.103491

Almond, S., & Healey, A. (2003). Mental health and absence from work: New evidence from the UK Quarterly Labour Force Survey. *Work, Employment and Society*, *17*(4), 731-742. https://doi.org/10.1177/0950017003174007

Altintas, E., & Sullivan, O. (2016). Fifty years of change updated: Cross-national gender convergence in housework. *Demographic Research*, *35*, 455-470. https://doi.org/10.4054/DemRes.2016.35.16

Altunbuken, U., Bukowski, P., Machin, S., & Slaughter, H. (2022). *Power Plays: The shifting balance of employer and worker power in the UK labour market*. Resolution Foundation. https://economy2030.resolutionfoundation.org/reports/power-plays/

Amiri, S., & Behnezhad, S. (2019). Depression and risk of disability pension: A systematic review and meta-analysis. *The International Journal of Psychiatry in Medicine*. https://doi.org/10.1177/0091217419837412

Amiri, S., & Behnezhad, S. (2021). Depression symptoms and risk of sick leave: a systematic review and meta-analysis. *International archives of occupational and environmental health*, *94*(7), 1495-1512. https://doi.org/10.1007/s00420-021-01703-0

Andersen, D. L., Luna-Reyes, L. F., Diker, V. G., Black, L., Rich, E., & Andersen, D. F. (2012). The disconfirmatory interview as a strategy for the assessment of system dynamics models. *System Dynamics Review*, *28*(3), 255-275. https://doi.org/10.1002/sdr.1479

Andrew, A., Bandiera, O., Costa-Dias, M., & Landais, C. (2021). *Women and men at work*. Institute for Fiscal Studies. https://ifs.org.uk/publications/women-and-men-work

Andrew, A., Bandiera, O., Dias, M. C., & Landais, C. (2020). *The careers and time use of mothers and fathers*. Institute for Fiscal Studies. https://ifs.org.uk/publications/careers-and-time-use-mothers-and-fathers

Apostolopoulos, Y. (2020). Harnessing Complex Systems. In Y. Apostolopoulos, K. H. Lich, & M. K. Lemke (Eds.), *Complex Systems and Population Health* (pp. 283-298). Oxford University Press. https://doi.org/10.1093/oso/9780190880743.003.0020

Arias-de la Torre, J., Vilagut, G., Ronaldson, A., Serrano-Blanco, A., Martin, V., Peters, M., Valderas, J. M., Dregan, A., & Alonso, J. (2021). Prevalence and variability of current depressive disorder in 27 European countries: a population-based study. *Lancet Public Health*, *6*(10), e729-e738. https://doi.org/10.1016/S2468-2667(21)00047-5

Asper, M., Osika, W., Dalman, C., Pollanen, E., Simonsson, O., Flodin, P., Sidorchuk, A., Marchetti, L., Awil, F., Castro, R., & Niemi, M. E. (2022). Effects of the COVID-19 pandemic and previous pandemics, epidemics and economic crises on mental health: systematic review. *BJPSYCH OPEN*, *8*(6), Article e181. https://doi.org/10.1192/bjo.2022.587

Atkinson, J. A., Wells, R., Page, A., Dominello, A., Haines, M., & Wilson, A. (2015). Applications of system dynamics modelling to support health policy. *Public Health Research & Practice*, *25*(3), e2531531. https://doi.org/10.17061/phrp2531531

Aydiner-Avsar, N., & Piovani, C. (2019). The Gender Impact of Unemployment on Mental Health: A Micro Analysis for the United States. *Forum for Social*

Economics, *50*(4), 505-529. https://doi.org/10.1080/07360932.2018.1535991

Bahr, S., Batinic, B., & Collischon, M. (2022). Heterogeneities in the latent functions of employment: New findings from a large-scale German survey. *Frontiers in Psychology*, *13*, 909558. https://doi.org/10.3389/fpsyg.2022.909558

Bambra, C. (2011). Work, worklessness and the political economy of health inequalities. *Journal of Epidemiology and Community Health*, *65*(9), 746-750. https://doi.org/10.1136/jech.2009.102103

Banos, R. M., & Miragall, M. (2024). Gender Matters: A Critical Piece in Mental Health. *The Spanish Journal of Psychology*, *27*, e28. https://doi.org/10.1017/SJP.2024.29

Barbalat, G., & Franck, N. (2020). Ecological study of the association between mental illness with human development, income inequalities and unemployment across OECD countries. *BMJ open*, *10*(4), e035055. https://doi.org/10.1136/bmjopen-2019-035055

Barbrook-Johnson, P., & Penn, A. S. (2022). *Systems Mapping: How to build and use causal models of systems*. Palgrave Macmillan.

Barlas, Y., & Carpenter, S. (1990). Philosophical roots of model validation Two paradigms. *System Dynamics Review*, *6*(2), 148-166. https://doi.org/10.1002/sdr.4260060203

Barr, B., Kinderman, P., & Whitehead, M. (2015). Trends in mental health inequalities in England during a period of recession, austerity and welfare reform 2004 to 2013. *Social Science & Medicine*, *147*, 324-331. https://doi.org/10.1016/j.socscimed.2015.11.009

Barr, B. R. (2021). The mental health impact of recession and welfare reform in England between 2008 and 2013 University of Liverpool]. psyh.

Baxter, A. J., Tindall, M., Wickham, S., Marimpi, M., Brown, H., Munford, L., Sutton, M., Richiardi, M., Cheetham, M., Amo-Agyei, S., Taylor-Robinson, D., Bambra, C., Katikireddi, S. V., & Craig, P. (2026). Impacts of the Universal Credit welfare reform on well-being: a natural experiment study using UK population survey data. *BMJ Public Health*, *4*(2), e003762. https://doi.org/10.1136/bmjph-2025-003762

Beck, V., Warren, T., & Lyonette, C. (2024). Is Any Job Better Than No Job? Utilising Jahoda's Latent Deprivation Theory to Reconceptualise Underemployment. *Work, Employment and Society*, *39*(2), 404-425. https://doi.org/10.1177/09500170241254794

Bell, D. N. F., & Blanchflower, D. G. (2019). The well-being of the overemployed and the underemployed and the rise in depression in the UK. *Journal of Economic Behavior & Organization*, *161*, 180-196. https://doi.org/10.1016/j.jebo.2019.03.018

Bell, T., Watson, M., Sharp, D., Lyons, I., & Lewis, G. (2005). Factors associated with being a false positive on the General Health Questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, *40*(5), 402-407. https://doi.org/10.1007/s00127-005-0881-6

Benach, J., Vives, A., Amable, M., Vanroelen, C., Tarafa, G., & Muntaner, C. (2014). Precarious employment: understanding an emerging social determinant of health. *Annual Review of Public Health*, *35*, 229-253. https://doi.org/10.1146/annurev-publhealth-032013-182500

Bilodeau, J., Marchand, A., & Demers, A. (2020). Psychological distress inequality between employed men and women: A gendered exposure model. *SSM - Population Health*, *11*, 100626. https://doi.org/10.1016/j.ssmph.2020.100626

Birt, C., Cebula, C., Evans, J., Hay, D., & McKenzie, A. (2023). *Poverty in Scotland 2023*. J. R. Foundation. https://www.jrf.org.uk/poverty-in-scotland-2023

Bittman, M., England, P., Sayer, L., Folbre, N., & Matheson, G. (2003). When Does Gender Trump Money? Bargaining and Time in Household Work. *American Journal of Sociology*, *109*(1), 186-214.

Blair, C. (2024). *The gender pay gap in Scotland and the UK*. Retrieved 15/12 from https://spice-spotlight.scot/2024/08/01/the-gender-pay-gap-in-scotland-and-the-uk/

Blake, C., Rigby, B. P., Armstrong-Moore, R., Barbrook-Johnson, P., Gilbert, N., Hassannezhad, M., Meier, P., Mitchell, K. R., Penn, A., Readman, B., Simpson, S. A., White, M., & Moore, L. (2024). *Participatory systems mapping for population health research, policy and practice: Guidance on method choice and design*. University of Glasgow.

Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, *55*(3), 789-865. https://doi.org/10.1257/jel.20160995

Bloom, D. E., & Canning, D. (2008). *Population Health and Economic Growth*. The World Bank on behalf of the Commission on Growth and Development.

Blundell, R., Costas Dias, M., Meghir, C., & Shaw, J. (2016). *Female Labour Supply, Human Capital and Welfare Reform*. Institute for Fiscal Studies. https://ifs.org.uk/publications/female-labour-supply-human-capital-and-welfare-reform-0

Borjas, G. J. (2013). *Labor Economics* (Sixth ed.). McGraw-Hill.

Braveman, P. (2025). Health Inequalities, Disparities, Equity: What's in a Name? *American journal of public health*, *115*(7).

Braveman, P., & Gottlieb, L. (2014). The social determinants of health: it's time to consider the causes of the causes. *Public Health Reports*, *129* 19-31. https://doi.org/10.1177/00333549141291S206

Brett, M., & Macfarlane, L. (2022). *Women, Work and Wealth in Scotland's changing economy*. Scottish Women's Budget Group. https://www.swbg.org.uk/content/publications/Women-Work-Wealth-in-a-changing-economy-report.pdf

Brouwers, E. P. M. (2020). Social stigma is an underestimated contributing factor to unemployment in people with mental illness or mental health issues: position paper and future directions. *BMC Psychology*, *8*(1), 36. https://doi.org/10.1186/s40359-020-00399-0

Bryan, M. L., Bryce, A. M., & Roberts, J. (2021). The effect of mental and physical health problems on sickness absence. *The European Journal of Health Economics*, *22*(9), 1519-1533. https://doi.org/10.1007/s10198-021-01379-w

Bryan, M. L., Bryce, A. M., & Roberts, J. (2022). Dysfunctional presenteeism: Effects of physical and mental health on work performance. *The Manchester School*, *90*(4), 409-438. https://doi.org/10.1111/manc.12402

Bubonya, M., Cobb-Clark, D. A., & Wooden, M. (2017). Mental health and productivity at work: Does what you do matter? *Labour Economics*, *46*, 150-165. https://doi.org/10.1016/j.labeco.2017.05.001

Buckman, J. E. J., Saunders, R., Stott, J., Cohen, Z. D., Arundell, L. L., Eley, T. C., Hollon, S. D., Kendrick, T., Ambler, G., Watkins, E., Gilbody, S., Kessler, D., Wiles, N., Richards, D., Brabyn, S., Littlewood, E., DeRubeis, R. J., Lewis, G., & Pilling, S. (2022). Socioeconomic Indicators of Treatment Prognosis for Adults With Depression: A Systematic Review and

Individual Patient Data Meta-analysis. *JAMA Psychiatry*, *79*(5), 406-416. https://doi.org/10.1001/jamapsychiatry.2022.0100

Budig, M., & England, P. (2001). The Wage Penalty for Motherhood. *American Sociological Review*, *66*(2). https://doi.org/10.1177/000312240106600203

Butterworth, P., Leach, L. S., Pirkis, J., & Kelaher, M. (2012). Poor mental health influences risk and duration of unemployment: a prospective study. *Social Psychiatry and Psychiatric Epidemiology*, *47*(6), 1013-1021. https://doi.org/10.1007/s00127-011-0409-1

Butterworth, P., Leach, L. S., Strazdins, L., Olesen, S. C., Rodgers, B., & Broom, D. H. (2011). The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. *Occupational and environmental medicine*, *68*(11), 806-812. https://doi.org/10.1136/oem.2010.059030

Byrne, D. (1998). Complexity theory and the social sciences: An introduction. Routledge.

Cabezas-Rodriguez, A., Utzet, M., & Bacigalupe, A. (2021). Which are the intermediate determinants of gender inequalities in mental health?: A scoping review. *The International journal of social psychiatry*, *67*(8), 1005-1025. https://doi.org/10.1177/00207640211015708

Calligaris, S., Calvino, F., Reinhard, M., & Verlhac, R. (2023). Is there a trade-off between productivity and employment? A cross-country micro-to-macro study. OECD.

Campaner, R. (2011). Mechanistic causality and counterfactual-manipulative causality: recent insights from philosophy of science. *Journal of Epidemiology & Community Health*, *65*(12), 1070-1074. https://doi.org/10.1136/jech.2011.134205

Campbell, D., Green, M. J., Davies, N., Demou, E., Howe, L. D., Harrison, S., Smith, D. J., Howard, D. M., McIntosh, A. M., Munafò, M., & Katikireddi, S. V. (2022). Effects of depression on employment and social outcomes: a Mendelian randomisation study. *Journal of Epidemiology & Community Health*, *76*(6), 563-571. https://doi.org/10.1136/jech-2021-218074

Carey, G., Malbon, E., Carey, N., Joyce, A., Crammond, B., & Carey, A. (2015). Systems science and systems thinking for public health: a systematic review of the field. *BMJ open*, *5*(12), e009002. https://doi.org/10.1136/bmjopen-2015-009002

Casarico, A., & Lattanzio, S. (2025). Exploring Gender Inequality: Firm Contribution and Policy Effects. *LSE Public Policy Review*, *3*(6), 1-14. https://doi.org/10.31389/lseppr.117

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A., Pizzinelli, C., Rockall, E., & Tavares, M. (2024). *Gen-AI: Artificial Intelligence and the Future of Work*. International Monetary Fund.

Cerda, M., Tracy, M., Ahern, J., & Galea, S. (2014). Addressing population health and health inequalities: the role of fundamental causes. *American journal of public health*, *104 Suppl 4*(Suppl 4), S609-619. https://doi.org/10.2105/AJPH.2014.302055

Chandola, T., & Zhang, N. (2018). Re-employment, job quality, health and allostatic load biomarkers: prospective evidence from the UK Household Longitudinal Study. *International journal of epidemiology*, *47*(1), 47-57. https://doi.org/10.1093/ije/dyx150

Chen, H. J., Xue, H., Liu, S., Huang, T. T. K., Wang, Y. C., & Wang, Y. (2018). Obesity trend in the United States and economic intervention options to change it: A simulation study linking ecological epidemiology and system

dynamics modeling. *Public Health*, *161*, 20-28. https://doi.org/10.1016/j.puhe.2018.01.013

Chen, P., & Semmler, W. (2018). Short and Long Effects of Productivity on Unemployment. *Open Economies Review*, *29*(4), 853-878. https://doi.org/10.1007/s11079-018-9486-z

Chum, A., Kaur, S., Teo, C., Nielsen, A., Muntaner, C., & O'Campo, P. (2022). The impact of changes in job security on mental health across gender and family responsibility: evidence from the UK Household Longitudinal Study. *Social Psychiatry and Psychiatric Epidemiology*, *57*(1), 25-36. https://doi.org/10.1007/s00127-021-02187-6

Churchill, B., Ervin, J., Ruppanner, L., Taouk, Y., & King, T. L. (2025). Underemployment and mental health amongst working-age Australians: a gendered analysis using the HILDA survey (2002-2022). *Health Promotion International*, *40*(2). https://doi.org/10.1093/heapro/daaf030

Cislaghi, B., Bhatia, A., Hallgren, E. S. T., Horanieh, N., Weber, A. M., & Darmstadt, G. L. (2022). Gender Norms and Gender Equality in Full-Time Employment and Health: A 97-Country Analysis of the World Values Survey. *Frontiers in Psychology*, *13*, 689815. https://doi.org/10.3389/fpsyg.2022.689815

Clancy, T., Langarudi, S. P., & Zaini, R. (2023). Never the strongest: reconciling the four schools of thought in system dynamics in the debate on quality. *System Dynamics Review*, *39*(3), 277-294. https://doi.org/10.1002/sdr.1740

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S., & Janecka, M. (2018). Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *The Lancet*, *391*(10121), 679-686. https://doi.org/10.1016/S0140-6736(17)32133-5

Compton, M. T., & Shim, R. S. (2015). The Social Determinants of Mental Health. *Focus*, *13*(4), 419-425. https://doi.org/10.1176/appi.focus.20150017

Connolly, S., & Gregory, M. (2008). The part-time pay penalty: earnings trajectories of British Women. *Oxford Economic Papers*, *61*(Supplement 1), i76-i97. https://doi.org/10.1093/oep/gpn043

Corlett, A. (2024). Get Britain's Stats Working: Exploring alternatives to Labour Force Survey estimates. Resolution Foundation.

Correll, S., Benard, S., & Paik, I. (2007). Getting a job: is there a motherhood penalty? *American Journal of Sociology*, *112*(5), 1297-1338.

Craigforth. (2026). Effective Government policies for advancing equality for women and girls: International Evidence Review.

Cruz, M. D. (2023). Labor Productivity, Real Wages, and Employment in OECD Economies. *Structural Change and Economic Dynamics*, *66*, 367-382. https://doi.org/10.1016/j.strueco.2023.05.007

Curran, C., Knapp, M., McDaid, D., & Tómasson, K. (2007). Mental health and employment: An overview of patterns and policies across Western Europe. *Journal of Mental Health*, *16*(2), 195-209. https://doi.org/10.1080/09638230701279865

Curtis, S., Pearce, J., Cherrie, M., Dibben, C., Cunningham, N., & Bambra, C. (2019). Changing labour market conditions during the 'Great Recession' and mental health in Scotland 2007-2011: an example using the Scottish Longitudinal Study and data for local areas in Scotland. *Social Science & Medicine*, *227*, 1-9. https://doi.org/10.1016/j.socscimed.2018.08.003

Cygan-Rehm, K., Kuehnle, D., & Oberfichtner, M. (2017). Bounding the Causal Effect of Unemployment on Mental Health: Nonparametric Evidence from Four Countries. Institute of Labor Economics.

Dahlgren, G., & Whitehead, M. (1991). Policies and strategies to promote social equity in health: Background document to WHO – Strategy paper for Europe. Institute for Futures Studies.

Daminger, A. (2019). The Cognitive Dimension of Household Labor. *American Sociological Review*, *84*(4), 609-633. https://doi.org/10.1177/0003122419859007

Darabi, N., & Hosseinichimeh, N. (2020). System dynamics modeling in health and medicine: a systematic literature review. *System Dynamics Review*, *36*(1), 29-73. https://doi.org/10.1002/sdr.1646

Davahli, M. R., Karwowski, W., & Taiar, R. (2020). A System Dynamics Simulation Applied to Healthcare: A Systematic Review. *International Journal of Environmental Research and Public Health*, *17*(16). https://doi.org/10.3390/ijerph17165741

Davlasheridze, M., Goetz, S. J., & Han, Y. (2018). The Effect of Mental Health on U.S. County Economic Growth. *The Review of Regional Studies*, *48*(2), 155-171.

De Witte, H., Pienaar, J., & De Cuyper, N. (2016). Review of 30 years of longitudinal studies on the association between job insecurity and health and well-being: Is there causal evidence? *Australian Psychologist*, *51*(1), 18-31. https://doi.org/10.1111/ap.12176

Department for Education. (2025). September 2025 early education and childcare entitlements expansion: Local authority system guidance.

Department for Work and Pensions. (2024). *Get Britain Working*. UK Government. https://www.gov.uk/government/publications/get-britain-working-white-paper/get-britain-working-white-paper

Department for Work and Pensions. (2026). *Young People and Work: Diagnostic Report*. UK Government. https://www.gov.uk/government/publications/young-people-and-work-interim-report/young-people-and-work-interim-report#introduction-the-review

Diez Roux, A. V. (2007). Integrating social and biologic factors in health research: a systems view. *Annuals of Epidemiology*, *17*(7), 569-574. https://doi.org/10.1016/j.annepidem.2007.03.001

Diez Roux, A. V. (2011). Complex systems thinking and current impasses in health disparities research. *American journal of public health*, *101*(9), 1627-1634. https://doi.org/10.2105/AJPH.2011.300149

Doan, T., LaBond, C., Banwell, C., Timmins, P., Butterworth, P., & Strazdins, L. (2022). Unencumbered and still unequal? Work hour - Health tipping points and gender inequality among older, employed Australian couples. *SSM - Population Health*, *18*, 101121. https://doi.org/10.1016/j.ssmph.2022.101121

Dominguez-Folgueras, M. (2022). It's about gender: A critical review of the literature on the domestic division of work. *Journal of Family Theory & Review*, *14*(1), 79-96. https://doi.org/10.1111/jftr.12447

Dykxhoorn, J., Osborn, D., Walters, K., Kirkbride, J. B., Gnani, S., & Lazzarino, A. I. (2024). Temporal patterns in the recorded annual incidence of common mental disorders over two decades in the United Kingdom: a primary care cohort study. *Psychological Medicine*, *54*(4), 663-674. https://doi.org/10.1017/S0033291723002349

Ervin, J., Taouk, Y., Alfonzo, L. F., Hewitt, B., & King, T. (2022). Gender differences in the association between unpaid labour and mental health in employed adults: a systematic review. *The Lancet Public health*, *7*(9), e775-e786. https://doi.org/10.1016/S2468-2667(22)00160-8

Ervin, J., Taouk, Y., Hewitt, B., & King, T. (2023a). The association between unpaid labour and mental health in working-age adults in Australia from 2002 to 2020: a longitudinal population-based cohort study. *The Lancet Public health*, *8*(4), e276-e285. https://doi.org/10.1016/S2468-2667(23)00030-0

Ervin, J., Taouk, Y., Hewitt, B., & King, T. (2023b). Trajectories of unpaid labour and the probability of employment precarity and labour force detachment among prime working-age Australian women. *Social Indicators Research*, *169*(3), 1033-1056. https://doi.org/10.1007/s11205-023-03197-4

Evans-Lacko, S., & Knapp, M. (2016). Global patterns of workplace productivity for people with depression: absenteeism and presenteeism costs across eight diverse countries. *Social Psychiatry and Psychiatric Epidemiology*, *51*(11), 1525-1537. https://doi.org/10.1007/s00127-016-1278-4

Fair Work Convention. (2016). Fair Work Framework.

Fauser, S. (2019). Time availability and housework: The effect of unemployment on couples' hours of household labor. *Social Science Research*, *83*, 102304. https://doi.org/10.1016/j.ssresearch.2019.04.017

Fink, D. S., Keyes, K. M., & Cerdá, M. (2017). Systems Science and the Social Determinants of Population Health. In A. M. El-Sayed & S. Galea (Eds.), *Systems Science and Population Health*. https://doi.org/10.1093/acprof:oso/9780190492397.001.0001

Fisher, M., & Baum, F. (2010). The social determinants of mental health: implications for research and health promotion. *Australian & New Zealand Journal of Psychiatry*, *44*(12), 1057-1063. https://doi.org/10.3109/00048674.2010.509311

Florisson, R., & Gable, O. (2022). *The gender gap: insecure work in the UK*. The Work Foundation at Lancaster University.

Forrester, J. (1994). System dynamics, systems thinking, and soft OR. *System Dynamics Review*, *10*(2-3).

Foster, G., & Stratton, L. S. (2018). Do Significant Labor Market Events Change Who Does the Chores? Paid Work, Housework, and Power in Mixed-Gender Australian Households. *Journal of Population Economics*, *31*(2), 483-519. https://doi.org/10.1007/s00148-017-0667-7

Frank, J., Mustard, C., Smith, P., Siddiqi, A., Cheng, Y., Burdorf, A., & Rugulies, R. (2023). Work as a social determinant of health in high-income countries: past, present, and future. *The Lancet*, *402*(10410), 1357-1367. https://doi.org/10.1016/S0140-6736(23)00871-1

Fraser, N. (2017). **Crisis of care? On the social-reproductive contradictions of contemporary capitalism**. In *Social reproduction theory: Remapping class, recentering oppression*. Pluto Press.

Fraser of Allander Institute. (2022). *Chapter 2: Employment, work, Chapter 2: Employment, work, and the labour market* (Health Inequalities in Scotland: Trends in the socio-economic determinants of health in Scotland Issue. The Health Foundation.

Frasquilho, D., Matos, M. G., Salonna, F., Guerreiro, D., Storti, C. C., Gaspar, T., & Caldas-de-Almeida, J. M. (2016). Mental health outcomes in times of economic recession: a systematic literature review. *BMC Public Health*, *16*(115). https://doi.org/10.1186/s12889-016-2720-y

Freebairn, L., Atkinson, J. A., Kelly, P. M., McDonnell, G., & Rychetnik, L. (2018). Decision makers' experience of participatory dynamic simulation modelling: methods for public health policy. *BMC Medical Informatics and Decision Making*, *18*(1), 131. https://doi.org/10.1186/s12911-018-0707-6

Fryer, D. (1992). Psychological Or Material Deprivation: Why Does Unemployment Have Mental Health Consequences? In E. Mclaughlin (Ed.), *Understanding Unemployment : New Perspectives on Active Labour Market Policies*. Taylor & Francis Group. https://doi.org/10.4324/9780203061459-6

Gagnon, N., Bosmans, K., & Riedl, A. (2025). The Effect of Gender Discrimination on Labor Supply. *Journal of Political Economy*, *133*(3), 1047-1081. https://doi.org/10.1086/733419

Galassi, G., Koll, D., & Mayr, L. (2024). The Intergenerational Correlation of Employment: Mothers as Role Models?

Galea, S., Riddle, M., & Kaplan, G. A. (2010). Causal thinking and complex system approaches in epidemiology. *International journal of epidemiology*, *39*(1), 97-106. https://doi.org/10.1093/ije/dyp296

GBD 2021 Diseases and Injuries Collaborators. (2024). Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*, *403*(10440), 2133-2161. https://doi.org/10.1016/S0140-6736(24)00757-8

Gedikli, C., Miraglia, M., Connolly, S., Bryan, M., & Watson, D. (2022). The relationship between unemployment and wellbeing: an updated meta-analysis of longitudinal evidence. *European Journal of Work and Organizational Psychology*, 1-17. https://doi.org/10.1080/1359432x.2022.2106855

Germinario, G., Amin, V., Flores, C. A., & Flores-Lagunes, A. (2022). What can we learn about the effect of mental health on labor market outcomes under weak assumptions? Evidence from the NLSY79. *Labour Economics*, *79*. https://doi.org/10.1016/j.labeco.2022.102258

Gershuny, J., & Kan, M. Y. (2012). Halfway to Gender Equality in Paid and Unpaid Work? Evidence from the Multinational Time-use Study. In *Gendered Lives*. https://doi.org/10.4337/9781849806275.00010

Global Burden of Disease Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry*, *9*(2), 137-150. https://doi.org/10.1016/S2215-0366(21)00395-3

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, *27*(1), 191-197. https://doi.org/10.1017/s0033291796004242

Graham, E., Gariepy, G., & Orpana, H. (2023). System dynamics models of depression at the population level: a scoping review. *Health Research Policy and Systems*, *21*(1), 50. https://doi.org/10.1186/s12961-023-00995-7

Gray, M., & Barford, A. (2018). The depths of the cuts: the uneven geography of local government austerity. *Cambridge Journal of Regions, Economy and Society*, *11*(3), 541-563. https://doi.org/10.1093/cjres/rsy019

Gueltzow, M., Bijlsma, M. J., van Lenthe, F. J., & Myrskyla, M. (2023). The role of labor market inequalities in explaining the gender gap in depression risk among older US adults. *Social Science & Medicine*, *332*, 116100. https://doi.org/10.1016/j.socscimed.2023.116100

Harkness, S., Borkowska, M., & Pelikh, A. (2019). *Employment pathways and occupational change after childbirth*. Government Equalities Office.

Heise, L., Greene, M. E., Opper, N., Stavropoulou, M., Harper, C., Nascimento, M., Zewdie, D., Gender Equality, N., & Health Steering, C. (2019). Gender inequality and restrictive gender norms: framing the challenges to health. *The Lancet*, *393*(10189), 2440-2454. https://doi.org/10.1016/S0140-6736(19)30652-X

Heitman, K. (2017). Reductionism at the Dawn of Population Health. In A. M. El-Sayed & S. Galea (Eds.), *Systems Science and Population Health*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190492397.001.0001

Henderson, K. H., Hammond, R. A., & Homer, J. B. (2023). Complex Adaptive Systems Simulation Modeling to Address Cardiovascular Disparities: Complex Science for a Complex Problem. *Circulation*, *148*(3), 201-203. https://doi.org/10.1161/CIRCULATIONAHA.122.061728

Himmelweit, S., & Sigala, M. (2004). Choice and the Relationship between Identities and Behaviour for Mothers with Pre-School Children: Some Implications for Policy from a UK Study. *Journal of Social Policy*, *33*(3), 455-478. https://doi.org/10.1017/s0047279404007779

Hohn, A., Stokes, J., Pollack, R., Boyd, J., Chueca Del Cerro, C., Elsenbroich, C., Heppenstall, A., Hjelmskog, A., Inyang, E., Kopasker, D., Sonthalia, S., Thomson, R. M., Zia, K., Katikireddi, S. V., & Meier, P. (2023). Systems science methods in public health: what can they contribute to our understanding of and response to the cost-of-living crisis? *Journal of Epidemiology & Community Health*, *77*(9), 610-616. https://doi.org/10.1136/jech-2023-220435

Homer, J. (2014). Levels of evidence in system dynamics modeling. *System Dynamics Review*, *30*(1-2), 75-80. https://doi.org/10.1002/sdr.1514

Homer, J., & Hirsch, G. (2006). System dynamics modeling for public health Background and opportunities. *American journal of public health*, *96*(3), 452-458. https://doi.org/10.2105/AJPH.2005

Hong, Q. N., Bangpan, M., Stansfield, C., Kneale, D., O'Mara-Eves, A., van Grootel, L., & Thomas, J. (2022). Using systems perspectives in evidence synthesis: A methodological mapping review. *Research Synthesis Methods*, 1-14. https://doi.org/10.1002/jrsm.1595

Huo, D., Rice, N., Roberts, J., & Sechel, C. (2022). *Mental health and productivity: evidence for the UK*. T. U. o. Sheffield.

Institute for the Future of Work. (2025). *The Pissarides Review into the Future of Work and Wellbeing*. Institute for the Future of Work.

Irvine, A., & Rose, N. (2022). How Does Precarious Employment Affect Mental Health? A Scoping Review and Thematic Synthesis of Qualitative Evidence from Western Economies. *Work, Employment and Society*, *38*(2). https://doi.org/10.1177/09500170221128698

Irvine, G. (2021). What next for Fair Work in Scotland? C. U. Trust.

Janke, K., Lee, K., Propper, C., Shields, K., & Shields, M. A. (2020). *Macroeconomic Conditions and Health in Britain: Aggregation, Dynamics and Local Area Heterogeneity*. Institute of Labor Economics.

Jenkins, J., & Chandler, M. (2010). Labour market gross flows data from the Labour Force Survey. *Economic & Labour Market Review*.

Jones, L. (2019). *Women's Progression in the Workplace*. https://www.kcl.ac.uk/giwl/assets/womens-progression-in-the-workplace.pdf

Jones, L., Cook, R., & Connolly, S. (2023). Parenthood and Job Quality: Is There a Motherhood Penalty in the UK? *Social Indicators Research*, *170*(2), 765-792. https://doi.org/10.1007/s11205-023-03214-6

Joyce, K., Pabayo, R., Critchley, J. A., & Bambra, C. (2010). Flexible working conditions and their effects on employee health and wellbeing. *Cochrane Database of Systematic Reviews*, *2010*(2), CD008009. https://doi.org/10.1002/14651858.CD008009.pub2

Joyce, R., & Xu, X. (2019). *The gender pay gap: women work for lower-paying firms than men (Comment)*. The Institute for Fiscal Studies. https://ifs.org.uk/articles/gender-pay-gap-women-work-lower-paying-firms-men

Kaiser, C., Muggleton, N., Quispe--Torreblanca, E., & De Neve, J. (2025). Two paradoxes in women's well-- being. *Science Advances*, *11*(10).

Kamerade, D., Wang, S., Burchell, B., Balderson, S. U., & Coutts, A. (2019). A shorter working week for everyone: How much paid work is needed for mental health and well-being? *Social Science & Medicine*, *241*, 112353. https://doi.org/10.1016/j.socscimed.2019.06.006

Kan, M. Y., Sullivan, O., & Gershuny, J. (2011). Gender Convergence in Domestic Work: Discerning the Effects of Interactional and Institutional Barriers from Large-scale Data. *Sociology*, *45*(2), 234-251. https://doi.org/10.1177/0038038510394014

Katikireddi, S. V., Niedzwiedz, C. L., & Popham, F. (2012). Trends in population mental health before and after the 2008 recession: a repeat cross-sectional analysis of the 1991-2010 Health Surveys of England. *BMJ open*, *2*(5). https://doi.org/10.1136/bmjopen-2012-001790

Kelly-Irving, M., Ball, W. P., Bambra, C., Delpierre, C., Dundas, R., Lynch, J., McCartney, G., & Smith, K. (2023). Falling down the rabbit hole? Methodological, conceptual and policy issues in current health inequalities research. *Critical Public Health*, *33*(1), 37-47. https://doi.org/10.1080/09581596.2022.2036701

Keyes, C. (2002). The Mental Health Continuum: From Languishing to Flourishing in Life. *Journal of Health and Social Behavior*, *43*(2), 207-222. https://doi.org/10.2307/3090197

Keyes, K., & Galea, S. (2015). What matters most: quantifying an epidemiology of consequence. *Annals of Epidemiology*, *25*(5), 305-311. https://doi.org/10.1016/j.annepidem.2015.01.016

Kim, T. J., & von dem Knesebeck, O. (2015). Is an insecure job better for health than having no job at all? A systematic review of studies investigating the health-related risks of both job insecurity and unemployment. *BMC Public Health*, *15*, 985. https://doi.org/10.1186/s12889-015-2313-1

Kim, T. J., & von dem Knesebeck, O. (2016). Perceived job insecurity, unemployment and depressive symptoms: a systematic review and meta-analysis of prospective observational studies. *International archives of occupational and environmental health*, *89*, 561-573. https://doi.org/10.1007/s00420-015-1107-1

Kleven, H., Landais, C., & Leite-Mariante, G. (2023). *The Child Penalty Atlas*. National Bureau of Economic Research.

Koopman, J., & Lynch, J. (1999). Individual Causal Models and Population System Models in Epidemiology. *American journal of public health*, *89*(8).

Kopasker, D., Bronka, P., Thomson, R. M., Khodygo, V., Kromydas, T., Meier, P., Heppenstall, A., Bambra, C., Lomax, N., Craig, P., Richiardi, M., & Katikireddi, S. V. (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. https://doi.org/10.1016/j.socscimed.2024.116953

Kreshpaj, B., Orellana, C., Burstrom, B., Davis, L., Hemmingsson, T., Johansson, G., Kjellberg, K., Jonsson, J., Wegman, D. H., & Bodin, T. (2020). What is precarious employment? A systematic review of definitions and operationalizations from quantitative and qualitative studies. *Scandinavian Journal of Work, Environment & Health*, *46*(3), 235-247. https://doi.org/10.5271/sjweh.3875

Krieger, N. (2008). Proximal, distal, and the politics of causation: what's level got to do with it? *American journal of public health*, *98*(2), 221-230. https://doi.org/10.2105/AJPH.2007.111278

Krieger, N. (2017). Health Equity and the Fallacy of Treating Causes of Population Health as if They Sum to 100. *American journal of public health*, *107*(4), 541-549. https://doi.org/10.2105/AJPH.2017.303655

Kromydas, T., Thomson, R. M., Pulford, A., Green, M. J., & Katikireddi, S. V. (2021). Which is most important for mental health: Money, poverty, or paid work? A fixed-effects analysis of the UK Household Longitudinal Study. *SSM - Population Health*, *15*, 100909. https://doi.org/10.1016/j.ssmph.2021.100909

Kuehner, C. (2017). Why is depression more common among women than among men? *Lancet Psychiatry*, *4*(2), 146-158. https://doi.org/10.1016/S2215-0366(16)30263-2

Langellier, B. A., Yang, Y., Purtle, J., Nelson, K. L., Stankov, I., & Diez Roux, A. V. (2019). Complex Systems Approaches to Understand Drivers of Mental Health and Inform Mental Health Policy: A Systematic Review. *Administration and Policy in Mental Health and Mental Health Services Research*, *46*(2), 128-144. https://doi.org/10.1007/s10488-018-0887-5

Layard, R. (2013). *Mental health: the new frontier for labour economics* (2193-9004). Centre for Economic Performance.

Lerner, D., & Henke, R. M. (2008). What does research tell us about depression, job performance, and work productivity? *Journal of Occupational and Environmental Medicine*, *50*(4), 401-410. https://doi.org/10.1097/JOM.0b013e31816bae50

Lewis, J. (1992). Gender and the Development of Welfare Regimes. *Journal of European Social Policy*, *2*(3). https://doi.org/10.1177/09589287920020030

Li, G., & Granados, C. M. (2023). *The Recent Decline in United Kingdom Labor Force Participation: Causes and Potential Remedies* (IMF Selected Issues Paper, Issue. International Monetary Fund.

Linder, A., Gerdtham, U. G., Trygg, N., Fritzell, S., & Saha, S. (2020). Inequalities in the economic consequences of depression and anxiety in Europe: a systematic scoping review. *European Journal of Public Health*, *30*(4), 767-777. https://doi.org/10.1093/eurpub/ckz127

Lindqvist, A. N., & Carnohan, S. (2023). Accuracy or alignment: a conflict in the participatory modeling process? International System Dynamics Conference, Chicago.

Lisi, D., & Malo, M. A. (2017). The impact of temporary employment on productivity. *Journal for Labour Market Research*, *50*(1), 91-112. https://doi.org/10.1007/s12651-017-0222-8

Long, K. M., & Meadows, G. N. (2017). Simulation modelling in mental health: A systematic review. *Journal of Simulation*, *12*(1), 76-85. https://doi.org/10.1057/s41273-017-0062-0

Lu, Z., Yan, S., Jones, J., He, Y., & She, Q. (2023). From Housewives to Employees, the Mental Benefits of Employment across Women with Different Gender Role Attitudes and Parenthood Status. *International*

Journal of Environmental Research and Public Health, *20*(5). https://doi.org/10.3390/ijerph20054364

Lundberg, O. (2020). Is lack of causal evidence linking socioeconomic position with health an 'inconvenient truth'? *European Journal of Public Health*, *30*(4), 619. https://doi.org/10.1093/eurpub/ckaa004

Mackley, A. (2020). *Social security powers in the UK*. H. o. C. Library. https://commonslibrary.parliament.uk/research-briefings/cbp-9048/

Mahamoud, A., Roche, B., & Homer, J. (2013). Modelling the social determinants of health and simulating short-term and long-term intervention impacts for the city of Toronto, Canada. *Social Science & Medicine*, *93*, 247-255. https://doi.org/10.1016/j.socscimed.2012.06.036

Manning, A., & Petrongolo, B. (2006). *The part-time pay penalty for women in Britain*. Institute for the Study of Labor (IZA).

Margerison-Zilko, C., Goldman-Mellor, S., Falconi, A., & Downing, J. (2016). Health Impacts of the Great Recession: A Critical Review. *Current epidemiology reports*, *3*(1), 81-91. https://doi.org/10.1007/s40471-016-0068-6

Marimpi, M., Barr, B., Baxter, A., Hugh-Jones, S., Taylor-Robinson, D., Katikireddi, S. V., Cheetam, M., Morris, S., Monford, L., Richiardi, M. G., Sutton, M., Bambra, C., Craig, P., & Wickham, S. (2025). *Estimating population mental health effects of the rollout of Universal Credit: Difference-in-Differences analyses using the UK Household Longitudinal Study, 2009 - 2019*. Centre for Microsimulation and Policy Analysis.

McAllister, A., Fritzell, S., Almroth, M., Harber-Aschan, L., Larsson, S., & Burström, B. (2018). How do macro-level structural determinants affect inequalities in mental health? - a systematic review of the literature. *International Journal for Equity in Health*, *17*(1), 180. https://doi.org/10.1186/s12939-018-0879-9

McCartney, G., Hearty, W., Arnot, J., Popham, F., Cumbers, A., & McMaster, R. (2019). Impact of Political Economy on Population Health: A Systematic Review of Reviews. *American journal of public health*, *109*(6), e1-e12. https://doi.org/10.2105/AJPH.2019.305001

McMunn, A. (2023). Unpaid labour is a neglected social determinant of health. *The Lancet Public health*, *8*(4), e250-e251. https://doi.org/10.1016/S2468-2667(23)00050-6

McMunn, A., Bird, L., Webb, E., & Sacker, A. (2020). Gender Divisions of Paid and Unpaid Work in Contemporary UK Couples. *Work, Employment and Society*, *34*(2), 155-173. https://doi.org/10.1177/0950017019862153

Meadows, D. (1999). *Leverage Points: Places to Intervene in a System*. The Sustainability Institute.

Mignon, K. (2025). Employee Jobs Paid Below the Real Living Wage 2024 Living Wage Foundation.

Mingers, J. (2014). Systems Thinking, Critical Realism and Philosophy: A confluence of ideas. Routledge.

Moriconi, S., & Rodríguez-Planas, N. (2021). *Gender Norms and the Motherhood Employment Gap*. Institute of Labor Economics

Murphy, L. (2022). Constrained choices: Understanding the prevalence of part-time work among low-paid workers in the UK. Resolution Foundation.

Naik, Y., Abbs, I., Elwell-Sutton, T., Bibby, J., & Spencelayh, E. (2020). *Using economic development to improve health and reduce health inequalities*. The Health Foundation.

Naik, Y., Baker, P., Ismail, S. A., Tillmann, T., Bash, K., Quantz, D., Hillier-Brown, F., Jayatunga, W., Kelly, G., Black, M., Gopfert, A., Roderick, P.,

Barr, B., & Bambra, C. (2019). Going upstream - an umbrella review of the macroeconomic determinants of health and health inequalities. *BMC Public Health*, *19*(1), 1678. https://doi.org/10.1186/s12889-019-7895-6

Nandi, A., & Platt, L. (2024). Gender, immigration, and ethnicity. *Oxford Open Economics*, *3*(Supplement_1), i335-i344. https://doi.org/10.1093/ooec/odad084

Nightingale, M. (2019). Part-time employment and the gender gap in low pay for UK employees: what changed over the period 1996–2016? *Community, Work & Family*, *24*(3), 272-290. https://doi.org/10.1080/13668803.2019.1681938

O'Halloran, J., & Thomas, C. (2024). *Healthy industry, prosperous economy,* (IPPR discussion paper, Issue. Institute for Public Policy Research. http://www.ippr.org/articles/healthy-industry-prosperous-economy

Occhipinti, J. A., Buchanan, J., Skinner, A., Song, Y. J. C., Tran, K., Rosenberg, S., Fels, A., Doraiswamy, P. M., Meier, P., Prodan, A., & Hickie, I. B. (2022). Measuring, Modeling, and Forecasting the Mental Wealth of Nations. *Frontiers in Public Health*, *10*, 879183. https://doi.org/10.3389/fpubh.2022.879183

Occhipinti, J. A., Skinner, A., Iorfino, F., Lawson, K., Sturgess, J., Burgess, W., Davenport, T., Hudson, D., & Hickie, I. (2021). Reducing youth suicide: systems modelling and simulation to guide targeted investments across the determinants. *BMC Medicine*, *19*(1), 61. https://doi.org/10.1186/s12916-021-01935-4

OECD. (2017). The Pursuit of Gender Equality: An Uphill Battle. OECD Publishing.

OECD. (2023a). *Measuring Population Mental Health*. OECD Publishing. https://doi.org/10.1787/5171eef8-en

OECD. (2023b). *Scotland's National Performance Framework and efforts to build a Wellbeing Economy*. https://www.oecd.org/en/publications/well-being-knowledge-exchange-platform-kep_93d45d63-en/scotland-s-national-performance-framework-and-efforts-to-build-a-wellbeing-economy_b58ad8e5-en.html

OECD. (n.d.). *Employment protection and minimum wages*. Organisation for Economic Cooperation and Development Retrieved 28 April from https://www.oecd.org/en/topics/employment-protection-and-minimum-wages.html

Office for National Statistics. (2023). *Rising ill-health and economic inactivity because of long-term sickness, UK 2019 to 2023*. Office for National Statistics. https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/economicinactivity/articles/risingillhealthandeconomicinactivitybecauseoflongtermsicknessuk/2019to2023#:~:text=Over%201.35%20million%20(53%25),rather%20than%20their%20main%20one.

Olsen, W., Gash, V., Sook, K., & Zhang, M. (2018). *The gender pay gap in the UK: evidence from the UKHLS* G. E. O. Department for Education.

Orchard, C., Lin, E., Rosella, L., & Smith, P. M. (2024). Using a causal decomposition approach to estimate the contribution of employment to differences in mental health profiles between men and women. *SSM - Population Health*, *28*, 101718. https://doi.org/10.1016/j.ssmph.2024.101718

Padrosa, E., Vanroelen, C., Muntaner, C., Benach, J., & Julia, M. (2022). Precarious employment and mental health across European welfare states: a gender perspective. *International archives of occupational and*

environmental health, *95*(7), 1463-1480. https://doi.org/10.1007/s00420-022-01839-7

Paina, L., & Peters, D. H. (2012). Understanding pathways for scaling up health services through the lens of complex adaptive systems. *Health Policy and Planning*, *27*(5), 365-373. https://doi.org/10.1093/heapol/czr054

Palmer, E. (2017). Structural Disadvantage: Evidence of Gender Disparities in the Norwegian Pension System. *Social Sciences*, *6*(1). https://doi.org/10.3390/socsci6010022

Patten, S. B., & Lee, R. C. (2004). Epidemiological theory, decision theory and mental health services research. *Social Psychiatry and Psychiatric Epidemiology*, *39*(11), 893-898. https://doi.org/10.1007/s00127-004-0872-z

Paul, K. I., & Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, *74*(3), 264-282. https://doi.org/10.1016/j.jvb.2009.01.001

Paul, K. I., Scholl, H., Moser, K., Zechmann, A., & Batinic, B. (2023). Employment status, psychological needs, and mental health: Meta-analytic findings concerning the latent deprivation model. *Frontiers in Psychology*, *14*, 1017358. https://doi.org/10.3389/fpsyg.2023.1017358

Pearce, N. (2011). Epidemiology in a changing world: variation, causation and ubiquitous risk factors. *International journal of epidemiology*, *40*(2), 503-512. https://doi.org/10.1093/ije/dyq257

Perales, F., Lersch, P. M., & Baxter, J. (2019). Birth cohort, ageing and gender ideology: Lessons from British panel data. *Social Science Research*, *79*, 85-100. https://doi.org/10.1016/j.ssresearch.2018.11.003

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression. Critical review. *British Journal of Psychiatry*, *177*, 486-492. https://doi.org/10.1192/bjp.177.6.486

Pinna Pintor, M., Fumagalli, E., & Suhrcke, M. (2024). The impact of health on labour market outcomes: A rapid systematic review. *Health Policy*, *143*, 105057. https://doi.org/10.1016/j.healthpol.2024.105057

Piovani, C., & Aydiner-Avsar, N. (2021). Work Time Matters for Mental Health: A Gender Analysis of Paid and Unpaid Labor in the United States. *Review of Radical Political Economics*, *53*(4), 579-589. https://doi.org/10.1177/04866134211035565

Pizzinelli, C., & Speigner, B. (2017). Matching efficiency and labour market heterogeneity in the United Kingdom. Bank of England.

Platt, J., Prins, S., Bates, L., & Keyes, K. (2016). Unequal depression for equal work? How the wage gap explains gendered disparities in mood disorders. *Social Science & Medicine*, *149*, 1-8. https://doi.org/10.1016/j.socscimed.2015.11.056

Platt, J. M., Bates, L. M., Jager, J., McLaughlin, K. A., & Keyes, K. M. (2020). Changes in the depression gender gap from 1992 to 2014: Cohort effects and mediation by gendered social position. *Social Science & Medicine*, *258*, 113088. https://doi.org/10.1016/j.socscimed.2020.113088

Powell, A. (2024). *Economic update: Inactivity due to illness reaches record*. UK Parliament. Retrieved 2-3-25 from https://commonslibrary.parliament.uk/economic-update-inactivity-due-to-illness-reaches-record/

Prosser, T. (2016). Dualization or liberalization? Investigating precarious work in eight European countries. *Work, Employment and Society*, *30*(6), 949-965. https://doi.org/10.1177/0950017015609036

Public Health Scotland. (2022). *Adult mental health indicators*. https://publichealthscotland.scot/media/12202/adult-mental-health-indicator-set.pdf

Pulford, A., Thapa, A., Thomson, R. M., Guilding, A., Green, M. J., Leyland, A., Popham, F., & Katikireddi, S. V. (2022). Does persistent precarious employment affect health outcomes among working age adults? A systematic review and meta-analysis. *Journal of Epidemiology & Community Health*, *76*(11), 909-917. https://doi.org/10.1136/jech-2022-219292

Randolph, H. (2024). *Economic inactivity and ill health in Scotland*. Scottish Parliament Information Centre.

Reich-Stiebert, N., Froehlich, L., & Voltmer, J. B. (2023). Gendered Mental Labor: A Systematic Literature Review on the Cognitive Dimension of Unpaid Work Within the Household and Childcare. *Sex Roles*, *88*(11-12), 475-494. https://doi.org/10.1007/s11199-023-01362-0

Rittel, H., & Webber, M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, *4*(2), 155-169.

Rönnblad, T., Gronholm, E., Jonsson, J., Koranyi, I., Orellana, C., Kreshpaj, B., Chen, L., Stockfelt, L., & Bodin, T. (2019). Precarious employment and mental health: a systematic review and meta-analysis of longitudinal studies. *Scandinavian Journal of Work, Environment & Health*, *45*(5), 429-443. https://doi.org/10.5271/sjweh.3797

Rose, G. (2001). Sick individuals and sick populations. *International journal of epidemiology*, *30*.

Rugulies, R., Aust, B., Greiner, B. A., Arensman, E., Kawakami, N., LaMontagne, A. D., & Madsen, I. E. H. (2023). Work-related causes of mental health conditions and interventions for their improvement in workplaces. *Lancet*, *402*(10410), 1368-1381. https://doi.org/10.1016/S0140-6736(23)00869-3

Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D. T., Greaves, F., Harper, L., Hawe, P., Moore, L., Petticrew, M., Rehfuess, E., Shiell, A., Thomas, J., & White, M. (2017). The need for a complex systems model of evidence for public health. *The Lancet*, *390*(10112), 2602-2604. https://doi.org/10.1016/S0140-6736(17)31267-9

Sahl, A. (2015). Gender, Employment Status, and Unpaid Household Labor: A Test of the Time Availability Perspective on Housework, Child Care, and Emotion Work Before, During, and After the 2007 U.S. Economic Recession University of Nevada]. UNL V Theses, Dissertations, Professional Papers, and Capstones.

Salis, S., Speckesser, S., Samek, L., & Bivand, P. (2021). *A study of work and health transitions: analysis of Understanding Society*. Department for Work and Pensions.

Saunders, M. (2017). The labour market [Speech]. In Bank of England (Ed.), *Speech*. London.

Savigny, D., & Adam, T. (2009). *Systems Thinking for Health Systems Strengthening*. W. H. O. Alliance for Health and Policy Research.

Schoenberg, W., Davidsen, P., & Eberlein, R. (2020). Understanding model behavior using the Loops that Matter method. *System Dynamics Review*, *36*(2), 158-190. https://doi.org/10.1002/sdr.1658

Schwaninger, M., & Grösser, S. (2008). System dynamics as model-based theory building. *Systems Research and Behavioral Science*, *25*(4), 447-465. https://doi.org/10.1002/sres.914

Schweitzer, M., & Tinsley, D. (2004). The UK labour force participation rate: business cycle and trend influences. Bank of England.

Scott, J., & Clery, E. (2013). *Gender roles: An incomplete revolution?* N. S. Research.
Scottish Government. (2019a). A Fairer Scotland for Women: Gender Pay Gap Action Plan.
Scottish Government. (2019b). *Gender and Work in Scotland: Findings from Understanding Society* (Equality and Welfare, Issue. https://www.gov.scot/publications/gender-work-scotland-findings-understanding-society/pages/1/
Scottish Government. (2021). Public Sector Pay Policy 2022-23.
Scottish Government. (2022a). *Delivering Economic Prosperity* (Scotland's National Strategy for Economic Transformation, Issue.
Scottish Government. (2022b). Fair Work Action Plan: Becoming a leading Fair Work Nation by 2025.
Scottish Government. (2022c). Scotland's Labour Market: People, Places and Regions – Protected Characteristics (Statistics, Issue.
Scottish Government. (2022d). *The Scottish Health Survey 2021*. https://www.gov.scot/collections/scottish-health-survey/
Scottish Government. (2023a). Evidence Narrative to inform the Scottish Government Mental Health and Wellbeing Strategy.
Scottish Government. (2023b). Mental Health and Wellbeing Strategy.
Scottish Government. (2024a). *Annual Survey of Hours and Earnings 2024* (Labour Market Statistics, Issue.
Scottish Government. (2024b). *Fair Work First - Guidance*.
Scottish Government. (2024c). Green Industrial Strategy.
Scottish Government. (2025a). Scotland's Population Health Framework 2025-2035.
Scottish Government. (2025b). *Scotland's Wellbeing Economy*.
Scottish Government. (2025c). Unpaid care and the gender pay gap: Literature review of recent evidence.
Scottish Government. (2026a). *Evaluation of Fair Work First: Final Report* (Social Research, Issue.
Scottish Government. (2026b). Mental Health and Wellbeing Delivery Plan.
Scottish Government. (2026c). Our Path Ahead: Chief Medical Officer for Scotland Annual Report 2025–2026.
Scottish Government. (2026d). *Programme for Government 2026-31*.
Seedat, S., Scott, K. M., Angermeyer, M. C., Berglund, P., Bromet, E. J., Brugha, T. S., Demyttenaere, K., de Girolamo, G., Haro, J. M., Jin, R., Karam, E. G., Kovess-Masfety, V., Levinson, D., Medina Mora, M. E., Ono, Y., Ormel, J., Pennell, B. E., Posada-Villa, J., Sampson, N. A., . . . Kessler, R. C. (2009). Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry*, *66*(7), 785-795. https://doi.org/10.1001/archgenpsychiatry.2009.36
Senge, P. (1990). The Fifth Discipline: The Art & Practice of the Learning Organization. Doubleday/Currency.
Shah, N., Walker, I. F., Naik, Y., Rajan, S., O'Hagan, K., Black, M., Cartwright, C., Tillmann, T., Pearce-Smith, N., & Stansfield, J. (2021). National or population level interventions addressing the social determinants of mental health - an umbrella review. *BMC Public Health*, *21*(1), 2118. https://doi.org/10.1186/s12889-021-12145-1
Siciliani, L. (2025). Waiting times for health services, health, and labour market outcomes. *European Journal of Public Health*, *36*(2). https://doi.org/10.1093/eurpub/ckaf213

Simpson, J., Albani, V., Bell, Z., Bambra, C., & Brown, H. (2021). Effects of social security policy reforms on mental health and inequalities: A systematic review of observational studies in high-income countries. *Social Science & Medicine*, *272*, 113717. https://doi.org/10.1016/j.socscimed.2021.113717

Slaughter, H. (2024). Firm foundations: Understanding why employers use flexible contracts. Resolution Foundation.

Solar, O., & Irwin, A. (2010). A Conceptual Framework For Action On The Social Determinants Of Health. World Health Organization.

Sterman, J. (2010). *Business dynamics*. Irwin/McGraw-Hill c2000..

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review*, *10*(2-3), 291-330.

Sterman, J. D. (2002). All models are wrong: reflections on becoming a systems scientist. *System Dynamics Review*, *18*(4), 501-531. https://doi.org/10.1002/sdr.261

Sterman, J. D. (2006). Learning from evidence in a complex world. *American journal of public health*, *96*(3), 505-514. https://doi.org/10.2105/AJPH.2005.066043

Stevenson, D., & Farmer, P. (2017). *Thriving at work: A review of mental health and employers*. https://www.gov.uk/government/publications/thriving-at-work-a-review-of-mental-health-and-employers

Stewart, W. F., Ricci, J. A., Chee, E., Hahn, S. R., & Morganstein, D. (2003). Cost of Lost Productive Work Time Among US Workers with Depression. *JAMA*, *290*(16).

Strandh, M., Hammarström, A., Nilsson, K., Nordenmark, M., & Russel, H. (2013). Unemployment, gender and mental health: the role of the gender regime. *Sociology of Health & Illness*, *35*(5), 649-665. https://doi.org/10.1111/j.1467-9566.2012.01517.x

Strazdins, L., Welsh, J., Korda, R., Broom, D., & Paolucci, F. (2016). Not all hours are equal: could time be a social determinant of health? *Sociology of Health & Illness*, *38*(1), 21-42. https://doi.org/10.1111/1467-9566.12300

Stuckler, D., Reeves, A., Loopstra, R., Karanikolos, M., & McKee, M. (2017). Austerity and health: the impact in the UK and Europe. *European Journal of Public Health*, *27*(suppl_4), 18-21. https://doi.org/10.1093/eurpub/ckx167

Suhrcke, M., McKee, M., Stuckler, D., Sauto Arce, R., Tsolova, S., & Mortensen, J. (2006). The contribution of health to the economy in the European Union. *Public Health*, *120*(11), 994-1001. https://doi.org/10.1016/j.puhe.2006.08.011

Sullivan, O. (2011). An End to Gender Display Through the Performance of Housework? A Review and Reassessment of the Quantitative Literature Using Insights From the Qualitative Literature. *Journal of Family Theory & Review*, *3*(1), 1-13. https://doi.org/10.1111/j.1756-2589.2010.00074.x

Thomas, C., Benzeval, M., & Stansfeld, S. A. (2005). Employment transitions and mental health: an analysis from the British household panel survey. *Journal of Epidemiology & Community Health*, *59*(3), 243-249. https://doi.org/10.1136/jech.2004.019778

Thomas, C., Jung, C., Statham, R., & Quilter-Pinner, H. (2023). *Healthy people, prosperous lives: The first interim report of the IPPR Commission on Health and Prosperity, IPPR*. Institute for Public Policy Research. http://www.ippr.org/research/publications/healthy-people-prosperous-lives

Thomson, R. M., Kopasker, D., Bronka, P., Richiardi, M., Khodygo, V., Baxter, A. J., Igelstrom, E., Pearce, A., Leyland, A. H., & Katikireddi, S. V. (2024). Short-term impacts of Universal Basic Income on population mental health inequalities in the UK: A microsimulation modelling study. *PLoS Medicine*, *21*(3), e1004358. https://doi.org/10.1371/journal.pmed.1004358

Thomson, R. M., Kopasker, D., Leyland, A., Pearce, A., & Katikireddi, S. V. (2023). To what extent does income explain the effect of unemployment on mental health? Mediation analysis in the UK Household Longitudinal Study. *Psychological Medicine*, *53*(13), 6271-6279. https://doi.org/10.1017/S0033291722003580

Thomson, R. M., Niedzwiedz, C. L., & Katikireddi, S. V. (2018). Trends in gender and socioeconomic inequalities in mental health following the Great Recession and subsequent austerity policies: a repeat cross-sectional analysis of the Health Surveys for England. *BMJ open*, *8*(8), e022924. https://doi.org/10.1136/bmjopen-2018-022924

Tomlinson, J. (2007). Employment regulation, welfare and gender regimes: a comparative analysis of women's working-time patterns and work–life balance in the UK and the US. *The International Journal of Human Resource Management*, *18*(3), 401-415. https://doi.org/10.1080/09585190601167466

Torrance, D. (2024). *Introduction to devolution in the United Kingdom* (Research Briefing Issue. UK Government.

Uleman, J. F., Petersen, A. H., & Rod, N. H. (2025). Bridging System Dynamics and Causal Epidemiology: An Opportunity for Both Fields. *System Dynamics Review*, *41*(1). https://doi.org/10.1002/sdr.1799

van Rijn, R. M., Robroek, S. J. W., Brouwer, S., & Burdorf, A. (2014). Influence of poor health on exit from paid employment: a systematic review. *Occupational and environmental medicine*, *71*(4), 295-301. https://doi.org/10.1136/oemed-2013-101591

Vriend, M., Rocks, C., & Finch, D. (2025). *Mental health trends among working-age people*. The Health Foundation. https://www.health.org.uk/reports-and-analysis/analysis/mental-health-trends-among-working-age-people

Wang, S., & Cheng, C. (2024). Opportunity or Exploitation? A Longitudinal Dyadic Analysis of Flexible Working Arrangements and Gender Household Labor Inequality. *Social Forces*, *102*(4), 1446-1466. https://doi.org/10.1093/sf/soad125

Wang, S., & Lu, Z. (2022). Is paid inflexible work better than unpaid housework for women's mental health? The moderating role of parenthood. *Applied Research in Quality of Life*. https://doi.org/10.1007/s11482-022-10091-7

Warr, P. (1994). A conceptual framework for the study of work and mental health. *Work & Stress*, *8*(2), 84-97. https://doi.org/10.1080/02678379408259982

Wheatley, D., Lawton, C., & Hardill, I. (2018). Gender Differences in Paid and Unpaid Work. In *Hidden Inequalities in the Workplace* (pp. 181-214). https://doi.org/10.1007/978-3-319-59686-0_8

Whittaker, M. (2019). Dead-end relationship? Exploring the link between productivity and workers' living standards. Resolution Foundation.

Wickham, S., Bentley, L., Rose, T., Whitehead, M., Taylor-Robinson, D., & Barr, B. (2020). Effects on mental health of a UK welfare reform, Universal Credit: a longitudinal controlled study. *Lancet Public Health*, *5*(3), e157-e164. https://doi.org/10.1016/S2468-2667(20)30026-8

Williams, M., & Dyer, W. (2017). Complex realism in social research. *Methodological Innovations*, *10*(2). https://doi.org/10.1177/2059799116683564

Winefield, H. R., Gill, T. K., Taylor, A. W., & Pilkington, R. M. (2012). Psychological well-being and psychological distress: is it necessary to measure both? *Psychology of Well-Being: Theory, Research and Practice*, *2*(3). https://hdl.handle.net/2440/76095

World Health Organization. (2004). Promoting mental health: concepts, emerging evidence, practice (summary report). World Health Organization.

World Health Organization. (2008). Closing the gap in a generation: health equity through action on the social determinants of health - Final report of the commission on social determinants of health. World Health Organization.

World Health Organization. (2022a). Systems thinking for noncommunicable disease prevention policy: Guidance to bring systems approaches into practice. WHO Regional Office for Europe.

World Health Organization. (2022b). *World mental health report: Transforming mental health for all*. World Health Organization. https://www.who.int/publications/i/item/9789240049338

World Health Organization, & Calouste Gulbenkian Foundation. (2014). *Social determinants of mental health*. World Health Organization.

Yasarcan, H. (2023). The four main elements of dynamic complexity. *System Dynamics Review*, *39*(2), 171-179. https://doi.org/10.1002/sdr.1731

Zhang, A., Gagne, T., Walsh, D., Ciancio, A., Proto, E., & McCartney, G. (2023). Trends in psychological distress in Great Britain, 1991-2019: evidence from three representative surveys. *Journal of Epidemiology & Community Health*, *77*(7), 468-473. https://doi.org/10.1136/jech-2022-219660

Zhou, M., & Kan, M.-Y. (2019). A new family equilibrium? Changing dynamics between the gender division of labor and fertility in Great Britain, 1991–2017. *Demographic Research*, *40*, 1455-1500. https://doi.org/10.4054/DemRes.2019.40.50

Zimmert, F. (2023). Early child care and the employment potential of mothers: evidence from semi-parametric difference-in-differences estimation. *Journal for Labour Market Research*, *57*(1). https://doi.org/10.1186/s12651-023-00344-9

Zivin, K., Paczkowski, M., & Galea, S. (2011). Economic downturns and population mental health: research findings, gaps, challenges and priorities. *Psychological Medicine*, *41*(7), 1343-1348. https://doi.org/10.1017/S003329171000173X

Zukowski, N., Davidson, S., & Yates, M. J. (2019). Systems approaches to population health in Canada: how have they been applied, and what are the insights and future implications for practice? *Canadian Journal of Public Health*, *110*(6), 741-751. https://doi.org/10.17269/s41997-019-00230-3