CARDIFF UNIVERSITY

DOCTORAL THESIS

Understanding the Impact of Social Networks on the Spread of Obesity

Author:
Mark TUSON

Supervisors:
Professor Paul HARPER
Dr Daniel GARTNER

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Operational Research Group
School of Mathematics

January 6, 2020

Declaration of Authorship

STATEMENT 1
This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.
Signed:________________
Date:__________________

STATEMENT 2
This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is it being submitted concurrently for any other degree or award (outside of any formal collaboration agreement between the University and a partner organisation).
Signed:________________
Date:__________________

STATEMENT 3
I hereby give consent for my thesis, if accepted, to be available in the University's Open Access repository (or, where approved, to be available in the University's library and for inter-library loan), and for the title and summary to be made available to outside organisations, subject to the expiry of a University-approved bar on access if applicable.
Signed:________________
Date:__________________

DECLARATION

This thesis is the result of my own independent work, except where otherwise stated, and the views expressed are my own. Other sources are acknowledged by explicit references. The thesis has not been edited by a third party beyond what is permitted by Cardiff University's Use of Third Party Editors by Research Degree Students Procedure.

Signed:________________

Date:__________________

WORD COUNT: ________________

(Excluding summary, acknowledgements, declarations, contents pages, appendices, tables, diagrams and figures, references, bibliography, footnotes and endnotes.)

CARDIFF UNIVERSITY

Abstract

Operational Research

School of Mathematics

Doctor of Philosophy

Understanding the Impact of Social Networks on the Spread of Obesity

by Mark TUSON

The spread of obesity through social networks has been well documented most notably by Christakis and Fowler in 2007. In this research we sought to understand the nature of the interaction between social networks, the spread of obesity and the behaviours that drive it. We applied this knowledge in a case study, seeking to evaluate the impact of these effects on different sub-groups of the population

These objectives were addressed in a hybrid systems modelling approach implemented in a hybrid simulation. An agent based model simulated the social network and embedded inside each agent was a system dynamics model replicating individual behaviour. The model was parameterised using a stochastic approximation algorithm. This approach allowed us to explore a range of scenarios and also evaluate the topology of the network generated by those scenarios.

The model allowed us to forecast BMI (Body Mass Index) issues for different age-groups and genders. We were also able to infer the network topography and its effects. We found that for the youngest population sub-groups the network magnified the impact of external factors on the individuals weight, conversely for the other sub groups it acted to reduce that impact. The magnitude of the network effect was inversely correlated with age.

Acknowledgements

I have been very grateful for the unstinting help, support and encouragement given to me by all of the academic and administrative staff of the School of Mathematics over the last 4 years. In particular I would like to thank Joanna Emery, Dr Vincent Knight and Professor Owen Jones. I would also like to thank Dr Roger Brooks from Lancaster University for his generosity, both with his time and advice throughout the research.

I would particularly like to thanks Professor Paul Harper, Dr Doris Behrens and Dr Daniel Gartner who between them guided me on this journey, each brought their own unique contribution. Throughout my time in Cardiff, Paul has created a positive, open, and above all enjoyable work environment. Without that understanding and trust it is not clear to me that I would necessarily have completed this journey. Doris was a constant source of ideas and stimulation and continued her support even after moving on. Daniel arrived at just the right moment to inject an element of structure and rigour. In addition, his unfailing generosity in giving his time and support whenever asked has been invaluable and much appreciated.

Thank you all.

Contents

List of Figures

List of Tables

Publications:

Tuson, M. England, T. Behrens, D. Edwards, D. Frankish, J. Jude, K. (2018). Modelling for the proposed roll-out of the '111' service in Wales: a case study, *Health Care Management Science*, **21**(2): 159-176.

List of Abbreviations

ABM	**A**gent **B**ased **M**odelling
AN	**A**norexia **N**ervosa
APCCC	**A**verage **P**er **C**apita **C**alorie **C**onsumption
BED	**B**inge **E**ating **D**isorder
BM	**B**ody **M**ass
BMI	**B**ody **M**ass **I**ndex
BMR	**B**asal **M**etabolic **R**ate
BN	**B**ulimia **N**ervosa
BW	**B**ody **W**eight
DALY	**D**isability **A**djusted **L**ife **Y**ear
DES	**D**iscrete **E**vent **S**imulation
ERG	**E**xponential **R**amdom **G**raph
EB	**E**nergy **B**alance
EI	**E**nergy **I**ntake
ESM	**E**lectronic **S**ocial **M**edia
FFM	**F**at **F**ree **M**ass
FM	**F**at **M**ass
FTA	**F**ollow**T**he **A**verage
GDP	**G**ross **D**omestic **P**roduct
HBM	**H**ealth **B**ehaviour **M**odel
HS	**H**ybrid **S**imulation
HSE	**H**ealth **S**urvey **E**ngland
HSM	**H**ybrid **S**ystems **M**odel
IDP	**I**nformation as a **D**ynamic **P**arameter
MET	**M**etabolic **E**quivalent **T**ask
NCD	**N**on-communicable **C**hronic **D**isease
NHS	**N**ational **H**ealth **S**ervice
NICE	**N**ational **I**nstitute of **C**linical **E**xcellence
OOD	**O**bject **O**riented**D**evelopment
OR	**O**perational **R**esearch
PAL	**P**hysical **A**ctivity **L**evel
PAPM	**P**recaution **A**doption **P**rocess **M**odel
PBC	**P**erceived **B**ehavioural **C**ontrol
POM	**P**attern **O**riented **M**odelling
QAP	**Q**uadratic **A**ssignment **P**rocess
R	Open source statistical software
SAB	**S**tochastic **A**ctor **B**ased (model)
SCM	**S**tages of **C**hange **M**odel
SD	**S**ystem **D**ynamics

SI	**S**atisfing **I**nterval
STC	**S**ocially **T**ransmitted **C**ondition
SPSS	**S**tatistical **P**ackage for **S**ocial **S**ciences
TEE	**T**otal **E**nergy **E**xpenditure
TPB	**T**heory of **P**lanned **B**ehaviour
WHO	**W**orld **H**ealth**O**rganisation
YFAS	**Y**ale **F**ood **A**ddiction **S**core

Dedicated to my mother who 50 years ago taught me to read (when the schools couldn't), ensuring I would eventually be able to embark on this journey. My wife who for the last 4 years has given me the 'space' (literally and metaphorically) to carry it through and my daughter whose example showed me that it might be fun…

Chapter 1

Introduction

"...the mathematical study of the reactions of human conglomerates in response to economic and social stimuli..."

'Psychohistory' as envisioned by Isaac Asimov in the 1940's.

1.1 Personal Motivation

The motivation to decide (at age 55) to spend four years undertaking a PhD, comes from the union of two elements.

The first of these is addressed in Chapter 2, which looks at the issues and consequences of the current obesity epidemic. The opportunity to do research which helps however minimally to address the issue, is a worthwhile use of time.

The other element is my own life-long term interest in understanding and predicting human behaviour. This started when as an adolescent I came across some of the science fiction of Isaac Asimov, in which he suggested that human behaviour could be understood and managed, using the mathematical tools I was just starting to encounter (calculus). My subsequent career built on aspects of that interest, initially as a junior officer in the Army and subsequent with roles in industry that included leadership and management development, organisational change and development, and more recently leadership roles in HR.

This came full circle during the MSc I undertook at Cardiff University, where I was introduced for the first time to the current simulation paradigms and associated software tools. These I found fascinating, in many ways starting to realise those ideas about anticipating and forecasting people behaviour (although on a rather more modest scale) I'd first encountered in my teens.

Once presented, the opportunity to undertake research that combined these areas and focused them on a worthwhile project was not one I was ever going to voluntarily relinquish.

1.2 Background and Aims

Non-communicable chronic diseases (NCD's) are not passed directly from individual to individual through physical contagion, but are driven by changing patterns of behaviour within a population. Mendis, Davis, and Norrving [1] suggest that taken together they accounted for 38 million of the 56 million deaths worldwide in 2012, 16 million of these occurred in people less than 70 years old and were probably preventable. According to the World Health Organisation [2] by 2017 the figure for those under 70 had reduced to 15 million. Their main impact is felt in low income countries where nearly 50% of the premature deaths are experienced.

The behaviours that drive NCD's are classified as health risk behaviours. These include smoking, substance abuse, avoidance of exercise, excessive consumption of alcohol, salt or food and avoidance of medical appointments.

The fact that behaviour can influence health is self-evident, eating in excess of calorie requirements will result in weight gain and eventually if continued, obesity and/or diabetes. Smoking, excessive consumption of alcohol and substance abuse, all have similar health penalties and associated NCD's. These include; cardiovascular and respiratory diseases, stroke, hypertension and diabetes.

The factors that drive that social contagion are not clear, either at the individual and the population level. What drives excessive alcohol or food consumption, why do individuals smoke or avoid exercise? The explanations offered range from physical addiction, through environmental, cultural and social issues, to inheritance and genetics. Finegood, Merth, and Rutter [3] cite all of these factors in the case of obesity. In order to reflect this Allen and Feigl [4] suggest that NCD's should be re-conceptualised as Socially Transmitted Conditions (STC's), in order to stress their anthropogenic and socially contagious nature.

In this thesis we explore the relationship between health risk behaviours, and the social environment in which they take place. We focus on obesity and in particular the way it appears to spread through face-to-face social networks (sometimes known as contact networks). This phenomenon was documented

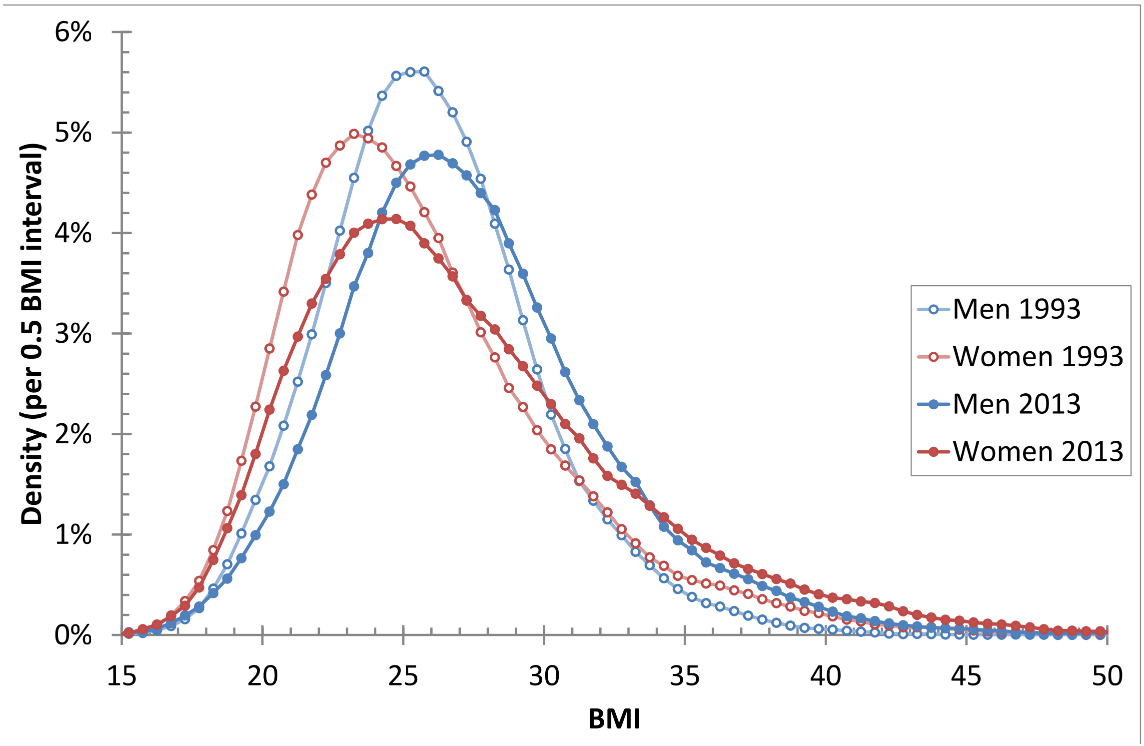


FIGURE 1.1: Changes in Body Mass Index Distribution
Brooks [6]

This Figure is eproduced with the kind permission of Roger Brooks, Lancaster University.

by Christakis and Fowler [5], whose work is briefly described in the next section. This is an important area because growth in levels of obesity are one of the major issues currently confronting the medical profession both in the UK and globally (this is explored in more detail in Chapter 2).

Whilst the growth is evident the mechanisms that drive it are not; the role played by advertising, food availability, sedentary lifestyles and other factors is not yet fully understood. How these then interact with the social network structure is also not clear and it is this that this research seeks to explore using a hybrid simulation model. For clarity, in the context of this document 'hybrid simulation' relates to the combination of 2 or more different simulation paradigms in a single simulation model.

The research is also informed by the work of Brooks [6], who examined the way that the distribution of obesity has changed within the broader population, as obesity levels have risen across the board. In essence he reports that as average BMI has risen the underlying distribution has changed, see Figure 1.1, creating a situation whereby for each sub group of the population, mean BMI values exceed median ones. One hypothesis for this distribution could be that whilst some parts of the population are relatively unaffected by the underlying causality, the adverse impact for others is increasing. There are other hypotheses, inherited susceptibility or a link between high body mass index (BMI) and unusually low basal metabolic rates (BMR) leading to an acceleration of weight gain with increasing BMI at very high levels.

If social mechanisms play a role in supporting or inhibiting this growth, a better understanding of that role would provide a number of benefits; a greater insight into the future impact of the current childhood obesity epidemic on the adult population, greater granularity in obesity forecasting for resource needs and opportunities to improve the design of future obesity interventions.

1.3 Context

[5] suggested that social networks could in themselves act as a mechanism for spreading obesity. Their findings are described briefly below in order to provide a context for the chapters that follow.

The authors describe how they examined a social network comprising some 12,067 adults (aged 21 or over), with data covering a period of 32 years. This data addressed the health and physical well-being of the individual including their height and weight, but also included social data including location, education and friendship ties. The data was obtained from the Framingham Heart Study, Dawber, Meadors, Moore, and Jr. [7].

Christakis and Fowler identified what appeared to be an evolving clusters of obese individuals within the social network. They considered 3 explanations for this phenomena [5, p. 371]:

- Individuals might choose to associate with others who shared similar physical characteristics (Homophily).
- There might be an unobserved set of confounding factors causing changes in weight for the individuals concerned.
- Individuals might be influenced by social or peer effects.

To examine the effect they mapped the network in greater detail, defining two types of relationship between individuals; directional ones where one individual claimed a 'friendship' (not reciprocated), or a mutual one (reciprocated).

They also differentiated between family relationships and friendship (which included marital ties).[5, p. 372-4].

They then constructed a longitudinal regression model incorporating; age, gender, educational level and friends' obesity status. This was used to examine the clustering in more detail [5, p. 376]. The conclusions included:

- The type of friendship was significant with a much increased risk of obesity where the relationship was mutual. For a relationship where only one individual perceived the friendship there was also a significant increase in risk for that individual, there was no statistically meaningful risk for the other individual concerned.
- Male friends of the same gender incurred increased risk, it is not clear that there was a significant risk between female friends.
- Amongst adult siblings the risk was increased, again between male siblings or female siblings but with no inter-gender effect.
- Married couples seemed to increase each other chances of obesity in a similar way.
- By contrast a neighbour becoming obese had no effect on the risk of an individual becoming obese.

In their discussion, they postulate that the mechanism driving the increased risk had less to do with behavioural imitation and was more likely to be driven by the change in perception of social norms (values) regarding the acceptability of obesity [5, p. 377], citing the importance of directionality of friendship in the magnitude of the effect.

Finally and of interest with regard to the purposes of this research, they suggest that social networks could be harnessed to slow the spread of obesity, and exploited to encourage beneficial health behaviours.[5, p. 378].

1.4 Evidence for the Network Effect in the Spread of Obesity

The findings described above provoked some interest and a number of researchers attempted to emulate the results using data from the National Longitudinal Study on Adolescent Health [8]. This provided panel data for adolescents over a 7 year period. Whilst 2 sets of researchers were able to replicate the effect, one set of authors Cohen-Cole and Fletcher [9], were not able to do so and questioned the findings.

This issue was addressed in Fowler and Christakis [10], who suggested that one of the key issues in the failure to reproduce the results was the assumption of a fixed social network (rather than a dynamic one), they then repeated the experiment using the same data to confirm this.

A number of the subsequent issues raised, centred around the issue of differentiating between effects caused by 'contagion' and those by homophily. These were addressed in a number of publications culminating with Christakis and Fowler [11].

Whilst the probabilities with which a social network affects an individuals likelihood of being obese and the degrees of separation over which this influence extends have now been quantified, discussed and to a greater or lesser extent validated [11], [5] and VanderWeele [12], the mechanism whereby this effect is achieved is less defined. In their original paper Christakis and Fowler proposed a sharing of social norms as the main mechanism. In an attempt to validate this hypothesis Hruschka, Brewis, Wutich, and Morin [13, p. 299] could only find limited support for this. Their results suggested that at most it accounted for 20% of the observed effect. However, when discussing the limitations of their study they do not discount a higher percentage effect, suggesting that improved data availability and improvements in methodology would clarify the issue.

The original hypothesis is further explored by Shakya, Christakis, and Fowler [14] where the effect of 'social norms (in this case explicitly defined as comparison to one's social contacts), is examined in the context of motivation to initiate weight reducing behaviours, specifically; dieting, exercise, and alcohol intake reduction. This was done in the context of a longitudinal study comprising 2 waves involving c.20,000 participants in which participants were asked to compare their own BMI and levels of fitness with those of their social groups.

The results supported the hypothesis that a negative comparison of BMI (the individual felt their BMI was significantly higher than their peers) was linked to an increased tendency to initiate dieting and that conversely if individuals felt themselves to be thinner they were less likely to diet. Perhaps surprisingly a negative comparison of BMI also resulted in a decreased likelihood of initiating an exercise programme. The authors note that the behaviour with regard to dieting is consistent with previous research suggesting that as the average weight of a population rises, individuals perception of 'normal' weight also rises, and hence the level at which dieting behaviour might be triggered. In effect acting as a goal seeking mechanism in which the individual seeks to 'normalise ' their weight based upon the norms displayed by their social group.

The behaviour around perceived levels of fitness was different, here a lower perceived level of fitness than one's peers was linked to a reduced tendency to initiate a programme of exercise, and conversely a higher perceived level of fitness was linked to higher levels of initiation of exercise programmes. There was also a positive link between self-esteem and tendency to exercise. Finally, a tendency to reduce alcohol intake was associated with positive self-perceptions of fitness.

Another possibility raised in [11, p. 563] is that individuals can act as 'carriers', in this context this means that an individual may become more accepting of a higher BMI amongst their peers, but not exhibit a higher BMI themselves. This may be a difficult concept to pursue, since its not immediately clear how you would test for such a phenomena.

The focus of the research described above, has been directed at separating out the networks impact from the perceived confounding impact of homophily. What does not seem to have been addressed in the research is the possibility that the network effect and homophily may combine to act in a systemic way, either as a reinforcing loop to amplify the spread of obesity, or alternatively as a balancing loop to slow that spread. Such a combination of network effect and homophily would tend to produce clusters within the network where the BMI was relatively uniform (high, medium or low).

Depending on the surrounding network effect, the clusters might then become isolated, reducing their impact on the network around them, or alternatively become a source of 'infection' for the surrounding network.

1.5 Research Objectives

Seeking to understand the impact of social networks on the spread of obesity is a broad objective. In order to focus the work and provide direction for the research, some specific research objectives were identified:

1. Explore the nature of the interaction between social networks, obesity and the behaviours that drive it, in particular to understand:

   (a) The topology of an 'obesity' network.

   (b) Whether that topology is modified by the spread of obesity?

   (c) Whether homophily play a role in that interaction?

2. Develop a generalisable model to facilitate that exploration, incorporating concepts from behavioural science, social network realisation and simulation to explore the impact of different external parameters on the interaction.

3. Apply that model to specific data for a region/country in a case study, in order to understand:

   (a) How that impact might vary for different sub-groups of the population?

   (b) Which sub-groups might make the most demand on healthcare resources in the future?

   (c) What are the managerial and theoretical insights in terms of both behaviour and social networks, that might be used to augment existing intervention strategies, or suggest new ones in the region/country under consideration?

(The availability of Health Survey England (HSE) data resulted in the use of England as the vehicle for the case study.)

1.6 Problem Classification

With these research aims in mind, it is appropriate to attempt to classify the problem described in the first research objective in terms of healthcare issues and Operational Research (OR) approaches. Both Hulshof, Kortbeek, Boucherie, *et al.* [15] and Brailsford, Harper, Patel, and Pitt [16] provide such taxonomies, classifying problems in the healthcare arena.

The former provides a matrix that references different levels of planning decision (Strategic, Tactical and Operational), with respect to six different care services (Ambulatory, Emergency, Surgical, Inpatient, Home and Residential).

In this context the research described in this paper is intended to support strategic decision making with regard to obesity which in turn may impact on several of the care services including Ambulatory, Surgical, Inpatient and Home.

[16] sets out to examine the current application of OR methodologies to different healthcare functions, ranging from finance, policy, governance and regulation, through risk management and forecasting, to workforce/staff management. The OR methodologies include qualitative and mathematical modelling, statistical analysis and modelling and simulation. Together they form a matrix cross-referencing application and methodology.

In this taxonomy the proposed research uses simulation to focus on Public Health service planning and Forecasting.

1.7 Approach classification

Within the context of a modelling and simulation study, Mustafee and Powell [17] examine the relationship between different modelling approaches and simulation, providing definitions and concepts with which to classify such approaches. A key concept is that of Hybrid Systems Modelling (HSM) in which simulation is combined with methods from other disciplines within Operational Research (OR) or from other fields (big data, machine learning etc.) These methods may be applied at any stage of the study and not just the implementation phase

The model developed to support this research uses theories from behavioural sciences (health related behaviour and social networks), and is implemented using hybrid simulation, which is in turn parameterised using stochastic optimisation techniques. Within the taxonomy suggested by [17] it is an example of Hybrid Systems Modelling.

1.8 The Role of the Simulation

At the heart of the approach is a Hybrid Simulation (HS). In their review of the use of agent based simulation of NCD's, Nianogo and Arah [18] characterise their use as modelling complex adaptive systems. 'Complex' because they cannot be fully explained simply by understanding each element of the system, 'adaptive' because the individuals within the model have the ability to adapt their behaviour according to circumstance. They suggest that the ability and flexibility of this approach to describe emergent phenomena and to represent complex systems in terms of their fundamental components, gives it some significant advantages over other methodologies.

Presciently (in the case of this research) they also identify the issues that may attend it's use:

- The need for large data sets.
- The need for significant computing power.
- Calibration, verification and validation can all be difficult to achieve.
- Models have potentially limited use elsewhere.

The issue of calibration, verification and validation was one of the significant challenges in this research. Whilst changes in obesity levels and food consumption levels are well documented, and there is a significant amount of theory to inform the 'architecture' of any model, much of that theory is qualitative not quantitative. This is particularly evident in behavioral theory and in the theory around social networks, where causality is identified but mathematical relationships are left undefined and where data is very sparse. This meant that the methodology used needed to facilitate a significant level of exploration and optimisation if a functional model was to be achieved.

The decision to use a simulation study enabled exploration and optimisation with regard to the three significant areas and their interdependence, namely:

- Individual norms and behaviours.
- The social network.
- External environment or confounding factors.

Individual norms and behaviours could be represented as either a Discrete Event process or a System Dynamics model, but given the continuous nature of the output of the functions involved (weight gain and loss) and the nature of the behavioural model selected, the latter was used. Social networks are often represented as graphs, and are formed using aggregated behaviours from multiple individuals, an agent based model was considered to be most effective. Modelling the impact of the external environment proved challenging (because of lack of data) and after a number of different methodologies were tried, a 'proxy' approach was settled on using data supplied by the Food and Agriculture Organisation of the United Nations on average daily per capita calorie intake.

The advantages of using a hybrid approach were:

- The ability to try different configurations and test different hypotheses. Hence delivering both qualitative and quantitative information about how the theories employed in its architecture.
- A replicable process that could be applied to other health behaviour scenarios.

The primary risk was whether a simulation model of such a complex system would be able to produce quantitative results that would be sufficiently accurate for resource forecasting. An alternative approach would have been to use a machine learning algorithms (supervised or unsupervised), these would have the potential to deliver a more accurate forecast, but without the qualitative supporting narrative.

1.9 Some Initial Hypotheses and an Analogy

Any hypothesis must address 2 sets of data, that developed in [5] and that developed in [6].

In order to address the first set of data, there are a number of possibilities:

- The social network acts as a system for communicating changing norms.
- The social network acts as a system for transmitting changed behaviours.
- The social network acts as a system for communicating (and transmitting) changing norms and behaviours.

From a practical perspective, differentiating between changes in norms and changes in behaviour may be complicated by the fact that a behavioural change is often triggered by the norms in question reaching a threshold value, thus a change in norms will often trigger a change in behaviour.

In order to gain some intuition about these hypotheses a useful analogy might be to consider a complex underground cave system, with multiple interconnected chambers of different sizes, randomly connected by passages of different capacities and lengths , with water continuously flowing through it. In this analogy each chamber would represent a cluster in the network accommodating a different number of people, with varying levels of connection to the other parts of the network (as is typical for such systems).

The water flowing through the system would be analogous to the flow of information leading to changes in norms/behaviours. The level of water in the system as a whole is driven by external factors, primarily recent rainfall. The level of water in each of the chambers is a function of their location, the size and nature of their connection to the rest of the system. With water levels in some chambers rising quickly, and in others more slowly, some apparently unaffected for a while before succumbing and others remaining unaffected throughout.

Similarly in our social network an individuals level of exposure to changing norms/behaviours would be dependent on the location of their cluster within the network, the connections with other clusters and their exposure to external factors. Again, some would be very vulnerable, others less so and some unaffected.

In order to address the second set of data, an additional hypothesis is needed:

- The topology of the social network is itself modified by the communication/transmission process leading to increased clustering and segmentation.

Extending our cave system analogy, the chambers and the connecting passages are now mutable, and are modified by the individuals within them so that those individuals who are still dry seek others in a similar situation and look to isolate themselves from those who have gotten wet and vice versa. This is achieved by changing location and modifying the chambers and the interconnecting passages.

1.10 Thesis Structure

The structure of the remainder of the thesis is described below:

- Chapter 2 looks at the causes of obesity in more detail, its impact and some of the current initiatives in place to mitigate this
- Chapter 3 provides a review of relevant literature.
- Chapter 4 uses the STRESS guidelines developed by Monks, Currie, Onggo, *et al.* [19] to describes the structure and logic of the final model. It also describes the 'evolutionary' process whereby it was developed and calibrated' and the data used within it.
- Chapter 5 provides analysis of the results generated in both the model development phase and the counterfactual scenarios.

- Chapter 6 discusses the strategic implications of the results in terms of the research questions and aims and presents our conclusions.

A number of Appendices provide context and more detailed information in support of Chapters 3, 4 and 5.

Chapter 2

Obesity; Causes and Impact

"...(childhood) obesity isn't some simple, discrete issue. There's no one cause we can pinpoint. There's no one program we can fund to make it go away. Rather it's an issue that touches on every aspect of how we live and how we work..."

Michelle Obama

2.1 Introduction

This chapter looks initially at the prevalence of obesity, before examining it's causes and consequences, both medical and financial. The final sections describes some of the strategies and initiatives which seek to address the issue.

2.2 What is Obesity and how is it Defined

The key medical factor in determining obesity is the ratio of fat mass (FM) to fat-free mass (FFM), the latter consisting of muscle, organs, connective tissue, bones in effect anything that isn't energy stored in the form of fat.

Measuring an individual's mass is simple, the complexity occurs when trying to ascertain what proportion of that mass is FM (or FFM).

Methods for measuring this ratio include the use of 'volume tanks', skin-fold thickness measurements, electrical impedance, ultra sound scanning and physical measurements such as the ratio of neck size to waist size. These are all effective and have been used in a range of studies, however for large scale studies the use of body mass index (BMI) remains ubiquitous. This is because it is by far the easiest measure to collect, requiring minimal equipment (scales and measuring tape), with minimal issues of calibration. As a consequence

BMI data is also much more widely available both historically and in terms of population coverage.

Whilst its availability is an advantage, the use of BMI as a measure creates a significant issue, as the actual FFM/FM ratio is inferred rather than directly measured. Unfortunately, it can be misleading when applied to athletes and other occupations where high levels of muscle mass are developed, or alternatively with individuals undertaking very low levels of physical activity and hence low muscle mass.

Nonetheless it remains the most commonly used measure and is the statistic used in the vast majority of National Health Service (NHS) data sets reviewed for this document and as a consequence it is also the descriptor used in the models and simulations developed for the thesis.

The actual measure for BMI is calculated as follows:

$$BodyMassIndex(BMI) = Mass(Kg)/Height(metres)^2$$

For adults the NHS applies the following definitions with regard to weight

- BMI less than 18.5 - underweight
- BMI of 18.5 to 24.9 - healthy weight
- BMI of 25 to 29.9 - overweight
- BMI of 30 to 39.9 - obese
- BMI of 40 or more - severely obese

2.3 Prevalence of Obesity: Global Trends

Bentham, Di Cesare, Bilano, *et al.* [20] published a review of worldwide trends in childhood and adolescent weight gain, covering the period from 1975 to 2016. The reported that globally in that period girls (5-19 years old) BMI had risen by $0.32kg/m^2$ each decade. The equivalent figures for boys was $0.42kg/m^2$ per decade. There was considerable variation with static or low increases in Eastern Europe and much higher rates of increase in Polynesia and South America. They also point out that whilst the increase in obesity in developed countries is slowing, this is not the case elsewhere, and that in many regions childhood and adult rates of increase are not correlated.

Finucane, Stevens, Cowan, *et al.* [21] published a similar review for the adult global population involving 199 countries and more than 9 million participants over the age of 19, the authors identified an increasing trend in almost every country. Overall the study identified an increase in (age standardised) BMI of $0.4kg/m^2$ for men and and $0.5kg/m^2$ for women. The report went on to estimate that in 2008, 1.46 billion adults were overweight and 297 million were obese

2.4 Prevalence of Obesity: United Kingdom

Obesity data for the UK is collected in different formats by each of the 4 countries and are not directly comparable, those in Wales being self-reported and not directly measured as in the other three. A comparison of all 4 can be found in Baker [22] from which the figures below are taken:

- In 2017 in England 40% of men were overweight and a further 27% were obese, compared 31% of women overweight and 30% obese. For children aged 4-5, 9.5% were obese and a further 12.8% were overweight.
- In 2018/19 23% of Adults in Wales reported themselves as obese and 36% as overweight. 66% of men were overweight as opposed to 52% of women. 11.9% of children aged 4-5 are obese with a further 26.5% being overweight.
- In 2017 in Scotland the equivalent figures for adults are 29% of adults are obese with 36% overweight, men are more likely to be overweight than women (40% v's 33%, but women are more likely to be obese (33% v's 30%). Of children aged 4-5, 12% are considered to be at risk of being overweight and 10% of being at risk of being obese.
- In 2017/18 in Northern Ireland, 27% of those over 16 were obese and 37% overweight, with men more likely to be overweight or obese than women (62% v's 57%). For children the sample sizes were too small for meaningful comparisons

In all 4 countries childhood obesity is directly linked to deprivation (in Wales the highest figures for childhood obesity are found in Merthyr Tydfil). Although the rate of increase has slowed since the 1990's, they are still rising across almost all age groups and genders.

2.5 A Population Perspective on Obesity

There have been relatively few attempts to forecast the rise in obesity within recent literature, the most comprehensive attempt is delivered by Finkelstein, Khavjou, Thompson, *et al.* [23], using logistic regression models in conjunction with a Behavioural Risk Factor Surveillance System from the District of Columbia, USA. This suggests that the percentage of obese individuals will rise to 39.5% by 2030 (compared to linear models which suggest a rate of 50.7%). This is broadly consistent with the view that obesity levels in the developed world are starting to level out.

[6] explored the way BMI distribution has changed over a period of 20 years (1993 - 2013), for the population as a whole. His results are shown in Figure 1.1. As can be seen from the figure, whilst in each case the mean values for BMI have increased between 1993 and 2013, the overall distribution has not simply moved to the right, but has extended to the right suggesting that there has been a greater increase in BMI amongst those whose values are already higher than the average. This effect is best illustrated by one of the approaches used in [6] to model this effect, applying a common scaling factor to all BMI values above 18.5 using an equation of the form below:

$$BMI_{2013} = BMI_{1993} + max(BMI - 18.5, 0) * (s - 1)$$

In his speculation as to the cause of effect he suggest that unhealthy lifestyle choices diffusing through the population via advertising, media and social networks may provide some explanation, and also that using these 'modes' to address them, may provide solutions.

2.6 Causes of Obesity

2.6.1 Food Consumption - Global Data and Trends

In a review of food consumption trends and drivers Kearney [24] looks at historical trends and changing patterns of consumption, before identifying and describing the impact of the elements driving them.

Perhaps the most startling comment he makes is taken from an earlier paper by Popkin [25], which suggests that the number of overweight and obese people in the world now exceeds those who are underweight or malnourished.

He notes that worldwide the per capita consumption of food (kcal per person per day) has risen from 2,411 in the period 1969/1971 to 2,789 in 1999/2001. Similar figures for the industrial countries are 3,046 rising to 3,446, broadly similar rises are recorded for the other global regions with the exception of sub-Saharan Africa where the rise has been much more modest, 2,100 to 2,194.

2.6.2 Causes of Obesity: A Medical Perspective

"Obesity is the consequence of a sustained positive energy balance with behaviour-associated (eating behaviour; activity behaviour) as well as biological factors (basal metabolic rate) playing a role in the regulation of both energy uptake and energy output. The phenotype obesity is considered today to be the result of an interaction between genetic - 'evolutionary' - predisposition and environmental factors."
Gleich, Lim, and Yu [26, p. 33]

In a report produced for the UK Government's foresight programme a visual representation of the causes of obesity within the United Kingdom was produced in the form of a causal loop diagram (or map). This was done in order to understand the complex and systemic nature of obesity and to help provide a framework for future policy decisions in this area the map and various versions of it, focusing on different areas and different levels of abstraction is given in Vandenbroek, P. Goossens, J. Clemens [27]. It comprises some 107 different factors and their dependencies, it is too complex to reproduce on a single A4 page so is reproduced in full at Appendix F. It was produced in a series of workshops attended by a range of NHS and external subject matter experts and were subsequently tested at a series of follow up workshops. This process, the rationale behind it and some interpretation of the output is described in IP Vandenbroeck, J Goossens [28]. Taken together the two documents provide a qualitative description of the issues underlying the rise of obesity, these are categorised into 7 interlinked thematic clusters.

- Physiology.
- Food Production.
- Food Consumption.
- Physical Activity Environment.
- Individual Physical Activity.
- Social Psychology.

- Individual Psychology.

Whilst the map gives a full and comprehensive picture of the inter-relationship between the factors because of the level of detail and complexity, it is sometimes less easy to decipher direct causality. To address this a graph structural analysis approach suggested by Oliva [29] was employed to create a 'causal tree diagram' where the factors preceding that which we are interested in were mapped and any redundancies removed. The resulting maps were then further edited to ensure that where the same factor appeared in more than one segment, it was only replicated in the most relevant of them. This process provided a feasible framework for the meaningful exploration of the more complex thematic clusters.
(The detailed process and the algorithms involved are described in Appendix D)

It must be emphasised that these are qualitative descriptions and that parts of the content have not been validated experimentally. Nonetheless the soft systems methodologies used to produce them; consultation and discussion involving a wide range of acknowledged experts, supported by review and revision processes, means that they are invaluable in providing a comprehensive view of the 'territory' as seen by those experts.

The sections that follow address the underlying causes of obesity using the thematic clusters.

Core Elements

Figure 2.1 describes the elements that form the basis of the physiological processes they are characterised as:

- The importance of physical need and the degree to which it can trigger calorie expenditure or intake.
- The effort needed to acquire bodily energy through food intake.
- The tendency of individuals to preserve or store energy.
- Level of available energy in the environment accessible to the individual.
- Strength of lock-in to accumulate energy (the degree to which behaviour with regard to consuming and burning calories is dictated by psychological, biological and institutional factors).

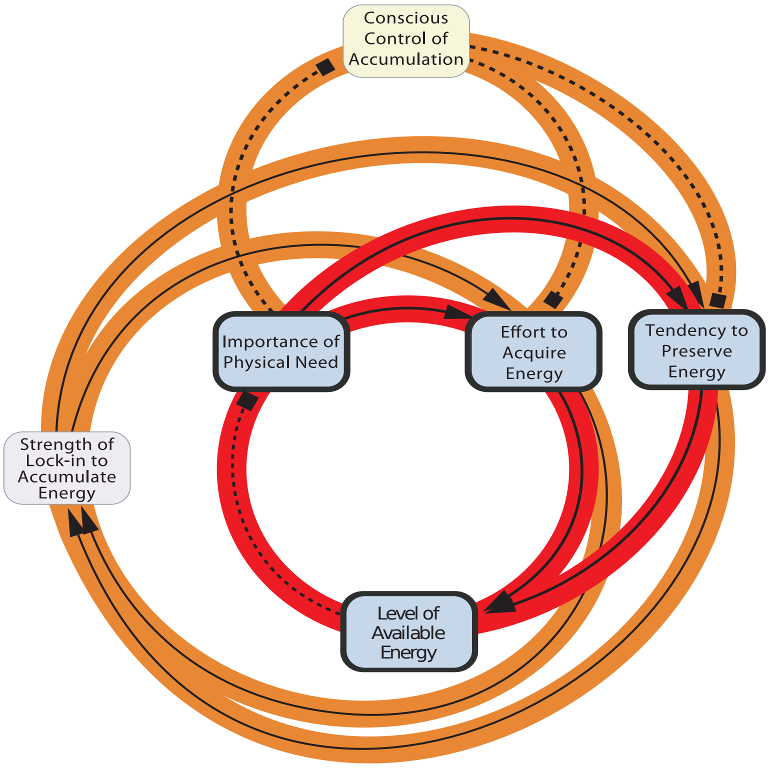


FIGURE 2.1: Foundation Loop
[27, Map 1]

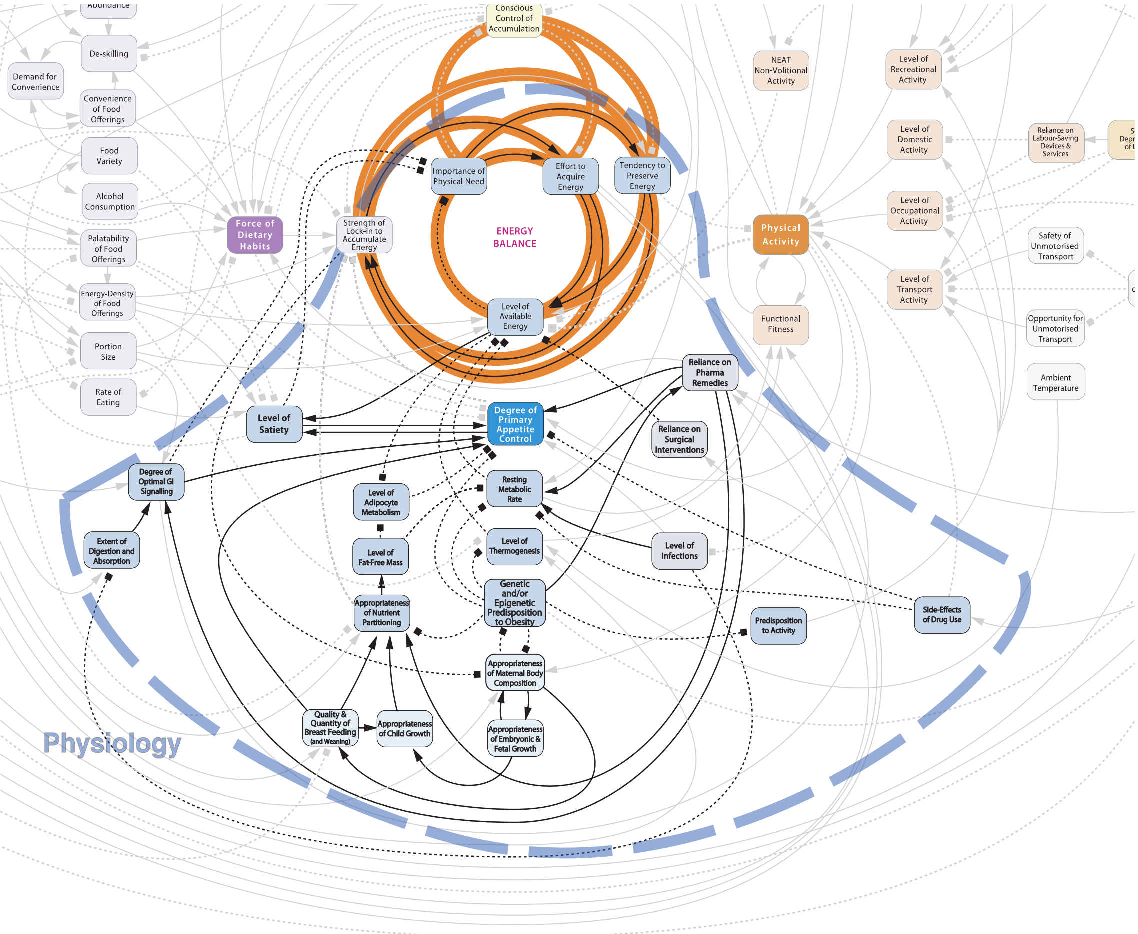


FIGURE 2.2: Physiological Factors
[27, Map 6]

This Figure contains public sector information licensed under the Open Government License v3.0.

- Conscious Control of accumulation (the level of control the individual can consciously exert to control intake of food).

Without some form of check to it's functioning the core loop would operate to constantly acquire energy, leading to continual increase in weight. There are two loops that act on this core loop, a reinforcing loop driven by the strength of the lock-in to acquiring energy and a balancing loop driven by the conscious control exerted by the individual to balance intake with expenditure. One might speculate that the level of available energy in the environment historically, might also have acted as a limiting factor, but that in the developed world its ability to moderate the process has been undermined by the increasing availability of calorie dense foods over the last 60 years, placing significantly more emphasis on the exertion of conscious control.

These elements are each then affected by a range of physiological, environmental and psychological issues. The primary physiological factors are described in Figure 2.2, where degree of primary appetite control is seen as the key factor, this in turn is affected by a range of childhood factors in parallel with any inherited predisposition to obesity.

It is interesting to note the range of factors impacted on by pharmaceutical remedies in comparison with surgical interventions and their perceived level of efficiency. The latter are seen to act directly on only one factor, the level of available energy but this is part of the core loop, by contrast pharmaceutical remedies act on 4 factors (all at one remove or more from the core loop). In spite of this surgical interventions are perceived to be considerably more effective than pharmaceutical (or dietary) options (Ostman, J. Britton, M. Jonsson [30]) in addressing obesity.

The lower section of the cluster also describes the process by which body composition is affected from one generation to another as maternal body composition has a direct impact on foetal growth and the offspring's ability to maintain 'appropriate nutrient partitioning' that is the storage of the correct mix of fat, protein and carbohydrates, and epigenetic effect (inherited non-genetic).

Food Production and Consumption

Food production and consumption account for some of the strongest perceived influences on the core loop (Figure 2.3) with high or very high impact on three of the elements in the core loop and key factors in three out of the remaining five clusters.

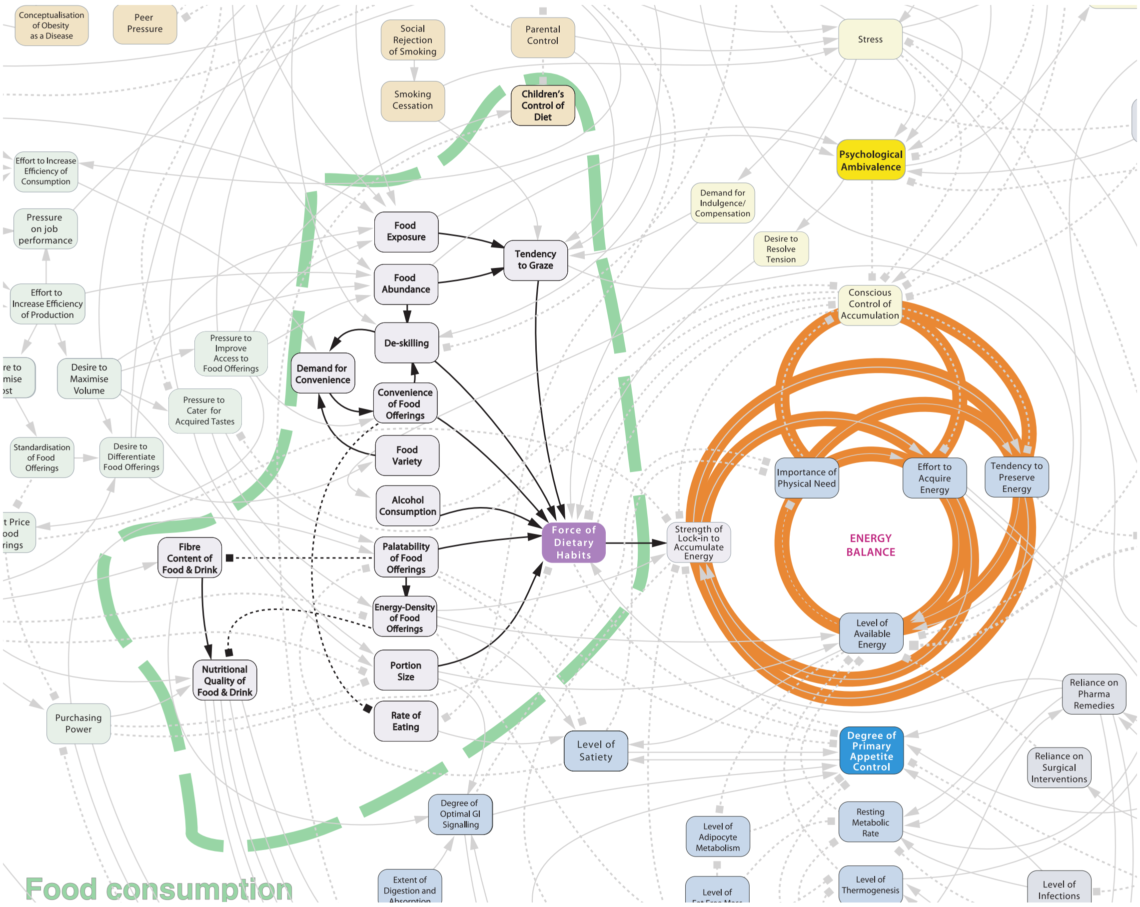


FIGURE 2.3: Food Consumption
[27, Map 12]

"... the food industry is part of an economic system in which trade and the means of production are largely or entirely operated for profit. According to any business enterprise logic, the food industry needs to grow. The exclusive confinement to the consumers saturation will not be sufficient. Instead food products have to be consumed beyond saturation which is achieved not only by increasing the palatability but also by the specific response to certain PT's (personality traits)."
Gerlach, Herpertz, and Loeber [31, p. 59]

The latter part of the quote refers to the use of media by the food industry to leverage personality traits such as impulsivity, neuroticism and hedonism in order to drive sales.

Food production has a high impact on a range of key factors in food consumption:

- Food exposure.

- Food abundance.
- Convenience of food offerings.
- Energy density of food offerings.
- Portion size.

Whilst a number of these act directly on the core loop, the remainder combine to create a set of dietary habits which in turn impact directly on the core loop.

Using 'Force of Dietary Habit' as the key factor of interest and the graph analysis process described in Appendix D generates Figure 2.4. This emphasises a number of factors, that are perhaps less obvious from the causal map, in particular 'purchasing power' and the 'desire to resolve tension'.

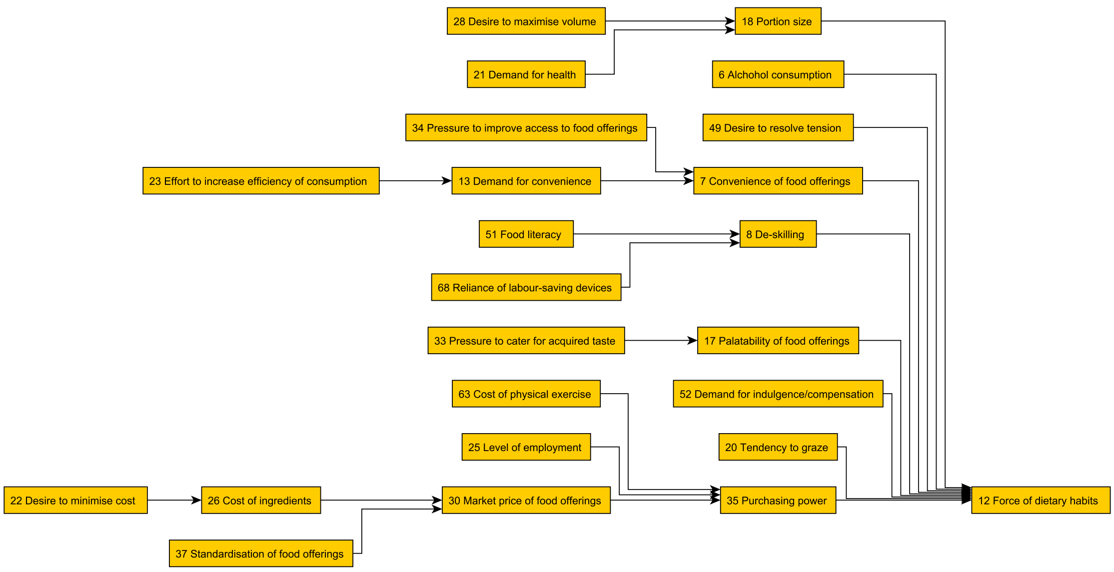


FIGURE 2.4: Force of Dietary Habit - Simplified Causal Map

Physical Activity

By contrast physical activity has less impact as a set of factors, simply acting to reduce the level of available energy, Kopelman, Jebb, and Butland [32, Map 4]. The factors representing the physical environment [32, Map 8] describe both the changes in the physical environment that act to reduce the need and/or opportunity for physical exercise. Separately a smaller cluster describes the cultural factors that affect physical activity (Figure 2.5).

Perhaps the most striking observation is the impact of recreational, occupational, domestic and transport activity, given that they have all been reduced significantly in the period that levels of obesity have been rising.

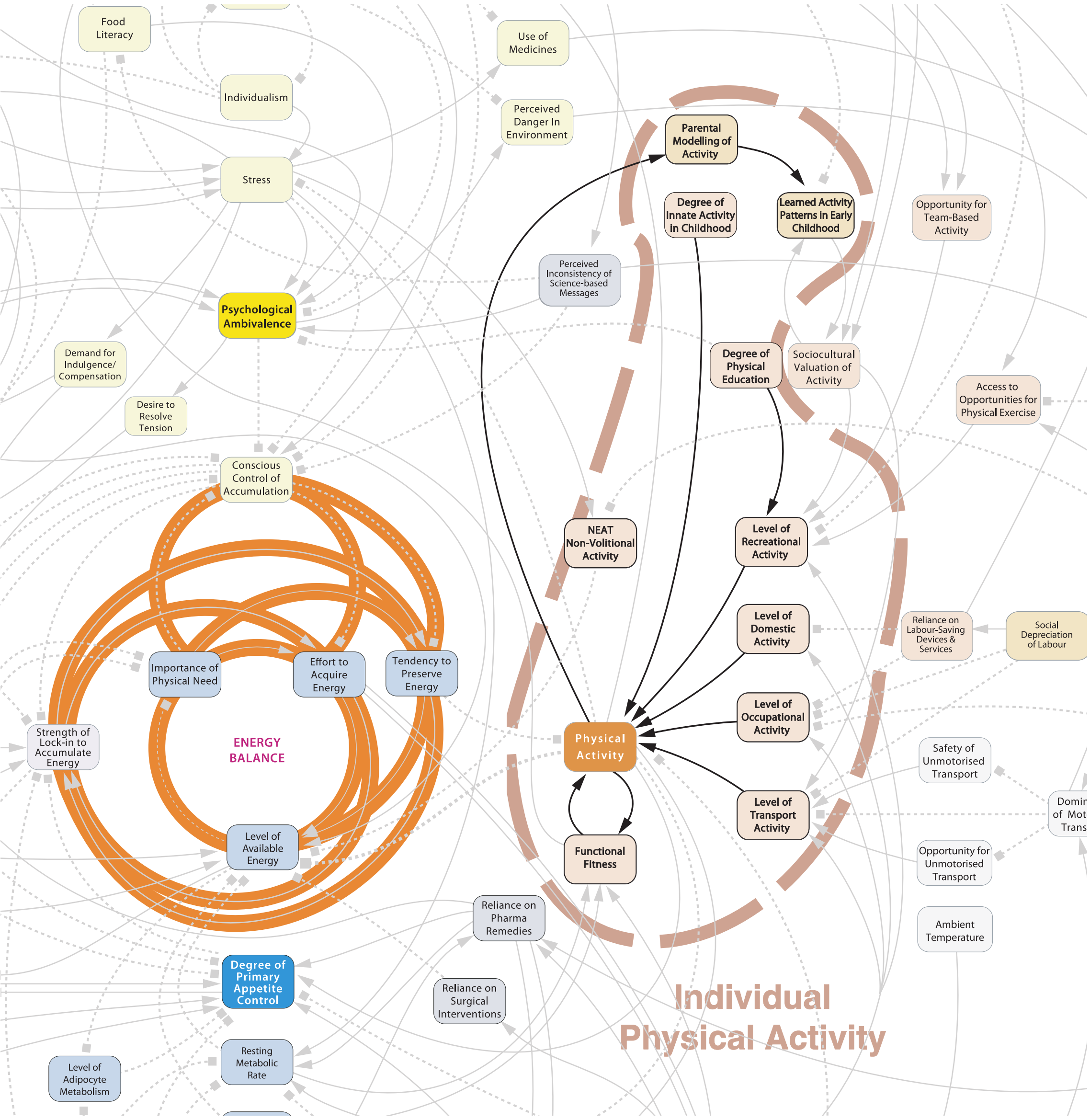


FIGURE 2.5: Individual Activity
[27, Map 7]

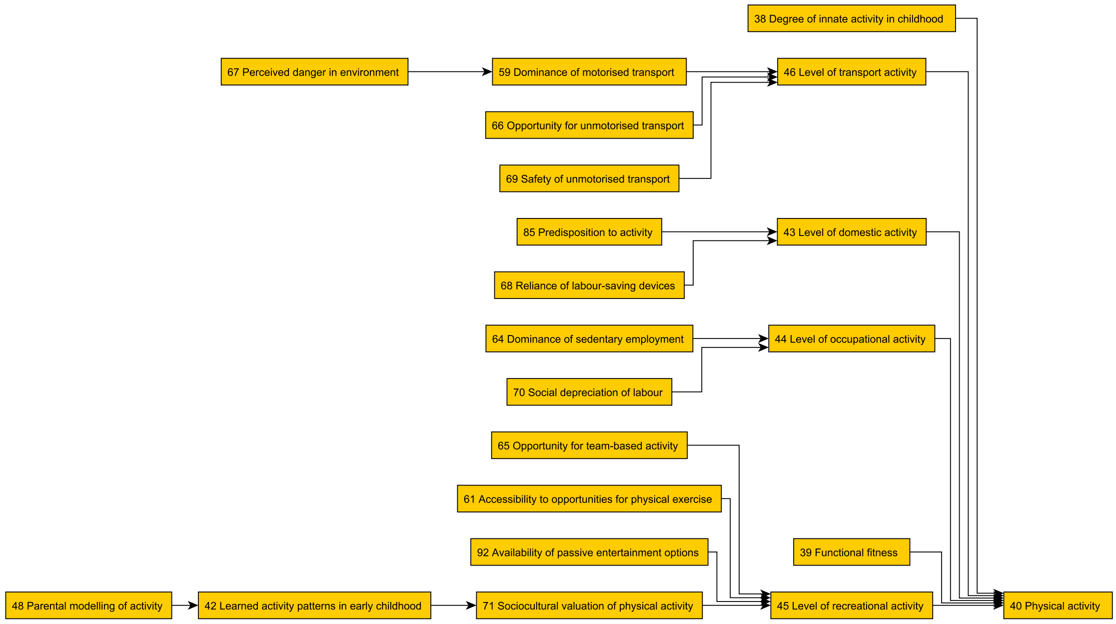


FIGURE 2.6: Physical Activity - Simplified Causal Map

Social and Individual Psychology

The two critical factors within the social element of the psychology seems to cluster around media and acculturation, with the latter including; body-size image, social acceptability, peer pressure and conceptualisation of obesity as a disease. These impact directly or indirectly on the individual psychology cluster which in turn is one of the two primary factors (the other being the food consumption and dietary habits) in influencing the conscious control of accumulation (appetite control). The media cluster links to a number of other areas but is the key/sole contributor 'sociocultural valuation of food' which drives a number of food consumption factors. One of the factors picked up on is the level of individualism within society, this refers to a trend towards more isolated individuals and is inversely linked to levels of 'face-to-face' contact and is a driver of 'stress', which in turn is a key factor for a number of issues and in particular is linked to one of the key variables 'Psychological Ambivalence'.

Figure 2.7 describes the architecture and the underlying causality that defines conflict between the desire to indulge in eating habits that are not healthy and the desire to for long term health. It is derived from Figure 2.8.

Psychological Ambivalence links to Conscious Control of accumulation which provides the key balancing loop within the broader core loop. Two clusters within the broader model relate to the way health and food related information is regarded, and the level and type of control dictating children's diets.

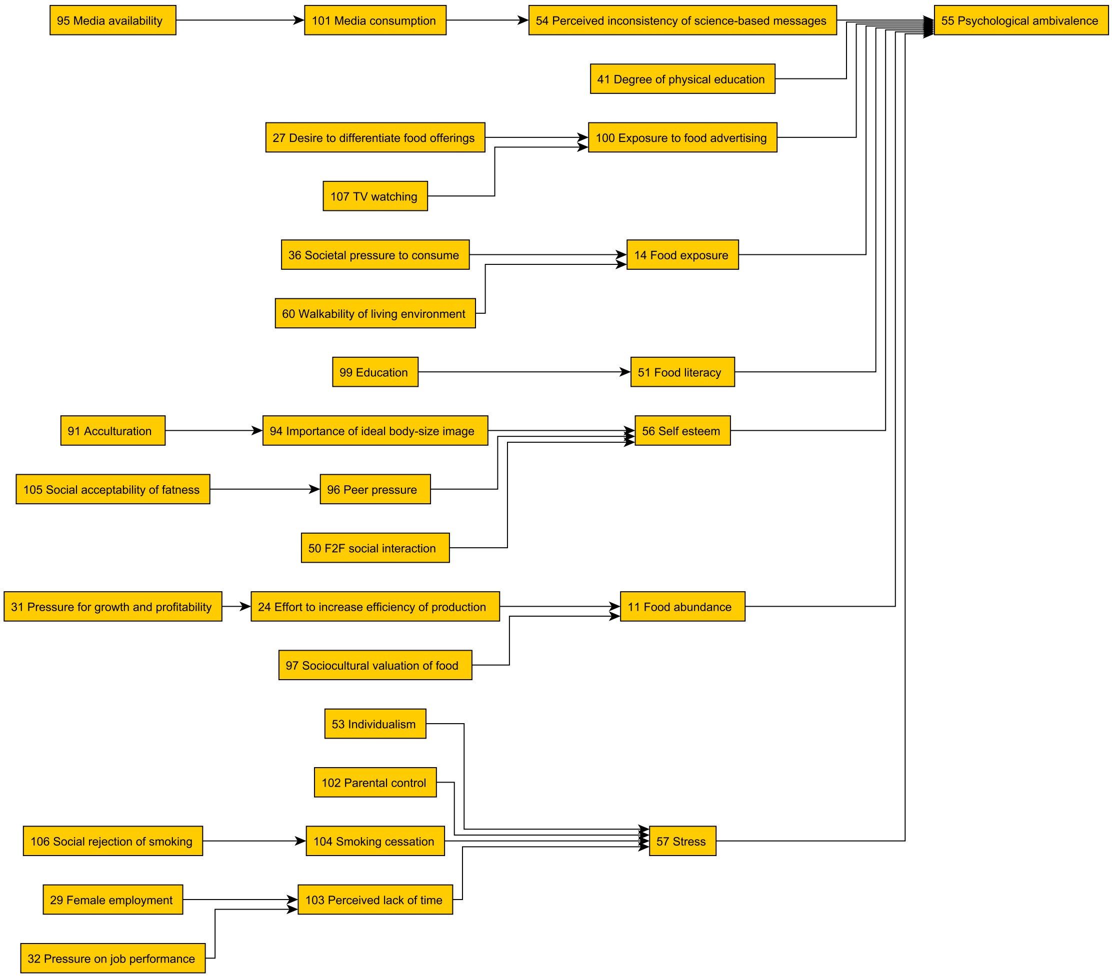


FIGURE 2.7: Psychological Ambivalence - Simplified Causal Map

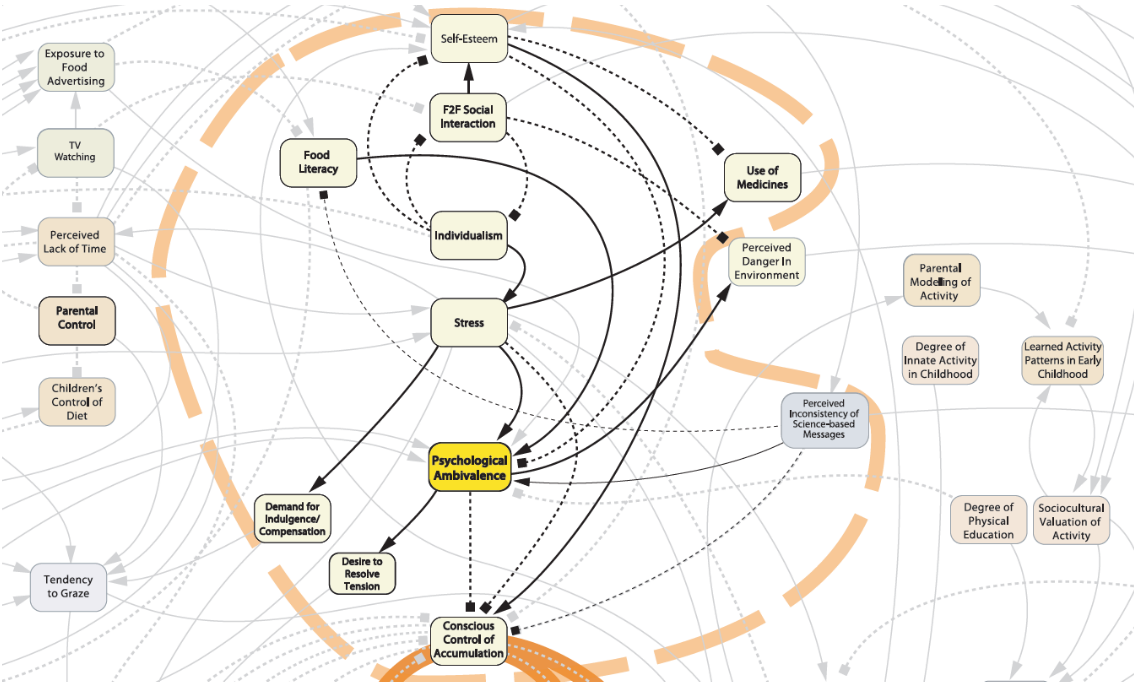


FIGURE 2.8: Psychological Ambivalence - Cluster

This Figure contains public sector information licensed under the Open Government License v3.0.

In a different approach to understanding the impact of individual psychology on obesity, Gerlach, Herpertz, and Loeber [31] conducted a review of the literature describing the links between individual personality traits and obesity, they concluded that in longitudinal studies there was a clear association between neuroticism and impulsivity with obesity and similarly between conscientiousness and restrained eating behaviour.

Vandenbroek, P. Goossens, J. Clemens have identified a number of key variables which in the view of the experts consulted in the preparation of the maps were identified as being most influential:

- Force of Dietary habits - the degree to which behavioural patterns related to food intake are dictated by routine and habit.
- Degree of Primary Appetite Control - the degree to which the brain (non-consciously) responds to signals of the digestive system to control the type and amount of food intake.
- Psychological Ambivalence - degree to which people experience a psychological conflict between what people desire (e.g. fatty, sweet foods) and need to stay healthy.

SES Variables	No controls for other SES variables			Controls for other SES variables		
	Currently smokes	No exercise	BMI obese	Currently smokes	No exercise	BMI obese
Education (years)						
0–11	3.7*	4.9*	1.8*	2.9*	2.8*	1.5*
12	2.7*	3.2*	1.8*	2.4*	2.1*	1.5*
13–15	2.3*	1.8*	1.6*	2.1*	1.4*	1.5*
16+	1.0	1.0	1.0	1.0	1.0	1.0
Occupation						
Labor-Farm	2.2*	3.1*	1.5*	1.2	1.7*	1.2
Protect-Service	1.9*	2.0*	1.7*	1.1	1.2	1.4*
Admin-Sales	1.6*	1.7*	1.4*	1.1	1.2	1.2
Prof-Manager	1.0	1.0	1.0	1.0	1.0	1.0
Income						
Low	2.6*	3.4*	1.5*	1.5*	1.9*	1.2
Middle low	1.5*	2.3*	1.2	1.1	1.6*	1.0
Middle high	1.4*	1.6*	1.3*	1.0	1.2	1.1
High	1.0	1.0	1.0	1.0	1.0	1.0
Unemployed						
Yes	1.6*	0.7*	0.2*	1.6*	0.7*	1.0
No	1.0	1.0	1.0	1.0	1.0	1.0
Housing						
Rent	1.9*	1.5*	1.1	1.5*	1.1	0.9
Own	1.0	1.0	1.0	1.0	1.0	1.0

[a]Controlling for age, age squared, gender, race, and foreign birth.
*$p < 0.001$.

TABLE 2.1: Odds Ratios from Logistic Regression of Health Behaviours on SES Variables
[33]

This Table is reproduced with the kind permission of the Annual Reviews publishing group (License ID 1007420-1).

- Physical Activity - level of physical activity.

2.6.3 Socio-Economic Status

The association between poor health behaviours and socio-economic status (SES) is well established empirically. Historically and in developing countries low socio-economic status has been associated with malnourishment and being underweight, conversely in developed countries it is now associated with obesity and (still) poor diet. Using data from the USA, the table given by Pampel, Krueger, and Denney [33] (Table 2.1) describes the output from a logistic regression approach exploring the relationship. This links obesity, smoking and lack of exercise to a range of socioeconomic indicators via odds ratios.

Whilst the effect is well defined, there are as yet few articles on the mechanisms that drive this disparity. In an effort to address this Cutler and Lleras-Muney

[34] focus on the relationship between education and the adoption and maintenance of healthy behaviours (smoking, drinking, drug use, excessive food consumption etc.) otherwise referred to as the education gradient.

They conclude that the economic resources that are contingent on education account for about 20% of the effect, with improved cognitive ability (as a consequence of education) accounting for 30% of the effect, specific knowledge accounted for 12%, and social integration (peer effects) accounting for 11%. Personality traits and discounting preferences/risk aversion only accounted for 4% of the effect, [34, p. 20] thus educational level accounted for nearly 2/3rds of the observed effect. It is worth noting that cognitive ability is also a factor in health behavioural models, in particular the theory of planned behaviour (TPB) Ajzen [35].

2.7 Clinical Consequences of Obesity

In their analysis for the Global Burden of disease study (2010), Lim, Vos, Flaxman, *et al.* [36] attempt to quantify the health impact of 67 different sets of health risks and how they have changed between 1990 and 2010. To do this they look at attributable deaths, Disability Adjusted Life Years (DALY's) and rankings, both by gender and by region.

The primary area of interest is of course high BMI, but they also categorise the risks for high fasting plasma glucose (a diagnostic for Diabetes) and Physical Inactivity. The latter is defined in terms of MET-minutes (Metabolic Equivalent of Task), with the level for physical inactivity set at < 600 MET-minutes per week.

The conditions list as associated with high BMI are:

- Oesophageal cancer.
- Gallbladder and biliary tract cancer.
- Pancreatic cancer.
- Kidney and other urinary organ cancers.
- Breast cancer.
- Uterine cancer.
- Colon and rectum cancers.

- Diabetes mellitus.
- IHD (coronary heart disease).
- Ischaemic stroke.
- HHD (Hypertensive heart disease).
- The aggregate of cardiomyopathy, myocarditis and endocarditis.
- The aggregate of atrial fibrillation and flutter.
- PVD other CVD (peripheral and cardiovascular disease).
- Chronic Kidney Disease.
- Osteoarthritis.
- Lower back pain.

In 1990 the deaths attributable to high body-mass index for men and women worldwide were estimated at 887,047 and 1,076,502 respectively. In 2010 the equivalent figures were 1,632,766 and 1,738,466 [36, p. 2238].

Similar figures for Disability Adjusted Life Years (1,000's) for men and women were; 25,391 and 26,174 for 1990, and 48,310 and 45,300 for 2010 [36, p. 2241].

In the intervening period high body-mass index moved from the 10th highest risk factor globally to the 6th [36, p. 2246].

Currently (2019) the World Health Organisation (WHO) cites NCD's as the second of it's top ten threats to global health, after 'Air pollution and climate change' WHO [37]. They suggest that NCD's are collectively responsible for 70% of deaths worldwide (41 million people) and 15 million dying prematurely, between the ages of 30 and 69. Of these premature deaths 13 million occur in low- and middle- income countries. Of the five major risk factors they cite as the causes of this issue, three (physical inactivity, the harmful use of alcohol and unhealthy diets) are directly linked to the rise in obesity.

2.8 Impact of Obesity on Health Costs

In 2004 a report estimated the direct economic cost to the Swedish Government for treatment of obesity and its associated conditions was approximately 2% of the total healthcare budget, with indirect costs that included lost earnings and early retirement. The authors also reported that the evidence to support this was confused and conflicting [30].

More generally data on the economic impact is sparse in their review of literature describing the economic impacts of adult obesity, Specchia, Veneziano, Cadeddu, *et al.* [38] found few articles that met their initial search criteria, and none of these met the basic criteria they felt appropriate to be considered robust. Of the articles they reviewed only one considered UK costs, this identified the illnesses potentially linked to obesity (and a number of other behaviours), the proportion of the incidence that could be attributed to obesity, and the cost of treatment of that proportion for 2006-7. For overweight and obesity related conditions the cost was estimated at £5.1 Billion, these are of course direct costs and take no account of indirect costs such as lost earnings, impact on quality of life, early retirement and so on.

In general there seems to be little research attempting to quantify the broader economic impact of obesity on the population. A further search of the literature by the authors yielded 52 results all of which were aimed solely at quantifying the benefit of specific interventions.

2.9 Obesity Initiatives

2.9.1 Medical Strategies

At a macro level [30] carried out a comprehensive literature review in 2004 for the Swedish government looking to understand how increasing levels of obesity within the national population could be addressed, both in terms of treatment and prevention.

They reviewed publications between 1966 and 2001, grading them according to their quality and considering their conclusions accordingly.

In their report they considered a number of elements:

- The context in which treatment and prevention needed to take place, in particular; causes, risks, quality of life and economic impact.
- The success of preventative initiatives.
- Treatment.
- The impact on related conditions.

They considered the causes of obesity amongst the population, which they characterised as a genetic predisposition to obesity amongst a large part of the population which could be triggered by behavioural, cultural or social factors. They also suggest that children are more susceptible to these triggers.

In terms of risks they surmised that the main risk was to those under the age of 64, and that abdominal obesity was a particular issue, being linked to a range of issues, including; type 2 diabetes, raised blood pressure, myocardial infection, gall stones, Sleep Apnea, joint problems, some cancers and infertility.

With regard to quality of life, they reported both physical issues and psychological issues (related to stigmatisation) which were positively correlated with excess weight.

In their review of treatment they discuss dietary, pharmaceutical, surgical and alternative approaches

- The discussion on dietary approaches looks at very low calorie diets, unlimited carbohydrate diets, protein rich diets, lacto-vegetarian approaches and high fibre diets and conclude that overall they are likely to deliver an initial weight loss of between 3-10 Kg in the first year but that this is unlikely to be sustained for the longer term. They also comment that exercise in conjunction with diet can create a greater weight loss. very low calorie diets have the ability to create much greater weight losses (15-20 Kg) but again with no longer term sustainability.
- Under pharmaceutical approaches they consider Orlistat and Sibutramine in comparison with placebos. Orlistat averaged an 8Kg weight loss over two years as opposed to 5kg using a placebo, however there were a number of side effects associated with it's use. For Sibutramine they note that after two years the weight loss was 4kg greater than that associated with the Placebo. (No evidence was available for studies greater than two years.)
- Their consideration of surgery identified a range of techniques, of which the most effective was perceived to be a Gastric Bypass. Typically triggered by a BMI of more than 40, surgery was identified with the most successful outcomes, with 30-40Kg weight losses being sustained over 5 years and 10kg losses being sustained over 10 years. surgery however did have drawbacks with 15% experiencing complications and 2% requiring further surgery. The mortality rate was characterised as less than 0.5%. The report didn't consider that there was sufficient evidence to draw conclusions about alternative approaches.

In the final part of the discussion of treatment the specific impact of initiatives directed at children and adolescents, combining dieting with exercise and counselling was addressed. The evidence suggested that these achieved a 10% weight loss over a year, but there was insufficient information to describe the longer term impacts.

Where data is available, the impact on related conditions is considered and it is suggested that a 5-10Kg loss can halve the onset rate of Type 2 Diabetes and create an improvement in the condition (provided the weight loss is sustained).

There is evidence of some impact on moderately elevated levels of blood pressure (if weight loss is sustained), but the report suggest that there are other more effective treatments available.

Cholesterol (blood lipids) can also be reduced by weight loss but according to the evidence this requires a sustained and significant decrease (20-30Kgs).

The overall picture presented in [30] is discouraging; with causes not fully understood, difficulty in prevention, a lack of effective treatment options (with the exception of surgery) and unclear but significant economic and social costs to contend with.

Linton and James [39] published 5 years later (2009) focuses on examining the treatment options in an American context, it offers a much more detailed breakdown of the bariatric options and the associated risks and issues, as well as the other treatment options. It also addresses their application when dealing with children and the elderly. However, with respect to the efficacy of the different treatment options it reaches similar conclusions to [30].

The National Institute of Clinical Excellence guidelines NICE [40], produced in 2014, do not identify any additional treatment options. Instead a structured approach is recommended, risk is assessed using BMI and waist size in conjunction with any co-morbidities. Treatment is an escalating process based on that risk assessment. Initially a diet and exercise regime is recommended, at higher risk levels this is then supported with pharmaceuticals, at higher risk levels again surgery is considered, until with very high BMI's (more than 50) surgery is the treatment option of choice.

The guidelines also give some guidance as to the cost to society of obesity. Suggesting that in 2007 this was £16 billion or 1% of the UK's GDP [40, p. 5], proportionately this seems considerably higher than that quoted by Ostman,

J. Britton, M. Jonsson [30] even taking into account the fact that the latter were quoting direct health care costs.

Some insight into the direction of travel for treatment is given by the guidelines' 5 research recommendations [40, p. 33], the first 3 of which relate to improving processes and options in relation to bariatric surgery.

2.9.2 Non-Medical Interventions

Examples of deliberate environmental interventions are rare, but include the taxing of drinks that are high in sugar Manyema, Veerman, Chola, *et al.* [41] and Briggs, Mytton, Kehlbacher, *et al.* [42] and the modification of urban environments to promote greater exercise in daily life.

[41] and [42] used econometric modelling, to assess the likely impact of additional taxation of sugar sweetened beverages. Both approaches used national level data to estimate consumption and then data for price-elasticity and cross price-elasticity to estimate the impact of a 20% tax increase on the overall consumption of calories from a range of beverages (sugar sweetened soft drinks, fruit juices, milk etc) for different ages, economic groups and genders. The conclusions were similar, specifically that the primary impact would be on young adults (children were not included in either piece of research), was likely to be independent of income and result in an overall drop in obesity (3.8% for South Africa and 1.3% for the UK).

Given the imposition in 2018 by the UK government of just such a tax, it will be interesting to see if the results bear out their findings.

The positive or negative impact of changes in other environmental factors related to obesity are hard to calibrate, some of the more obvious include:

- The decreasing cost of food with respect to the average household budget [27].
- Rising levels of education [34].
- The decrease in physical effort needed to carry out most occupations [30].
- The changing age profile means that there are more older people, which since age is a factor in BMI, drives up the average BMI of the population accordingly, Thomas, Das, Levine, *et al.* [43].
- Changes in types of food consumed (more calorie dense) and increasing portion sizes, Ello-Martin, Ledikwe, and Rolls [44].

- Media stereotypes and promotion of specific body image ideals [27].
- Increasing social norms with regard to weight and size, Cruwys, Bevelander, and Hermans [45].

2.10 Chapter Summary

When this research was started there was a perception that the rise in obesity was levelling off or had peaked for the adult populations of high-income countries, in fact the rate of change appears to have remained constant.

The underlying causes are complex and inter-linked; the decreasing cost of foods, increased availability, the commercial context in which food is produced, varying perceptions of what is and isn't a healthy weight, changes in the environmental demands of our work and where we live, genetic and epigenetic issues and our innate drive to conserve energy all contribute to the issue. The role played by social networks is also potentially a factor.

The health consequences of obesity are far-reaching both in terms of impact on the individual and society, with a host of associated medical issues for the former and a spiralling demand on healthcare resources for the latter.

Leaving aside the issue of resources, it is also a difficult problem to address success rates for the treatment options are variable and sometimes costly. This may in part be due to the fact that the vast majority of resources are focused on initiatives that address the symptoms rather than the underlying causes.

Chapter 3

Literature Review - Obesity and Simulation

"If you only read the books that everyone else is reading, you can only think what everyone else is thinking."

Haruki Murakami

3.1 Introduction

At the core of the hybrid systems modelling (HSM) approach used to address the research objectives is an hybrid simulation (HS). This chapter reviews current literature, and covers the three key topics (see Figure 3.1) which taken together are critical to successfully addressing those research objectives with such a simulation:

- Social network models in healthcare; looking at the current utilisation and realisation of contact networks in healthcare research.
- Hybrid simulation in healthcare; reviewing the use of hybrid simulation in relation to NCD's generally and obesity in particular.
- Modelling individual behaviour in relation to NCD's; reviewing behavioural modelling approaches, with particular reference to obesity, and the impact of that behaviour.
- The application of these three in combination.

The impact of individual behaviour on their BMI, is addressed by a set of models and equations used to forecast the impact of changes in calorie intake on that BMI. These are described in a separate section along with the rationale for their choice.

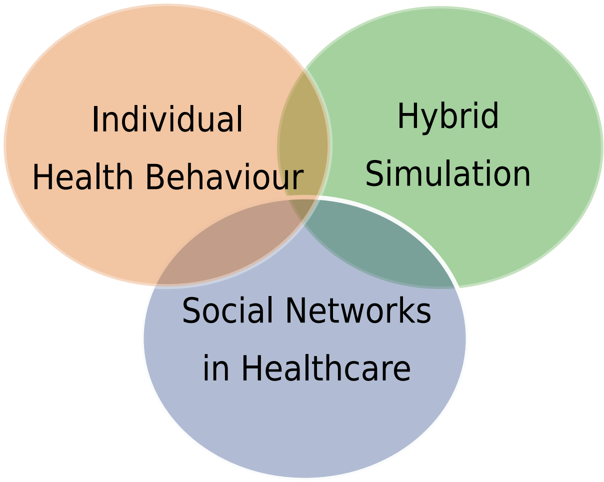


FIGURE 3.1: Literature Review: Focus.

The review was undertaken in order to:

- Provide a foundation of knowledge with which to underpin the work described in this thesis.
- Critically consider the different approaches previously used by researchers when addressing one or more of the key topics.
- Using the output from the critiques above, identify, extend or develop appropriate approaches for the research.
- Suggest future directions for research.

3.2 Methodology

The searches for each of the three key topics identified above, were carried out using Scopus, a search engine provided by Elsevier. In the initial searches only articles were considered (as they are generally subject to a higher level of peer review) and a limited list of journals. This developed a very small database of articles, so in the subsequent 'forward' and 'backwards' searches the restrictions on journal sources were removed. In the case of hybrid simulation, conference proceedings were also considered as they made up a considerable part of the literature. The initial search terms and journal lists are given in Appendix A (Section 2).

A total of 93 articles were identified for the review, addressing one or more of the three themes identified above. These are also listed in Appendix A (Section 1), in alphabetical order with a reference number and are referred to in the remainder of this chapter by their reference number (00), as opposed to articles [00] from the Bibliography. Of the publications 36 had hybrid simulation as their primary focus, 16 health behaviours and the remaining 41 focused on social contact networks.

Every article was then considered from two perspectives, the research aims and the methodology and paradigms used to address those aims. A common set of categories was used to classify the research aims for the articles, where relevant a secondary aim was also identified thus in the case of (87), which explores the use of hybrid simulation in the context of triple bottom line (economic, environmental and societal) sustainability, the primary research aim was 'Hybrid simulation' and the secondary was 'Policy development'.

The categories used, and their frequency of appearance as primary or secondary aims are given (in order of frequency), in Table 3.1.

Research Area	Primary	Secondary
Disease transmission	26	5
Hybrid simulation	13	1
Healthcare behaviour	12	2
Obesity	9	7
Policy development	9	3
Impact of social influence on healthcare issues	7	4
Optimising/Forecasting use of resources	7	2
Medical research	7	1
NCD's (not including obesity)	3	5

TABLE 3.1: Research Categories

A bespoke taxonomy was developed because it was felt that neither of those commonly used, [15] or [16], provided sufficient granularity in the key areas necessary for the purpose of this review.

As is implied by the variation in research aims, there was some overlap between the key topics, Figure 3.2 illustrates this with a Venn diagram. Thus 36 of the articles reviewed addressed hybrid simulation and one of these also addressed individual health behaviour (57), this examined the way linguistic variables affected decision making in healthcare issues. It used a combination of SD and ABM in which the agents are realised as 'fuzzy' entities (using

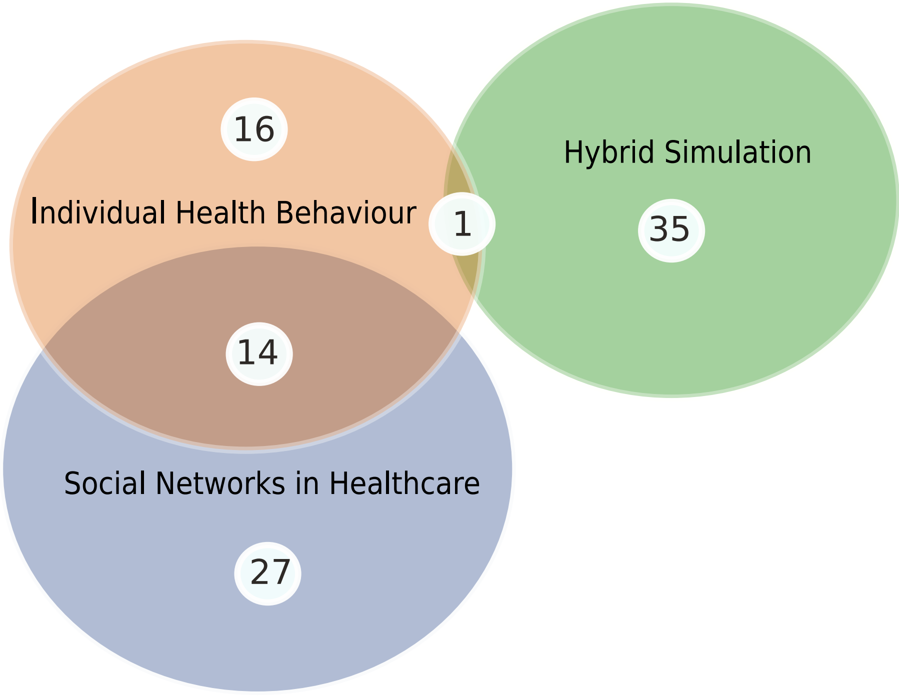


FIGURE 3.2: Overlap of Key Topics.

principles from the paradigms of fuzzy logic) and modelled the impact of aggregate interventions on individuals behaviour.None of the articles reviewed addressed both hybrid simulation and social networks in healthcare.

The methodology and paradigms are specific to the three key topics, and are described in the relevant sections below. Each section describes the relevant theory using a range of references, this is then used to contextualise the methodology and paradigms used in the 93 review articles, the final part of each sections provides a critique.

3.3 Modelling Social Networks

(Social networks differ in a number of ways from other networks. A brief review of their features, relevant metrics,intervention strategies and analysis is given in Appendix E.)

As Table 3.1 shows, research into disease transmission (primarily modelling the progression of epidemics and vaccination strategies) accounts for more than 25% of the primary research aims of the articles. As a consequence the main area of interest in many healthcare applications is contact networks describing face-to-face contacts between individuals, as opposed to those facilitated through electronic means; snapchat, instagram, facebook etc.

As social networks differ in a number of ways from other networks, the features, topology and metrics relevant to them are described briefly in Appendix E.

Modelling social contact networks is often perceived as challenging, Bernstein and O'Brien [46] state the problem succinctly in their paper describing a stochastic ABM replicating a social network:

"Researchers face a trilemma of inadequate data from real world datasets, statistical simulation models, and agent-based simulation models. Large-scale real world data sets are expensive to collect and difficult to obtain high fidelity ground truth for. Statistical models, such as Erdös-Renyi, Chung-Lu, and blockmodels, have parameters that are easy to specify and allow for simple replication of large-scale data sets. What is often missing, however, is the ability to encode narratives into the data because there is no sense of individual agents, just interactions between nodes."

In addition social contact networks vary considerably in topology according to type, Newman and Park [47] compare a scientific collaboration network with that of a board of directors and get a variation of 58% in assortativity values. Anyone attempting to model a social network must address a numberr of issues; apart from the studies already referenced there may be little or no information on the topology of 'obesity' contact networks, they may vary significantly in terms of that topology from other types of contact network and there may be significant variance amongst 'obesity' networks.

Within the literature surveyed 4 broad approaches to network realisation were identified, these are discussed in the following sections.

3.3.1 Geographical Networks

The geographical paradigm is used in nine of the 41 articles on social networks of which 8 (4, 56, 59, 61, 67, 68, 78, 91) looked at disease transmission, the 9th (44) looked at the spread of obesity. In this approach nodes represent locations and/or activities. Edges represent the routes between them. Agents follow schedules which involve spending time at different nodes, with probabilities of interaction/infection varying according to the amount of time and risk associated with each node. In one iteration (56) a geographical network was combined with family network data to form a 'hybrid' network.

3.3.2 Graph Methods

An alternative approach and the most popular (17 out of 41) is to create a graph that replicates the expected network topology. Common algorithms used for this are Erdös-Renyi, the Watts-Strogatz small world model and scale-free models. (See Figure 3.3.)

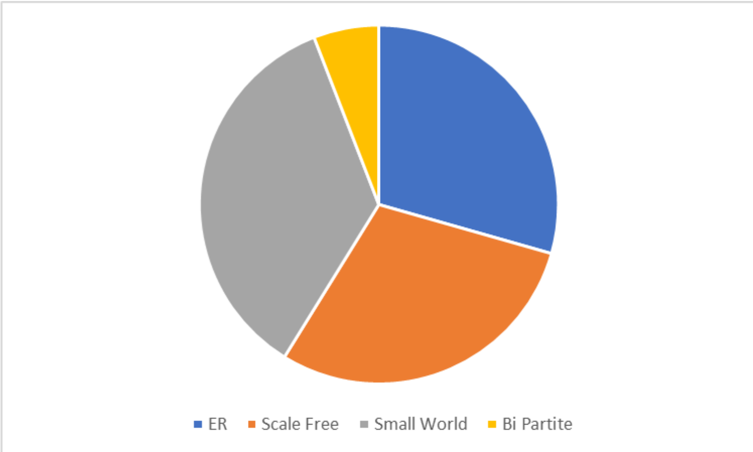


FIGURE 3.3: Frequency of Use - Graph Algorithms

Erdös-Renyi has two manifestations, in the first the number of nodes (n) and edges (m) is specified and the graph is selected at random from all the possible permutations of n and m. In the second (and more common approach) the number of nodes is defined (n) along with the probability (p) of a connection between them, the graph is then generated from these. The latter is commonly used in diffusion models and as a baseline comparator for other models. They are good for mimicking the small diameters often associated with social networks but lack the clustering described in Kiesling, Günther, Stummer, and Wakolbinger [48] that such networks often exhibit . This approach was used by five of the articles (22, 33, 69, 71, 70).

In a Watts-Strogatz small world model, a specified number of nodes (n) are linked to their (k) nearest neighbours, these connections are then re-routed to random nodes with a probability (p), creating a network of loosely linked neighbourhoods of nodes. As (p) approaches 1, this tends to an Erdös-Renyi structure. As described in Watts and Strogatz [49] these are more effective at mimicking the small diameter and clustering of many social networks. Six of the articles used this approach (18, 37, 62, 64, 75, 83).

Many networks also exhibit a characteristic where a small number of individuals (often called hubs) are connected to a much larger number of individuals, with the node degree (number of links connecting a node to other nodes) distribution of the network following a power law distribution [48]. These are known as scale free networks with the degree distribution of the nodes defined by:

$$P(k) \propto k^{-\tau}$$

($P(k)$ is the probability of a node having degree k, and τ is an appropriate exponent, typically with a value between 2.0 and 3.0.)
5 of the documents reviewed (10, 28, 29, 30, 36) followed this approach

In (20) a bi-partite graph (a graph where the nodes make up two disjoint subsets) is used to model intimate contacts in the context of sexually transmitted diseases. Extending the bi-partite approach are stochastic block models, these are not mentioned in any of the articles reviewed, but they capture the communities element described in [47]. They do this by separating the nodes into disjoint subsets, the nodes are then sampled in pairs and a matrix of edge-probabilities (specifying the probability of an edge existing between and within the subsets) is used to generate the edges between those nodes

3.3.3 Real Data

In relatively simple situations where interactions are easy to map, real data can be collected and used to operationalise the network. In all, 10 of the articles used this approach (9, 17, 19, 21, 40, 45, 47, 58, 76, 85). Typically the data was derived from organisational structure and patterns of activity; (85) looked at nosocomial infection and used data on health worker shifts and locations in conjunction with observed data on number of interactions, (40) used classroom data, (58) used interaction data from RFID's, and (45) used data from the Framingham heart study.

3.3.4 Emergent Models

Four of the articles use agent based models to create networks through emergent behaviour (46, 54, 66, 84). A social network friend recommendation algorithm is used in (66), utility/cost functions are used in (54) and (84), (46) used a 'hybrid' approach combining a set of stochastic rules with a small world rewiring process.

3.3.5 Critique - Modelling Social Networks

Where real world panel or longitudinal data exists, then using that data is perhaps the most credible option but it does still raise a number of issues. (46) used data from the Framingham heart study, which as the name implies was actually collected in relation to a study of heart disease. This is sub-optimal since the process artificially limited the number and type of connections that were recorded, and was collected for a different purpose (individuals were asked who they turned to for 'advice').

The graph models also have a number of issues when realising social networks:

- With the possible exception of blockmodels, the 'realisation' mechanism bears little relation to social network constructs. A nodes tendency to form connections is based on probability and distance, and does not take account of constructs such as homophily and propinquity. None of the articles identified made use of blockmodels, and only three referenced the concept of homophily (28, 30, 84).
- There is no obvious mechanism for exhibiting the dynamic behaviour that social networks exhibit over time, with individual connections being made and broken and sometimes remade. Thus for a simulation that runs over an extended time period they are less effective in representing network behaviours. Of the 41 network models reviewed only 6 incorporated this type of behaviour into the model, and only one used a graph model to achieve this (by randomly 're-wiring' connections) (83). The remainder used real data (21, 40, 45), emergent models (46, 66) or a geographic network (67).

As a consequence of these issues, graph models have less face validity, and are perhaps best suited to applications where the time periods are relatively short.

Geographical models provide an alternative way of delivering the dynamic pattern of contacts described above, but are subject to some obvious limitations when applied to large simulations and where the key issue is influence as opposed to contact.

Emergent models seem to offer the most flexibility in delivering networks. Of the four reviewed, three delivered the required dynamic behaviour through the use of utility functions and stochastic rules, the fourth (21) used a similar approach but reflecting some paradigms from social theory; social reach, synergy etc. An approach proposed by Erbach-Schoenberg, Bullock, and Brailsford [50] (studying the formation and longevity of social networks) and cited in (21) takes this further. Agents are given a set of behaviours and parameters which aggregate together to deliver a stable network, which nonetheless exhibits the characteristic dynamic behaviour of 'individuals' within the network. This is achieved by using the concepts of social range, affinity, and memory to define the algorithm's parameters. The algorithm uses these to generate the topology of the emergent network, which remains stable at a 'macro' level, but at the 'micro' (individual) level remains fluid with connections being made and broken continuously as in a real social network. Whilst the approach is based on an homogeneous population, it does have the potential to address a number of the issues raised above.

A more fundamental issue that faces any attempt to model/simulate a large social network, as part of a study into a specific issue, is the difficulty in defining that network. Social networks are often open-system, with fuzzy boundaries, and their topologies vary. Two individuals may both belong to several of the same networks simultaneously, but have different relationships and roles according to the specific network (multiplexity). Additionally such networks can be facilitated using a number of modes; purely face-to-face (contact networks), using electronic communication (email, phone etc.) and across social media.

It is noticeable that in almost all of the articles reviewed for this section, the research either uses real world data and as a consequence is limited specifically to one mode such as face-to-face, or social media, or else there is little or no discussion of the network at all and it is treated as a simple directed or undirected graph with metrics derived from literature.

In reality, as is pointed out in [46], there is very little likelihood of there being literature describing the topography of the specific social network (or even network type) relevant to the subject under investigation. This results in more generalised data being used, with networks using broadly similar parameters. This is in effect a default assumption of homogeneity between networks, when the the limited evidence suggests that this is not the case.

3.4 Simulation in Healthcare

To provide context for the discussion on the role of hybrid simulation in healthcare applications it is worth considering the current applications of the individual simulation paradigms. This is a broad discipline with multiple examples from the four simulation paradigms; Monte Carlo Simulation (MCS), Agent Based Modelling (ABM), System Dynamics (SD) and Discrete Event Simulation (DES) Mustafee, Katsaliaki, and Taylor [51], [16]. In the context of our research we are interested in the applications of SD, ABM and DES.

In 2018 a review of systems modelling in obesity research by Xue, Slivka, Igusa, *et al.* [52] (defining systems modelling as either ABM or SD) identified 35 papers that fulfilled their definition of systems modelling, and addressed some element of obesity research. They also noted that the number of publications in this area has been rising steadily from 2010 peaking in 2016, and that the rate for ABM studies was rising twice as fast as that for SD.

The SD applications included:

- Metabolism simulation Abdel-Hamid [53].
- Body weight dynamics Homer and Hirsch [54].
- Health behaviours Abidin, Mamat, Dangerfield, *et al.* [55].
- Policy analysis Liu, Osgood, Gao, *et al.* [56].

Similarly for ABM:

- Social influence on obesity related outcomes Beheshti, Igusa, and Jones-Smith [57].
- Eating behaviours and food environment Zhang, Shoham, Tesdahl, and Gesell [58].
- Physical activity and the built environment Yang and Diez-Roux [59].

The first of these is an example of the use of ABM to model health related behaviours within the context of a (fixed) social network, as discussed in Section 3.3.

DES is perhaps the most commonly used paradigm in healthcare, some examples of its use include:

- Modelling the flow of patients through emergency departments, Hurwitz, Lee, Lopiano, *et al.* [60].

- Comparing the benefits of two treatment pathways with respect to mother to child transmission of the HIV virus, Rauner, Brailsford, and Flessa [61].
- Comparing the effect of different patient behaviours with respect to attendance at breast cancer screening appointments, on the overall effectiveness of the process, Brailsford, Harper, and Sykes [62].

In the latter it's worth noting that the behaviour under consideration was the choice to attend or not attend (there was no underlying behavioural model). Vieira, Cheng, Harper, and Senna [63] is the only example found in our literature search in which an healthcare application of DES is used in the context of a social network, modelling the impact of decisions made by individuals within that network on the spread of HIV.

3.5 Hybrid Simulation in Healthcare

With respect to hybrid modelling in healthcare, our search identified 36 relevant articles, of which 26 described real applications. Of the remainder two reviewed applications of hybrid modelling in healthcare (1, 14), 2 looked at technical aspects of implementing hybrid simulation (11, 12), and the remainder proposed frameworks or gave guidance for implementing hybrid simulation in healthcare settings (2, 16, 23, 27, 72, 93)

The 26 articles describing applications used the full range of combination options as described in Figure 3.4, these are discussed in more detail below.

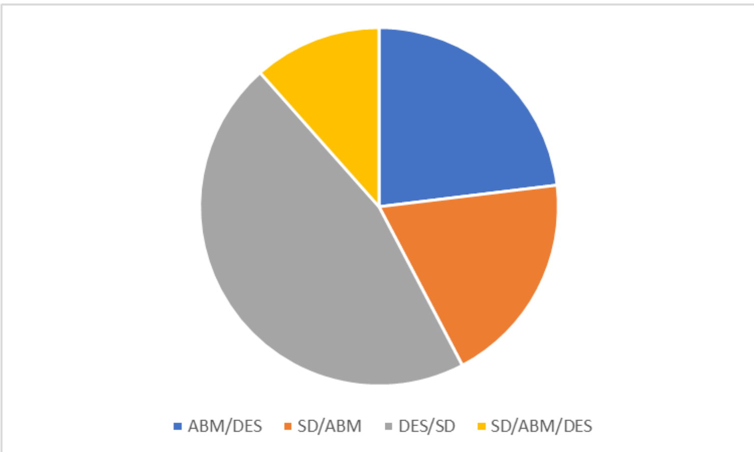


FIGURE 3.4: Frequency of Use - Hybrid Options

3.5.1 ABM/DES Models

Six of the articles described applications combining ABM with DES (3, 7, 38, 49, 55, 80). The research aims looked at utilisation of resources (3, 49, 80), policy (7, 38) and medical outcomes (80). Typically the DES element of the simulation replicated departmental and treatment processes, with the agent based element modelling patient behaviour. The exception to this approach was (55) where the combined elements were used to model heart failure and the surrounding treatment processes.

3.5.2 DES/SD Models

The most common structure for a DES/SD model uses system dynamics to model a changing environment and DES to model the relationship between that environment and the operation of a department/organisation within it e.g.(88). As Figure 3.4 indicates, it is the most popular hybrid paradigm used within the healthcare sector, with 11 of the 26 applications using it (5, 6, 13, 39, 43, 53, 63, 74, 82, 88, 92), one of the reviews (16) focusing solely on it, and with the earliest paper appearing in 2007 (74).

3.5.3 SD/ABM Models

Usefully, Swinerd and Mcnaught [64] define a specific taxonomy for hybrid SD/ABM models, within which they identify three classes:

- Integrated, the design incorporates feedback between the SD and ABM modules in a continuous dynamic process.
- Interfaced, modules run in parallel, with their results combined to produce the required output.
- Sequential, the results from one module once complete are fed to the second, which then delivers the required output

Within the Integrated class, they describe three further sub classes:

- Stocked agents; whereby one or more stock levels within an SD model are defined by an aggregated output from an ABM.
- Parameters with emergent behaviour; one or more parameters are dictated by the emergent behaviour of an ABM.

- Agents with rich internal structure; an SD module or modules are built into each agent to help direct its actions, which are then aggregated to deliver the emergent behaviour.

For the purposes of this review articles that described ABM in conjunction with ordinary differential equations (ODE's) were also considered in this category, two (15, 51) of the five used ODE's rather than explicitly referencing SD models (25, 26, 57). Whilst not completely clear from the model descriptions in the document all five models seem to be integrated with one or more parameters dictated by emergent behaviour.

3.5.4 SD/ABM/DES Models

With only 2 articles this is the least represented architecture in the papers, the first (41) investigates the health and cost impacts and intervention trade-offs for diabetic end-stage renal disease, the second is a generic model for investigating the impact of health technologies (24).

3.5.5 Critique - Hybrid Simulation Models in Healthcare

Historically much of the debate has been about the relevance and practicality of HS in health care Brailsford [65]. With at least one of the common software packages offering the ability to integrate different paradigms relatively easily (AnyLogic) and increasing frequency of use, the debate has moved on to the efficacy of different hybrid combinations in relation to different healthcare applications. Examination of researchers choice with regard to the combination of paradigms (ABM, DES and SD) used in relation to the area being modelled and the effectiveness of that implementation is difficult. The detailed rationale is rarely addressed directly in the articles, similarly other paradigm combinations are rarely discussed. However, reading the hybrid simulation reviews, of which there are a significant number (10 out of the 26 articles addressing hybrid simulation), this is obviously an area of considerable discussion. Whilst two (11, 12) look at the technicalities and software involved in implementing different combinations, the remaining eight (1, 2, 14, 16, 23, 27, 72, 93) seek to provide guidance or frameworks for the selection of hybrid combinations with reference to different healthcare issues.

3.6 Modelling Individual Behaviour

In their review of the implementation of behavioural OR in the healthcare sector Kunc, Harper, and Katsikopoulos [66] divide the field into three areas; behaviour in models, behaviour with models and behaviour beyond models. Consistent with the HSM approach adopted, our area of interest is the first of these, since our desired approach is to represent the behaviours as individuals within the model, responding to a range of external factors by modifying or changing their behaviours (reducing or increasing their calorie intake). In order to achieve this some form of theoretical paradigm supported by operationalisable models is needed. Six such paradigms were identified in the articles reviewed:

1. Information as a dynamic parameter (IDP) was one of the most popular paradigm with 10 examples from the documents reviewed. These used mechanisms that linked behaviour directly to information flowing through the network, thus as the information reaching the individual is modified the behaviour follows. In its simplest iteration this involves 'Follow The Average' (FTA), and thresholds, and in more complex iterations diffusion and cascade models taken from electronic social media (ESM) analysis.

2. Behavioural economics using utility or cost functions were used in three of the articles.

3. Game theory was used in two of the articles.

4. Health behaviour theory was also used in 10 of the articles, it uses a range of models taken from the social sciences in order to explain and/or forecast individual behaviour in relation to healthcare issues.

5. A medical model of addiction was used in one of the articles.

6. Electronic social media (ESM) constructs, analysis of linguistic variables in relation to behaviour and stochastic actor models were used in two of the articles.

Their distribution is described in Figure 3.5.

Figure 3.2 shows a significant overlap between the articles identified addressing the use of social networks in healthcare and those identified addressing health behaviour, 14 articles in total. In reality this overlap is driven almost entirely by the articles on social networks, all bar one of which use IDP and/or

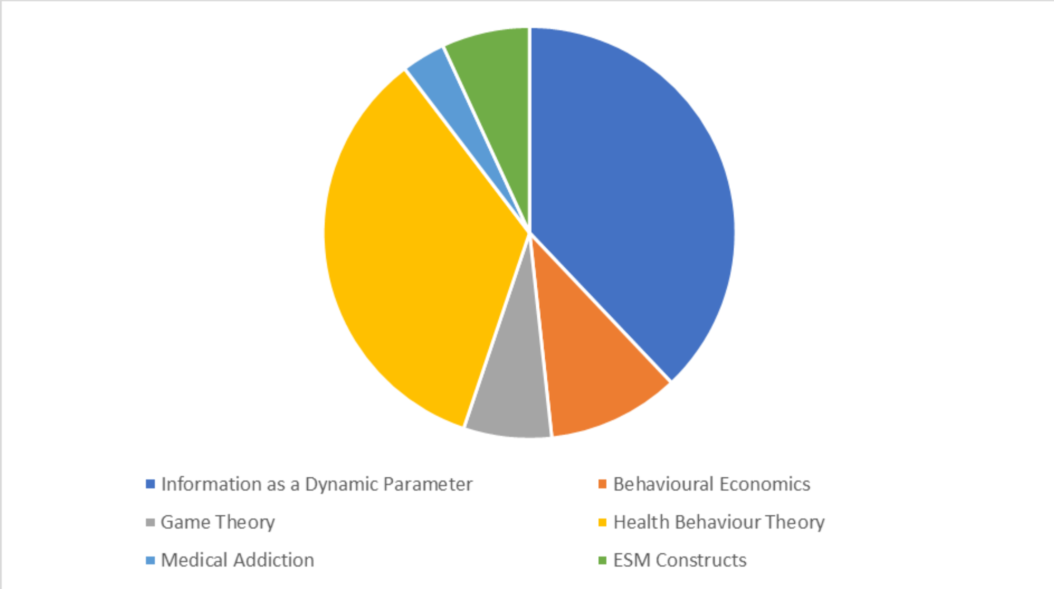


FIGURE 3.5: Use of Different Behavioural Approaches.

utility and objective functions to model individual behaviour. Conversely of the remainder of the articles reviewed, only two use utility functions (8, 81), and none use IDP as an approach.

3.6.1 Information as a Dynamic Parameter

Researchers seeking to understand the spread of information through social networks and its subsequent effect, have used a number of approaches.

Hammond and Ornstein [67] explore the idea of averaging as a mechanism for behavioural adoption, they use the idea of network averaging, whereby each agent exhibits goal seeking behaviour aimed at the network average (FTA), they then refined this by modifying this behaviour to 'local' conformity (averaging across their local network).

However there is little evidence to support its use as a generaliseable rule, the data is taken from research into the spread of obesity amongst children and adolescents, and there is no reason to assume that the same mechanism would apply in other social contexts.

Whilst not directly applicable, the field of Social Decision Theory (which addresses conscious consensus processes amongst groups) suggests a range of strategies including (but not limited to) the two identified above Busemeyer and Diederich [68], again suggesting that there may be other options to consider.

More complex models have been developed by Higham, Grindrod, Mantzaris, *et al.* [69] and Fang, Jen-Hwa Hu, Li, and Tsai [70] in the context of social networks on social media. Both seek to extend the factors considered beyond social influence and include issues such as structural equivalence Burt [71] , confounding factors, interaction intensity and individual attributes. It is clear that much of the conceptual and mathematical frameworks used by them could be applied more broadly.

One algorithm proposed in [70] is:

$$p_v = 1 - \Pi_{u \in U}(1 - p_{u,v})$$

Where p_v is the probability of node v adopting at time $t+1$ and U is the set of v's neighbours who have adopted at time t, and $p_{u,v}$ describes the probability that node u influences v to adopt.

Another branch of the research on network diffusion and cascades, focuses on network topology and its impact on dissemination as described in Centola [72] and Centola [73], the critical issues seem to be:

- The existence of 'weak ties'.
- Transitivity/clustering.
- Group heterogeneity.
- Homophily.
- Consolidation.

The bulk of the research in this area comes from examination of networks enabled by social media, which can provide much more detailed data on the network involved. However there is sufficient early work (eg. [71]) to support the implicit assumption that the same concepts and effects will apply in other social network 'modes'. It's also worth noting that in much of the research described above there is little differentiation between diffusion of information and behaviour in networks, they are treated as similar or conjoined effects.

The default amongst the 11 articles (21, 28, 30, 42, 54, 61, 62, 64, 67, 76, 84) reviewed using this approach, was some variation on FTA, a number used this approach as part of a combination to model the health behaviours. Thus (30) used the probabilities from the research by Christakis and Fowler [5] in combination with categorical information from network neighbours (obese, not obese), (62) uses a more sophisticated neighbourhood averaging function (Management Information Fields) in conjunction with thresholds and a 'relative agreement' algorithm, (28) uses FTA in conjunction with a threshold for the network influence (this is combined with an environmental influence and threshold). (76) used diffusion algorithms derived from ESM's.

3.6.2 Behavioural Economics

The field of behavioural economics provides some alternative approaches to understanding how individuals make many decisions on a day to day basis. It uses insights from the social and behavioural sciences to more accurately describe the processes used by individuals and groups to make decisions in an imperfect world with limited abilities.

The simplest behavioural economics approach, is that of 'heuristics'. An heuristic is a procedure habitually used by an individual to reach a decision in the context of cognitive or information processing constraints, imperfect information and less than perfect ability to access relative probabilities. In essence it is a judgement made about the likelihood of an event or it's complement occurring. It emphasises pragmatic problem solving (i.e. good enough...) rather than optimisation.

West and Brown [74] identify four common heuristics:

- Representation
- Availability
- Affect
- Anchoring and Adjustment

Associated with each of these are issues or biases that can undermine or restrict their effectiveness in that estimation process.

In **Representation** a process of analogy is used comparing the scenario with similar ones to judge likelihood. In practise this heuristic is insensitive to sample size, prior probabilities and on occasion previous evidence, it is also vulnerable to inappropriate use of regression models, overconfidence and the gambler's fallacy.

In the **Availability** heuristic an event is considered more likely if easily recalled, this tends to distort the estimate of likelihood for larger and more frequent events. It is also undermined if certain types of event are more eye-catching, easier to recall or envisage, creating the illusion of greater frequency. Another issue can be the assumption of correlation, wher ein fact it doesn't exist.

In the **Affected** heuristic emotion is used to guide the decision making process, leading to distorted judgement of probability and risk.

Finally in the **Anchoring and Adjustment** heuristic an initial assessment is made which is then re-adjusted in light of subsequent events. This can be undermined by biases in evaluation and assessment, leading to insufficient or over-adjustment accordingly.

Heuristics are an interesting paradigm, and one can envision them being applied in an intervention to influence immediate decisions about food choice, perhaps in an intervention utilising "Nudge Theory", Thaler and Sunstein [75]). Unfortunately there seems to be little literature linking them to decisions around broader health behaviour.

A more complex approach treats behaviour as a series of rational choices driven by current and expected utility for the individual (or cost). Perhaps the most general of these is the Subjective Expected Utility Theory, "The foundations of statistics. By Leonard J. Savage, John Wiley & Sons, Inc., 1954, 294 pp" [76]. This suggests that individuals (and groups) make their choices by considering the immediate and future (discounted appropriately) benefits or utilities to the individual of each option in conjunction with the likelihood of occurrence, and then making the choice accordingly.

The Theory of rational addiction model uses these concepts to derive its descriptive equations and apply them to the specific circumstances of addiction. Proposed by Becker and Murphy [77], a set of equations are derived linking consumption of goods (addiction) to a range of variables derived from concepts used in market economics. Thus developing tolerance to a substance or activity is described as a reduction in utility of those goods.the equations are complex and allow the user to make some sophisticated predictions, however real world data does not consistently reflect the predictions and is sometimes directly contradicted [74, p's. 48-49].

In practise people's ability to process decision making in this way is severely limited both by ability and lack of information, and the fact that individuals actions are more influenced by their current situation than the models allow for [74, p. 78].

Amongst the articles reviewed only 3 used an approach based on behavioural economics (8, 69, 81), of these (69) used the model in an attempt to forecast the take-up of vaccination, and the other 2 related it to weight loss decisions.

3.6.3 Game Theory

Game theory provides another mechanism for modelling decisions, here the environment is seen as a competitive, with an individuals choices predicated on achieving the best outcome for themselves, within the context of the decisions made by the other individuals in that environment. To of the documents reviewed used this as a mechanism for describing behaviour, (50) used it to model the impact of patient choice on resource requirements with regard to knee surgery, (70) used it to model the dynamics of voluntary uptake of vaccination and it's impact on disease eradicability.

3.6.4 Health Behaviour Theory

Health behaviour theory is a body of social theory used to guide the health community in maximising the effect of health interventions, and provides tools for planning and implementing interventions, 10 of the documents reviewed used this approach (31, 32, 34, 35, 48, 60, 65, 73, 77, 79).

Glanz and Rimer [78] describes three levels of influence, see Table 3.2:

Concept	Definition
Intrapersonal Level	Individual characteristics that influence behavior, such as knowledge, attitudes, beliefs, and personality traits
Interpersonal Level	Interpersonal processes and primary groups, including family, friends, and peers that provide social identity, support, and role definition
Community Level	
Institutional Factors	Rules, regulations, policies, and informal structures, which may constrain or promote recommended behaviors
Community Factors	Social networks and norms, or standards, which exist as formal or informal among individuals, groups, and organizations
Public Policy	Local, state, and federal policies and laws that regulate or support healthy actions and practices for disease prevention, early detection, control, and management

TABLE 3.2: Levels of Influence.
[78]
This Table is taken from material developed by the US Department of Health and Human Services (National Institutes of Health) and its use does not constitute endorsement of the material in this thesis or its conclusions by the US Government, Department of Health and Human Services or the National Institutes for Health.

Four of the most widely used models within the intrapersonal (individual) level are [78]:

- Health Belief Model; this examines the perceived threats posed by the health issue, the benefits of avoiding it and the factors that influence decision making
- Stages of Change; looks at individuals motivation and readiness to change behaviours
- Theory of Planned Behaviour; looks at the persons beliefs, attitudes, intentions and perceived control over behaviour
- Precaution Adoption Process Model; describes the stages and individual experiences moving from awareness through behavioural adoption to maintenance

Health belief Model (HBM); this argues that a number of criteria affect an individuals willingness/ability take action to prevent health issues. In its original iteration there were five criteria proposed by Rosenstock [79]:

- Perceived Susceptibility; beliefs about the chances of getting a condition.
- Perceived Severity; beliefs about the seriousness of a condition and its consequences.
- Perceived Benefits; beliefs about the effectiveness of taking action to reduce risk or seriousness.
- Perceived Barriers; beliefs about the material and psychological costs of taking action.
- Cues to Action; factors that 'cue' or encourage initiation of the required actions.

Subsequently a sixth has been added [78]:

- Self-Efficacy; confidence in one's ability to take action.

To use this effectively the practitioner needs to understand how susceptible the individual feels to the issue, whether they view it as serious, whether the action needed to address it is effective and is not unacceptable in terms of 'cost'.

Stages of Change Model (SCM) - This is a model that looks at changing behaviour as a process, defined by different stages. A series of questions are used to ascertain what stage individuals are at in that process, and as a consequence what actions are likely to prove beneficial in moving forward Diclemente [80]. The stages and associated questions are:

1. Precontemplation; are you interested in changing the behaviour?
2. Contemplation; are you thinking about changing the behaviour?
3. Preparation; are you ready to plan how you will change this behaviour?
4. Action; are you trying to change the behaviour?
5. Maintenance; are you sustaining the change in behaviour?

Theory of Planned Behaviour (TPB) - This is an extension of the Theory of Reasoned Action (TRA), and is proposed in Ajzen [35].

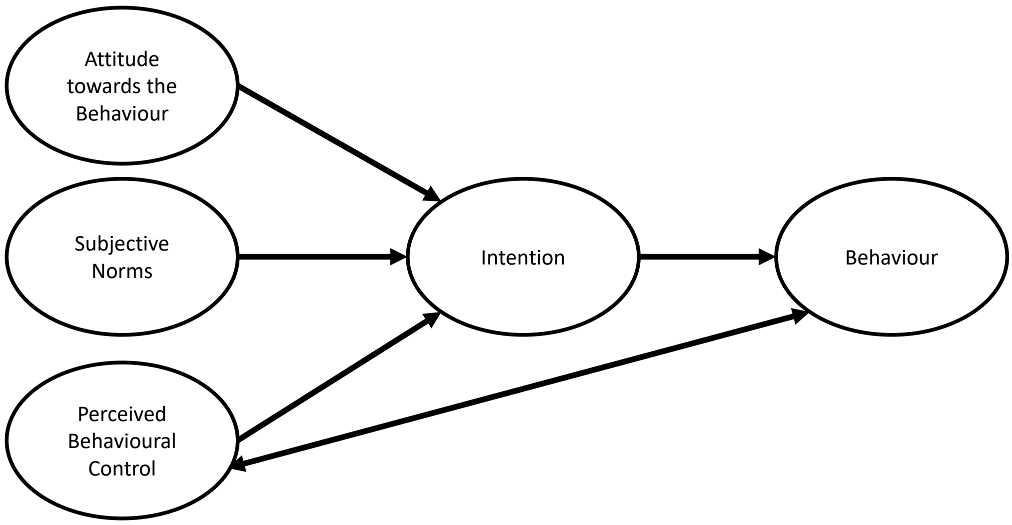


FIGURE 3.6: Theory of Planned Behaviour.

"Intentions to perform behaviours of different kinds can be predicted with high accuracy from attitudes toward the behaviour, subjective norms, and perceived behavioural control; and these intentions, together with perceptions of behavioural control, account for considerable variance in actual behaviour."

Figure 3.6 represents a simplified version of the model first presented by Ajzen, who examines behavioural intention in terms of:

- Intention; the likelihood of the individual carrying out the behaviour.
- Attitude towards the Behaviour; which describes the likelihood of carrying out the behaviour which in turn is driven by the individuals beliefs about what is involved in carrying out the behaviour and its likely outcomes.
- Subjective Norms; beliefs about whether members of the individuals network people would approve of the propose behaviour and whether the individual is motivated to gain their approval.
- Perceived Behavioural Control; the extent to which the individual believes that they have the ability to implement the behaviour. (Note: When the perceived control is high there is an expectation that the behaviour will be implemented)

Precaution Adoption Process Model(PAPM); this describes seven stages in a progression from lack of awareness to adoption or maintenance of a behaviour [78, p. -18], briefly those stages are:

1. Unaware of issue.
2. Un-engaged by issue.
3. Deciding about issue.
4. Decide not to act (stop).
5. Decided to act.
6. Acting.
7. Maintenance.

At first sight this appears broadly similar to the stages of change model, in practise it is more flexible as there is an assumption that once an individual has reached the 4th stage they can move backwards as well as forwards. Separately the stages of change model is primarily focused on the hard to change behaviours frequently linked with NCD's, whereas the PAPM is a more general model looking at things like the decision to have surgery or take medication with difficult side-effects (as well as changing behaviours).

When reviewing the individual models that are grouped under the banner of health behaviour theory, there seem to be two distinct groups. The SCM and the PAPM bear obvious similarities, both in their structure and application. Similarly, the HBM, the TPB (and the TRA from which it was developed), share a number of similar constructs described in the form of beliefs. Given the chronology it is tempting to assume that the latter was in part derived from the former, although if that is the case it is not immediately clear from the references in the relevant papers.

A review of the literature over the last five years would suggest that the TPB and the HBM have attracted the most attention, with broadly similar numbers of results in a SCOPUS search (1,202 v's 1,158) when limited to the healthcare sector and a pattern of slowly increasing numbers of publications. However a broader search makes it clear that the TPB is also used much more widely (3,859 in the same period) compared to the others. The results for SCM and PAPM were 523 and 31 respectively.

The documents identified for this review follow the same pattern with seven of them (32, 35, 60, 65, 73, 77, 79) using implementations of TPB, (32) uses

HBM and the remaining two (34, 48) comparing the efficacy of different health behaviour models. Tellingly all 10 articles used the models in the context of health risk behaviours and NCD's.

3.6.5 Addiction - A medical Perspective

Addiction provides a potentially compelling paradigm for modelling over-eating. In the introduction to [74, p. 7] addiction is described as; "a chronic condition... ...in which there is an abnormally and damagingly high priority given to a particular activity" They go on to describe three underlying pathologies:

- Abnormalities in the motivational system caused by issues not related to the actual activity; stress, depression, low impulse-control, chronic anxiety and so on.
- Abnormalities in the motivational system caused by the addictive substance/activity itself; acting on the subject; tolerance and withdrawal symptoms, sensitisation to the effects of substances.
- 'Toxic' environments acting on normal motivational systems that are not equipped to deal with them.

(In this context motivational system refers to the set of brain processes that energise and direct our actions.)

In their commentary Taylor, Curtis, and Davis [81] describe the rationale for classifying some manifestations of compulsive overeating as an addiction. The two key points that they make are the similarity between the symptoms that characterise compulsive over-eating and those described in the Diagnostic and Statistical Manual of Mental Disorders (4th edition) and the phenomena of 'transfer of addictions' (whereby the patient then starts to exhibit other compulsive behaviours such as gambling or over-spending) exhibited by a subset of patients who have undergone gastric surgery [81, p. 327]

The concept of food addiction is explored further by Gearhardt, Corbin, and Brownell [82] who developed a behavioural questionnaire - The Yale Food Addiction Scale (YFAS).

Their purpose in doing this was to validate a mechanism for identifying those amongst the obese population who have lost control of their eating behaviour as opposed to those who simply indulge in unhealthy foods. This was done with a view to ensuring that where relevant treatment takes account of the pathological behaviour [82, p. 435].

It's worth noting that in their discussion they also reference the potentially negative impact of constant food advertising and the ubiquitous nature of unhealthy foods as significant factors in reducing the impact of public health interventions.

The YFAS has a number of limitations; it was validated in a population of college students and thus needs to be examined in other populations, the sample was to a certain extent self selecting, there was a relatively low number of obese respondents within the sample and BMI and height were self-reported and hence potentially under and over-reported accordingly.

The concept of addiction as the main driver in Binge Eating Disorders (BED) and as a contributing factor for obesity more generally is developed by Davis, Curtis, Levitan, *et al.* [83] who cite a range of evidence to support the hypothesis, and then describe the process they used to extend the validity of the YFAS.

Using a cohort of 72 obese individuals and after considerable statistical analysis the YFAS classified 18 as food addicts [83, p. 714]. The analysis involved looking for correlation between demographic features, clinical features, personality traits and eating behaviours. whilst the p values for demographic features were all $> .05$ the remainder all showed some statistical relevance. There was also a considerable overlap with 50% of those exhibiting binge eating behaviours also classified as food addicts, and 70% of food addicts exhibiting binge eating behaviours.

Regression analysis of the data yielded a model that included; addictive traits, hedonic eating, snacking on sweets and binge eating. The model had an R^2 value of 0.56. They concluded that whilst binge eating is one mode of food addiction it is not the only one and that there are other consumption patterns that lead to food dependence and impairment [83].

The validity of the medical model of addiction in relation to obesity is tested by Wilson [84] and Ziauddeen, Farooqi, and Fletcher [85]. The latter suggests that there are five key pieces of evidence cited in support of the addiction model that need to be considered.

- A clinical overlap between obesity(or, more specifically, BED) and drug addiction
- Evidence of shared vulnerability to both obesity and substance addiction

- Evidence of tolerance, withdrawal and compulsive food-seeking in animal models of overexposure to high-sugar and/or high-fat diets
- Evidence of lower levels of striatal dopamine receptors (similar to findings in patients with drug addiction) in obese humans
- Evidence of altered brain responses to food-related stimuli in obese individuals compared with non-obese controls in functional imaging studies.

These are then examined in turn. In the first a comparison is made with the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) which defines substance abuse by the presence of characteristic patterns of behaviour in Table 3.3 in order to ascertain the level of clinical overlap between the two conditions.

After examining each of the seven criteria relevant to substance abuse they concluded that only three of the criteria can be mapped across to the general behaviours that drive obesity and that this is insufficient for the assumption of addiction and it's inclusion in clinical consideration of treatment. However, in the specific instance of BED, they suggested that this may be appropriate in those cases where the severity and impairment thresholds mandated by the YFAS are met.

With regard to the concept of **shared vulnerabilities** between obesity and substance addiction, one set of supporting evidence comes from family studies indicating a common genetic susceptibility to both issues. An initial study looking at genetic variation in dopamine receptors in association with various types of substance abuse suggest that this is the case, but subsequent studies have failed to replicate this.

Research into personality traits in particular impulsivity, does show a "modest association" with the same genes, and has also been shown to be higher in obese individuals and those with BED.

Studies into reward sensitivity where it is suggested that poor reward sensitivity in the brain is associated with substance addiction is also applicable to obesity, fail to address the fact that some people overeat as a consequence of "enhanced sensitivity to the hedonic aspect of food".

There is then some modest evidence to suggest shared vulnerabilities between substance abuse and obesity, but it is not at all clear that the underlying mechanisms are similar and that therefore similar clinical approaches would have similar effects.

DSM-IV criteria for substance dependence	Proposed food-addiction equivalent*	Comment
Tolerance: increasing amounts of drug are required to reach intoxication	Tolerance: increasing amounts of food are required to reach satiety	Not a convincing equivalent to drug tolerance because it assumes an equivalence between satiety and intoxication. In addition, key characteristics of binges are eating in the absence of hunger and to the point of physical discomfort (beyond satiety)
Withdrawal symptoms on drug discontinuation, including dysphoria and autonomic symptoms such as shakes and sweats	Distress and dysphoria during dieting	No convincing evidence of a human withdrawal syndrome for foods
Persistent desire for and unsuccessful attempts to cut drug use	Persistent desire for food and unsuccessful attempts to curtail the amount of food eaten	This criterion requires the application of severity and impairment thresholds to be meaningful
Larger amounts of drug taken than intended	Larger amounts of food eaten than intended	This criterion requires the application of severity and impairment thresholds to be meaningful
A great deal of time is spent on getting the drug, using the substance or recovering from it	A great deal of time is spent eating	It is difficult to apply this criterion because of the easy availability of foods in most developed societies
Important social, occupational or recreational activities are given up or reduced because of substance abuse	Activities are given up through fear of rejection because of obesity	A strict equivalence would require engagement in eating to the exclusion of other activities
Substance use is continued despite knowledge of having a persistent or recurrent physical or psychological problem caused or exacerbated by the drug	Overeating is maintained despite knowledge of adverse physical and psychological consequences caused by excessive food consumption	This criterion requires the application of severity and impairment thresholds to be meaningful

TABLE 3.3: Modelling Food addiction on Substance Dependence. [85, p. 280]

Evidence from animal models is frequently cited as evidence of food addiction, [85, p. 282] describes the process whereby rats can be induced to compulsively overeat when presented with high sugar and high fat diets, and subsequently exhibit withdrawal symptoms consistent with addiction. There are however some significant differences, not least the need for very specific food presentation to develop the addictive behaviours, which is very different to the variety of environments encountered by people. They concluded that the degree to which such models can be extended to human obesity needs to be explored further.

An initial **Dopamine receptor study** seemed to suggest that reduced dopamine ligand binding was associated with obesity this has not subsequently been replicated [85, p. 283], instead the picture developed by succeeding studies is much more complex, with no clear picture emerging even when the focus is narrowed to those with BED.

Functional neuroimaging is a useful mechanism for testing the obesity as addiction hypothesis, predicting that responses to foods and food related stimuli should result in consistent perturbation in reward related regions of the brain. This does not seem to be the case with increasing numbers of studies failing to find evidence to support a common view of obesity and overeating and even when the studies are limited to those with BED there is no consistent abnormal pattern of response [85, p. 283].

[85, p. 285] argues that the vast majority of obese individuals do not exhibit the behavioural or neurobiological profiles that go with addiction and that obesity as a health issue is characterised by high levels of heterogenity. as a consequence the attempt to apply a single model is likely to be ineffective. They also suggest that even if the focus is purely on BED the the evidence is at best inconsistent and weak.

[84] makes many of the same arguments, additionally he points out the difference in relapse patterns where the chances of relapse are inversely correlated with the time since the last relapse in cases involving heroin, nicotine and alcohol by comparison to obesity where the pattern of relapse is remarkably consistent, initial rapid rate of weight loss declining over a six month period, then slow regain until the weight stabilises again at a place somewhat lower than the start point [84, p. 345-6]. This pattern is contrasted with those suffering from eating disorders, Bulimia Nervosa (BN) and BED. where after appropriate treatment weights are frequently maintained without relapse, suggesting

that BN and BED are the consequence of different mechanisms to those inducing obesity. Similarly he cites the difference in relapse patterns as further evidence of the differences between addiction and obesity.

In conclusion it's clear that addiction does not of itself provide a clear model for obesity as a whole, it may provide models of treatment for part of the BED population but even here the application is limited. The issues are summarised in Table 3.3

Despite the high level of research that has gone into addiction as a mechanism for understanding behaviours in the context of non-communicable chronic diseases (NCD's), surprisingly only one of the documents appeared to use it (90). This was in the context of exploring policy interventions with respect to non-medical opioid use.

3.6.6 ESM Constructs

Perhaps inevitably the growth of algorithms to analyse electronic social media and internet applications has had an impact on this area, cascade and diffusion algorithms are addressed in the section on information as a dynamic parameter, but there are a range of other algorithms and applications to consider. Two of the documents reviewed use such algorithms. (40) develops new friendship selection algorithms with which to forecast the spread of smoking within an adolescent network. By contrast, (57) analyses linguistic variables with a view to understanding how perceptions then impact on the implementation of health policy

3.6.7 Critique - Models of Behaviour

As the options for modelling decision behaviour described in the preceding sections suggest, there are a number of approaches that could be used to further the research aims, their advantages and disadvantages are discussed below.

It's clear that whilst the medical perspective on addiction addresses the behavioural patterns of some of the elements of the obese population (those suffering from BN and BED) it is not effective at describing the behaviours of the majority [84] [85].

Behavioural economics and the theory of rational addiction are hindered by their assumptions of complete information and rational decision making and also by their failure to make sufficient allowance for current circumstances. As Baddeley [86] suggests, the rational and logical processes are based on an assumption of full knowledge and the ability to process and optimise that information, and fail to address the imperfect world and abilities of the people in the world in which we live. Heuristics are an attractive set of tools, but operate at too granular a level to be easily incorporated into a simulation involving multiple individuals running over a decade.

Within the health behaviour category given the publication numbers, it is clear that the TPB and the HBM are currently perceived to be of more utility than SCM. PAPM seems to have gained little traction, perhaps because of the crossover with SCM.

TPB is a flexible approach, aside from the healthcare examples listed in the review documents, it has also been used for a range of other applications, for example:

- Investigating academic dishonesty amongst business school students Hendy and Montargot [87].
- Modelling consumer behaviour with respect to plastic waste Khan, Ahmed, and Najmi [88].

Examples of the application of the HBM not picked up in our document search but illustrative of it's applications are:

- Exploration of the cultural factors and attitudes amongst nurses affecting their take up of influenza vaccinations Kwok, Li, Lee, *et al.* [89].
- Analysing instagram posts about Zika virus to understand their impact on the recipients behaviourGuidry, Carlyle, LaRose, *et al.* [90].
- A study exploring the attraction of Yoga for veterans with PTSD symptoms Cushing, Braun, and Alden [91].
- Providing the framework for an investigation into behavioural intentions amongst high school students in terms of avoiding cervical cancer Shida, Kuwana, and Takahashi [92].

The nature of the models describing both the TPB and HBM may well account for both their popularity and their usage.

The HBM provides a series of criteria or a checklist with which to benchmark an intervention, identifying 'blockages' or areas of weakness which can be addressed through the appropriate remedial action.

By contrast the TPB offers an additional element, the different factors can be used in the same way as the HBM, but the model also describes a set of relationships amongst those factors, lending itself to modelling and forecasting as well as analysis. This is reflected in the examples of usage listed above.

Information as a dynamic parameter is another credible approach, with multiple examples in the literature and a strong body of current research primarily driven by investigation of diffusion and cascade processes in ESM (eg. [72], [73] and Centola and Rijt [93]). Where the qualitative investigation of behaviour is not the main focus of interest, but the quantitative impact of it is, then these are probably the most effective tools to use (in accordance with the principle of minimising unnecessary complexity). Conversely where some form of narrative about behaviour is of interest then health belief models might be more appropriate.

Where used appropriately game theory and ESM constructs, seem to offer a range of more specific applications, that may also be of use in this field.

3.7 Modelling Changes in BMI

A necessary adjunct to the behavioural element of the model is the impact of changing behaviour on individual weight gain or weight loss. This is typically addressed through energy balance equations.

In contrast to some of the other areas in this review there is a great deal of literature addressing the formulation of these equations, with much of the initial work based on that done by Forbes [94]. More recently key contributors include Hall, Thomas and Henry, all three of whom have used data to produce regression models designed to forecast individual weight gain or loss based on changes in calorie intake.
(Thomas, Martin, Heymsfield, *et al.* [95], Thomas, Ivanescu, Martin, *et al.* [96], Hall [97], Hall [98], Hall, Chandramohan, Chow, *et al.* [99], Westerterp, Donkers, Fredrix, *et al.* [100], and Henry C J [101].)

The critical issue with much of the work is obtaining data sets on which to base the modelling, many articles still use data collected during trials in the 1940's.

The terms of reference for the report produced by the UK Scientific Advisory Committee on Nutrition [102] included:

- Reviewing and agreeing methods for defining human energy requirements.
- Agreeing a framework for arriving at the energy requirements for the UK population and its subgroups.
- Agree dietary reference values for the population taking into account age, weight, height, levels of activity, gender and physiological state (pregnancy etc.)

To deliver these the report authors uses the equations and the recalculated basal metabolic rate (BMR), data presented in [101]. This updated earlier values taking into account the factors listed above. Henry then combined this with data on physical activity levels (PAL) to calculate total energy expenditure (TEE).

$$TEE = PAL * BMR$$

$$BMR = (\alpha * weight) + (\beta * height)$$

(Where α is a weight coefficient and β is a height coefficient and both vary with age and gender)

These are the equations that are used to model changing BMI for the research.

Whilst the recommendations of the SACN cannot be ignored, there still remain an number of concerns/shortcomings that need to be acknowledged when modelling BMI.

Any health based research into the obesity is rendered more complex by the fact that much of the medical work done in this field considers the ratio of Fat Mass(FM) to Fat Free Mass (FFM), whilst as previously discussed much if not all of the general data on obesity is recorded as BMI, which only takes account of Body Weight (BW). This is problematic, both because the ratio of FM to BW is not a constant but varies with BW, gender and age [97] and also because for a number of reasons BMI is not necessarily an adequate measure of clinical risk. Meeuwsen, Horgan, and Elia [103, p. 560], suggest a number of issues:

- BMI focuses on body mass, since muscle mass is denser than fat mass then for a number of professional groups where development of musculature is important (athletes, manual workers, servicemen) this can lead to false diagnoses of over-weight and obesity.
- As individuals age the proportion of FM to FFM changes, thus for an ageing individual who maintains a constant BMI their proportion of FM to FFM is increasing, this effect also varies between genders
- In children and adolescents patterns of short term growth means that BMI can on occasion be misleading.

There are also a number of additional elements that need to be considered when modelling weight loss for individuals over an extended time period:

- Thomas, Martin, Redman, *et al.* [104] addresses the issue that a conscious decision to reduce calorie intake (diet) is often only intermittently adhered to , resulting in the cyclical weight loss patterns and plateaux observed in practise.
- [67] introduces the concept of ’Satisficing Behaviour’ and a related Satisfaction Interval (SI), in the form of a rule that an individual will only attempt to gain or lose weight if their BMI and the target BMI are separated by a value greater than some specified amount.
- A rule of thumb used in many weight loss studies and scenarios equates a daily shortfall of 2Mj per day to a weekly loss of 0.5 Kg (alternatively a weekly shortfall of 3,500 kcal equates to a 1lb weight loss). However as is pointed out in [99], this approach fails to compensate for the reduction in metabolic rate associated with weight loss and so starts to over-estimate energy expenditure resulting and hence likely weight loss. As an alternative he provides a comprehensive and complex set of equations that address this, modelling an initial weight loss phase related to changes in fluid and glycogen levels and then a second phase focusing on body mass (FFM and FM), Chow and Hall [105].

Any weight loss modelling must consider the implications of these elements.

3.8 Summary

This review has of necessity looked at a broad range of subjects, both as a consequence of the research aims and the methodology chosen to address them.

The critiques from each of the relevant sections are reviewed below, along with their impact on the choices made in relation to the model and its architecture:

- A default approach when modelling the spread of disease through a network, is to generate a large number of random graphs using an appropriate algorithm (e-r, small-world, scale-free...), and then simulate the spread of the disease through each network, creating a database from which statistical valid conclusions can be drawn. This assumes that the time periods are short enough that the networks can be treated as static. Given the length of the timeframes under consideration, this is not valid when addressing the issue of obesity [10], so an alternative approach is needed. [50] provides a basis from which a more appropriate mechanism could be developed.

- Identifying an appropriate network topology constitutes a significant challenge, the multiplexity exhibited by real-world networks means that neither social contact networks or those facilitated by social media can be assumed to be exclusively involved in the effect under examination. Geographical networks would enable us to side-step the issue, but are not feasible with the type of data available to us. Pragmatically the actual network topology defined by the model development process must be assumed to be a representation of the combined networks and their effects.

- The theory of planned behaviour [35] is the most used approach for modelling individual behaviour in relation to healthcare, but there is very little quantitative information on its application. Its use in the simulation model would offer the opportunity to add to the body of knowledge in this area.

- Hybrid simulation is still relatively new to healthcare OR, however it potentially offers the ability to model the interaction of two quite different processes operating simultaneously which is critical to the research. Given the emphasis on flow and potential for feedback loops, TPB lends itself most appropriately to the use of SD as a representative methodology. ABM lends itself to the realisation of an emergent (social) network model, with its emphasis on individual 'rules of behaviour' it is able to deliver the dual aspects of that network, dynamic at an individual level whilst remaining topologically stable at a macro level. As yet there are few published examples of this configuration, however examples from other sectors suggest that the chosen configuration is suitable to address the research aims.

- The impact of calorie intake reduction on body mass is well-researched but is hindered by issues with data. The most current relevant data (UK population) is probably that used by SACN. However, any modelling attempt must also address the issues of intermittency in dietary adherence and 'satisficing behaviour', identified in [104] and [67] respectively.
- The challenges of building a useful simulation both in terms of lack of data and complexity are clear. The approach used to address these difficulties can be broken into two parts:
  - Building a simulation that is capable of recreating the key 'behavioural' patterns of the system at different levels of granularity, to ensure that it is capable of capturing sufficient of the system complexity to address the research questions.
  - Using a 'supervised learning' approach to parameterise the model utilising an appropriate stochastic approximation algorithm.

3.9 Future Directions

This review has highlighted a number of areas where future research may be of significant value:

- Multiplexity is a clearly acknowledged feature of social interaction, individuals belong to many of the same networks, with potentially differing roles and responsibilities in each, and with varying methods of interaction. However, the concept remains unacknowledged in any of the articles reviewed, and is clearly relevant to many of them. Some working hypotheses about how to address the issue, would certainly add value.
- Network realisations using emergent behaviour clearly have the potential to deliver much more sophisticated social networks than graph models, but as yet are relatively rare. The development, testing and publication of emergent models aimed at different categories of use would provide a structure in which progress in this field could be made.

Chapter 4

Simulation Model

"Essentially, all models are wrong, but some are useful..."

George E. P. Box

4.1 Introduction

Robinson [106] suggests that a simulation study can be thought of as four activities which deliver the outputs required for a successful outcome. The four activities are:

- Conceptual modelling - using data from the real world to build a model that addresses the research aims
- Model coding - translating the conceptual model into a computer model
- Experimentation - using the computer model to explore the problem underlying the research aims, perhaps through a range of scenarios
- Implementation - implementing the learning from the experimentation phase back in the real world

Whilst it would be tempting to view these activities as a linear process, [106] makes the point that it is reversible, with occasions when the output from one activity requires re-consideration of its predecessors (e.g. the output from a computer model doesn't match that of the world, requiring re-consideration of the conceptual model on which it is based). It is also cyclic, often with a number of iterations required before the improvement in the real world issue under consideration is achieved.

The remainder of this chapter seeks to describe the output from the first two activities, giving a comprehensive description of the conceptual model and in conjunction with Appendix B, the computer model used to realise it. The results of the experimentation are described in Chapter 5 and the implications for implementation are addressed in Chapter 6.

Historically the descriptions of many ABM's described in the literature was incomplete making reproducability an issue for other researchers as commented on in Railsback and Grimm [107, p. 36], who go on to describe an appropriate protocol to standardise such descriptions. [19] builds on this and other approaches to develop standardised protocols for ABM, DES and SD. We seek to provide a comprehensive description of our model by combining their protocols for both SD and ABM.

4.2 Conceptual Approach

It is a generally accepted principle of simulation design that the model should be as simple as possible, whilst retaining the ability to address the research questions that motivated it Pidd [108] and [106]. The research questions act as the filter to decide what should be replicated in the model and what should be omitted.

In more complex systems where the questions themselves may not provide enough information to adequately specify the model design, [107] suggests an approach from the field of ecology modelling referred to as pattern oriented modelling (POM). This is the use of patterns observed in the system being modelled as additional information to make the models more structurally realistic, and hence better able to address the research questions. They are clear that this isn't a new approach but suggest that as modelling of progressively more complex systems is attempted, the concept is increasingly relevant. This approach requires the modeller to identify relevant patterns in the system at different levels of granularity and ensure that the various components of the simulation replicate them, reproducing the necessary level of complexity needed for an effective simulation.

The issue of calibration is also likely to be significant (Nianogo and Arah [18]), especially in cases where the output is stochastic. Spall [109]) suggests a number of approaches to address these include the use of heuristics, machine learning algorithms and stochastic optimisation methods.

In addition to the more accepted iterative modeling processes described in a range of simulation literature [108] [106], as exploration of more complex systems is attempted the modelling approach also needs to involve:

- Identification of critical patterns at different levels of granularity within the system under study.
- Testing of components individually and in combination to confirm that they are capable of reproducing those patterns.
- Model development using an appropriate calibration algorithm to identify the most appropriate set of parameters (and associated values).

4.3 Model Overview

The model described in the remainder of this chapter is a hybrid simulation model with two components, the first uses the emergent behaviour from an ABM (where each agent represents an individual), to deliver a social network. The second generates individual behaviour and is delivered through an SD model, a copy of which is internalised within each agent.

The first component delivers a topologically stable social network, maintaining a set of consistent global metrics throughout the functioning of the simulation. The critical metrics were:

- Clustering coefficient.
- Transitivity.
- Assortativity.
- Average Degree.

At the individual level the network exhibits the characteristic dynamic behaviour of a social network, with individuals making new contacts, maintaining relationships and occasionally losing contact with individuals). This is critical to the modelling of obesity related social networks [10].

Both effects are achieved with a set of probabilistic relationship rules used by every 'living' agent within the simulation, the process by which they are implemented is described in the section on Model Logic later in this chapter.

Agents in the model have only two states; not restricting calorie intake (1), and restricting calorie intake (2), see Figure 4.1.

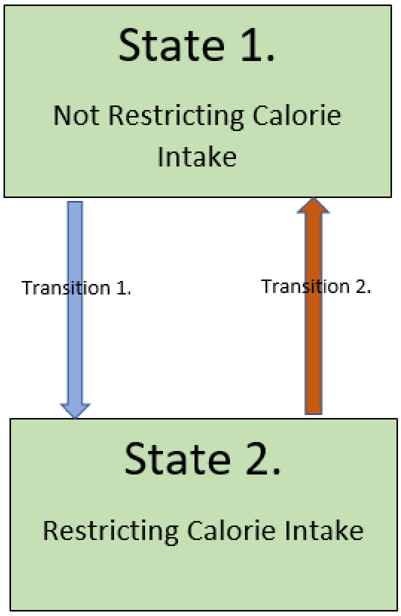


FIGURE 4.1: Agent State Chart

Transition 1. in Figure 4.1 is triggered by the second component, an SD model (Figure 4.2), a copy of which is internalised within each agent. This uses Average Per Capita Calorie Consumption (APCCC) data (a 'global' model parameter) and BMI data from the agents current immediate neighbours (it's network neighbourhood for that time step) to drive behaviour around calorie consumption and more specifically the decision whether or not to restrict calorie intake. This in turn impacts on the agents weight and consequently its BMI, acting in turn on its network neighbourhood in subsequent time steps. This model varies individually in its implementation, according to each agents height, age, current weight, gender and physical activity levels (PAL). It's functioning is also described in more detail in the section on Model Logic.

Transition 2. is time related and defined by a probabilistic function which determines the length of time for which the dieting behaviour is to be maintained.

4.3.1 Purpose of Model

The model is designed to explore the interaction between social networks and the spread of obesity, with a view to addressing the research questions identified in Chapter 1.

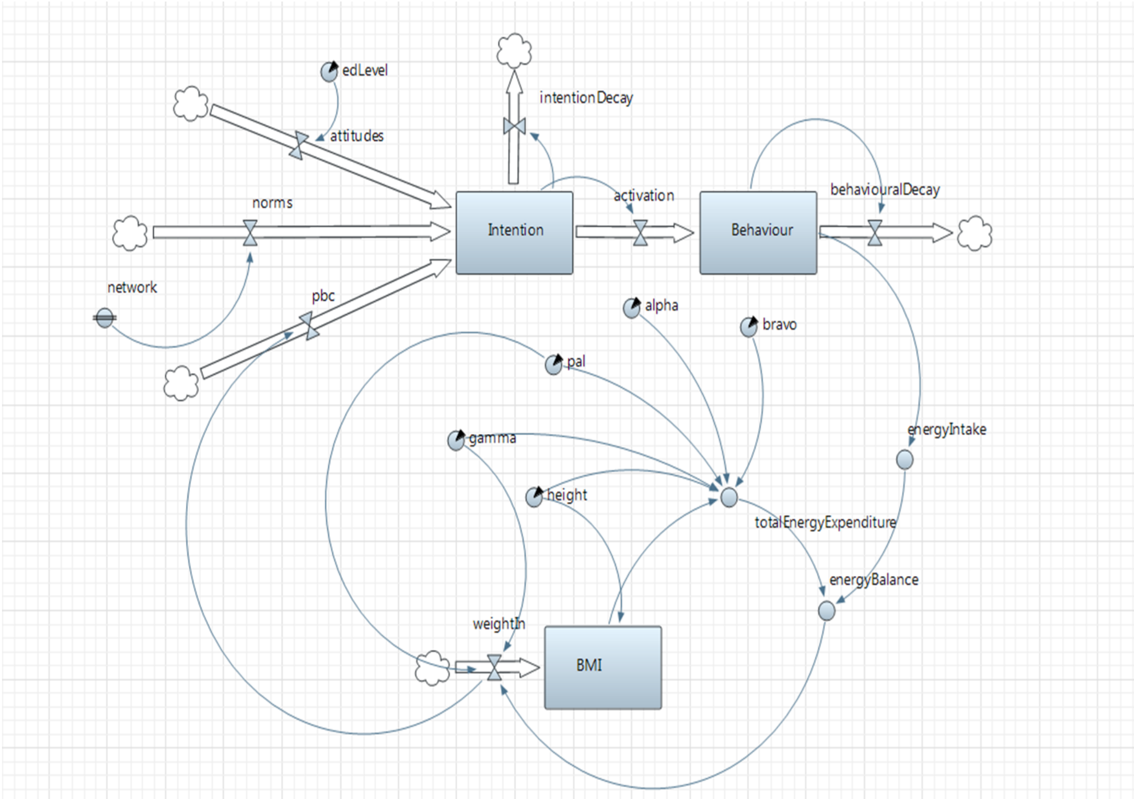


FIGURE 4.2: The Individual System Dynamics Model

4.3.2 Model Outputs

The model runs over a 10 year time span, using time steps of one month and involves a 1,000 agents. In the initial stages of its development the focus was on replicating critical patterns of behaviour, Grimm and Railsback [110]. In the network component these were:

- Consistent metrics with regard to clustering, transitivity, number of components and so on, creating a stable network topology
- At an individual level, continually varying network neighbourhood size and composition, maintaining contacts with some individuals, varying with others.
- Occasional persistent 'Weak' ties or connections across network clusters.

In the SD component, the critical patterns were of a different order. Whilst operating within a network, different individuals needed to exhibit one or more of a range of BMI trends. This included:

- Steadily rising BMI.
- Constant BMI.

- Cyclical BMI, rising and falling with varying frequencies.
- Various combinations of these.

Additionally at a macro level:

- In aggregate the distribution of BMI's needed to be non-symmetrical with a long 'right hand' tail, with median values consistently lower than means [6].

Once this was achieved, multiple runs using a stochastic optimisation algorithm were made. These used a range of parameter permutations and training and development data sets. The output of these runs was then tested using a third (test) data set, with a view to deriving the parameter values that gave a 'best fit' model with which to address the research aims.

The best fit was assessed using a Loss function that summed the (squared) difference between data forecast by the simulation and actual data over a 10 year period. This was compared on an annual basis for mean and median values of BMI across gender and five age categories. An additional penalty was imposed if the median value for a specific year, age and gender was higher than the equivalent mean, supporting a distribution similar to that described in [6]. This is detailed in Appendix B, Algorithm 24.

Output from the development phase identified an 'optimum' set of parameters, and described their relative values and gave some indication of the relationships between them. Providing additional (quantitative and qualitative) insight into the underlying theories and research on which those models were based.

Of particular interest was the relationships between the various parameters, and where appropriate how they varied with age and gender, including:

- The relative values of the three inputs to the theory of planned behaviour model.
- The impact of the network neighbourhood on individual changes in behaviour.
- The impact of global factors.

Separately the network parameters were also of interest, providing insight into the topology of a social network in the contex of NCD's.

However, the key question addressed in this phase was the role payed by homophily (related to commonality of BMI) and it's interaction with network formation. This was examined through the parameter values (from the best fit model) and the associated network topology.

Output from the second phase was of a different order, with a differing methodology. Here the best fit model was used in conjunction with a number of scenarios, varying the obesity levels of the population joining the simulation (at age 16) in conjunction with the proxy measures for environmental influence, and assessing the impact of those variations. This was averaged over multiple runs. The output comprised a 10 year (annual) forecast of mean and median BMI's, by gender and age; 16-20, 21-30, 31-45, 46-60, 61-75, 76+. The different scenarios used in this phase allowed us to address issues related to impact on resources.

4.3.3 Experimentation Aims

The research questions are listed in Chapter 1.

4.4 Hybrid Simulation Development

Developing a simulation to address the questions listed above involved a number of processes; parameterisation, sensitivity analysis and experimental scenarios. These are described below.

4.4.1 Simulation Parameterisation

The development phase involved finding the parameter values that when implemented in the model, gave the most accurate forecast. This presented a number of challenges:

- Potentially large numbers of parameters.
- Stochastic input and output.
- No gradient function.
- A computationally expensive loss function.

After considering a range of options including heuristics such as particle swarms, simulated annealing and genetic algorithms as well as approaches like finite-difference and Nelder-Mead, the algorithm used to address this was simultaneous perturbation stochastic approximation (SPSA) described in [109] and Spall [111]. This is in effect a gradient descent method, where the gradient is estimated, and rather than only modifying one parameter at each evaluation (as in the finite difference method), every parameter is modified simultaneously using a random perturbation vector. The basic algorithm is recursive and takes the general form:

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{k+1}} = \hat{\boldsymbol{\theta}}_{\boldsymbol{k}} - a_k \hat{\boldsymbol{g}}_{\boldsymbol{k}}(\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}) \tag{4.1}$$

Where $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ is a vector of parameters a_k represents a scalar gain coefficient, $\hat{\boldsymbol{g}}_{\boldsymbol{k}}$ represents the gradient approximation and k is the iteration count. This takes the general form:

$$\boldsymbol{g}(\hat{\boldsymbol{\theta}}) \equiv \frac{\boldsymbol{\delta L(\theta)}}{\boldsymbol{\delta\theta}} \tag{4.2}$$

Where $\boldsymbol{L(\theta)}$ represents a loss function

More specifically the gradient function is calculated using a simultaneous random perturbation vector $\triangle_{\boldsymbol{k}}$, often in the form of a binomial distribution (-1, 1) with equal probabilities.

$$\hat{\boldsymbol{g}}_{\boldsymbol{k}}(\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}) = \frac{L(\hat{\boldsymbol{\theta}}_k + c_k\triangle_k) - L(\hat{\boldsymbol{\theta}}_k - c_k\triangle_k)}{2c_k}[\triangle_{k1}^{-1}, \triangle_{k2}^{-1}, ...\triangle_{kp}^{-1},] \tag{4.3}$$

(c is a scalar coefficient)

The scalar coefficients are updated after each iteration:

$$a_k = \frac{a}{(k+A)^\alpha} \tag{4.4}$$

$$c_k = \frac{c}{k^\gamma} \tag{4.5}$$

.

Each SPSA implementation consisted of 3,000 iterations, using the values below:

$a = 0.16$

$A = 100$

$c = 0.1$

$\alpha = 0.602$

$\gamma = 0.101$

As described above a binomial distribution $(1, -1)$ was used to realise the simultaneous perturbation vector $\triangle$.

11 parameters or sets of parameters were used, these are described in Table 4.1. In later iterations to achieve greater granularity more parameters were added to some of the sets. Thus in the first iteration there would have been one parameter for norms, in later versions 10 were used, and in the final version 12 (breaking it down by gender and age). In earlier versions a total of 30 parameters were considered, the most complex version tested used 46.

Parameter	Effect
θ_{norms}	Set of parameters varying by age and gender, mediating the impact of local BMI on the individual
$\theta_{BMIfactor}$	Set of parameters varying by age and gender, mediating the impact of global factors on the individual
θ_{BMIadj}	Parameter varying by gender, controlling the impact of BMI difference on homophily
θ_{mem}	Parameter controlling network memory
θ_{range}	Parameter controlling network range
θ_{sn}	Parameter controlling satisficing number
θ_{lag}	Parameter controlling lag duration (TPB model)
$\theta_{dietTime}$	Parameter controlling diet duration
$\theta_{trigger}$	Parameter mediating network modification threshold
θ_{pbc}	Parameter modifying perceived behavioural control
$\theta_{edLevel}$	Parameter controlling impact of educational level

TABLE 4.1: SPSA Parameters

The initial population for the training and development sets used a fixed set of 1,000 individuals (balanced for age, gender and BMI), randomly selected from a data set of 4,000. (The data sets were obtained from Health Survey England e.g. [112], more generally the data sets used in the model are defined in detail in Section 4.4.) The selection process was modified to ensure that the distribution of ages and gender was correct for the start year. Thereafter individuals were added randomly from a pool at an overall rate of 12.0 per year (again randomly selected 16 year olds from data for the correct year). 'Deaths' occurred at an overall rate of 7.0 per year. The test set operated differently randomly selecting a different set of 1,000 agents (from the 4,000) for each iteration.

4.4.2 Sensitivity Analysis

The parameters were divided into sets, some comprising several similar parameters (e.g. the 'norms' parameters which involved separate value for both genders and age categories), or individual parameters such as 'Range' or 'PBC'.

Each run used a similar methodology to the test runs described above, using the same data, and the modified loss score mechanism (accounting for the additional age groups). After an initial run using the optimum parameters to provide a benchmark, a series of runs, were carried out varying the parameters by plus or minus 5%, or in the case of Memory and Lag (integer values) by plus or minus 1. This generated a loss score which was compared to the benchmark, recording the % difference.

4.4.3 Model Scenarios

The second phase used scenarios to compare the impact of changing levels of childhood obesity. The parameters were fixed at the values suggested by the development phase. In each run a different population of 1,000 individuals was selected randomly from a broader (age and gender balanced) data set of 4,000 for the relevant year and run for 10 years. To produce the final output the results were averaged over multiple runs.

The scenarios involved varying two elements; the average BMI of the 16 year olds entering the scenario year on year, and varying the environmental factors influencing the simulation year on year. The latter is represented by a proxy measure, the average daily per capita calorie consumption (APCCC). To get realistic but useful values the data was scanned to identify the highest sustained rate of change for each, in the case of BMI this equated to an annual rise in BMI of 0.1, in the case of APCCC this equated to a rise of 6 Kcal, these values were used to set rising and falling rates for each scenario. Nine scenarios were run in total see Table 4.2

	Falling BMI	Static BMI	Rising BMI
Falling APCCC	2	3	4
Static APCCC	5	1	6
Rising APCCC	7	8	9

TABLE 4.2: Scenarios

4.5 Model Logic

In this section the two elements of the model are described in more detail, the final subsection details the inter-dependencies between the two models.

4.5.1 Hybrid Model Overview

Each individual is represented by an agent. They are heterogeneous with differing characteristic variables; gender, age, weight, height, BMI, and educational level. Within the network, nodes are agents and the edges represent links or connections to their current contacts (network neighbourhood). There are two processes that run concurrently (but on different time scales) during a simulation run, updating the network connections and the internal processes that take place within each agent, modelled using an SD approach.

The nominal unit of model time is one month and to facilitate the SD element which uses numerical methods (Euler), this is further subdivided into 1,000 time steps. For a simulation run of 1 year there would be 12,000 time steps.

At each time step (12,000 times in a one year simulation run) each agent runs their own specific version of the SD model, which takes into account all the characteristic variables listed above (and the relevant parameters listed in Table 4.2) along with input from the network neighbourhood. The model is continuously considering all of this information and deciding whether or not to reduce calorie intake, and then assessing the impact of that decision on the individuals BMI.

The dynamic behaviour of the network (making new connection and letting others lapse) is driven by the emergent behaviour of the agents. This uses a range of stochastic rules in conjunction with data taken from each agents SD model. This is updated once a month (model time) or every 1,000 time steps. Thus in a one year simulation run it would be updated 12 times.

4.5.2 ABM Model Element

The purpose of the ABM model is to generate a stable but dynamic network that can deliver different topologies in a way that is consistent with social network constructs. (This is achieved by modifying the relevant parameters.)

The dynamic element is important given the 10 year timeframe over which the model runs, in this context use of a static network (or one that is modified randomly) would necessitate two assumptions:

- In the case of a purely static network, the characteristic dynamic behaviour exhibited by social networks has no effect on the spread of obesity. Whilst there does not seem to be any research on the issue, the possibility that a contagious effect might be affected by contact with a changing set of individuals cannot be discounted.
- In a randomly modified network, homophily plays no role in the network effect on obesity. The alternative would be that obesity plays some role in homophily and hence influences the spread (positively or negatively). Understanding whether this is the case, is a key element of the research questions.

[50] describes such a model in the context of an exploration into the fragmentation of social networks.

Key concepts in this algorithm are:

- **Spatial constraint** - nodes are located at random fixed point in a two dimensional environment.
- **Range** - the network is spatially constrained and thus distance between node is a critical element of the algorithm. Two nodes cannot form a direct connection unless they are within the specified range.
- **Affinity** - a boolean function determined probabilistically, describing whether 2 nodes can connect directly with each other, independently of whether they are in range or not.
- **Interactions** - the algorithm restricts the nodes to a specific number of interactions in each time period.
- **History** - the likelihood of a connection between two nodes is also dependent on their previous history, thus nodes that have interacted more

frequently in the past are more likely to connect than those with fewer interactions. This is subject to the constraints of a parameter which dictates the length of time considered for that history.

The algorithm was described briefly in [50, p. 378], and the outputs were presented in detail, making it possible to reproduce the original work. Using the information detailed algorithms were produced and coded, and the outputs from these were validated against the original results.

A number of extensions were then incorporated in order to deliver the social network algorithm used in the simulation.

The key extensions are:

1. Heterogeneous (rather than homogeneous) agents, across a range of characteristic variables.
2. Replacing the concept of Affinity with Homophily and using appropriate theory to underpin it.
3. Incorporating influence from agent variables (via an SD model) into the dynamic network processes.
4. Linking the time-steps to specific time-units and calibrating the parameters accordingly.

In addition a range of more minor modifications were implemented, to enable it's use in the context of the simulation.

A key element of the network model was the dual nature of the network which was required to remain topologically stable whilst exhibiting dynamic behaviour at the individual level. This was assessed by running the extended network model over sustained periods of time and taking regular 'snapshots' of the topology using the metrics previously described in Section 4.3. The network model was run for extended periods of time and the metrics remained stable over the longest time period considered relevant for the simulation.

The initialisation and general functioning of the network is described below, detailed pseudocode for the algorithms used to update the network at each time step and sub step, are given in Appendix B.

Initialisation:

1. Each agent is allocated random (fixed) x and y co-ordinates, within a defined area.

2. Parameters are set for range(r), memory(h) and interactions(s).
3. A set of characteristic variables are allocated to each agent; gender, age, BMI, height and educational level. Each set is randomly drawn from an external data source.
4. Each agent generates lists of agents who are within range and for whom the homophily boolean is $true$.
5. An initial random history data set is created for each agent, detailing the contacts with other agents over h time steps. The other agents are chosen probabilistically from those within range (euclidean distance) and for home the boolean homphily is $true$.

Warm up

After some experimentation, the default warm up period for the network element of the simulation was set at $2h$ time steps.

Time step

From an agent perspective:

1. Each agent in the network issues an 'invitation' to initiate contact to every other agent within distance r (euclidean distance calculated using coordinates).
2. Each agent accepts a maximum of s 'invitations', using a combination of boolean and probabilistic functions:
   (a) In the boolean function, the agent checks for the existence of Homophily between the two agents, if this is true then the probability of accepting the invitation is proportional to the number of times the agent issuing the 'invitation' appears in its history data set. It is also potentially modified by the absolute difference in BMI between the two agents.

   (Appendix B: Algorithm 4.)
3. An 'Acceptee' list of those who have accepted the agent's 'invitations' is recorded.

   (Appendix B: Algorithm 8.)
4. Each agent then compiles an 'Accepted' list of all the 'invitations' it has chosen to accept.

5. Each agent then compiles an 'Attendees list of all the other agents who have accepted one or more of the same invitations as it did. An agent may appear more than once in the 'Attendees' list.

6. Each agent then combines its 'Acceptees', 'Accepted' and 'Attendees' lists to form a single 'Latest Contacts' list. An agent may appear several times in the 'Latest Contacts' list, not only from repititions in the 'Acceptees' list, but also from mutual acceptance of each others 'invitations'.

   (Appendix B: Algorithm 9.)

7. Each agents history data set is then updated by adding it's 'Latest Contacts' list, and removing the one from h time steps ago.

   (Appendix B: Algorithm 10.)

8. Separately each agent removes the repetitions from the 'Latest Contacts' list to form a new list; 'Network Neighbourhood' which defines its current set of contacts.

   (Appendix B: Algorithm 9.)

From a network perspective:

1. All agent ages are updated by one month.

2. New agents (age 16) are added at a rate equivalent to the national birth rate 16 years previously.

3. Older agents (age > 80) are removed randomly at a rate equivalent to current death rate.

ABM Algorithms

The description of the model logic above describes a boolean function for determining the existence of homophily, and the probabilistic function for accepting an 'invitation'.

A review of the literature on homophily in social networks featuring face-to-face contact (as opposed to online networks) reveals a range of factors that can drive the effect. In our context probably the most useful characteristics to consider are age, gender and education (given that we don't have data on ethnicity). McPherson, Smith-Lovin, and Cook [113] suggest that in an adult population age is the most powerful predictor and the algorithm for determining the Homophily boolean between two agents a and b (h_{ab}) uses this as a differentiator:

$$h_{ab} = true$$

if:

$$random(0,1) < (p_{range} - |a_{age} - b_{age}|)\frac{p_{ap}}{p_{range}}$$

else:

$$h_{ab} = false$$

and

$$h_{ab} = h_{ba}$$

Where h_{ab} denotes the boolean for homophily between agents a and b, p_{range} describes the euclidean distance between the two agents and p_{ap} is a constant. (Appendix B: Algorithm 12.)

The values used in the simulation are $p_{range} = 50$ and $p_{ap} = 1.152$, these are chosen to deliver an overall probability for $h_{ab} = true$ of 0.75 (suggested in [50]), and a probability distribution that approximates that described in McPherson, Smith-Lovin, and Cook [113]. This is described in Figure 4.3.

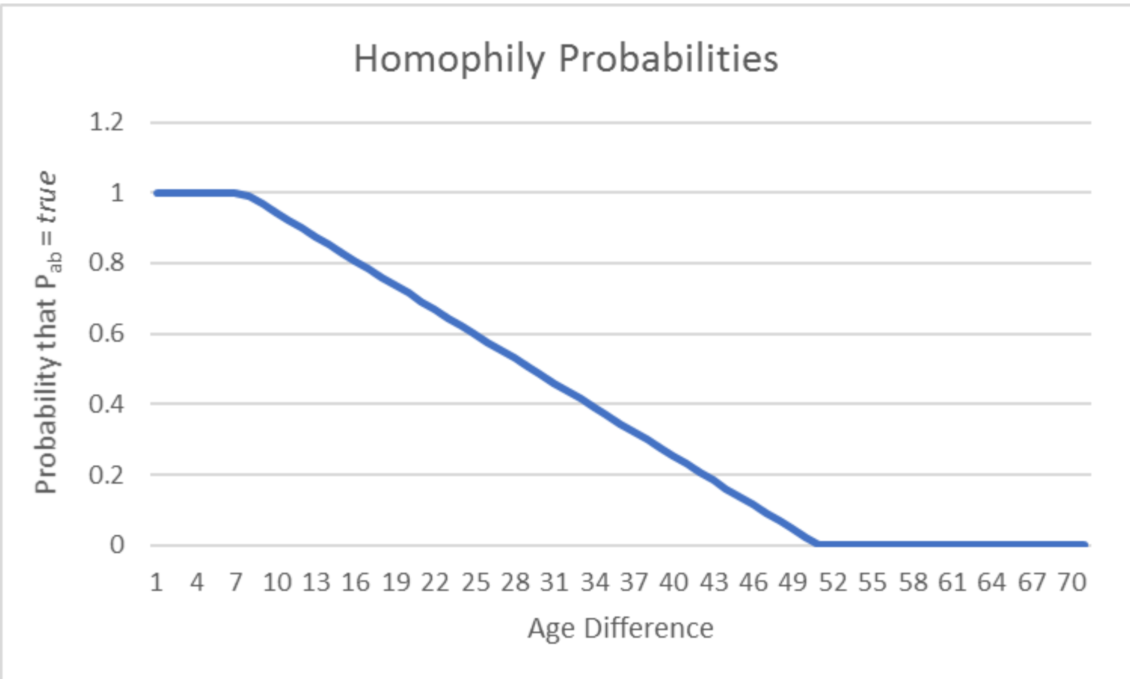


FIGURE 4.3: The Impact of Age Difference

The second algorithm describes the probability of accepting an 'invitation' (see

2.b above). Each agent maintains a history data set, in the form of a set of h lists, where h defined the number of time steps stored in the data set and each list comprises the 'Latest Contacts' from that time step as described in 6. above.

In the first stage of the algorithm, the history data set is reviewed and an edgeweight list is created identifying each agent that appears in the data and it's frequency of occurrence.

The BMI of each agent in that edgeweight list is then compared with the original agent, and if the BMI difference exceeds a threshold value (controlled by parameter $\theta_{trigger}$) then the value is reduced by a percentage (controlled by θ_{BMIadj}).

The updated edgeweight list is then used to calculate the probability by cycling through each agent on the list, and:

1. Confirming that the homophily boolean for that agent and the original agent is $true$.
2. Confirming that the agent is within range of the original agent
3. Counting the total number of times that agent appears in the history data set and setting the edgeweight for that agent equal to that count.
4. Adding the edgeweight value for the agent to a a total

Once the process is complete it cycles through the agent list again creating a list of agent probabilities, where the probability of accepting an 'invitation' from an agent a is:

$$P_{agent\ a} = \frac{Edgeweight_{agent\ a}}{Sum\ of\ Edgeweights} \tag{4.6}$$

In the final stage of the algorithm, the agent cycles through its probability list, accepting an 'invitation' if:

$$random(0,1) < P_{agent\ a}$$

Continuing for a maximum of 20 iterations, or until s (the number of interactions) 'invitations' have been accepted. (Appendix B: Algorithm 5.)

4.5.3 ABM Components

The purpose of the agent-based element of the model is to create and maintain the network, in the context of a cycle in which information from the immediate network neighbourhood is fed to the internal SD model incorporated into each agent, processed and then re-broadcast back to that network neighbourhood.

Environment

Agent are best envisioned as nodes in a two dimensional graph, their locations are fixed and defined by randomly generated x and y co-ordinates during the initialisation phase of the simulation.

The density of the nodes is defined by the maximum value A of the x or y co-ordinates (in this iteration the maximum x or y values are always the same) and the number of nodes (agents) in the simulation..

Agents

Agents have a number of characteristic variables:

- Age
- Gender
- Height
- BMI
- Physical Activity Level (PAL)
- Educational level (edLevel)

Initially, with the exception of PAL (which is drawn from [101]) this is drawn from external data using data from individuals who took part in different iterations of the Health Survey England for the relevant year.

Subsequently the data for agents joining the network is drawn from HSE data for 16 year olds for the appropriate year.

Each agents objective is to minimise the difference between their BMI and the average BMI of the other agents in their network neighbourhood. The detailed interactions are given in the description of the SD model.

Interaction Framework

Agents only interact directly with agents in their network neighbourhood as defined above. So, at each time step, each agent carries out two actions:

- It computes the average BMI ($networklIn$) of its network neighbourhood (for the previous timestep) and this information is then the input to its SD model, for processing and further action.

  (Appendix B: Algorithm 18.)

- It's BMI is shared with its network neighbourhood to enable them to compute their own value for $networkIn$ as described above.

4.5.4 SD Model Element

The SD model is implemented in Java, using numerical methods (Euler) where $\delta t = 1,000$. Thus for each network time step in the simulation the SD model described below runs 1,000 sub steps (for each agent). The approach and the settings were chosen after some experimentation, comparing the performance of the Java implementation and a bespoke software package (AnyLogic) on an early iteration of the SD element, in order to achieve an acceptable level of accuracy.

SD model Overview

The SD model is shown in Figure 4.2. The flows stocks and variables illustrated in the top half of the model are an implementation of TPB) [35], whilst taking into account the satisficing effect noted in [67] . The lower half implements the Henry equations [101] and also takes account of the intermittancy effect described in [104]. The impact of the surrounding environment is also addressed by incorporating a proxy measure in the form of a figure for Annual Per Capita Calorie Consumption (APCCC), modified by an appropriate parameter.

As with the ABM, the model structure remained constant throughout the development process. With the level of parameter specificity increasing as the phase progressed. For example in the initial implementations of the model their was a single parameter for modifying the impact of the APCCC on the agent, in later versions there was a separate parameter applied for each gender and age group.

4.5.5 SD Model Logic

The stock and flow diagram that comprises the upper half of Figure 4.2 presents a simplified implementation of TPB as shown in Figure. **??**.

In the SD model Intention and Behaviour are represented as stocks, with Intention driving Behaviour through a flow (activation), Intention is itself driven by three flows; attitude, norms and PBC (perceived behavioural control).

Intention and Behvaiour both have decay flows (intentionDecay and behaviourDecay).

In essence Intention stocks build as a consequence of flows from attitude, norms and perceived behavioural control (PBC), subject to the negative effect of intentionDecay. It is passed on to the Behaviour stock through the activation flow (which is subject to a lag). When the behaviour stock reaches a threshold level, dieting behaviour is triggered, and the Behaviour stock is reset to 0.

The decay flows are intended to provide some additional flexibility on the model (not yet utilised), activation is also subject to time lag dictated by parameter θ_{17}.

The other flows are also subject to parameterisation; norms θ_{norms}, attitudes $\theta_{edLevel}$ and PBC θ_{pbc}.

Attitudes is linked to educational level, since in this context it represents the ability to reach considered views (beliefs) about the long and short term consequences of health behaviours. This is the effect described in [34, p. 20], which describes a number of statistical analyses attempting to understand the impact of socio-economic factors on health behaviours, in the conclusion education is identified as one of the key factors. It goes on to suggest that education (to degree level) directly influences cognitive ability, and that this may account for approximately 30% of health behaviours.

Similarly research suggests that a significant factor in PBC is a reinforcement loop whereby success drives higher levels of PBC and conversely failure reduces PBC. Thus in our model the level of weight loss or gain derived from the weightIn flow described below, has a direct impact on the PBC flow [78].

The norms flow is driven by a positive difference between the individuals BMI and the average BMI of the individuals in the network neighbourhood. It is also subject to a parameterised threshold ($\theta_s n$), to represent the satisficing effect [67].

The lower section of the model has only one stock (BMI) and one flow (weightIn). It does however have a range of variables and associated formulae. Those visible in Figure.4.2 include:

- **energyIntake (EI)** - the daily amount of calories consumed by the agent.
- **totalEnergyEpenditure (TEE)** - the total energy expended by that agent in the day (calculated using the Henry equations).
- **energyBalance (EB)** - the difference between energyIntake and totalEnergExpenditure.
- **alpha** - Henry equation coefficient [101].
- **bravo** - Henry equation coefficient [101].
- **gamma** - Henry equation coefficient [101].
- **pal** - physical activity level (derived from values given in [102]).
- **height** - agents' height in metres.

Additionally there are also:

- **dietTime** - a value derived probabilistically which identifies how long the individual will maintain the diet (modified by $\theta_{dietTime}$), this in a simplified way provides for the intermittancy effect identified by [104]. (In reality two modes of intermittancy were identified adherence to daily calorie intake and duration.)
- **APCCC** - average per capita calorie consumption for the year in question, this modified by the appropriate parameter ($\theta_{BMIfactor}$), forms the basis for the default energy intake value.

- **behaviourTrigger** - the behaviour stock value that triggers dieting behaviour.

- **calorie restriction** - after consultation with subject matter experts this was set at 480Kcal per day. Thus whilst an agent was in a dieting phase its calorie intake was reduced by 480Kcal per day.

Thus in this model a period of dieting is triggered when the Behavior stock reaches 50, the calorie intake is reduced by 480 Kcal, and the impact on weight is calculated, along with the subsequent impact on the BMI stock.

4.5.6 SD Model Components

For each δt the SD model carried out 3 processes; updating stocks, updating variables and updating flows.
(Appendix B: Algorithm 14

Stocks

The equations used to update the stocks in the SD component are:

$$Intention_t = Intention_{(t-1)} + \frac{\delta}{\delta t} attitudes_{(t-1)} + \frac{\delta}{\delta t} norms_{(t-1)} + \frac{\delta}{\delta t} pbc_{(t-1)} - \frac{\delta}{\delta t} intentionDecay_{(t-1)} \tag{4.7}$$

$$Behaviour_t = Behaviour_{t-1} + \frac{\delta}{\delta t} activation_{(t-1)} - \frac{\delta}{\delta t} behaviourDecay_{(t-1)} \tag{4.8}$$

$$BMI_t = BMI_{t-1} + \frac{\delta}{\delta t}\left(\frac{weightIn_{t-1}}{height^2}\right) \tag{4.9}$$

(Appendix B: Algorithm 15

Flows

The process used to define the flow values in the TPB area of the model, required some consideration, a completely parameterised set of flows would deliver a huge solution space on the other hand there is very little data with which to set values within the model. After some experimentation a behavioural trigger value of 0 and $activatio$ and $intentionDecay$ and $behaviouralDecay$ flow values as described below, seemed to give realistic values when combined with $attitudes$, $norm$s and pbc values of 10 and a lag value of 2.

In terms of the research objectives insight into the relative effect of the $attitudes$, $norms$ and pbc flows (as well as any lag) were deemed to be of more interest. Therefore the behavioural trigger value and $activation$ and $intentionDecay$ and $behaviouralDecay$ flows were fixed, and the parameterisation was focused on $attitudes$, $norms$, pbc and lag.

- **weightIn**

$$\boldsymbol{p.weightIn_t} = \frac{p.EB_{t-1}}{p.pal * p.gamma * 12} \tag{4.10}$$

- **PBC**

$$\boldsymbol{p.pbc_t} = p.weightIn_{t-1} * \theta_{pbc} \tag{4.11}$$

- **norms**

$$\boldsymbol{p.norms_t} = \frac{p.networkIn_{t-1}\theta_{norms}}{2} \tag{4.12}$$

  (Where θ varies with age group and gender.)

- **attitudes**
  If $p.edLevel = 0\ or 1$ then $\boldsymbol{p.attitudes} = 10$, else $\boldsymbol{p.attitudes} = 4\theta_{21}$

- **intentionDecay**

$$\boldsymbol{p.intentionDecay_t} = \frac{p.Intention_{t-1}}{2} \tag{4.13}$$

- **p.behaviourDecay**

$$\boldsymbol{p.behaviourDecay_t} = \frac{p.Behaviour}{2} \tag{4.14}$$

- **p.activation**

$$\boldsymbol{p.activation_t} = \frac{p.Intention_{t-1}}{3} \tag{4.15}$$

  (The process for updating flows is described in Appendix B: Algorithm 16)

- **activationLagged**
  The activation flow is of course subject to a lag and this is determined by the following:

$$j = \lfloor \theta_{lag} * \delta t \rfloor$$

$$\boldsymbol{p.activationLagged} = p.activation_{t-j} \tag{4.16}$$

Variables and Constants

The updating process for a number of the variables ($networkIn, TEE, EI$ and EB is described in Appendix B: Algorithm 17, the remaining constants are updated or reviewed individually

- **networkIn** - this is calculated by taking the average BMI of the individuals of the same gender, in the agents network neighbourhood.
  (Appendix B: Algorithm 18.)

- **TEE**

$$\begin{aligned}\boldsymbol{p.TEE_t} &= p.pal(p.alpha + (p.bravo * p.height) \\ &+ (p.gamma * p.height^2 * p.BMI_{(t-1)})) \end{aligned} \tag{4.17}$$

- **EI**

$$\boldsymbol{p.EI_t} = p.TEE_{(t-1)} + (\frac{APCCC}{3400})(\frac{p.TEE_{(t-1)}}{(p.BMI_{(t-1)} * p.height^2 * \theta_{BMIfactor})}) \tag{4.18}$$

  However if dieting was triggered at a time t, then energyIntake is fixed for the duration of the diet ($dietTime$) as folows:

$$\boldsymbol{p.EI_{(t \rightarrow (t+p.dietTime))}} = p.EI_{(t-1)} - 480 \tag{4.19}$$

  (Appendix B: Algorithm 19.)

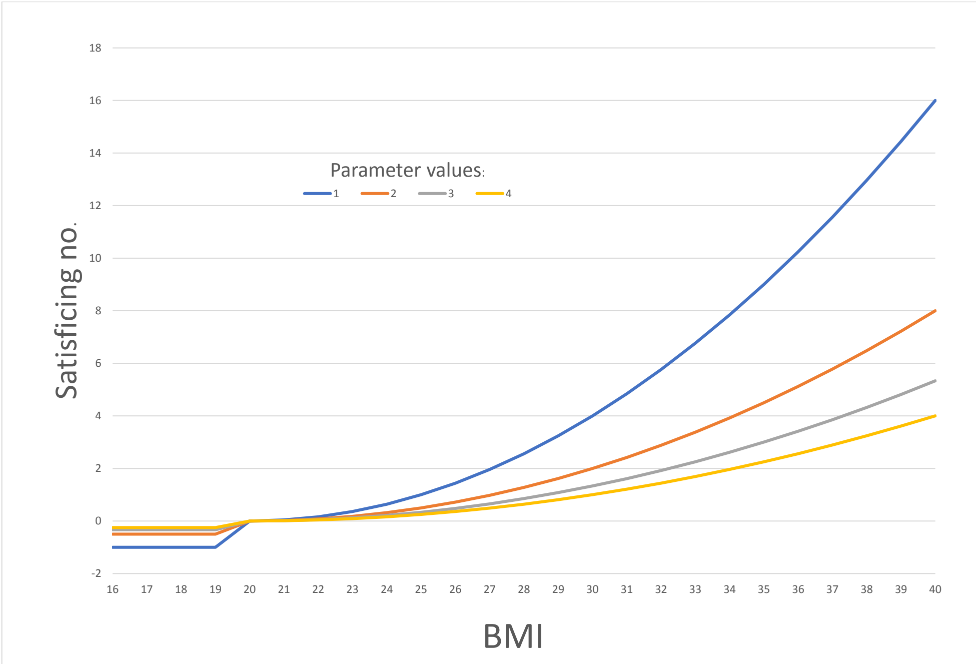


FIGURE 4.4: The Impact of Different Parameter Values on the Satisficing Number

- **satisficingNumber** - the satisficing number algorithm is designed to produce an exponential increase in line with increasing BMI, it is parameterised (θ_{sn}), and Figure 4.4 describes the shape of the function in relation to different parameter values. If the agents BMI is less than 20 the satisficing no. is set to -1, otherwis:

$$\boldsymbol{a.satsficingNo} = \frac{(a.bmi - 20)^2}{\theta_{sn}} \tag{4.20}$$

  (Appendix B: Algorithm 20.)

- **pal** - this uses a range of values according to current BMI, they are taken from the report published in [102].
  (Appendix B: Algorithm 21.)

- **edLevel** - this is initially set at 0 and then revised probabilistically when the agent reaches age 22, either to 1 (not degree level) or 2 (degree level) the probability of the latter is currently set at 30%, although there is some uncertainty about the specific number with estimates ranging from 27.2%

to 34.4%.
(Appendix B: Algorithm 22.)

- **alpha, bravo** and **gamma** - these are constants related to age and gender, they are checked and updated at each time step.
(Appendix B: Algorithm 23.)

- **APCCC** - this is the figure for average daily per capita calorie consumption for the UK, for the year in question, it is updated annually within the simulation (taken from data published by the Food and Agricultural Organisation of the UN).

- **dietTime** - this is set using a parameter (θ_{18}) and function that returns a a random number (v) based on a gaussian distribution with mean of0 and standard deviation of 1.

$$\boldsymbol{p.dietTime} = max(1, v\frac{\theta_{dietTime}}{4} + \theta_{dietTime}) \tag{4.21}$$

Dieting is triggered when a number of criteria are met:

  - The difference between the agent BMI and the average of the (same gender) network neighbourhood BMI is exceeded by a value greater than the satisficing index

    AND

  - The Behaviour stock exceeds 50.

    AND

  - The agent is not already dieting

4.5.7 Element Inter-Dependencies

There are two sets of inter-dependencies that exist between the elements of the simulation. The first takes place during the simulation run and is intrinsic to its operation, involving the exchanging of BMI information between individual agents and its subsequent use to drive the SD element of the model. The second assumes that $\theta_{BMIadj} > 0$, this set of parameters relates to the BMI related adjustment made in the algorithm determining the probability of homophily(Appendix B: Algorithm 6) . If that is the case then during the warm up phase there is some interaction between the network component and the BMI data stored in the SD component. This is because the difference in BMI values will be considered during each time step of that warm up period, modifying the initial random structure so that it reflects to some extent the BMI values of the individuals concerned. (The default warm up period is set at $2h$ time steps, and h is defined by $int|\theta_{mem}|$)

4.5.8 Experimental Logic

As previously described there were two phases to the experimental process. In the model development phase training and development data sets were used on increasingly granular versions of the models to ascertain the best fit model, and provide information about the likely relationship between different parameters. Each of scenarios were then run using the same set of (best fit) parameters, so that from an ABM/Social network perspective the only variation between scenarios was in the weight profile of the individuals being added at each time step and the values for APCCC.

Throughout the process described above, the model structure remained constant, however in the first phase (obtaining the best fit model) the level of parameter granularity was increased as the phase progressed. In the first iteration the BMI adjustment in the Homophily algorithm (described below) was set to neutral (no effect), subsequently it was allowed to vary, then a separate parameter was introduced for each gender, and in the final iterations there was a separate parameter for each age group and gender.

4.6 Data

With the exception of that obtained directly from subject matter experts, all the data used in this research was taken from the public domain. The data included:

- data describing individual characteristics (age, gender height, weight, BMI and educational level) 1993 - 2013
- Data describing the distribution of the UK population by age and gender between 1993 and 2013.
- Average daily per capita calorie consumption (APCCC) for the UK between 1993 and 2013.
- Average physical activity levels (PAL) by BMI, age and gender.
- Average reduction in calorie intake when dieting

- The nature of the relationship between satisficing index and BMI (from subject matter experts).
- Number of degree educated individuals within the population

The raw data was then processed to deliver the various inputs needed to run the simulation and optimisation functions. This was then combined into a 'data book' available as an excel spreadsheet with the simulation code.

4.6.1 Data Sources

Individual Characteristics

Data describing individual characteristics for the years 1993 - 2015 was downloaded from the Health Survey for England on the UK Data Service website:
https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000021

Population Distribution

Data describing the UK population distribution by age and gender was obtained from the Office for National Statistics (ONS) the data sheet was titled:
Mid-1971 to Mid-2015 Population Estimates for Regions in England and Wales, Quinary age groups and Single year of age and sex. Available via:
https://www.ons.gov.uk/peoplepopulationandcommunity/
populationandmigration/populationestimates/datasets/
populationestimatesforukenglandandwalesscotl
andandnorthernireland
Given that the HSE data relates specifically to England the population data was also taken only from the English values in the data set.

Average per Capita Calorie Consumption

Data for the APCCC for the UK was obtained from the Food and Agriculture Organization of the Unied Nations via their website:
http://www.fao.org/faostat/en/#search/Food%20supply%20kcal%2
Fcapita%2Fday

Physical Activity Level

Data for average PAL by age BMI and gender was obtained from [102].

Degree Level Education

Number of degree educated individuals within the population from https://www.theguardian.com/higher-education-network/blog/2013/jun/04/higher-education-participation-data-analysis in the event there was considerable conflict between estimates so a default of 30% was adopted.

Calorie Intake Reduction

There was no definitive information on the typical amount by which individuals reduce their calorie intake when dieting, a provisional figure of 480Kcal was used and subsequently confirmed as realistic with subject matter experts from Aneurin University Health Board.

4.6.2 Pre-Processing

A significant amount of pre-processing was carried out on the data collected for the simulation. The primary source was the HSE data, in combinations with that from the ONS. The process for each year was the same, and once the initial data cleansing was complete was automated with a short piece of Java code

1. Initially the HSE data was reviewed and unwanted columns were removed, leaving age, gender, height, weight and/or BMI and educational level.
2. Entries with missing values were removed, leaving on average 6-8,000 entries.
3. The data was then sorted into categories defined by age range (initially 16-30, 31-45, 46-60, 61-75, 76+) and gender.
4. The ONS data for the relevant year was reviewed to ascertain the % of the population falling into each gender/range.
5. Each category was randomly sorted and then, the correct number of entries to form a total population of 4,000 was selected from it, this was the highest population that could be realistically generated from the data numbers.
6. New entries for the year were created by taking the whole population of 16 year old's, randomly shuffling them and creating a new list from which the simulation picks in order.

7. Mean and median BMI values were calculated for each category (used in calculating the loss function).

4.6.3 Input Parameters

Given that the the optimisation method was based on a gradient descent approach, the initial set of 19 optimisation parameters (and the loss function) were all tailored so that their theoretical value might realistically be expected to fall somewhere between 1 and 10 (ideally 5), thus ensuring that the solution space topology was as conducive as it could be to the descent process.

The simulation was then run for 5,000 iterations in a 'random walk' through the solution space and the best performing set of parameter values were used as the starting parameters in the initial phases of the optimisation process.

4.6.4 Assumptions

There are a number of assumptions and simplifications made in the implementation of models and theories within the simulation:

- The use of intermittant diet duration to represent the intermittancy described in [104], is a simplification of the actual effect, which also involves intermittent adherence to calorie intake restrictions within the diet duration, in practise it was felt that this could be approximated by the one variable (given that it wasn't a key area in the research aims).
- Initially a single Satisficing number was used to parameterise the simulation as opposed to other population parameters that were broken down by age and gender.
- Using age as the main predictor for homophily was as a consequence of not having data on ethnicity, and the fact that for the majority of the population under consideration gender was not a significant issue [113].

The representation of TPB required considerable simplification, the original model (Figure. **??**) suggests inter-dependencies between each of the inputs, there is also the issue of putting realistic data behind each of the input flows. An attempt to parameterise the whole function would have created an infinite solution set, with no distinguishable relationships, and thus no useable information about the relationship between the parameters. After reviewing the research aims, it was determined that the most useful relationship to explore was that between the three input flows (attitudes, norms and PBC) and in particular the relative value/impact of norms.

After some initial experimentation with sample values it was found that a trigger level of 50 and input flows of 10 for PBC and attitudes (along with 50% decay rates and a short time lag), generated appropriate output patterns (see 4.1.2). So in the optimisation phase the trigger level, activation and decay flows were fixed, and the input flows and time lag were allowed to vary.

There was also the issue of linking these input flows to appropriate data sources

- PBC - the literature does not suggest any quantifiable data in relation to PBC, but it does clearly describe the reinforcing loop between actual success and perceived behavioural control, this is represented in the model [35].A default flow of 10 was assumed, which was subsequently modified by a feedback loop controlled by θ_{pbc}.
- Attitudes - as discussed previously the concept of attitudes as described in [35] is similar to the cognitive ability described in [34], which they link to general educational level, and specifically to degree level education.
- Norms - this is clearly quantifiable and is represented by the average BMI of individuals in the agents network neighbourhood, the additional element considered in this is the satisficing index [67] or threshold value.

4.7 Experimentation

4.7.1 Initialisation

The SD component did not utilise a warm-up period. The network component required a warm-up period to establish the network, this was set to a period equal to twice the value of the memory parameter used in the simulation run. This was decided on after experimentation in the early stages of the model build.

Each run for each scenario involved 1,000 agents selected randomly from a pool of 4,000 agents that reflected the age, gender, educational level and BMI distribution for the start year. Agents were added from similar pools of 16 year olds, for each of the subsequent years.

4.7.2 Run Length

The run length for both the development and the experimentation phases was 10 years, broken into one month intervals, the SD component subdivided each month into a 1,000 time intervals. Thus the total number of time steps addressed in the simulation was 120,000.

4.7.3 Estimation Approach

Given the stochastic nature of the output, the second phase involved the use of multiple runs to estimate the results. The results quoted are based on the median values from those runs.

4.8 Software and Hardware

4.8.1 Programming Language

The computer which was primarily used for coding and running the simulation used Windows7 Professional as its operating system.

The simulation was implemented in Java with Java Development Kit 1.8.0_44, using Eclipse Java Photon as an IDE.

Analysis of the network topology was done in Python 3.6.3, using the Networkx 2.2 package, implemented in a Jupyter notebook.

4.8.2 Random Sampling

Random sampling was carried out using the appropriate java classes. For uniform random distributions between O and 1, the $Math.random()$ method was used from the Math class. For Gaussian random distributions an instance of the Random class was created and then used in conjunction with the $.nextGaussian()$ method to generate a value (the method generate a value using a mean of 0 and a standard deviation of 1).

Since the development phase of the process did not use common random numbers as part of the optimisation process, no seeds were defined.

4.8.3 Model Execution

In the network component the sequence of actions at each time step is defined by the algorithm described at Appendix B Algorithm 2. Similarly the sequence of actions for each sub step in the SD component is defined by the algorithm described at B Algorithm 14. The interaction between the two components is defined by the algorithm described at Appendix B Algorithm 1.

4.8.4 System Specification

The computer used to run the majority of simulation runs had an Intel(R)Core(TM)i7-2600 cpu @3.4GHz processor with 12.0GB of installed memory (RAM), and it used a 64-bit Operating System.

Parallel processing was enabled within the execution of each procedure/method (using the The Java function 'parallelStream'), but not across them. This minimised the run time while maintaining the integrity of the process.

The SD update consists of three main processes (updating stocks, flows and variables) for each agent, since these are relatively independent of the other agents, they are combined into one function (PhasesSDS.runSD) for each agent, to which the parallelisation process was then applied, splitting the agents amongst the cpu's:

The network update consists of five processes, each of which needs to be completed for all agents before the next is started so the parallelisation is applied within each process, and the next process wasn't started until all the current agents had completed the predecessor.

Using the hardware and software described above the execution of a single iteration of the model varied between 45 and 120 seconds, depending on the specific parameter values implemented. In a descent run (3,000 iterations) the average time per run was taken as 60 seconds, which proved to be slightly conservative, with actual values in the range 55-60 seconds.

4.9 Summary

This Chapter has sought to provide a comprehensive description of the simulation and the model it realises. It has described its purpose, objectives, outputs, development and underlying logic and architecture. It has also described the sources of the data used within the simulation, the experimental approach and the software and hardware used to run it. The next Chapter describes the results produced by that simulation.

Chapter 5

Results and Parameter Analysis

"...grant me the serenity to accept that many parameters must be estimated, the intelligence to recognise that some can be calibrated, and the wisdom to know the difference..."

with apologies to Reinhold Niebuhr

5.1 Introduction

The development and implementation of the simulation model involved:

- The model development process, which included; training, testing and sensitivity analysis.
- An experimentation process; using appropriate scenarios.

These processes generated the results and data used in respect of the research aims:

- The combined output of the Model Training and Testing phases delivered a 'best fit' model, providing insights into the theory underpinning the model structure and the relationships between the different model elements and their variables.
- This was augmented by the data generated from the sensitivity analysis.
- Data from the experimentation phase directly addressed a number of the research aims.

Two sets of data (2004-13, 2003-12) were used during the training process, an initial set to create the 'base' model and a development set to further refine it. A third (2002-11), was used for the testing phase. Model selection was based on performance (loss score) in the testing phase.

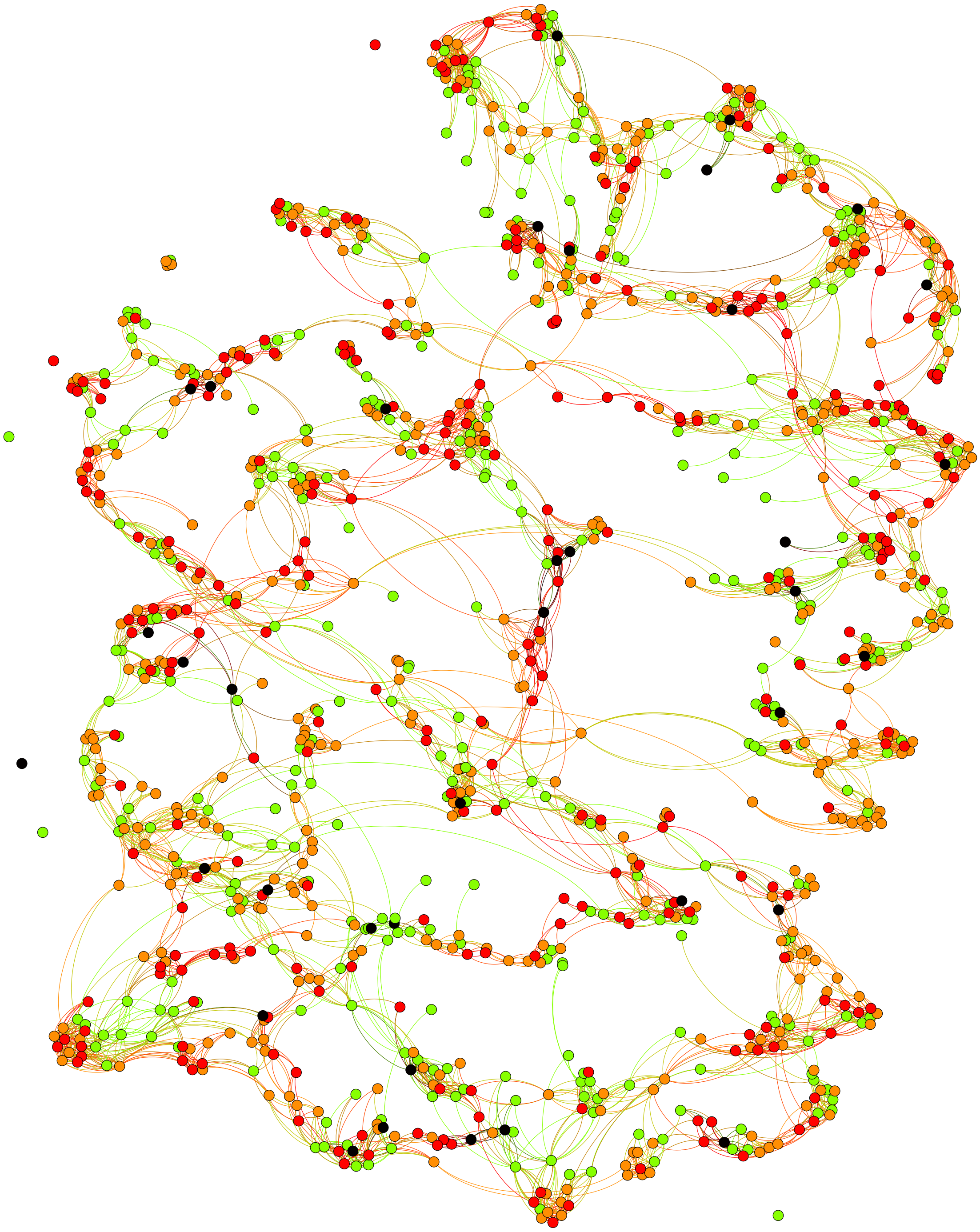


FIGURE 5.1: Model Output - Network Diagram

Figure 5.1, represents the social network at the end of 10 years, using the final simulation model and a scenario that assumed that the external environmental factors remain static. Green nodes represent individuals with a BMI of less than 25, amber individuals with a bmi less than 30 and greater than 25, red those with a BMI greater than 30 and black those individuals with a BMI greater than 40.

After a description of the methodologies used to generate the results, the remainder of the chapter describes the following:

- The output from the model development phase, specifically; parameter values for the best fit model and where relevant the results of the sensitivity analysis.
- Network topology for the best fit model.
- The choice of scenario configurations and their outputs.

5.1.1 Model Training Methodology

A series of descent runs were carried out on two parallel tracks, in the first BMI played no part in the algorithms that influenced the formation and dissolution of friendship ties (homophily), in the second a pair of parameters were inserted to allow BMI difference to play a role in that process, the first defined a threshold value above which the effect would take place, the second the magnitude of that effect. Otherwise at each stage of the process the changes were similar.

The output from each descent run comprised a revised set of parameters, and a set of loss scores measuring the improvement in model accuracy. Figure 5.2 documents the loss scores chart from Run 1.

The runs increased in model complexity as they progressed this was achieved by adding additional parameters, or by dividing existing parameter sets into larger sets. Thus in early runs a parameter might take a single value, in later runs it might be subdivided by gender and later still by age. Each parallel run took as the starting values for it's parameters, the final parameter values from the preceding run in that parallel track, creating a cascade effect. Each descent run comprised 3,000 iterations

The initial runs in each track were carried out using the training data set, before switching to the development set for the later runs.

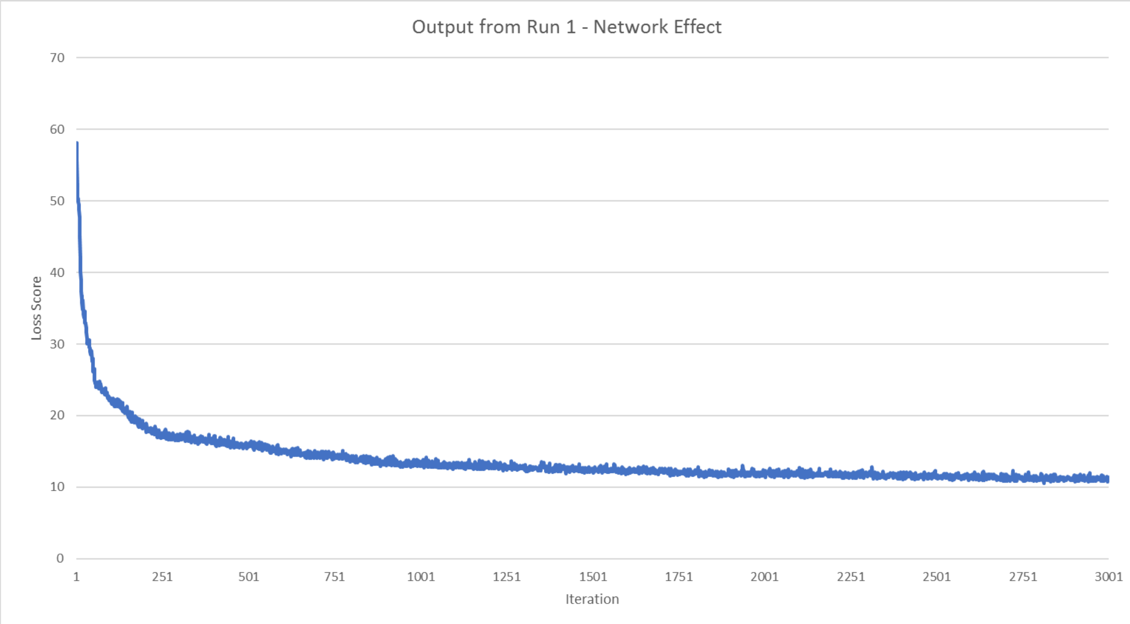


FIGURE 5.2: Descent Run - Output

Both the training and development data set used a fixed set of 1,000 agents, chosen at random from the the available data for the year, but modified to reflect the relevant age and gender ratios within that population.

5.1.2 Model Testing Methodology

The testing methodology used a fixed set of parameters taken from the output of the relevant descent run. A random population was selected from a data set of 4,000 agents for the start year (reflecting the relevant age and gender ratios within the population for that year). It was run for 10 years, and the data recorded comprised the mean and median BMI for each age group, gender and year. This was repeated 1,000 time with a new population selected for each iteration. The overall results were then combined to produce a single set of data which was compared with the real data for that 10 year period. The loss scores were also combined to give a single representative value. It was found that median (rather than mean) values were more effective for the comparison process, generating more accurate BMI's and a lower loss score.

Separately each run was also examined topographically, with each iteration generating an adjacency matrix for the network, the largest component of each was then measured in the following domains:

- Number of nodes.
- Number of edges.

- Clustering coefficient.
- Transitivity.
- Assortativity.
- Shortest path.
- Network Diameter.
- Average node degree.
- Number of components.

These values were collected for each of the 1,000 iterations, and median, mean and standard deviation values were calculated and reported.

5.2 Sensitivity Analysis

Having identified and quantified an optimum parameter set, and reviewed the associated network topography, the focus then moved to sensitivity analysis. This was considered important in the context of the research aims, in order to assess the impact of variation in the parameter sets on the forecast accuracy.

The parameters were divided into 11 sets. A series of runs, were carried out varying the parameters by plus or minus 5%, or in the case of Memory and Lag (integer values) by plus or minus 1. Each run generated a loss score and compared it to the benchmark, reporting the % difference (Table C.1.) additionally it generated a more detailed heat map to indicate where the disparities occurred. These are given in full in Appendix C

The group of 'BMI Factor' parameters, illustrate the process well, an increase of 5% in the parameters generated a 0.32%increase in the loss score. (Given that the loss score is non-linear, it would be unwise to compare the % variation with the % increase directly.) Conversely a 5% decrease in parameter values generated a 1.22% increase in loss score.

Reviewing the data across the parameters, the area which showed the most variation was the male population in the 61-75 age group. Broadly there were two patterns of issue presentation either accuracy decreased with age and model run-time as illustrated in Tables C.2, or a more diffuse pattern where the inaccuracies were more evenly spread.

Each set of parameters are addressed in more detail as part of the section on Parameter Values and Analysis, the overall impact is shown in Table 5.1.

Parameter	Variation	Plus	Minus
Norms	5%	-0.41%	-0.78%
BMIAdjustment	5%	-0.22%	-1.36%
BMIFactor	5%	-0.32%	-1.22%
DietTime	5%	-1.81%	-1.36%
EdLevel	5%	-1.60%	-0.78%
'PBC'	5%	0.10%	-0.81%
Range	5%	-1.1%	-1.89%
Satisficing No.	5%	-0.81%	-1.39%
Network Trigger	5%	-1.03%	-2.71%
Memory	1	-0.48%	-1.34%
Lag	1	-1.02%	-1.19%

TABLE 5.1: Summary of Sensitivity Analysis

5.3 Implementation Process

The starting parameter set was arrived at after some initial experimentation and consideration of the research aims. They are described in Table 5.2. θ_{11} and θ_{12} were only relevant to the parallel run where BMI affected network dynamics.

The process became an iterative one where each step involved examining the results from the previous one in conjunction with theory in order to ascertain what parameter modifications might improve the model, these improvements were then run using the descent algorithm and the resulting parameter values were then tested.

For the parallel run where BMI had no effect on network dynamics, the parameter evolution followed the steps describe below (each run incorporates the parameter changes from its predecessors) :

Run 1 - Global influence on individuals represented by the BMI factor parameter was split into two, with a parameter for males and females.

Run 2 - Influence from the network in the form of BMI data from network neighbours was made gender specific, so males only considered

Parameter	Effect
θ_0	norms male 16-30
θ_1	norms male 31-45
θ_2	norms male 46-60
θ_3	norms male 61-75
θ_4	norms male 76+
θ_5	norms female 16-30
θ_6	norms female 31-45
θ_7	norms female 46 - 60
θ_8	norms female 61-75
θ_9	norms female 76+
θ_{10}	BMI factor
θ_{11}	male BMI adjustment
θ_{12}	female BMI adjustment
θ_{13}	network memory
θ_{13}	network range
θ_{15}	satisficing number
θ_{16}	lag duration
θ_{17}	diet duration

TABLE 5.2: Initial Parameters

BMI data from other males in their network, and females that of other females.

Run 3 - PAL levels were made variable (previously set at an average value) and updated regularly according to age, gender and current BMI.

Run 4 - Parameters were introduced to modify the impact of educational level on the flow 'attitudes' and to parameterise the flow 'PBC'.

Run 5 - The BMI factor parameters were expanded to take account of five age ranges and gender.

For the parallel run where BMI was allowed to affect network dynamics, the stages in parameter evolution were:

Run 1 - Global influence on individuals represented by the BMI factor parameter was split into two, with a parameter for males and females.Network modification threshold parameterised.

Run 2 - Influence from the network in the form of BMI data from network neighbours was made gender specific, so males only considered BMI data from other males in their network, and females that of other females.

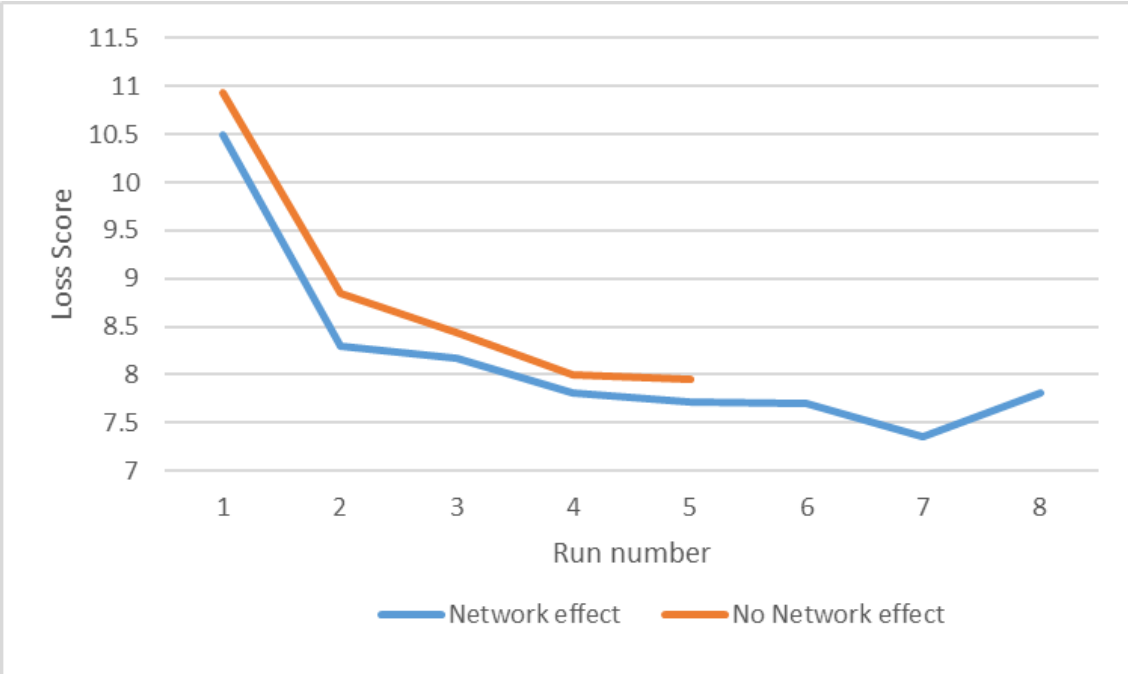


FIGURE 5.3: Test Loss Scores

Run 3 - PAL levels were made variable (previously set at an average value) and updated regularly according to age, gender and current BMI.

Run 4 - Parameters were introduced to modify the impact of educational level on the flow 'attitudes' and to parameterise the flow 'PBC'.

Run 5 - The BMI factor parameters were expanded to take account of 5 age ranges and gender.

Run 6 - The BMI factor parameters and those parameterising the 'norms' flow, were split into six age groups (previously 5) and by gender.

Run 7 - The BMI adjustment was dis-aggregated so that male-male, male-female and female-female, female-male comparisons were made

Run 8 - Using the parameter settings from Run 6, the parameter for the satisficing number was split into six age ranges and by gender.

The first parallel run (no network effect) was abandoned after five runs, when it became apparent through the test run loss scores, that it was consistently performing less well than the second. The second run was abandoned when the test loss score started to increase. See Figure 5.3

The addition of an additional age group to both genders in run 6, was triggered by a review of the test data from previous runs, where it became apparent that the model was consistently performing less well for the 16 -30 group in both genders than for the remaining groups.

A number of sources suggest that adolescent behaviour may well vary from that of young adults in terms of the impact of social networks Shoham, Tong, Lamberson, *et al.* [114] Daw, Margolis, and Verdery [115], and in terms of metabolism [101]. It was hypothesised that splitting the 16-30 age group into an adolescent group(16-20) and a young adult group (21-30) would improve model performance which turned out to be the case, although the initial improvement in test loss score was relatively small.

The 8th option involving splitting the satisficing number parameter using the same age and gender categories. This delivered an improved loss score in the descent run, but the test score was higher than its predecessor, suggesting overfitting. In retrospect this is unsurprising given that this version involved 46 parameter values.

The final parameter set is shown in Table 5.3

5.4 Forecast Accuracy

To gain some insight into the level of accuracy provide by the parameters, the BMI figures (means and medians) generated by the final parameter set were compared with the actual figures for the same time period (2004-2013), the results are given in Table 5.4 The negative values in the table indicate an issue where the model is forecasting low, whereas the positive values indicate over forecasting.

As one would expect there is some variation in the results for the highest and lowest age groups where weight change models are generally less accurate. The changes made in Run 6 have improved the accuracy for the younger age groups in the model, but there is still under-forecasting in the 21-30 age range for both genders, particularly for males. The variation is relatively consistent and therefore unlikely to be the consequence of a specific issue occuring in the real world within the time frame, but more likely an issue with the model itself, perhaps failing to address a specific factor (for example changing attitudes to physical activity) relevant to that age group. This is clearly an issue to be considered when looking at a forecast, but less so in the scenarios where the main data is derived from comparisons between forecasts (cancelling out any model issues).

Parameter	Effect
θ_0	norms male 16-20
θ_1	norms male 21-30
θ_2	norms male 31-45
θ_3	norms male 46-60
θ_4	norms male 61-75
θ_5	norms male 76+
θ_6	norms female 16-20
θ_7	norms female 21-30
θ_8	norms female 31-45
θ_9	norms female 46 - 60
θ_{10}	norms female 61-75
θ_{11}	norms female 76+
θ_{12}	BMI Factors male 16-20
θ_{13}	BMI Factor male 21-30
θ_{14}	BMI Factor male 31-45
θ_{15}	BMI Factor male 46-60
θ_{16}	BMI Factor male 61-75
θ_{17}	BMI Factor male 76+
θ_{18}	BMI Factor female 16-20
θ_{19}	BMI Factor female 21-30
θ_{20}	BMI Factor female 31-45
θ_{21}	BMI Factor female 46 - 60
θ_{22}	BMI Factor female 61-75
θ_{23}	BMI Factor female 76+
θ_{24}	male-male BMI adjustment
θ_{25}	male-female BMI adjustment
θ_{26}	female-female BMI adjustment
θ_{27}	female-male BMI adjustment
θ_{28}	network memory
θ_{29}	network range
θ_{31}	satisficing number
θ_{32}	lag duration
θ_{33}	diet duration
θ_{34}	network effect trigger
θ_{35}	PBC
θ_{36}	Education Level

TABLE 5.3: Final Parameter Set

Forecast v's Actual

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1	-0.01547	-0.02232	-0.0396	-0.04577	0.000761	0.015762	-0.01239	-0.00808	-0.00208	-0.00933	0.022224	0.018806
2	-0.6733	-1.55439	-0.60452	-0.86649	0.593904	0.963444	0.214294	0.291132	0.428351	0.209336	1.155675	0.779914
3	-0.52722	-0.63497	-1.48293	-1.65862	0.163773	0.476619	-0.13697	0.271434	0.255891	0.265046	1.10292	0.933673
4	-1.18944	-1.39932	-1.47372	-1.71882	0.13398	0.591926	-0.09625	0.162094	0.014692	0.05033	1.046688	0.930763
5	-0.7246	0.280609	-1.00534	-0.92512	0.296802	0.541565	-0.17287	0.06678	0.30399	0.500925	1.092522	0.940977
6	-0.28701	0.498358	-1.46277	-1.67804	0.015358	0.014489	0.065707	0.237121	0.132992	0.387329	1.074851	0.85332
7	-0.77901	-0.10012	-1.38683	-1.80527	0.240998	0.134695	0.083991	0.313956	0.470021	0.688869	1.241143	0.753171
8	-0.44505	0.722343	-2.25397	-2.26254	-0.32957	-0.42646	0.101484	0.399559	0.283388	0.643356	0.907157	0.580027
9	-0.35229	0.477596	-1.9285	-2.18167	-0.3096	-0.34205	0.058491	0.634221	0.443849	0.771513	1.004756	0.73497
10	-1.48217	0.040307	-2.08801	-2.32902	-0.80434	-0.69219	0.279668	0.738345	0.64557	0.967057	1.402621	1.280575

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.71898	-0.74272	-0.98557	-0.8093	0.059864	-0.07938	0.284752	0.25875	0.102935	-0.26733	-0.09394	-0.42076
-1.94446	-1.37651	-0.20784	-0.18275	-0.06833	-0.32785	0.330596	0.419383	-0.12207	-0.58521	0.707136	0.463007
-2.17324	-1.15468	-0.56482	-0.56243	0.030428	0.084126	0.376588	0.408625	0.255255	-0.15964	0.684539	0.494739
-1.47303	-1.04743	-0.31612	-0.40917	0.365854	0.203949	0.661834	0.587071	0.205363	0.077989	0.897377	0.487163
-0.57974	-0.42931	-0.75859	-0.78211	0.57754	0.653384	0.714584	0.516961	0.57402	0.203402	1.457556	1.379623
-0.47536	-0.42645	-0.44461	-0.39582	0.085857	0.102685	0.635749	0.385466	0.589538	0.42511	1.350186	1.198709
-0.90396	-1.5356	-1.12539	-0.72934	0.218131	0.356555	0.559206	0.449822	0.555533	0.746014	1.673341	1.258787
-0.06836	-0.40808	-1.32743	-1.04596	-0.12294	-0.20087	0.625574	0.50643	0.583235	0.441108	1.39143	1.310616
-0.80265	-0.53723	-0.96675	-0.57276	-0.21611	-0.09102	0.699666	0.537192	1.110931	1.195067	1.302622	0.889222
-0.40801	-0.62572	-1.52507	-0.90816	0.122452	0.244184	0.808156	0.696548	1.468835	1.58511	2.094633	1.731533

TABLE 5.4: Forecast BMI's v's Actual BMI's (2004-2013)

5.5 Parameter Values and Analysis

This section describes the parameter values derived from Run 7.

5.5.1 Norms

The norms flow value is derived from the the difference between the agents BMI and the average BMI's of the network neighbours (of the same gender), the parameter modifies it directly by acting as a multiplier for the value.

This initially started as a two values, one for males and one for females. In it's final iteration this was subdivided into 6 age groups, as shown. The values are shown in Table 5.5.

Age Range:	16-20	21-30	31-45	46-60	61-75	76+
Male	1.95	1.87	4.91	2.99	3.69	1.52
Female	2.84	3.85	1.02	4.99	0.87	3.725

TABLE 5.5: Norm Parameter Values

Given the variation in values it is clear why the subdivision of age groups aided the accuracy of the model with regard to the female population (it is less clear that it affected the male population).

Sensitivity analysis did broadly support the expectation that results would be less accurate with age and time, although the main impact on both positive and negative variation was in the 61-75 age group, males in the positive variation and female in the negative variation. (Table C.7.)

5.5.2 Educational Level

The educational level parameter defines the size of the attitudes flow in the TPB model, the final value for this was **5.56**, meaning that if they have a degree level qualification, then the flow becomes 22.24 $(4 * 5.56)$.

The rationale for this is described in 2.5.3, but in essence [34] relates educational level to SES and obesity through statistical analysis, citing cognitive ability (developed through exposure to degree level education) as the key determinant. Their definition of cognitive ability tallies closely with that used to define attitudes in the theory of planned behaviour [35].

Perhaps unsurprisingly the impact of varying this parameter in the sensitivity analysis was more diffuse, with less relation to age and time. (Table C.4.)

5.5.3 Perceived Behavioural Control

The perceived behavioural control flow is defined by the value of the weightIn flow multiplied by the PBC parameter, since the parameter value is negative, if the individual is losing weight, then the PBC flow is positive, conversely if they are gaining weight the PBC flow is negative. A single value is used for all individuals within the simulation in this case **-5.74**.

Unusually increasing PBC by 5% had no significant impact on the loss score (actually increasing it by 0.1%), decreasing it however produced one of the largest variations (3.5%). The issue was primarily with the male population, particularly the 61-75 age group. (Table C.8.)

5.5.4 Lag

The lag between Intention and Behaviour in the model was parameterised and the value returned was **1** (month).

Sensitivity analysis revealed a diffuse pattern. (Table C.5.)

5.5.5 Satisficing Number

The satisficing number defines the threshold value above which the input from the network neighbourhood (the difference in BMI between the individual and the network neighbourhood average for the same gender), registers with the individual and starts to affect behaviour. It is positively correlated with BMI (see Table 4.4, but non-linear. The parameter value is **3.5**, which implies a satisficing number of 0 for an individual whose BMI is 20, rising to 1 for an individual whose BMI is 30 and 4 for an individual with BMI 40.

The pattern revealed in the sensitivity analysis was of the more diffuse type. (Table C.10.)

5.5.6 Diet Time

The diet time addresses the issue of lack of consistency in dieting behaviour, both in duration and in adherence, for simplicity this was simplified to duration using a standard figure for daily shortfall of 480Kcal. Each time a diet is initiated the time is chosen at random from a parameterised gaussian distribution. The final parameter values for the distribution gave an average diet time of **7.93** months with a standard deviation of **1.98** months.

Unsurprisingly given the non-specific nature of the parameter, sensitivity analysis gave a diffuse pattern of impact. (Table C.3.)

5.5.7 Modelling Changes in Weight

Established models and data were used in this element of the model with the addition of 2 elements; the satisficing number and intermittancy in dieting behaviour.

With regard to the satisficing number it was notable that a single value provided the best model and that subdividing it by gender and age resulted in overfitting. Anecdotally, the diet time figure seems high, but in retrospect it could be predicated on a relatively conservative daily calorie shortfall (480Kcal). It may well be that the majority of people dieting undertake more demanding regimes and consequently their adherence may be shorter lived [104].

5.5.8 BMI Factor (Environmental Impact

APCCC is the proxy for the environmental issues described in [27] in the segments relating to 'Food Consumption' (food exposure, energy density, portion size, alcohol consumption etc.), 'Food Production' (market price, increased production efficiency, purchasing power etc.) and 'Social Psychology' (primarily exposure to media and the impact of food advertising).

The BMI factor modifies the effect of the APCCC, by modifying the EI for that individual in light of the parameter. The calculations involve a range of constants specific to the individual's age, gender, current BMI and so on. The individuals EI is inversely proportional to the parameter, so a larger parameter value indicates a reduced environmental influence.

Age Range:	16-20	21-30	31-45	46-60	61-75	76+
Male	3.5	3.27	4.24	3.9	4.5	4.53
Female	4.9	4.52	4.65	5.12	4.8	5.1

TABLE 5.6: BMI Factor Parameter Values

The values given in Table 5.6 suggest that as age increases the impact of environmental issues such as food availability, advertising and so on decrease. More tellingly they consistently indicate that the impact of environmental factors is higher on males than females.the average value of the parameters for each gender. The sensitivity results are shown in Table C.2.

5.5.9 Network Trigger

This controls the impact of homophily on network formation, and defines the threshold value (BMI difference between agents) above which network formation is affected by differences in BMI between individuals, the value is **4.31**. This suggests a fairly broad tolerance before difference in BMI start to affect the network dynamics.

The pattern revealed in the sensitivity analysis was also of the more diffuse type.

5.5.10 BMI Adjustment (Homophily)

Male-Male	0.74
Male-Female	0.76
Female-Female	1.27
Female-Male	1.25

TABLE 5.7: BMI Adjustment Parameter Values

Once the threshold referred to above has been passed then BMI adjustment pertains to the level of influence on the network (forming connections) exerted by that difference in BMI. The values are given in Table 5.7.

Looking at the values, if a male perceives a difference between their BMI and that of a potential contact the edgeweight in the history file is reduced by 42.6% or 42.4% (depending on the gender of that contact), a female reduces the edgeweight by 37.3% or 37.5% (according to gender). The impact in the sensitivity testing was fairly diffuse. (Table C.1.)

A value of 5 would be neutral, signalling that there is no effect. Given the difference in the values between genders, an additional pair of tests were run. The tests checked for BMI and gender based assortativity within the social networks.

The test design involved running scenario 1 (see below), and collecting data on individual's network neighbourhoods, and using this to determine the likelihood of the assortativity effects:

- Test 1 - data was collected describing the gender of each individual and the proportion of same gender to different gender in their network neighbourhood after 10 years model time, and aggregated over a 1,000 iterations. The mean, median and standard deviation values were collected for each gender. A significant deviation from the expected values (based on the gender numbers in the original data) would indicate the existence of a gender effect.
- Test 2 - After a 10 year model run, 2 sets of data were collected, the first recording the difference between each agents BMI, the average BMI for the agents in their network neighbourhood, and the second the difference between that agents BMI and the average BMI for agents of the same gender in that network neighbourhood. Both data sets were further averaged over 100 runs. A value of 0 with minimal variation would indicate a high level of BMI based assortativity (or alternatively a completely random situation).

The results from Test 1 indicated that there was no significant gender effect.

The results from Test 2 are more complex and are described in Table 5.8:

	16-20	**21-30**	**31-45**	**46-60**	**61-75**	**76+**
Network Neighbourhood	-4.79	-3.45	-0.50	1.03	1.17	0.79
Same Gender	-4.91	-3.36	-0.55	1.01	1.24	0.76

TABLE 5.8: Test 2 - BMI Assortativity

A negative value suggests that the average BMI against which the agent's BMI was compared was higher that that of the agent, a positive value indicates that it was lower.

What the results seem to illustrate is an effect that varies according to age, describing a situation where the average BMI of the network neighbours of the youngest agents in the network is considerably higher than that of the agent themselves. At the other end of the age spectrum the difference is both lower and positive, suggesting a much higher level of BMI based assortativity and a network neighbourhood where the BMI is genrally lower than that of the agent themselves.

More detailed investigation of the data suggests that this effect is slightly different when comparing across genders with the youngest female age group having a larger (more negative) effect than the males in the same age group.

5.5.11 Range

Range within the network model defines the maximum number of individuals an agent can link to, the higher the value the more possible connections. The value returned is **34**, this marginally higher than 30, the value used in modelling idealised social networks [50], but broadly consistent. It may also explain in part the denser topology (higher clustering coefficients and transitivity) exhibited by the final model.

Sensitivity analysis revealed a pattern of increasing inaccuracies with age (but not time), this may be connected to the formation of connections where homophily based on age also plays a part. (Table C.9.)

5.5.12 Memory

Memory is also a function of the network algorithm, and defines the dynamic behaviour of the network at an individual level, the value derived from the parameter is **32**.

Memory as a parameter applies equally to all individuals, and the sensitivity analysis revealed a diffuse pattern of impact. (Table C.6.)

5.6 Network Topography

Within the model, the parameters for Network Trigger, BMIAdjustment, Range and Memory all had the potential to affect the network topography. The topographical data for the test runs incorporating a 'network effect' are given in Table 5.9.

Metric	1	2	3	4	5	6	7	8
Loss Score	10.49	8.29	8.18	7.81	7.71	7.60	7.35	7.81
No. of Nodes	1,037	1,047	1,047	1,047	1,061	1,063	1,062	1,038
No. of Edges	3,638	4,046	4,049	4,049	4,965	5,097	4,960	3,663.5
Clustering Coefficient	0.68	0.67	0.67	0.67	0.67	0.67	0.67	0.68
Transitivity	0.64	0.64	0.64	0.64	0.63	0.63	0.63	0.64
Assortativity	0.44	0.42	0.42	0.42	0.34	0.33	0.34	0.44
Shortest path	17.4	16.37	16.23	16.34	12.21	11.8	12.18	17.22
Network Diameter	45	42	42	42	31	29	30	45
Mean Node Degree	7.01	7.73	7.74	7.74	9.35	9.59	9.35	7.06
No. of Components	20	11	11	11	3	3	3	19

TABLE 5.9: Topographical Results (Medians) - Network Effect Runs

It's interesting to note the step change that seems to occur in the network structure between the 4th and the 5th run for nearly all the values except the clustering coefficient and transitivity, whilst the loss score seems to continue decreasing at a relatively steady rate. This step change delivers a more densely connected network, with fewer components, higher mean degree, smaller diameter and smaller shortest path. Looking at the parameter values that impact on the network structure, there is no obvious single parameter variation driving this (perhaps indicating a saddle point of some description).

Assortativity (correlation between node degree) decreases at the same point, suggesting a more heterogeneous network. The clustering coefficient and transitivity remain pretty stable with only minor variations, the former is very similar to the clustering coefficient (0.66) calculated for the network inferred from the Framingham heart study data [7] in [57].

The detailed results for the final model are given in Table 5.10.

The relatively small differences between the means and medians suggest a symmetrical distribution and the relatively low standard deviations suggest consistent data. The exceptions are the shortest path, network diameter and components, these are of course interlinked, since they are measured in the largest component. More components would suggest smaller component sizes and so shorter path lengths and diameters. At first sight, the data here suggests that relationship seems to be reversed with less components delivering reduced diameter and shortest path. To achieve this requires a more densely

Metrics	Median	Mean	Std. Deviation
Nodes	1,062	1,061.69	5.48
Edges	4,960	4,964.31	100.78
Clustering	0.67	0.67	0.01
Transitivity	0.63	0.63	0.01
Assortativity	0.34	0.34	0.05
Shortest path	12.18	12.31	0.8
Diameter	30	30.86	2.85
Average Degree	9.35	9.35	0.19
Components	3	3.46	1.62

TABLE 5.10: Topographical Results - BestFit Model

connected network (as is indicated by the increased mean node degree), this is also accompanied by a small reduction in transitivity (the number of closed triplets as a proportion of the total number of triplets in the network) and a more significant reduction in assortativity, suggesting a more heterogeneous network with fewer dense clusters than previously.

The distribution of the component numbers suggested by the data is atypical with high standard deviation (54% of the mean) and a significant difference between median and mean, suggesting much higher variability.

An analysis of data from networks generated in Scenario 1, also suggested that there is a significant negative correlation between an agents BMI and its 'network neighbourhood' size. For the male population the average size varied from 9.40 for those with a BMI under 25 to 8.82 for those with a BMI over 30. For the female population the equivalent values were 9.2 and 8.56.

Ultimately however we are interested in the metrics which would define the likely social network (independent of network size) involved in the spread of obesity. This data would suggest that they are:

Clustering Coefficient of 0.67.

Transitivity of 0.63.

Assortativity of 0.34.

Mean Node Degree of 9.35.

These are consistent with data collected from other social networks [47]. The network would also exhibit:

BMI based assortativity

An inverse correlation between 'network neighbourhood' size and BMI

5.7 Scenarios

The scenarios sought to investigate the influence of two factors; the impact of childhood obesity and that of changing environmental factors. This were represented by the mean BMI of 16 year olds joining the simulation as it runs and the APCCC.

In both cases the data was scanned to identify the fastest period of sustained growth, and this became the default rate of change in both the growth and decline scenarios. The different permutations are described in Table 5.11.

Scenario 1. was run for 10 years from 01/01/2014 (the last year for which age specific data was available) to 31/12/2023, and it assumed no change in the two factors. The output comprised a raw forecast (derived from the model with no amendments) and an adjusted forecast where the raw forecast was modified by incorporating the data from Table 5.4. A range of other scenarios were also considered; different combinations of the two factors, comparisons between scenarios, extreme scenarios and scenarios with and without network effect. The full set of outputs are given in Appendix C.

The initial scenarios addressing different combinations of the two factors, are described in Table 5.11

	Falling APCCC	**Static APCCC**	**Rising APCCC**
Falling BMI	4	6	9
Static BMI	3	1	8
Rising BMI	2	5	7

TABLE 5.11: Scenario Permutations: Adolescent (16 year old) BMI v's Environmental influence (APCCC)

For scenarios 1 - 9, the first of the two pairs of tables (in the 'Raw Forecast' Table) forecasts the expected mean and median BMI's for each age range and gender over the period, the second pair of tables record the difference between the mean and median values in that forecast. The latter provides an indication of how the BMI distribution in that gender and age group might be changing. To aid interpretation, the difference is also displayed in a graph, a generally positive gradient potentially indicates an extending right tail and vice versa.

In the second set of tables 'Adjusted Forecast and Comparison', the first pair present an adjusted forecast using the data from Table 5.4, the second pair are a comparison with the output from Scenario 1 (Table C.12.), which models a scenario where there the two factors under examination continue at their current level. For obvious reasons this is omitted for scenario 1 itself.

5.7.1 Scenario 1 - Static BMI and Static APCCC

As the tables in Tables C.12 and C.13 illustrate whilst the BMI of 16 year olds entering the simulation remain stable (along with the APCCC), the situation itself is not stable with respect to comparator data taken from 2013. In the 'Mean v's Median' tables all the male age categories show a value increasing with time, suggesting that the right tailed distribution is increasing. Conversely the youngest female age category shows an improving situation with a potential reduction in the number of very high BMI values. This problem becomes more apparent in the adjusted forecast.

Comparing the adjusted forecast with the data from 2013, we see a complex pattern for both genders. In the 16-20 and 21-30 age groups there is some improvement with time (if only in the decrease in the range of BMI values in the male population). Conversely in the older population groups for both genders we see a worsening scenario, this time more evident in the female population. The trends displayed do not seem to have stabilised by the end of the time period, but may be slowing a little.

5.7.2 Scenario 2 - Rising BMI & Falling APCCC

Tables C.14 and C.15 presents a less extreme picture, the 'Mean v's Median' tables suggest a much less negative scenario in the younger age ranges in both genders, and in the comparison with scenario 1, the values for the older age ranges remain stable and relatively static. show a less marked increase than previously.

The comparison with Scenario 1, shows issues in the youngest age group (which also appears after 4-5 years in the next age group) for both genders, but otherwise is neutral or marginally positive.

5.7.3 Scenario 3 - Static BMI & Falling APCCC

Unsurprisingly the story portrayed in Tables C.16 and C.17 represents a small general improvement across the board when compared to Scenario 1.

5.7.4 Scenario 4 - Falling BMI & Falling APCCC

Scenario 4 represents the best case and Tables C.18 and C.19 and creates a complex conundrum, when compared to scenario 1 it is very consistent overall, with very few significant differences, the main impact within the male population is to increase the BMI distribution in the youngest age group, within the female population there is a similar effect but less pronounced. There is a general improvement but it's impact is small.

5.7.5 Scenario 5 - Rising BMI & Static APCCC

As one would expect in the comparison with scenario 1, shows little variation except in the initial age groups, with the male data decreasing initially but with increasing BMI disparity, the female disparity also rises as does the overall BMI. Tables C.20 and C.21.

5.7.6 Scenario 6 - Falling BMI & Static APCCC

Scenario 6 replicates the patterns of scenario 5, but with the lower values you would expect given that the BMI is falling not rising. Tables C.22 and C.23.

5.7.7 Scenario 7 - Rising BMI & Rising APCCC

Scenario 7 confirms a pattern of non-elasticity in the results, with most (but not all) of the values rising but only a minor amount. Tables C.24 and C.25.

5.7.8 Scenario 8 - Static BMI & Rising APCCC

Scenario 8 indicates a minor impact increasing across the board on BMI and it's distribution. Tables C.26 and C.27.

5.7.9 Scenario 9- Falling BMI & RisingAPCCC

Scenario 9 reverts to a relatively neutral pattern except in the younger age groups which are impacted by the falling BMI. Tables C.28 and C.29.

5.7.10 Extreme Scenarios

Given the relatively small impact of the rates of change, 2 additional scenarios were run, doubling the rates of change, and comparing rising and falling BMI (for new agents) and rising and falling APCCC. The results are given in Tables C.30 and C.31. The impacts of changing BMI are as expected, with the impact directly evident in the youngest age groups and after five years in the next age group. The impact of changing APCCC was evident in the older age groups and in particular the female population

5.7.11 Cross Scenario Comparisons

In the tables referred to in this section, values highlighted in red indicate the areas of most impact.

Table C.32 compares the output from Scenario 7 (rising BMI and rising APCCC) with that from Scenario 4 (falling APCCC and falling BMI). It also compares Scenario 9 (rising APCCC and falling BMI) with Scenario 2 (falling APCCC and rising BMI).

Table C.33 seeks to isolate the different effects by comparing rising and falling BMI (Scenarios 5 and 6) whilst the APCCC remains static, and vice versa (Scenarios 8 and 3)

The first of these confirms the impact of BMI to be primarily on the youngest age groups, with little or no evidence to suggest they influence other age groups in the population.

APCCC has a more general (but not large) effect across the population, perhaps slightly more evident in the older female age groups.

A cross comparison for both these scenarios showed that impact on distribution was negligible. There was a positive impact in the 21-30 age group as the simulation progressed but that was directly attributable to the BMI profiles of the individuals joining the simulation.

5.7.12 Counterfactual Comparisons

The model assumes that the there is feedback between the network dynamics and the way obesity spreads, a set of comparison were also made to assess the impact of weakening or removing that effect using scenarios 1, 4 and 7. These were re-run with the network effect removed, and the results compared with the original scenarios. The results are given in Tables C.34, C.35 and C.36.

These create a complex picture, the scenario 1 comparison (Table C.34.) suggests a broadly similar impact across the genders and age ranges, marginally negative in the male population and perhaps more neutral in the female population, but none of the figures represent variations of more than 0.5%

The scenario 4 (falling BMI and APCCC) comparison gives a much clearer picture (Table C.35.) with removal of the network effect delivering an improved result across the board for the older female population and to a lesser extent the male population. However, the maximum gain is still less than 1%. However, for the 16-20 age group in each gender (and in the 21-30 age groups after 5 years) there is an increasingly strong negative effect, with a maximum impact of 6%. This effect is almost precisely reversed with scenario 9 (rising BMI and APCC) with figures of similar magnitude.

In summary the network effect is most apparent in a dynamic rather than a static situation. For the majority of the population the network effect seems to be relatively low (consistent with the in-elasticity of the network already commented on) and acts to maintain the 'staus quo' by mitigating changes in obesity levels, reducing the impact of falling or rising BMI and APCCC. This effect seems to be reversed for the youngest age group where the effect is to amplify changes in the BMI and APCCC, it is also much more significant. .

5.8 Summary

This chapter has reviewed the model output under four headings, forecast accuracy, parameter values, network topology and scenarios.

In terms of accuracy the 'raw' forecast produced by the model performs well for the majority of the population, with the only significant issue arising in respect of the male population aged between 21 and 30. This was ascertained by comparing the 'raw' forecast with actual data from the same period (2004 - 2013). Two strategies are used to mitigate the issue.

In the first strategy data from the comparison was used to adjust the raw forecast. Under- and over- forecasts for each age group and gender were recorded and over-laid onto the raw scenario forecasts. This approach suggested that the current situation was not stable, worsening over the forecast period, with a particular issue for the younger male population.

In the second the 'raw' scenario forecast was contrasted with the 'raw' forecast from scenario 1, potentially cancelling out any model issues. The data generated from this was of a different order but useful for understanding the impact of the different scenarios. It served to highlight the 'in-elasticity' of the situation to significant changes in adolescent BMI and/or environmental factors.

The primary value of understanding the parameter values lies in their joint ability to throw some light onto the behavioural model (discussed in the following chapter). The other key area addressed by these was the impact of environmental factors on different sub-groups of the population. The model suggests that in general males are more susceptible than females and that this susceptibility decreases with age.

The model clearly defines the topology of a 'representative' network, and the metrics are broadly consistent with the expected values of a social contact network.

In all of the scenarios a common element is the variation in the level of impact on different age groups with older age groups typically experiencing a maximum impact of 1% on Mean BMI, with the younger age groups (<31) this can be as high as 6%. The extreme scenarios doubled this impact.

Reducing or raising the BMI profile of individuals primarily affects the 2 younger age groups in each gender, whereas reducing or raising the level of APCCC has a much broader impact across the population.

The scenarios are surprisingly in-elastic to changes in those inputs, these mirrored the fastest sustained rises in the inputs (0.1 BMI per year and 6Kcal per year) on the assumption that this would be the realistic maximum that a sustained intervention could achieve. In reality the impact of either or indeed both was minimal. Doubling the rate of change produced an increased effect, but the gains were still minimal when compared to the effort likely to involved in achieving those rates of change.

The effect of changes in BMI and APCC are mitigated by the network effect for the majority of the population. However, for the youngest part of the population this effect seems to be reversed acting to amplify the effect of those changes.

Scenario 1 demonstrated that the 'system' is not in a state of equilibrium. The data over a 10 year period did not suggest that equilibrium would be established within that period. The prognosis for the female population over the age of 45 is poor with fluctuating Mean BMI's and distributions. The prognosis for the Male population over the age of 60 is similar.

This chapter has described the relevant outputs from the development, testing and experimentation phases of the simulation, the next looks at interpreting these in light of the research questions posed in Chapter 1.

Chapter 6

Research Outcomes

"A wise man proportions his belief to the evidence."

David Hume (1748)

6.1 Introduction

The purpose of this chapter is to interpret the results reported in the previous chapter in light of the research objectives outlined in Chapter 1. An initial section reviews the limitations imposed on the results and interpretation by the methodology and availability of data, subsequent sections address each of the three objectives.

6.2 Limitations

To understand the limitations of the results reported in the previous chapter, it is worth recalling again the process by which the model was developed.

In terms of the model development:

- Theoretical models were combined together to form a hybrid simulation that was capable of reproducing the behaviours observed in the data, and by subject matter experts.
- Initial parameter settings were based on theory (where available), expert advice or experimentation.
- A training data set was used in conjunction with a gradient descent methodology (incorporating a simultaneous perturbation stochastic approximation algorithm), in order to refine the initial parameter settings.

- Increasingly complex settings were refined with a second (development) data set.
- Each iteration of the revised model was tested using a third test data set, in order to select the best performing model.

Thus the parameter values have not been experimentally verified, rather they define a model that can achieve a certain level of accuracy in forecasting obesity, and that therefore (subject to that level of accuracy) are **likely** to represent many of the key relationships amongst the parameters.

The study has a number of other limitations:

- Translating social theory into mathematical relationships is always challenging, and whilst the TPB is comparatively well researched from a social science perspective, there is very little data available on which to base the initial parameter settings, or with which to operationalise the model.
- The issue of variability in dieting behaviour is a complex one, it may be that the method chosen in this model to represent it was not sufficiently flexible to address the behaviours of the population as a whole.
- The data set available for the 16-20 age group was small in comparison to the remainder and suggested large variations in BMI. It is not clear whether this reflects the actual situation, or is a consequence of the small size of the data set and collection issues.
- The simulation size and duration were defined by the available computer processing power this dictated a population of a 1,000 agents in the network and a 10 year time frame.
- The PAL data used in the model, it is taken from [102]. It involves taking one of three values dependent on BMI and age. This probably fails to adequately represent the variation in individual physical activity/exercise likely to occur within the network.
- The simulation had issues forecasting for the 21-30 age groups. The raw forecast for the final version was able to forecast out to 10 years and to within 0.5 BMI (average error) for both genders in the remaining age groups. An adjustment strategy was used to improve the forecast accuracy.
- The definition of an 'obesity' network is problematic, the multiplex nature of networks, the diffuse boundaries of such a network and the fact

of constant evolution (as the role of social media changes) make it difficult to represent such a network precisely. This research has identified a representative network that approximates the impact of these in a single undirected network. Thus the network parameters derived from the model do not represent a single network, but an amalgam of several different networks (enabled through social contact networks, technology and social media platforms). This proxy network approximates the overall impact of the different elements of that amalgam during the time-frame used by the training/testing phases (2002-2014). Our forecasts assume that this approximation remains valid for the period of the forecast (2014-2023)

- Over the lifetimes of the individuals represented in the simulation (and in the relevant literature), the process of social networking has changed radically. It is possible that different age groups represented within the simulation interact with their networks using different mixes of the available methods.
- Much of the original research in this area used directed networks, our approach has necessitated the use of an un-directed network, making comparison of results complex.
- Much of the following discussion references age groups, it's worth recalling that the age groups in the model were chosen arbitrarily.

6.3 Research Findings

To recap, the research objectives identified in Chapter 1 were:

1. Explore the nature of the interaction between social networks, obesity and the behaviours that drive it, in particular to understand:
   (a) The topology of an 'obesity' network.
   (b) Whether that topology is modified by the spread of obesity?
   (c) Whether homophily play a role in that interaction?
2. Develop a generalisable model to facilitate that exploration, incorporating concepts from behavioural science, social network realisation and simulation to explore the impact of different external parameters on the interaction.

3. Apply that model to specific data for a region/country in a case study, in order to understand:
   (a) How that impact might vary for different sub-groups of the population?
   (b) Which sub-groups might make the most demand on healthcare resources in the future?
   (c) What are the managerial and theoretical insights in terms of both behaviour and social networks, that might be used to augment existing intervention strategies, or suggest new ones in the region/country under consideration?

Given the processes and limitations described above, the insights described in the sections that follow are best considered as **hypothesised** conclusions, or insights that need to be evidenced by further research.

6.4 The Nature of the Interaction

As a consequence of previous research [5] the model assumed that social networks had an effect on the spread of obesity, one aspect of the research looked to understand whether the network topology was itself affected by that facilitation process (network effect).

In the research models, incorporating a network effect (BMI assortativity), consistently outperformed identical models without that network effect. It is reasonable therefore to infer that the process of transmission modifies the network, which in turn affects that transmission process in a process similar to that of a reinforcing or balancing loop.

Revisiting the cave system analogy used in Chapter 1. the chambers and the connecting passages of our caves are mutable, and are modified by the individuals within them.

Unlike our analogy, the data from the counterfactual scenarios suggests that this network effect is experienced differently by different parts of the population, in the older population it acts to inhibit the impact of changes in the external environment (APCCC) and adolescent BMI levels (a balancing loop). In the youngest part of the population (<21), it appears to amplify their effect (a reinforcing loop). There is also a significant difference in impact with the younger population much more susceptible to that effect.

Metrics	Median	Mean	Std. Deviation
Clustering	0.67	0.67	0.01
Transitivity	0.63	0.63	0.01
Assortativity	0.34	0.34	0.05
Average Degree	9.35	9.35	0.19

TABLE 6.1: Topology of an 'Obesity' Network

6.4.1 The Topology of an 'Obesity' Network

The standard metrics (clustering coefficient, transitivity, degree assortativity and average degree) defining the network that enable these effects are broadly consistent with data collected from other social networks [47]. See Figure 6.1, which gives the relevant data for the case study in our model.

However our network does exhibit some other features that may not be so consistent:

- The BMI assortativity data described in Table 5.8, suggests that the difference between agent BMI and the average of their networks BMI, changes significantly with age. The youngest agents generally associating with individuals (in their network neighbourhoods) whose BMI's are much larger than their own, transitioning to a situation where the older agents are associating with individuals hose BMI's are marginally lower than their own.
- BMI and network neighbourhood size are inversely correlated.

Hypothesising:

- The magnitude of the network effect on an agent is correlated to the size of the difference between their BMI and the average BMI of their network neighbourhood, and possibly the number of agents within that neighbourhood.
- Where that difference is negative (their BMI is less than of their network neighbourhood), the impact of changes in the external environment is reinforced, where positive they are minimised.

It should be noted that the BMI assortativity data suggests a gradual transition across the age range, and the relatively high levels of network effect are confined to the youngest age groups, so if the hypotheses are correct then the relationship is complex.

Is the Topology Modified by the Spread of Obesity

It's clear from the narrative in Section 5.3 and above that the model performs more accurately when there is an interaction between the network and the spread of obesity, suggesting that the the topology is itself modified by its role in the spread of obesity.

The Role of Homophily

The theory is clear that if that the topology is modified by the spread of obesity, then Homophily is likely to provide. the vehicle to facilitate this. Our model reflect this, and gives us some insight into the relative impact of that influence, with a significantly greater effect in males than females (this is detailed in Section 5.5.10).

6.5 A Generalisable Model

In this research we have developed a model using an HSM approach combining HS, concepts from behavioural science and methods from machine learning (stochastic optimisation). The HS element involves two components one reproducing an individual behavioural/decision making process and the other the social network in which that process takes place.

The behavioural component is based on a specific health behaviour model that we considered appropriate in the context of obesity

The network generated is in effect a proxy network, mimicing the combined effect with regard to obesity of the multiple networks (social contact and social media) that an individual may belong to.

We can consider its generalisability at two levels:

- Could this modelling approach be applied to other countries/regions to address the same research objectives?
- Could this modelling approach be applied to explore the impact of social networks on other NCD's?

With regard to the first level, the answer is clearly in the affirmative. There is no evidence to suggest that TPB is specific to a single culture or population segment, although it is likely that the relative impact of the parameter inputs may well vary with cultural differences. Given this, the optimisation process will ensure that the relevant parameters are 'tuned' appropriately. Similarly for the network element the topology may well vary according to location and culture, but again the optimisation process can address that. The critical issue with this level of generalisation is ensuring that there is enough data with which to train the model.

With regard to the second level of generalisation, the answer is more complex. The network component remains viable (given enough data to facilitate the optimisation process). The behavioural model used in the behavioural/decision making component would have to be reconsidered, specifically whether TPB was an appropriate choice. In issues like smoking, drug use and certain types of eating disorder (BN and BED) it is unlikely to be the best option. In these circumstances generalisability depends on the type of behavioural model considered most appropriate to the issue and its suitability for incorporation in an HS using an appropriate simulation paradigm.

6.6 Case Study

The results described below relate to the adolescent (16+) and adult population of England.

6.6.1 Variation Within the Population

Table 6.2 gives two pairs of tables, the upper pair are the adjusted forecast data based on Scenario 1 (incorporating the data from Table 5.4), the lower tables compares these values with 2013 levels.

Looking at these in more detail:

- Male 16-20; the rises in mean BMI and the fall in mean-median difference stabilise after 4 years , the final mean figure is 25.335 which is above the upper boundary of the healthy BMI category suggested by the NHS (25). The decreasing mean-median difference suggests that the number of severely obese individuals will decrease over the same period before it too stabilises.
- Male 21-30; a generally static value for the mean BMI, with a decrease in the median suggests a slightly improving position overall, with a small increase in BMI offset by a reduction in severely obese individuals.

Scenario 1: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.992	21.828	26.071	25.336	27.572	26.316	28.869	27.842	28.916	28.485	27.408	27.507
2015	23.397	22.572	25.413	24.651	26.831	25.536	28.608	27.627	28.624	28.328	26.956	27.213
2016	24.250	21.652	25.797	25.187	27.342	26.124	28.988	27.721	28.917	28.345	27.272	27.329
2017	24.977	22.488	25.921	25.341	27.291	26.008	29.113	27.910	29.201	28.562	27.568	27.563
2018	24.577	20.880	25.295	24.307	27.199	26.089	29.238	28.056	29.137	28.349	27.634	27.654
2019	24.139	20.663	25.764	24.819	27.337	26.608	29.071	27.917	29.603	28.792	27.795	27.847
2020	24.631	21.261	25.549	24.441	27.066	26.470	29.153	27.906	29.342	28.535	27.894	28.114
2021	24.298	20.439	26.409	24.659	27.602	26.929	29.235	27.941	29.599	28.671	28.346	28.396
2022	24.205	20.683	26.264	24.578	27.421	26.684	29.307	27.694	29.566	28.524	28.573	28.553
2023	25.335	21.121	26.392	24.531	27.847	26.974	29.213	27.697	29.476	28.363	28.466	28.217

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
	24.436	23.452	26.483	24.859	26.813	25.678	28.266	26.732	28.627	28.290	28.374	28.114
	25.668	24.136	25.545	24.178	27.127	26.085	28.247	26.633	28.957	28.662	27.965	27.725
	24.776	23.169	25.804	24.504	27.218	25.724	28.346	26.581	28.742	28.395	28.251	28.018
	24.166	23.163	25.822	24.592	26.697	25.514	28.277	26.539	29.035	28.408	28.152	28.072
	23.362	22.646	26.240	25.043	26.631	25.159	28.407	27.007	28.666	28.226	27.856	27.303
	23.258	22.643	25.867	24.747	27.073	25.648	28.553	27.067	28.865	28.329	28.064	27.645
	23.687	23.752	26.481	24.899	26.998	25.409	28.784	27.180	28.996	28.020	27.861	27.681
	22.851	22.625	26.431	24.996	27.468	26.075	28.822	27.250	29.057	28.393	28.362	27.812
	23.585	22.754	25.933	24.273	27.775	26.220	28.724	27.029	28.703	27.744	28.584	28.253
	23.191	22.842	26.225	24.488	27.570	25.880	28.601	26.968	28.491	27.265	28.115	27.716

Adjusted Scenario 1 v's Comparator Data(2013):

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.400	-0.493	0.177	0.145	-0.008	-0.605	0.128	-0.339	0.367	0.501	-0.399	0.039
2015	0.005	0.251	-0.481	-0.541	-0.749	-1.384	-0.132	-0.554	0.075	0.344	-0.850	-0.256
2016	0.858	-0.668	-0.097	-0.005	-0.238	-0.797	0.247	-0.460	0.367	0.360	-0.534	-0.140
2017	1.585	0.168	0.027	0.149	-0.289	-0.913	0.372	-0.272	0.652	0.577	-0.238	0.094
2018	1.185	-1.440	-0.600	-0.885	-0.381	-0.832	0.497	-0.125	0.587	0.365	-0.172	0.186
2019	0.748	-1.658	-0.130	-0.373	-0.243	-0.313	0.330	-0.264	1.053	0.808	-0.011	0.378
2020	1.240	-1.059	-0.345	-0.751	-0.514	-0.451	0.412	-0.275	0.792	0.550	0.088	0.645
2021	0.906	-1.882	0.514	-0.533	0.021	0.008	0.495	-0.241	1.049	0.686	0.540	0.927
2022	0.813	-1.637	0.370	-0.614	-0.160	-0.237	0.566	-0.488	1.016	0.540	0.767	1.085
2023	1.943	-1.200	0.497	-0.661	0.266	0.053	0.472	-0.484	0.926	0.378	0.659	0.748

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
	0.629	0.716	0.630	0.170	-0.326	-0.043	0.445	0.048	0.378	1.140	1.136	1.075
	1.861	1.401	-0.309	-0.511	-0.012	0.365	0.426	-0.051	0.708	1.513	0.727	0.686
	0.969	0.433	-0.049	-0.185	0.079	0.004	0.524	-0.103	0.493	1.245	1.014	0.979
	0.359	0.427	-0.032	-0.097	-0.442	-0.206	0.456	-0.145	0.786	1.258	0.914	1.033
	-0.445	-0.090	0.387	0.355	-0.508	-0.562	0.586	0.323	0.417	1.076	0.619	0.264
	-0.549	-0.093	0.013	0.058	-0.066	-0.072	0.732	0.383	0.616	1.180	0.826	0.606
	-0.121	1.016	0.628	0.210	-0.142	-0.311	0.963	0.495	0.747	0.871	0.623	0.642
	-0.956	-0.111	0.578	0.307	0.329	0.355	1.001	0.565	0.808	1.243	1.125	0.773
	-0.222	0.018	0.080	-0.416	0.636	0.500	0.903	0.344	0.454	0.595	1.347	1.214
	-0.616	0.107	0.372	-0.201	0.431	0.160	0.780	0.284	0.242	0.115	0.878	0.677

TABLE 6.2: Initial Forecasts

- Male 31-45; initial significant decreases in both mean and median values are not maintained, with final values showing a deterioration in BMI and a marginal improvement in distribution.
- Male 46-60; an increasing mean BMI (28.9 to 29.2) firmly in the unhealthy weight category, combined with mean-median values that are stable again denotes a worsening scenario, but with little change in the number of extreme BMI cases.
- Males 61-75; the final mean value for this age group is 30.1 (having increased from 28.9), moving the group into the obese weight category (>30). There is also a small increase in the mean-median differences suggesting a small increase in the number of severely obese individuals.
- Male 76+; with an initial mean of 27.4 rising to 28.5 and a similar growth in mean-median difference, suggesting a worsening situation.
- Female 16-20; unlike the males in the same age category, the falls in mean BMI and mean-median difference indicates a general improvement in the situation. The final value is lower than the males and the mean-median values suggest a decreasing number of severely obese individuals.
- Female 21-30; broadly static values for mean and medians suggest a stable and relatively unchanging picture.
- Female 31-45; these follow a broadly similar pattern to their male counterparts, but with an increased difference between mean and medians resulting in a worsening distribution (higher levels of severe obesity).
- Female 46-60; very similar values to the males in this age group although with a marginally increasing mean-median difference.
- Female 61-75; an initial value for mean BMI of 28.6 remains relatively constant, this is combined with a decrease in the mean-median difference over the time period, suggesting a marginally improving situation.
- Female 76+; reducing mean and median values suggesting an improving situation.

Looking at these there seem to be a number of areas that are likely to change significantly over the course of the forecast:

- Mean obesity levels in the male population 61-75 and 76+ will continue to rise significantly, along with an increase in the proportion of severely obese individuals.

- Mean obesity levels in the female population 46-60, 61-75 and 76+ will rise at a higher rate along with a greater increase in the proportion of severely obese individuals
- Whilst the mean obesity levels for both genders in the 16-20 and 21-30 age groups will remain relatively stable or decrease, the level of severely obese individuals will rise significantly, particularly amongst the males.

6.6.2 Impact on Healthcare Resources

Changes in demand are likely to come from four groups:

- Males and females above the age of 61 are all adversely affected, with obesity rates continuing to rise and in many cases a widening distribution suggesting more cases presenting with severe obesity.
- For the female population this is also apparent for the 46-60 age group.
- The situation is improving for younger female age groups with decreasing means and narrowing distributions, suggesting fewer cases presenting with severe obesity from these age groups (16-20 and 21-30), and as a consequence a reduced impact on resources.
- The situation for younger males (<21) is less positive, increasing means are offset by narrowing distributions, suggesting an increase in the overall level of obesity but fewer extreme cases.

6.6.3 Managerial and Theoretical Insights

This section seeks to identify the insights and learning from the previous sections that might be used to augment or inform current and future obesity interventions.

Historically whilst there has been a great deal of concern about the rise in adolescent obesity levels as described in Viner, Kinra, Nicholls, *et al.* [116], the NHS workload has primarily come from the older population [112]. The data from the model forecast suggests that even if the current situation doesn't worsen, the workload from the younger male population is likely to increase significantly. The apparent greater susceptibility of the male population (in particular the younger age groups) to external factors, may provide both an explanation of the effect and a lever to help address the issue.

The model proved surprisingly inelastic when presented with global interventions (Falling adolescent BMI, or reducing APCCC) at what were designed to be the maximum realistic values for the population as a whole, even when these rates were doubled the impact was not great. This may at first seem to contradict the forecasts associated with initiatives such as sugar tax [42] [41], where significantly greater benefits are expected. However, it's worth recalling that these anticipate much higher reductions in calorie intake and are focused on specific elements of the population. Given that the levels used in our model were based on historical levels and hence realistic, this suggests the need to avoid whole population initiatives, but instead to focus on specific issues and segments of the population.

One of the more striking effects was the impact of the network on the youngest age groups, who make up 24% of the network population. When implementing weight loss strategies targeted at that age group consideration could be given to manipulating that effect to amplify the impact of the strategies . For this age group this would entail attempting to strengthen the network effect. Valente [117] identifies four strategies in relation to network interventions in healthcare settings:

- Individual - identifying champions or opinion leaders
- Segmentation - identifying cliques or groups on which to focus the attention.
- Induction - deliberately creating interactions to spread information.
- Modification - adding new elements into the network or re-wiring them to deliver the objectives

The first three strategies increase the impact of a network on the individual. The last can be used either to reduce or increase the network impact and perhaps has more relevance to the older portion of the population. Given that an effect in the older populations was of a much smaller degree, initiatives focused on those age groups will have to consider whether the effort involved in manipulating the network is consistent with the expected return.

More fundamentally this research suggests a variation in the network's impact across the age groups, not recognised in previous research. The data based statistical research in this area including the original paper [11] used either the Framingham data [7] or the Longitudinal Study on Adolescent Health [8].

These had a number of limitations which would have made it difficult to identify the effect:

- The two data sets recorded information from two different age groups. [8] recorded the data of adolescents between the ages of 14 and 19, [7] recorded data for those over the age of 21. This research was able to encompass both.
- In both cases the data either pre-dates the existence of social media or fails to address it in the collection methodology.

Unsurprisingly, their results indicate a single coherent population wide model of network interaction. This research suggests that the nature of the network effect, the scale of impact, and the structure of the social network may differ significantly for different elements of that network. The key determinant seems to be age. A hypothesis consistent with the age determinant (and the type of data used in the initial studies) would be that the advent of social media has changed, and is continuing to change, the nature and impact of the 'obesity' network, and that this effect varies amongst the different sub-groups that make up the population.

The model also provided insights into the The impact of the external environment on individuals as represented by the parameter values described in 5.6 (BMIFactor). This encompasses a range of issues including; force of dietary habit, media pressures, availability of food, income and so on. The values suggest that the population as a whole is less susceptible to the external factors as it ages, but that males across the board remain more susceptible than their female counterparts.

Four parameters define the flows that in turn drive the decision making component of the model which is based on the TPB [35]. How they combine within the model to drive the Intention stock (which in turn drives the behaviour stock) is clearly of some interest:

- The 'attitudes' flow is dependent on educational level and is constant.
- The 'norms' flow is more complex, it is initially triggered by a universal threshold parameter (satisficing number) describing the difference between the individuals BMI and the average BMI of it's same-gender network neighbours. Once triggered that difference is then further modified by age and gender specific parameters t produce the flow.

- The 'PBC' flow has a default value which is then modified by a feedback loop acting as a multiplier for weight gain or loss, to produce the flow.

For context, it's worth recalling that the 'behaviour' stock value in the model that triggers dieting behaviour is 50, and that it is subject to a number of additional 'draining' flows. A 'behaviour' stock increase rate of 5, requires a combined flow rate into the 'intention' stock of 60 (discounting the effect the 1 month lag)

The parameters for norms are given in Table. 5.5. One hypothesis would be that age plays some role in the narrative, the generations under consideration in the model, represent a wide range of experiences in terms of peer interaction, those in their 60's at the start of the simulation will have grown up in an era of almost exclusively face-to-face communication, with relatively little exposure to mass media, this will have increased constantly until those in their 40's who will have experienced a much greater exposure to such media. For those under 30 the advent of social media will also play a significant role. The variation in the male parameter does seem to support this hypothesis, with a narrative of initial susceptibility to peer norms, decreasing in early adulthood, then increasing to peak in middle age, before declining again with age. The female narrative however is more complex, with generally higher values than the male ones (suggesting more susceptibility) , but with two distinct major drops in early middle age and again in older age, this may represent the influence of some other factor not yet represented in the model. Differentiating between the genders is made more complex in that the difference between the mean values for each are very similar (less than 2%).

Putting these factors together in a couple of examples, we can gain some insight into how they might interact.

If we look at a 25 year old male who left full time education after completing A-levels whose BMI is 29 (5.5 higher than the average of his male network neighbours), is currently dieting and losing 2Kg per month. The flows would be:

- Attitudes: 0
- Norms: 7.87 (Note: 5.5 exceeds the satisficing number in these conditions.)
- PBC: 10.93
- Total: 18.8

It's worth noting that if the individual were 10 years older the norms flow would be three times higher, giving a total flow of 34.54. Alternatively if they were educated to degree level then the flow would be 28.89.

Conversely for a 50 year old female with a degree whose BMI is 26.8 (1 lower than her same gender peers), and is gaining weight at a rate of 1Kg per year. The flows would be:

- Attitudes: 10.68
- Norms: 0 (Note: the BMNI difference is below the satisficing number.)
- PBC: 9.62
- Total: 20.3

Here the attitudes flow comprises just over half of the total flow.

Given the bi-modal nature of the attitudes flow and the apparently low variation in the impact of pbc, it would seem that in our model, peer influence when triggered can be a significant influence.

6.7 Summary

The HSM approach has generated a wealth of information to consider and from which to make inferences, much of it potentially very useful in addressing the issues driving this research. However before acting on these it is worth recalling the caveat issued at the beginning of the chapter and considering them as **hypothesised** conclusions, or insights that need to be evidenced by further research.

Chapter 7

Discussion and Conclusions

"At the end of reasons comes persuasion."

Ludwig Wittgenstein

7.1 Introduction

One of the key aims of this research was to understand the future impact of the current obesity situation on different sub groups of the population, and the associated resource implications. To achieve this we used an HSM approach to develop a model primarily based on theory and 'trained' it to mimic reality. We have then used our model to forecast future trends and values over a 10 year period. As previously discussed, we have sought to address any shortcomings in the model by comparing the forecasts for a set time period with real data from the same time period to identify over and under forecasting, this information was then used to modify the raw model forecasts for the future.

7.2 Reflections on the Research

One of the key aims of the research was to understand the interaction between social networks and the spread of obesity. To address this we used a flexible, dynamic but topologically stable network model to realise a 'representative' social network. 'Representative', because it needed to reflect the issue of multiplexity and the nature of individuals involvement in the different social networks, that in combination act to enable or hinder the spread of obesity. In practical terms this approach has significant advantages, it simplifies analysis, experimentation and intervention strategies. It is fortunate that the data that drove the insights in this area come primarily from comparisons between different model scenarios, and hence were less susceptible to the accuracy issues discussed above.

To achieve our aims, the research also had to address the issues involved in operationalising a health behaviour model (TPB) as part of a hybrid simulation. This proved to be one of the more challenging elements of the model development phase. The difficulty lay in the lack of numerical data with which to realise this element of the model. In the end the most effective approach was the progressive strategy used to optimise the parameters (described in Chapter 5).

Perhaps the most significant finding was that the way individuals interact with the network and it's level of impact on their behaviour is not fixed. There is certainly an inter-generational effect and it may well be that the behaviours within the age cohort are evolving too. In retrospect this ought not to be surprising, but it was not suggested in any of the literature reviewed for this research.

7.3 Reflections on the Implementation

The number of agents used in the model was a reflection of three factors; the implementation language (Java), the available computational power, the size of the data sets available. Initially it was envisaged that the simulation would be implemented in AnyLogic and augmented by specific classes/functions (written in Java) as necessary. However the process of embedding the classes into the AnyLogic proved complex and the decision was taken to programme the whole simulation in Java. This allowed a more sophisticated implementation but had the unintended consequence of making it difficult to use shared servers. Separately the number of people addressed in the HSE data varied considerably from year to year, with one year having as many as 15,000 and another less than 5,000. In the end the network size was set at 1,000 on the basis that with an augmented desktop computer the run time for each iteration was less than one minute. This enabled the gradient descent methods (typically 3,000 iterations) used in model training and development feasible.

In retrospect once the decision was made to step way from AnyLogic it might have been more appropriate to start afresh in Python, this would have offered the option to access shared servers. Python also offers the possibility of building larger ABM models by moving away from the OOD approach, whereby agents are usually instances of a class with various methods associated with them. Using the appropriate Python libraries an alternative approach uses multiple matrices to store agent variables, each time step is then facilitated by a number of matrix operations leading to significant computational savings especially in larger models.

One of the significant features of the implementation was the process used to refine the model parameters. A number of approaches were tried ranging from a greedy search to simulated annealing, in the end SPSA provided a simple and effective approach. A literature search (for search string see Appendix A.) identified seven articles and thre conference papers describing case studies of the use of SPSA to calibrate simulations, of these eight addressed its use in the context of ABM, and two in the context of DES, no instances of its use in a hybrid simulation were found. It proved to be an effective tool for the purpose.

With some caveats our model was able to give accurate 10 year forecasts. The limit on forecast range was a function of both the model itself and the data available with which to parameterise it. Whilst 20 years of data (1993-2013) were available, it could not be considered to represent a consistent or continuous situation. The period encompasses a number of step changes both in communication technology and social interaction, mostly occurring in the first decade of this century. This presented us with a dichotomy in that we could either take 10 years of relatively stable data (coincidentally the period in which much of the data for the original research was collected). Alternatively we could use the later data which whilst encompassing a number of 'faultlines' allowed the development of a model with more current relevance.

One of the key issues our research highlighted was the variation in how different age groups operate within the social network, the hypothesis put forward to address this suggests an inter-generational difference in the way that individuals interact with their networks. As this started to become apparent the case for switching to a cohort based model became stronger. However, to address this would require a longer model run-time than that over which our model was able to operate credibly. If we assume a cohort comprising all the individuals whose birth dates fall within a five year time period, then it's hard to envisage developing useful insights from a model with a run time of less than 20-25 years

7.4 Future Research

The literature review highlighted 2 areas where further research would add value:

- Multiplexity is a clearly acknowledged feature of social interaction, individuals belong to a number of networks, with potentially differing roles and expectations in each, and with varying methods of interaction. However, the concept remains unacknowledged in any of the articles reviewed and is clearly relevant to many of them. Our approach was to accept that our network was a representative one in that it mimicked the combined effect those networks. Further research into this concept and alternative approaches would certainly add value.
- Network realisations using emergent behaviour clearly have the potential to deliver much more sophisticated social network simulations than graph models, but as yet are relatively rare.Further research in this area would clearly asdd value.

The results and parameter analysis also highlighted a number of areas where further research may be of value:

- We have suggested that network behaviour and impact varies significantly across the age domain, and that it is significant for the youngest age groups. This suggests two areas for further research;
  - Given the current priority placed on mitigating adolescent obesity, focusing on examining the network effect as it applies to adolescents and young adults is an area where further research has the potential to add significant value. Our findings suggest that this may have the potential for significant impact and that the behaviour and level of impact may well have changed since the initial research was done.
  - A more general exploration into the way different age groups interact with their social networks and how or if it is evolving.
- Our findings suggest that the impact of external factors varies across different age groups and genders. This has implications both for health behavioural models and more generally. Research to confirm the finding and quantify the level of impact, would add value in the design of future health care interventions.

- TPB is an attractive model when simulating health related behaviour, our research has added to the literature describing it's use and suggested some parameters that might be of value in future such projects, but it remains an area where further research would be valuable

7.5 Contributions

We believe our research has made contributions in three areas; providing additional insights supporting health care professionals in strategy planning and decision making, extending the body of knowledge with regard to hybrid simulation in healthcare settings and replicating and extending existing work by other researchers.

Insights to support healthcare professionals:

- The insights described in the preceding sections will be of value in the formulation of healthcare strategy and allocation of resources, in terms of identifying sub-groups of the population that are particularly at risk, and in quantifying the level of impact and the broad time frames.
- None of the research documented in the literature review or in our wider research has identified the transition in social network impact from adolescent to young adult. We have generated a number of insights into this process and identified the need for further research.
- Our model suggests that in general the male population is more susceptible to external influences than the female one.
- Insights derived from the model can help further understanding of the functioning of TPB and it's application to the spread of obesity and potentially other NCD's.

Extending the body of knowledge with regard to hybrid systems modelling and hybrid simulation in healthcare settings

- Insights derived from the model development process have helped clarify the likely topology of 'obesity' networks.
- Incorporating health behaviour models into hybrid simulations, our literature review only identified two other examples of this approach.

- Parameterisation of healthcare hybrid simulations using SPSA, we were only able to identify 10 examples of the use of SPSA in optimising simulations (ABM or DES), and we found none in the healthcare field or addressing the complexities of a hybrid simulation. More generally SPSA is an example of a 'supervised learning' algorithm frequently used in machine learning applications, none of the 36 hybrid health simulations identified in our literature review used this type of methodology.

Replicating existing work and extensions:

- The network algorithm described in [50] was designed to study social network formation and stability. We adapted and extended the algorithm and used it model the social network in our simulation. The process involved replicating their original work and confirming the results, then extending it to account for agent heterogeneity, homophily and chronology (the original model did not relate model time-steps to a specific time-frame).
- The ideas described by [29] proved extremely useful in understanding the complex causality underlying rising obesity levels. in re-implementing them we were able to extend the functionality to classify loops as balancing and reinforcing loops.

Appendix A

Literature Review Documents

A.1 Literature Review - Documents

(1) M. Abdelghany and A. B. Eltawil, "Linking approaches for multi-methods simulation in healthcare systems planning and management", *International Journal of Industrial and Systems Engineering*, vol. 26, no. 2, pp. 275-290, 2017.

(2) M. Abdelghany and A. B. Eltawil, "Individual versus integrated simulation techniques in healthcare applications", *IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 1214, 1218, 2014.

(3) M. Abdelghany, A. B. Eltawil, and S. F. Abdou, "A discrete-event and agent-based hybrid simulation approach for healthcare systems modeling and analysis", *Proceedings of the International Conference on Industrial Engineering and Operations Management*, pp. 1921-1928, 2016.

(4) A. Adiga, S. Chu, S. Eubank, *et al.*, "Disparities in spread and control of influenza in slums of Delhi: Findings from an agent-based modelling study", *BMJ*, vol. 8, 2018.

(5) N. Ahmad, N. A. Ghani, A. A. Kamil, *et al.*, "Modeling emergency department using a hybrid simulation approach", *Lecture Notes in Electrical Engineering*, pp. 360-371, 2013.

(6) N. Ahmad, N. A. Ghani, A. A. Kamil, and R. Tahar, "Modelling the complexity of emergency department operations using hybrid simulation", *International Journal of Simulation and Process Modelling*, pp. 360-371, 2015.

(7) R. Aringhieri, "An integrated DE and AB simulation model for EMS management", *2010 IEEE Workshop on Health Care Management*, pp. 1-6, 2010.

(8) A. Aswani, P. Kaminsky,Y. Mintz, and R. Tahar, "Behavioral modeling in weight loss interventions", *European Journal of Operational Research*, vol. 272, pp. 1058-1072, 2019.

(9) S. Barnes, B. Golden, and E. Wasil, "Exploring the effects of network structure and healthcare worker behavior on the transmission of hospital-acquired infections", *IIE Transactions on Healthcare Systems Engineering*, vol. 2, no. 4, pp. 259-273, 2012.

(10) R. Beheshti, M. Jalalpour, and T.A. Glass, "Comparing methods of targeting obesity interventions in populations: An agent-based simulation", *SSM - Population Health*, vol. 3, pp. 211-218, 2017.

(11) A. Borshchev, "Multi-method modeling", *1st International Simulation Tools Conference and Expo 2013*, pp. 59-67, 2013.

(12) S. C. Brailsford, "Hybrid simulation in healthcare: New concepts and new tools", *Proceedings - Winter Simulation Conference*, pp. 1645-1653, 2016.

(13) S. C. Brailsford, S. M. Desai, and J. Viana, "Towards the holy grail: Combining system dynamics and discrete-event simulation in healthcare", *Proceedings - Winter Simulation Conference*, pp. 2293-2303, 2010.

(14) S. C. Brailsford, , T. Eldabi, M. Kunc, a*et al.,* "Hybrid simulation modelling in operational research: A state-of-the-art review", *European Journal of Operational Research*, vol 26, no 1, pp. 721-727, 2019.

(15) L. Caudill and B. Lawson, "A hybrid agent-based and differential equations model for simulating antibiotic resistance in a hospital ward", *Proceedings of the 2013 Winter Simulation Conference - Simulation: Making Decisions in a Complex World*, pp. 1419-1430, 2013.

(16) K. Chahal, T. Eldabi, and T. Young, "A conceptual framework for hybrid system dynamics and discrete event simulation for healthcare", *Journal of Enterprise Information Management*, vol. 26, no. 1, pp. 50-74, 2013.

(17) B. Chen, Y. Ge, L. Zhang, *et al.,* "A modeling and experiment framework for the emergency management in AHC transmission", *Computational and Mathematical Methods in Medicine*, pp. 1-18, 2014.

(18) H. Y. Chung, C. Y. Chung, and S.C. Ou, "Analysis of a bio-dynamic model via Lyapunov principle and small-world network for tuberculosis", *IET Systems Biology*, vol. 6, no. 5, pp. 196-206, 2012.

(19) P. Cooley, B. Y. Lee, S. Brown, *et al.*, "Protecting health care workers: A pandemic simulation based on Allegheny County", *Influenza and other Respiratory Viruses*, vol. 4, no. 2, pp. 61-72, 2010.

(20) C. D. Corley, A. R. Mikler, D. J. Cook, and K. P. Singh, "Dynamic intimate contact social networks and epidemic interventions", *International Journal of Functional Informatics and Personalised Medicine*, vol. 1, no. 2, pp. 171-188, 2008.

(21) J. Demongeot, M. Jelassi, and C. Taramasco, "From susceptibility to frailty in social networks: The case of obesity", *Mathematical Population Studies*, vol. 24, no. 4, pp. 219-245, 2017.

(22) J. Demongeot and C. Taramasco, "Evolution of social networks: the example of obesity", *Biogerontology*, vol. 15, no. 6, pp. 611-626, 2014.

(23) A. Djanatliev and R. German, "Towards a guide to domain-specific hybrid simulation", *Proceedings - Winter Simulation Conference*, pp. 1609-1620, 2016.

(24) A. Djanatliev and R. German, "Prospective healthcare decision-making by combined system dynamics, discrete-event and agent-based simulation", *Proceedings of the 2013 Winter Simulation Conference - Simulation: Making Decisions in a Complex World*, pp. 270-281, 2013.

(25) A. Djanatliev, P. Kolominsky-Rabas, B. M. Hofmann, *et al.*, "System dynamics and agent-based simulation for prospective health technology assessments", *Advances in Intelligent Systems and Computing*, vol. 256, pp. 85-96, 2014.

(26) A. Djanatliev, P. Kolominsky-Rabas, B. M. Hofmann, *et al.*, "Hybrid simulation approach for prospective assessment of Mobile Stroke Units", *SIMULTECH 2012 - Proceedings of the 2nd International Conference on Simulation and Modeling Methodologies, Technologies and Applications*, pp. 357-366, 2012.

(27) A. Djanatliev and F. Meier, "Hospital processes within an integrated system view: A hybrid simulation approach", *Proceedings - Winter Simulation Conference*, pp. 1364-1375, 2017.

(28) W. Duan, X. Qiu, Z. Cao, *et al.*,. "Heterogeneous and stochastic agent-based models for analyzing infectious diseases' super spreaders", *IEEE Intelligent Systems*, vol. 28, no. 4, 2013.

(29) W. Duan, R. Quax, M. Lees, *at al.,* "Topology dependent epidemic spreading velocity in weighted networks", *Journal of Statistical Mechanics: Theory and Experiment*, pp. 18-25, 2014.

(30) H. P.Duerr, M. Schwehm, C. C. Leary, *et al.,* "The impact of contact structure on infectious disease control: Influenza and antiviral agents", *Epidemiology and Infection*, vol 135, no. 7, pp.1124-1132, 2007.

(31) D. P. Durham and E. A. Casman, "Incorporating individual health-protective decisions into disease transmission models: A mathematical framework", *Journal of the Royal Society Interface*, vol. 9, no. 68, pp. 562-570, 2012.

(32) D. P. Durham, E. A. Casman, and S. M. Albert, "Deriving Behavior Model Parameters from Survey Data: Self-Protective Behavior Adoption During the 2009-2010 Influenza A(H1N1) Pandemic", *Risk Analysis*, vol. 32, no. 12, pp. 2020-2031, 2012.

(33) K. Eames, J. M. Read, and W. J. Edmunds, "Epidemic prediction and control in weighted networks", *Epidemics*, pp. 70-76, 2009.

(34) M. P. Eccles, J. M. Grimshaw, G. MacLennan, *et al.,* "Explaining clinical behaviors using multiple theoretical models", *Implementation Science*, vol. 7, no. 1, 2012.

(35) M. A. Elliott and K. Ainsworth, "Predicting university undergraduates' binge-drinking behavior: A comparative test of the one- and two-component theories of planned behavior", *Addictive Behaviors*, vol. 37, no. 1, pp. 92-101, 2012.

(36) A. M. El-Sayed, L. Seemann, P. Scarborough, and S. Galea, "Are network-based interventions a useful antiobesity strategy? An application of simulation models for causal inference in epidemiology", *American Journal of Epidemiology*, vol 178, no. 2, pp. 287-295, 2013.

(37) S. Eubank, H. Guclu, , V. Kumar, *et al.,* "Modelling disease outbreaks in realistic urban social networks", *Nature*, vol. 429, pp. 180-184, 2004.

(38) M. Fakhimi, A. Anagnostou, L. Stergioulas, and S. Taylor, "A hybrid agent-based and Discrete Event Simulation approach for sustainable strategic planning and simulation analytics", *Proceedings - Winter Simulation Conference*, pp. 1573-1584, 2015.

(39) M. Fakhimi, L. K. Stergioulas and N. Mustafee, "An investigation of Hybrid Simulation for modeling sustainability in healthcare", *Proceedings - Winter Simulation Conference*, pp. 185-1596, 2016.

(40) A. Fetta, P. Harper, V. Knight, and J. Williams, "Predicting adolescent social networks to stop smoking in secondary schools", *European Journal of Operational Research*, vol. 265, no. 1, pp. 263-276, 2018.

(41) A. Gao, N. D. Osgood, W. An, and R. F. Dyck, "A tripartite hybrid model architecture for investigating health and cost impacts and intervention tradeoffs for diabetic end-stage renal disease", *Proceedings - Winter Simulation Conference*, pp. 1676-1687, 2015.

(42) R. A. Hammond and J. T. Ornstein, "A model of social influence on body mass index", *Annals of the New York Academy of Sciences*, 2014.

(43) Y. He, M. Li, S. Sala-Diakanda, *et al.,* "A hybrid modeling and simulation methodology for formulating overbooking policies", *IIE Annual Conference and Expo*, pp. 34-42, 2013.

(44) E. Hennessy, J. T. Ornstein, C. D. Economos, *et al.,* "Designing an agent-based model for childhood obesity interventions: A case study of child obesity", *Preventing Chronic Disease*, vol. 13, no. 1, 2016.

(45) A. L. Hill, D. G. Rand, M. A. Nowak, N. A. Christakis, "Infectious disease modeling of social contagion in networks". *PLoS Computational Biology*, vol. 6, no. 11, 2010.

(46) B. Iotti, A. Antonioni, S. Bullock, *at al.,* "Infection dynamics on spatial small-world network models", *Physical Review E*, vol. 96, no. 5, 2017.

(47) B. Kaya and M. Poyraz, "Age-series based link prediction in evolving disease networks.", *Computers in Biology and Medicine*, vol. 63, pp. 1-10, 2015.

(48) K. Garcia and T. Mann, "From 'I Wish' to 'I Will': Social-cognitive Predictors of Behavioral Intentions", *Journal of Health Psychology*, vol. 8, no. 3, 2003.

(49) C. Kittipittayakorn and K. Ying, "Using the integration of discrete event and agent-based simulation to enhance outpatient service quality in an orthopedic department", *Journal of Healthcare Engineering*, pp. 1-8, 2016.

(50) V. A. Knight, J. E. Williams, and I. Reynolds, "Modelling patient choice in healthcare systems: Development and application of a discrete event

simulation with agent-based decision making.", *Journal of Simulation*, vol. 6, no. 2, pp. 92-102, 2012.

(51) C. J. Kuhlman, Y. Ren, B. Lewis, and J. Schlitt, "Hybrid Agent-based modeling of Zika in the united states", *Proceedings - Winter Simulation Conference*, pp. 1085-1096, 2018.

(52) M. Kunc, P. Harper, and K. Katsikopoulos, "A review of implementation of behavioural aspects in the application of OR in healthcare", *Journal of the Operational Research Society*, 2018. Available: https://doi.org/10.1080/01605682.2018.148935 (visited on 03/01/2019)

(53) P. Landa, M. Sonnessa, M. Resta, *et al.*, "A hybrid simulation approach to analyse patient boarding in emergency departments", *Springer Proceedings in Mathematics and Statistics*, vol. 210, pp. 133-144, 2017.

(54) B.A. Langellier, "An agent-based simulation of persistent inequalities in health behavior: Understanding the interdependent roles of segregation, clustering, and social influence", I*SSM - Population Health*, vol. 2, pp. 757-769, 2016.

(55) A. Lassnig, C. Baumgartner, and J. Schröttner, "A modeling approach to heart failure treatment", *31st Annual European Simulation and Modelling Conference*, pp. 404-406, 2017.

(56) B. Y. Lee, V. L. Bedford, M. S. Roberts, *et al.*, "Virtual epidemic in a virtual city: simulating the spread of influenza in a US metropolitan area", *Translational Research*, vol. 151, no. 6, pp. 275-287, 2008.

(57) S. Liu, K. P. Triantis, L. Zhao, and Y. Wang, "Capturing multi-stage fuzzy uncertainties in hybrid system dynamics and agent-based models for enhancing policy implementation in health systems research", *PLoS ONE*, vol. 13, no. 4, 2018.

(58) W. Luo, "Visual analytics of geo-social interaction patterns for epidemic control", *International Journal of Health Geographics*, vol. 15, no. 1, 2016.

(59) W. Luo, P. Gao, and S. Cassels, "A large-scale location-based social network to understanding the impact of human geo-social interaction patterns on vaccination strategies in an urbanized area", *Computers, Environment and Urban Systems* vol. 72, pp. 78-87, 2018.

(60) M. Lawental, A. Kipnis, and K. Rigg, "Binge drinking among young adults in Israel: application of the theory of planned behaviour", *Psychology, Health & Medecine*, in press, 2018.

(61) L. Mao, "Modeling triple-diffusions of infectious diseases, information, and preventive behaviors through a metropolitan social network-An agent-based simulation", *Applied Geography*, vol. 50, pp. 31-39, 2014.

(62) L. Mao and Y. Yang, "Coupling infectious diseases, human preventive behavior, and networks - A conceptual framework for epidemic modeling", *Social Science and Medicine*, vol. 74, no. 2, pp. 167-175, 2012.

(63) B. Mielczarek and J. Zabawa, "Simulation model for studying impact of demographic, temporal, and geographic factors on hospital demand", *Proceedings - Winter Simulation Conference*, pp. 4498-4500, 2018.

(64) S. J. Mooney and A. M. El-Sayed, "Stigma and the etiology of depression among the obese: An agent-based exploration", *Social Science and Medicine*, vol. 148, pp. 1-7, 2016.

(65) J. E. Navarro-Barrientos, D. E. Rivera, and L. M. Collins, "A dynamical model for describing behavioural interventions for weight loss and body composition change", *Mathematical and Computer Modelling of Dynamical Systems*, vol. 17, no. 2, pp.182-203, 2011.

(66) A. Nikakhtar, , A. Abbasian-Hosseini, H. Gazula, and S. M. Hsiang, "Social Network based sensitivity analysis for patient flow using computer simulation", *Computers and Industrial Engineering*, vol. 88, pp. 264-272, 2015.

(67) M. G. Orr, S. Galea, M. Riddle, and G. A. Kaplan, "Reducing racial disparities in obesity: Simulating the effects of improved education and social network influence on diet behavior", *Annals of Epidemiology*, vol. 24, no. 8, pp. 563-569, 2014.

(68) X. Peng, Z. Zhang, J. Yang, and Z. Jin, "An SIS epidemic model with vaccination in a dynamical contact network of mobile individuals with heterogeneous spatial constraints", *Communications in Nonlinear Science and Numerical Simulation*, vol. 73, pp. 52-73, 2019.

(69) A. Perisic and C. T. Bauch, "A simulation analysis to characterize the dynamics of vaccinating behaviour on contact networks", *BMC Infectious Diseases*, vol. 9, 2009.

(70) A. Perisic and C. T. Bauch, "Social contact networks and disease eradicability under voluntary vaccination", *PLoS Computational Biology*, vol. 5, no. 2, 2009.

(71) P. M. Polgreen, T. L. Tassier, S. V. Pemmaraju, A. M. Segre, "Prioritizing healthcare worker vaccinations on the basis of social network analysis", *Infection Control and Hospital Epidemiology*, vol. 31, no. 9, pp. 893-900, 2010.

(72) J. H. Powell and N. Mustafee, "Widening requirements capture with soft methods: An investigation of hybrid M&S studies in health care", *Journal of the Operational Research Society*, vol. 68, no. 10, pp. 1211-1222, 2017.

(73) W. T. Riley, C. A. Martin, and D. E. Rivera, *et al.,* "Development of a dynamic computational model of social cognitive theory", *Translational Behavioral Medicine*, vol. 6, no. 4, pp. 483-495, 2016.

(74) T. R. Rohleder, D. P. Bischak, and L. B. Baskin, "Modeling patient service centers with simulation and system dynamics", *Health Care Management Science*, vol. 10, no. 1, pp. 1-12, 2007.

(75) M. Salathé and J. H. Jones, "Dynamics and control of diseases in networks with community structure", *PLoS Computational Biology*, vol. 6, no. 4, 2010.

(76) M. G. Sangachin, M. Samadi, and L. A. Cavuoto, "Modeling the spread of an obesity intervention through a social network", *Journal of Healthcare Engineering*, vol. 5, no. 3, pp. 293-312, 2014.

(77) J. S. Savage, D. S. Downs, Y. Dong, and D. E. Rivera, "Control systems engineering for optimizing a prenatal weight gain intervention to regulate infant birth weight", *American Journal of Public Health*, vol. 104, no. 7, pp. 1247-1254, 2014.

(78) P. Schumm, W. Schumm, and C. Scoglio, "Impact of Preventive Responses to Epidemics in Rural Regions", *PLoS ONE*, vol. 8, no. 3, 2013.

(79) H. Shu-Chen and R. Lanese, "The applicability of the theory of planned behaviour to the intention to quit smoking across workplaces in southern Taiwan", *Addictive Behaviors*, vol. 23, no. 2, 1998.

(80) E. Sulis and A. Di Leva, "An agent-based model of a business process: The use case of a hospital emergency department", *Lecture Notes in Business Information Processing 2016-February*, vol. 308, pp. 124-132, 2018.

(81) R. Sun, "Optimal weight based on energy imbalance and utility maximization", *Physica A: Statistical Mechanics and its Applications*, vol. 442, pp. 429-435, 2015.

(82) J. J. Tejada, J. S. Ivy, R. E. King, *at al.*, "Combined DES/SD model of breast cancer screening for older women, II: Screening-and-treatment simulation", *IIE Transactions (Institute of Industrial Engineers)* vol. 46, no. 7, pp. 707-727, 2014.

(83) Z. Toroczkai and H. Guclu, "Proximity networks and epidemics.", *Physica A: Statistical Mechanics and its Applications*, vol. 378, no. 1, pp. 68-75, 2007.

(84) J. G. Trogdon and B. T. Allaire, "The effect of friend selection on social influences in obesity.", *Economics and Human Biology*, vol. 15, pp. 153-164, 2014.

(85) T. Ueno and N. Masuda, "Controlling nosocomial infection based on structure of hospital social networks", *Journal of Theoretical Biology*, vol. 254, no. 3, pp. 655-666, 2008.

(86) F. Verelst, L. Willem, and P. Beutels, "Behavioural change models for infectious disease transmission: A systematic review (2010-2015)", *Journal of the Royal Society Interface*, vol. 13, no. 125, 2016.

(87) J. Viana, "Reflections on two approaches to hybrid simulation in healthcare", *Proceedings - Winter Simulation Conference*, pp. 1585-1596, 2015.

(88) J. Viana, S. C. Brailsford, V. Harindra, and P. R. Harper, "Combining discrete-event simulation and system dynamics in a healthcare setting: A composite model for Chlamydia infection", *European Journal of Operational Research*, vol. 237, no. 1, pp. 196-206, 2014.

(89) J. Viana, V. M. Ziener, M. S. Holhjem, *et al.*, "Optimizing home hospital health service delivery in Norway using a combined geographical information system, agent based, discrete event simulation model.", *Proceedings - Winter Simulation Conference*, pp. 1658-1669, 2018.

(90) W. Wakeland, A. Nielsen, and P. Geissert, "Dynamic model of non-medical opioid use trajectories and potential policy interventions", *American Journal of Drug and Alcohol Abuse*, vol. 41, no. 6, pp. 508-518, 2015.

(91) B. Wang, Y. Han, and G. Tanaka, "Interplay between epidemic spread and information propagation on metapopulation networks", *Journal of Theoretical Biology*, vol. 420, pp. 18-25, 2017.

(92) J. Zabawa and B. Mielczarek, "Overcoming challenges in hybrid simulation design and experiment", *Advances in Intelligent Systems and Computing*, vol. 853, pp. 207-217, 2019.

(93) J. Zulkepli and T. Eldabi, "Towards a framework for conceptual model hybridization in healthcare", *Proceedings - Winter Simulation Conference*, pp. 1597-1608, 2016.

A.2 Search Terms and Journal Lists

A.2.1 Literature Review Search Terms

Listed below are the terms used in each of the searches used to develop the initial list of documents from which the Literature Review documents listed above were derived.

> (TITLE-ABS-KEY ("hybrid simulation") AND TITLE-ABS-KEY ("system dynamic*" OR "discrete event" OR "agent")
>
> (KEY(simulation) AND KEY ("social network") AND KEY (health OR healthcare))
>
> (KEY (simulation) AND KEY ("social network") AND KEY (health OR healthcare)) AND (LIMIT-TO (SUBJAREA , "MEDI") OR LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "MATH") OR LIMIT-TO (SUBJAREA , "SOCI") OR LIMIT-TO (SUBJAREA , "HEAL") OR LIMIT-TO (SUBJAREA , "DECI") OR LIMIT-TO (SUBJAREA , "MULT") OR LIMIT-TO (SUBJAREA , "PSYC")) AND (LIMIT-TO (SRCTYPE , "j"))
>
> (TITLE-ABS-KEY ("health" OR "health-care") AND TITLE-ABS-KEY ("discrete event" OR "system dynamics" OR "agent" OR "simulation" OR "hybrid simulation" OR "network") AND TITLE-ABS-KEY ("smoking*" OR "alcohol*" OR "obesity" OR "exercise" OR "physical activity" OR "system*" OR "tobacco" OR "drug*" OR "addiction" OR "sex*" OR "chronic")
>
> (TITLE-ABS-KEY ("health" OR "health-care") AND TITLE-ABS-KEY ("health behaviour" OR "health behavior" OR "behavioural" OR "hybrid simulation") AND TITLE-ABS-KEY ("smoking*" OR "alcohol*" OR "obesity*" OR "exercise" OR "physical activity" OR "system*" OR "tobacco" OR "drug*" OR "addiction" OR "sex*" OR "chronic")

A.2.2 Journal List

The journal list below was used in conjunction with the search terms above, as the initial set of journals from which to develop the Literature Review documents.

ISSN Number	Journal
0001-5172	Acta Anaesthesiologica Scandinavica
0003-2409	Anaesthesia.
0003-2999	Anesthesia and Analgesia
0003-6870	Applied Ergonomics
0003-2409	Anaesthesia
0003-3022	Anesthesiology
0007-0912	British Journal of Anesthesia
0007-8506	CIRP Annals
0010-3853	Community Mental Health Journal
0013-791x	The Engineering Ergonomist
0014-0139	Ergonomics
0017-9124	Health Services Research
0018-9391	IEEE Transaction on Engineering Managemnt
0020-7314	International Journal of Health Services
0020-7543	International Journal of Production Research
0020-7721	International Journal of System Science
0020-7543	International Journal of Production Research
0022-3239	Journal of Optimization Theory and Applications
0022-4065	Journal of Quality Technology
0025-1909	Management Science
0025-5610	Mathematical Programming
0025-7079	Medical Care
0026-1270	Methods of Information in Medecine
0028-3045	Networks
0030-364x	Operations Research
0041-1655	Transportation Science
0046-9580	The Journal of Healthcare Organization, Provision and Financing
0092-2102	INFORMS Journal on Applied Analytics
0094-5145	Journal of Community Health
0140-0118	Medical and Biological Engineering and Computing
0143-2087	Optimal Control Applications and Methods
0143-991x	Industrial Robot
0148-5598	Journal of Medical Systems
0160-5682	Journal of Operational Research Society
0163-2787	Evaluation and the Health Professions
0166-4972	Technovation

ISSN Number	Journal
0167-6296	Journal of Health Economics
0167-6377	Operations Research Letters
0167-9236	Decision Support Systems
0167-6911	Systems and Control Letters
0168-8510	Health Policy and Planning
0169-2607	Computer Methods and Programs in Biomedicine
0169-8141	International Journal of Industrial Ergonomics
0171-6468	OR Spectrum
0191-2615	Transportation Research Part B: Methodological
0195-8631	Health Care Financing Review
0197-5897	Journal of Public Health Policy
0217-5959	Asia-Pacific Journal of Operational Research
0219-6220	International Journal of Information Technology and Decision Making
0233-1934	Optimization
0254-5330	Annals of Operations Research
0257-0130	Queueing Systems
0263-5577	Industrial Management and Data Systems.
0265-0215	European Journal of Anaesthiology
0266-4623	International Journal of Technology Assessment in Health Care
0267-5730	International Journal of Technology Management
0268-1080	Health Policy and Planning
0269-9648	Probability in the Engineering and Informational Sciences
0272-6963	Journal of Operations Management
0272-989x	Medical Decision Making
0275-5823	Military Operations Research
0277-6715	Statistics in Medicine
0278-2715	Health Affairs
0278-6125	Journal of Manufacturing Systems
0278-6125	Journal of Manufacturing Systems
0304-3959	Pain
0305-0483	Omega

ISSN Number	Journal
0305-0548	Computers and Operations Research
0305-215x	Engineering Optimization
0305-0548	Computers and Operations Research
0310-057x	Anaesthesia and Intensive Care
0315-5986	Information Systems and Operational Research (Journal)
0360-8352	Computers and Industrial Engineering
0361-6878	Journal of Health Politics, Policy and Law
0361-6274	Health Care Management Review
0364-765x	Mathematics of Operations Research
0377-2217	European Jopurnal of Operational Research
0399-0559	RAIRO Operations Research
0453-4514	Journal of the Operations Research Society of Japan
0733-9364	Journal of Construction Engineering and Management
0737-6782	Journal of Product Innovation Management
0739-5175	IEEE Engineering in Medecine and Biology
0740-817x	IIE Transactions
0740-817x	IIE Transactions
0742-597x	Journal of Management in Engineering
0746-9179	Research and Devlopment
0748-8017	Quality andf Reliability Engineering International
0748-5492	Issues in Science and Technology
0748-8017	Quality and Reliability Engineering International
0749-8047	The Clinical Journal of Pain
0749-6753	The International Journal of Health Planning and Management
0825-8597	Journal of Palliative Care
0832-610x	Canadian Journal of Anesthesia
0885-8195	Journal of Cancer Education
0887-378x	The Milbank Quarterly
0890-765x	The journal of Rural Health
0891-5245	Journal of Pediatric Health
0894-069x	Naval Research Logistics
0894-587x	Administration and Policy in Mental Health and Mental Health Services Research
0895-6308	Research Technology Management
0898-2112	Quality Engineering
0898-2643	Journal of Aging and Health
0898-4921	Journal of Neurosurgical Anesthiology

ISSN Number	Journal
0913-8668	Journal of Anesthesia
0923-4748	Journal of Engineering and Technology Management
0924-6703	Discrete Event Dynamic Systems
0924-0136	Journal of Materials Processing Technology
0925-5001	Journal of Global Optimization
0925-5273	International Journal of Production Economics
0925-7535	Safety Science
0925-5273	International Journal of Production Economics
0925-7535	Safety Science
0926-6003	Computational Optimization and Applications.
0933-3657	Artifical Intelligence in Medecine
0934-9839	Research in Engineering Design
0951-8320	Reliability Engineering and System Safety
0951-192x	International Journal of Computer Integrated Manufacturing
0952-7907	Current Opinion in Anaesthiology
0952-8180	Journal of Clinical Anaesthesia
0953-7287	Production Planning and Control
0954-0121	AIDS Care
0957-4174	Expert Systems with Applications
0957-4824	Health Promotion International
0959-289x	The International Journal of Obstetric Anesthesia
0962-2802	Statistical Methods in Medical Research
0963-1801	Cambridge Quarterly of Healthcare Ethics
0969-6016	International Transaction in Operational Research
1004-3756	Journal of Systems Science and Systems Engineering
1004-4132	Journal of Systems Engineering and Electronics
1012-277x	South African Journal of Industrial Engineering
1041-0236	Health Communication
1049-7323	Qualitative Health Research
1049-2089	Journal of Health Care for the Poor and Underserved
1052-6234	SIAM Journal on Optimization
1053-0770	Journal of Cardiothoracic and Vascular Anesthesia
1054-8289	The Future of children
1055-6788	Optimization Methods and Software
1057-9230	Health Economics
1059-1478	Production and Operations Management

ISSN Number	Journal
1063-293x	Concurrent Engineering
1065-3058	Health Care Analysis
1067-5027	Journal of the American Medical Informatics Association.
1075-2730	Psychiatric services
1076-8971	Psychology, Public Policy, and Law
1077-2618	IEEE Industry Applications Magazine
1077-5587	Medical care research and review
1088-0224	The American Journal of Managed Care
1089-7771	IEEE Transactions on Information Technology in Biomedicine
1090-3801	European Journal of Pain
1091-4358	Journal of Mental Health Policy and Economics
1091-9856	INFORMS Journal on Computing
1094-3412	The Journal of Behavioral Health Services and Research
1094-6136	Journal of Sceduling
1096-9012	Journal of Healthcare Management
1098-1241	Systems Engineering
1098-3015	Value in Health
1098-7339	Regional Anesthesia and Pain Medicine
1134-5764	TOP
1155-5645	Paediatric Anaesthesia
1178-1653	The Patient - Patient-Centered Outcomes Research
1220-1766	Studies in Informatics and Control
1348-9151	Pacific Journal of Optimization
1353-4505	International Journal for Quality in Healthcare
1355-8196	Journal of Health Services Research and Policy
1356-1294	Journal of Evaluation in Clinical Practice
1356-1820	Journal of Interprofessional Care
1366-5545	Transportation Research Part E: Logistics and Transportation Review
1369-6513	Health Expectations
1386-5056	International Journal of Medical Informatics
1386-9620	Health Care Management Science
1387-1307	Journal of Clinical Monitoring and Computing
1389-6563	International Journal of Health Care Finance and Economics
1389-4420	Optimization and Engineering
1432-2994	Mathematical Methods of Operations Research
1435-246x	Central European Journal of Operations Research

ISSN Number	Journal
1438-8871	Journal of Medical Internet Research
1446-1242	Health Sociology Review
1448-7527	Australian Journal of Primary Health
1460-4582	Health Informatics Journal
1472-6947	BMC Medical Informatics and Decision Making
1472-698x	BMC International Health and Human Rights
1477-7525	Health and Quality of Life Outcomes
1478-4491	Human Resources for Health
1523-4614	Manufacturing and Service Operations Management
1524-1904	Applied Stochastic Models in Business and Industry
1530-7085	Pain Practice
1532-0464	Journal of Biomedical Informatics
1538-2931	Computers, Informatics, Nursing
1542-894x	Industrial Engineer
1547-5816	Journal of Industrial and Management Optimization
1551-3203	IEEE Transactions on Industrial Informatics
1566-113x	Networks and Spatial Economics
1568-4539	Fuzzy Optimization and Decision Making
1572-5286	Discrete Optimization
1618-7598	The European Journal of Health Economics
1619-4500	4OR
1684-3703	Quality Technology and Quantitative Management
1696-2281	SORT (Statistics and Operations Research Transactions)
1741-1122	Journal of Policy and Practice in Intellectual Disabilities
1744-1331	Health Economics, Policy and law
1748-006x	Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability
1748-5908	Implementation Science
1751-5254	European Journal of Industrial Engineering.
1753-8157	Informatics for Health and Social Care
1819-5164	Asian Journal of WTO and International Health, Law and Policy
1833-3583	Health Information Management Journal
1862-4472	Optimization Letters
1932-8184	IEEE Systems Journal
1936-6574	Disability and Health Journal

ISSN Number	Journal
1936-6582	Flexible Services and Manufacturing Journal
1943-670x	International Journal of Industrial Engineering: Theory, Applications and Practice
1948-8319	IIE Transactions on Healthcare Systems Engineering
2044-5415	BMJ Quality and Safety
2047-6973	Health Systems

A.2.3 SPSA - Simulation Applications

Listed below are the search terms to identify applications of SPSA in simulations parameterisation.

(TITLE-ABS-KEY ("simultaneous perturbation stochastic optimisation" OR "simultaneous perturbation stochastic optimization" OR spsa) AND TITLE-ABS-KEY ("discrete event simulation" OR "DES" OR "system dynamic" OR "agent"))

Appendix B

Pseudocode for the Simulation

B.1 Running the Simulation

The network algorithms and the SD algorithms are linked together by a number of parameters and sets. The current agent population is critical to both, and the interaction of timings between the two components needs to be clearly defined. These are described Table B.1

$\mathcal{P}$	The set of all current agents.
$\mathcal{T}^m$	The set of all model times, one unit of model time in the simulation is the equivalent to 1 month. The Network model updates itself after each unit of model time.
δt	The number of sub steps ($1,000$) of each unit of model time created, in order to facilitate the numerical methods (Euler) that underpin the SD component. Thus the SD component updates $1,000$ times in each unit of model time
$\mathcal{T}$	The set of **ALL** times t, such that: $\mathcal{T} = \mathcal{T}^m * \delta t$.

TABLE B.1: Simulation: Parameters and Sets

The set of SD algorithms is triggered at each new time $t \in \mathcal{T}$, and the network algorithms are triggered at each model time step (where $t \mod \delta t = 0$). This is described in Algorithm 1.

Additional functions are run at different time intervals to update a range of agent specific parameters. In initial versions these were fixed, but are varied in later versions to potentially extend the functionality of the model:

- Satisficing number.
- Physical activity levels.

- Educational levels
- Henry equation constants (alpha, bravo, gamma).

In the algorithms below occasional reference is made to optimisation parameters (θ_n) these are described in full at Table 4.1.

B.2 Running the Network Model

Running the network algorithm involves a number of phases; initialisation, warm-up, running and the addition or removal of agents.

The parameters and sets needed to do this are given in Table B.2. In the initialisation phase agents are given characteristic variables, an initial history and the agent specific range and homophily variables are generated. The warm-up and simulation phases share an identical process (the time step process described below), in the case of the latter there are additional processes for initialising individual agents to add to the simulation and also to remove 'dead' agents.

The pseudocode describes the core element of network model (the time step process) in detail, for completeness the pseudocode for generating the homophily and range variables are also included (Algorithms 12 and 13).

The process for updating the network at each time step ($t \in T$) is described in Algorithms 2 to 13. The critical procedure is that used by each agent to determine which invitations to accept (Algorithm 4).

The subsequent procedures used to derive the list of those who have accepted each agents invitation (Algorithm 8) and hence the latest set of contacts (Algorithm 9) are relatively simple. The latter is then used to update each agents history (J_p^{hist}) which in turn is used to derive a set of network neighbours for each agent. With the exception of the set recording the history of each agents interactions ($\mathcal{J}^{hist}$), the sets are re-calculated for each time step, hence the first step of the update process is to clear these ready for the algorithms that follow. The set of network neighbours ($\mathcal{P}^{nn}$) enables the broader simulation, by providing the set of network neighbours or current contacts for each agent.

Ra	Network parameter representing social range between agents
S	Network parameter representing no. of interactions in each time slot
M	Network parameter representing 'memory' in history of interactions
Z	Network parameter describing the maximum value of $p.x$ & $p.y$
H	Boolean variable $(1, 0)$ describing whether a relationship can or can't exist between 2 agents.
$p.x$ & $p.y$	Characteristic variable, randomly allocated x and y axis positions for agent p.
$p.bmi$	Characteristic variable current BMI value for agent p.
$p.age$	Characteristic variable current age of agent p.
$p.gender$	Characteristic describing gender of agent p, $(0, 1)$.
$\mathcal{P}_p^{a'ted}$	The set of agent invitations($p' \subset P$) accepted by an agent (p).
$\mathcal{P}_p^{a'tees}$	The set of agents ($p' \subset P$) accepting an agent(p's) invitations.
$\mathcal{P}_p^{lc}$	The set of agents ($p' \subset P$) interacting with agent p in current time slot, may include repetitions.
$\mathcal{P}_p^{nn}$	The set of agents ($p' \subset P$) with which agent p is currently connected (no repetitions).
$\mathcal{P}_p^{r}$	The set of agents ($p' \subset P$) whose distance from p is less than Ra.
$\mathcal{P}_p^{h}$	The set of agents ($p' \subset P$) for which $Hm(p, p') = 1$.
$\mathcal{R}$	Set of tuples $(p, \mathcal{P}_p^{range}), \mathcal{P}_p^{range}$ is the set of Agents($p' \in \mathcal{P}$) within social range of p.
$\mathcal{H}$	Set of tuples $(p, \mathcal{P}_p^{H}), \mathcal{P}_p^{H}$ is the set of Agents($p' \in \mathcal{P}$) who share affinity with p.
$\mathcal{J}_p^{prob}$	A set of tuples $(p'_\pi, p) \in \mathcal{J}_p^{prob}$ with agent $p' \in \mathcal{P}$ and probability $\pi \in \mathbb{R}_{[0,1]}$ describing the likelihood of accepting an invitation from agent $p \in \mathcal{P}$. $\mathcal{J}_p^{prob}(p')$ returns the value π associated with p'.
$\mathcal{J}_p^{edge}$	A set of tuples $(p', \alpha) \in \mathcal{J}_p^{edge}$, with agent $p' \in \mathcal{P}$ and edge weight $\alpha \in \mathbb{R}_+$ describing the edge weight between p and p' over the time period H ($\alpha =$ (no.of contacts in time $H)/H$). $\mathcal{J}_p^{edge}(p')$returns the value α associated with p'
$\mathcal{J}_p^{hist}$	A set of tuples (β, P_p^{lc}) with $\beta \in \mathbb{Z}$, and $\mathcal{P}_p^{lc} \in \mathcal{P}$, describing the history of interactions of agent p with other agents $p' \in \mathcal{P}$. $J_p^{hist}(\beta)$ returns the set $\mathcal{P}_p^{lc}$ associated with β.

TABLE B.2: Network Algorithms: Parameters and Sets

B.3 Running the SD Model

Running the SD algorithm involves an initialisation phase, and then for each δt, three processes; updating stocks, updating flows and updating variables. Additionally at each network time step there are also processes for the addition or removal of agents.

In the initialisation phase agents are given a set of characteristic variables; age, height, gender, BMI and educational level ($edLevel$)

The pseudocode in Algorithms 14 to 19 below describe the three updating processes carried out for each δt, and uses the parameters and sets shown in Table B.3.

$p.trigger$	Common agent variable, defining the point at which dieting behaviour is triggered.
$p.networkIn$	Variable derived from $\mathcal{P}_p^{nn}$.
$p.pal$	Agent variable describing its physical activity level.
$p.tee$	Agent variable describing its total energy expenditure.
$p.ei$	Agent variable describing its energy intake.
$p.dietTime$	Agent variable describing the period of dieting.
$p.eb$	Agent variable its energy balance, defined by $p.tee - p.ei$.
$p.bmi$	Characteristic variable current BMI value for agent p.
$p.age$	Characteristic variable current age of agent p.
$p.gender$	Characteristic describing gender of agent p, $(0, 1)$.
$p.height$	Characteristic variable height of agent p.
$p.alpha$	Agent specific constant (Henry Equations).
$p.bravo$	Agent specific constant (Henry Equations.)
$p.gamma$	Agent specific constant (Henry Equations).
$p.edLevel$	Description of agents educational level.
$p.degreeProbability$	Probability of an individual; being educated to degree level(set to same value for all agents).
$p.intention$	Intention stock value for agent p.
$p.behaviour$	Behaviour stock value for agent p.
$p.bmi$	BMI stock value for agent p.
$p.norms$	Agent value $norms$ flow.
$p.attitudes$	Agent value for $attitudes$ flow.
$p.pbc$	Agent value for pbc flow.
$p.inDecay$	Agent value for $intentionDecay$ flow.
$p.beDecay$	Agent value for $behavioralDecay$ flow.
$p.activation$	Agent value for $activation$ flow.
$p.actLag$	Agent value for lagged $activation$ flow.
$\theta_{BMIfactor}$	Agent parameter for global influence (using APCCC).
θ_{bmiAdj}	Agent parameter for bmi adjustment in relation to disparities in BMI.
Lag_1	The lag (model time) in the flow between intention and behaviour.

TABLE B.3: SD Algorithms: Parameters

B.4 Running the Loss Function

Running the loss function involves comparing two sets of data; the output from the simulation and historical data for the same time period, each set comrpises the same information, as described in Table B.4.

$\mathcal{O}$	The output set produced by the simulation in the optimisation phase, comprising mean and median BMI's by age category and gender for each year .
$\mathcal{C}$	The comparator set produced from historical data, comprising mean and median BMI's by age category and gender for each year.

TABLE B.4: Loss Function:Sets

The process is described in Algorithm 24.

B.5 Algorithms

Algorithm 1 Running the Simulation ($\mathcal{P}$) for $t \in \mathcal{T}$

procedure RUNSIMULATION($\mathcal{P}, \mathcal{T}$)
  for all $t \in \mathcal{T}$ **do**
    $updateSD(\mathcal{P})$ ▷ Algorithm 14
    $updateSatisficingNo(\mathcal{P})$ ▷ Algorithm 20
    $updatePAL(\mathcal{P})$ ▷ Algorithm 21
    if $(t \mod \delta t = 0)$ **then**
      $updateNetwork(\mathcal{P})$ ▷ Algorithm 2
    if $(t \mod (\delta t * 12) = 0)$ **then**
      $updateEdLevel(\mathcal{P})$ ▷ Algorithm 22
      $updateAlphaBravoGamma(\mathcal{P})$ ▷ Algorithm 23

Algorithm 2 Updating the Network ($\mathcal{P}$) for $t^m \in \mathcal{T}^m$

procedure UPDATENETWORK($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    $p.age += 0.083333$
  for all $p \in \mathcal{P}$ **do**
    $clearSets(p)$ ▷ Algorithm 3
  for all $p \in \mathcal{P}$ **do**
    $updateAccepted(p)$ ▷ Algorithm 4
  for all $p \in \mathcal{P}$ **do**
    $updateAcceptees(p)$ ▷ Algorithm 8
  for all $p \in \mathcal{P}$ **do**
    $updateLatestContacts(p)$ ▷ Algorithm 9
  for all $p \in \mathcal{P}$ **do**
    $updateHistory(p)$ ▷ Algorithm 10
  for all $p \in \mathcal{P}$ **do**
    $updateNetworkNeighbours(p)$ ▷ Algorithm 11

Algorithm 3 Clearing the temporary Sets associated with individual Agents

procedure CLEARSETS($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    $\mathcal{P}_p^{a'ted} = \{\}$
    $\mathcal{P}_p^{a'tees} = \{\}$
    $\mathcal{P}_p^{lc} = \{\}$
    $\mathcal{P}_p^{nn} = \{\}$
    $\mathcal{J}_p^{prob} = \{\}$
    $\mathcal{J}_p^{edge} = \{\}$

Algorithm 4 Update Accepted Invitations for each Agent

procedure UPDATEACCEPTED(Agent p)
  $updateProbabilities(p)$ ▷ Algorithm 5
  for all Agent $p' \in \mathcal{J}_p^{prob}$ **do**
    for (int $i = 0; i < 20; i++$) **do**
      if $(Rand_u(0,1) \leq \mathcal{J}_p^{prob}(p')) \wedge (p' \notin \mathcal{P}_p^{a'ted}) \wedge (\mathcal{P}_p^{a'ted} \leq S)$) **then**
        $\mathcal{P}_p^{a'ted} \leftarrow p'$
  return $\mathcal{P}_p^{a'ted}$

Algorithm 5 Update Probabilities for each Agent - Equation 4.6.

procedure UPDATEPROBABILITIES(Agent p)
  int $total = 0$
  for all Agent $p' \in J_p^{edge}$ **do**
    if $(\mathcal{H}(p,p') = true \wedge Range(p,p') \leq Ra)$ **then**
      ▷ Algorithms 12, 13
      $total+ = \mathcal{J}_p^{edge}(p')$
  for Agent $p' \in \mathcal{J}_p^{edge}$ **do**
    if $(\mathcal{H}(p,p') = true \wedge \mathcal{R}(p,p') = true)$ **then**
      $\mathcal{J}_p^{prob} \leftarrow (p', \mathcal{J}_p^{ew}(p')/total)$
  return $\mathcal{J}_p^{prob}$

Algorithm 6 Update the EdgeWeights of each Agent

procedure UPDATEEDGEWEIGHTS(Agent p)
  for Agent $p' \in \mathcal{P}$ **do**
    if $p \neq p'$ **then**
      double $count = contactsInM(p,p')$ ▷ Algorithm 7
      if $|p.bmi - p'.bmi| < \theta_{trigger}$ **then**
        $\mathcal{J}_p^{edge} \leftarrow (p', count/H)$
      else
        $\mathcal{J}_p^{edge} \leftarrow (p', count/H)\theta_{BMIadj}$
  return $\mathcal{J}_p^{edge}$

Algorithm 7 Contacts in History Time Frame (M)

procedure CONTACTSINM(Agent p, Agent p')

int $count = 0$

for (int $i = 0; i < M; i++$) **do**

for Agent $p'' \in \mathcal{J}_p^{hist}(i)$ **do**

if $p' == p''$ **then**

$count += 1$

return $count$

Algorithm 8 Update list of those accepting each Agent's invitation

procedure UPDATEACCEPTEES(Agent p)

for Agent $p' \in \mathcal{P}$ **do**

if $p \in P_{p'}^{a'ted}$ **then**

$P_p^{a'tees} \leftarrow p'$

return $\mathcal{J}_p^{a'tees}$

Algorithm 9 Update Latest Contacts for each Agent

procedure UPDATELATESTCONTACTS(Agent p)

$\mathcal{P}_p^{lc} \leftarrow \mathcal{P}_p^{a'ted}$

for all Agent $p' \in P_p^{a'ted}$ **do**

$\mathcal{P}_p^{lc} \leftarrow \mathcal{P}_{p'}^{a'tees}$

$\mathcal{P}_p^{lc}$ remove p

return $\mathcal{P}_p^{lc}$

Algorithm 10 Update Agent History

procedure UPDATEHISTORY(Agent p)

for (int $i = 0; i < M; i++$) **do**

$\mathcal{J}_p^{hist}(i) := \mathcal{J}_p^{hist}(i) - J_p^{hist}(i - M)$

$\mathcal{J}_p^{hist}(i) \leftarrow \mathcal{P}_p^{lc})$

return $\mathcal{J}_p^{hist}$

Algorithm 11 Update Network Neighbours for each Agent

procedure UPDATENETWORKNEIGHBOURS(Agent p)

for (int $i = 0; i < H; i++$) **do**

for all $p' \in \mathcal{J}_p^{hist}(i)$ **do**

if $(p \neq p' \wedge p' \notin \mathcal{P}_p^{nn})$ **then**

$\mathcal{P}_p^{nn} \leftarrow p'$

return $\mathcal{P}_p^{nn}$

Algorithm 12 Procedure to determine Homophily between Agents

procedure HOMOPHILY(Agent p)
  for Agent $p' \in \mathcal{P}$ **do**
    if $p' \neq p$ **then**
      if $p' \notin \mathcal{P}_p^h || p \notin \mathcal{P}_{p'}^h$ **then**
        $double\ d = |p.age - p'.age|$
        $double\ c = (ageRange - d) * homophilParameter/agerange$
        if $Rand_u(0,1) < c$ **then**
          $\mathcal{P}_p^h \leftarrow p'$
          $\mathcal{P}_{p'}^h \leftarrow p$
  return $\mathcal{P}_p^h$

Note: $ageRange$ and $homophilParameter$ are network constants currently set at 50 and 1.152 respectively.

Algorithm 13 Procedure to determine whether agent is within range (Ra).

procedure INRANGE(Agent p)
  for all Agent $p' \in \mathcal{P}$ **do**
    if $p' \neq p$ **then**
      if $\sqrt{(p.x - p'.x)^2 + (p.y - p'.y)^2} \leq Ra$ **then**
        $\mathcal{P}_p^{range} \leftarrow (p')$
        $\mathcal{P}_{p'}^{range} \leftarrow (p)$
  return $\mathcal{P}_p^{range}$

Algorithm 14 Update the SD model

procedure UPDATESD($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    $updateStocks(p)$ ▷ Algorithm 15
  for all $p \in \mathcal{P}$ **do**
    $updateVariables(p)$ ▷ Algorithm 17
  for all $p \in \mathcal{P}$ **do**
    $updateFlows(p)$ ▷ Algorithm 16

Algorithm 15 Update Stocks - Equations 4.7, 4.8 and 4.9.

procedure UPDATESTOCKS(Agent p)
  for all $p \in \mathcal{P}$ **do**
    $p.intention_t = p.intention_{(t-1)} + (p.pbc_{(t-1)} + p.norms_{(t-1)} + p.attitudes_{(t-1)} - p.inDecay_{(t-1)} - p.activation_{(t-1)} - p.inDecay_{(t-1)})/\delta t$
    $p.behaviour_t = p.behaviour_{(t-1)} + (p.actLagged_{(t-1)} - p.beDecay_{(t-1)})/\delta t$
    $p.bmi_t = p.bmi_{(t-1)} + (p.weightIn_{(t-1)}/p.height^2)/\delta t$
  return $(p.intention_t, p.behaviour_t, p.bmi_t)$

Algorithm 16 Update Flows - Equations 4.10, 4.11, 4.12, 4.13, 4.14 and 4.15.

procedure UPDATEFLOWS(Agent p)
  for all $p \in \mathcal{P}$ **do**
    $p.weightIn_t = p.eb_{(t-1)}/(p.pal * p.gamma * 12)$
    $p.pbc_t = p.weightIn_{(t-1)} * \theta_{pbc}$
    $p.norms_t = (\theta_{(norms} * p.networkIn_{(t-1)})/2$ ▷ See note.
    if $p.edLevel == 2$ **then**
      $p.attitudes_t = 4 * \theta_{edLevel}$
    else
      $p.attitudes_t = 10$
    $p.inDecay_t = p.intention_{(t-1)}/2$
    $p.beDecay_t = p.behaviour_{(t-1)}/2$
    $p.activation_t = p.intention_{(t-1)}/3$
    $p.actLagged_t = p.activation_{(t-(Lag_1/\delta t)}$
  return $(p.weightIn_t, p.pbc_t, p.norms_t, p.attitudes_t, p.inDecay_t, p.beDecay_t,$
  $p.activation_t, p.actLagged_t)$

Algorithm 17 Update Variables - Equations 4.17 and 4.19.

procedure UPDATEVARIABLES(Agent p)
  for all $p \in \mathcal{P}$ **do**
    $updateNetworkIn(p)$ ▷ Algorithm 18
    $p.tee_t = p.pal(p.alpha + (p.bravo * p.height) + (p.gamma * p.height^2 * p.bmi_{(t-1)}))$
    $updateEnergyIntake(p)$ ▷ Algorithm 19
    $p.eb_t = p.energyIntake_{(t-1)} - p.tee_{(t-1)}$

Algorithm 18 Update Network In

procedure UPDATENETWORKIN(Agent p)
  $doubletotalBmi = 0$
  $double\ p.networkIn = p.bmi$
  $int\ i = 0$
  for all $p' \in \mathcal{P}_p^{nn}$ **do**
    if $p'.gender == p.gender$ **then**
      $totalBMI += p'.bmi$
      $i += 1$
    if $i > 0$ **then**
      $p.networkIn_t = totalBMI/i$
  return $p.networkIn_t$

Algorithm 19 Update Energy Intake - Equations 4.18 and 4.21.

procedure UPDATEENERGYINTAKE(Agent p)
  if $(p.behaviour_{(t-1)} < p.trigger \wedge (p.bmi_{(t-1)} - p.networkIn) > p.satisNo \wedge p.dietTime < modelTime)$ **then**
    $p.calorieDeficit = a.tee_{(t-1)} - 480$
    $double\ val = Rand_N(0,1) * \theta_{dietTime}/4 + \theta_{dietTime}$
    if $val > 1$ **then**
      $a.dietTime = modelTime + val$
    else
      $dietTime = modelTime + 1$
      $p.behaviour_t = 0$
  else if $(p.bmi_{(t-1)} - p.networkIn) > 0 \wedge p.dietTime <= modelTime$ **then**
    $p.ei_t = p.calorieDeficit$
  else
    $p.ei_t = p.tee_{(t-1)} + (apccc/3400)(p.tee_{(t-1)}/(p.bmi(t-1) * p.height^2 * (0.5 + \theta_{BMIfactor}))$

Algorithm 20 Update Satisficing Number - Equation 20.

procedure UPDATESATISFICINGNO($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    $double\ p.satisficingNo = -1$
    if $p.bmi > 20$ **then**
      $p.satisficingNo = (((p.bmi - 20)/5)^2)\theta_{sn}$
    return $p.satisficingNo$

Algorithm 21 Update PAL

procedure UPDATEPAL($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    $double\ p.pal = 1.63$
    if $p.bmi < 25$ **then**
      $p.pal = 1.61$
    else if $(p.bmi >= 25 \wedge p.bmi < 30)$ **then**
      $p.pal = 1.65$
    return $p.pal$

Algorithm 22 Update EdLevel

procedure UPDATEEDLEVEL($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    if $(p.edLevel == 0 \wedge p.age > 21)$ **then**
      if $Random_u(0,1) < p.degreeProbability$ **then**
        $p.edLevel = 2$
      else
        $p.edLevel = 1$
    return $p.edLevel$

Note. $p.edLevel$ is set to 0 for all new agents joining the network, and those under the age of 21 at the start of the simulation run.

Algorithm 23 Update AlphaBravoGamma

procedure UPDATEABG($\mathcal{P}$)
  for all $p \in \mathcal{P}$ **do**
    if $(p.age \leq 30 \wedge p.gender == 1)$ **then**
      $p.alpha - -113$
      $p.bravo = 313$
      $p.gamma = 14.4$
    else if $(p.age \leq 60 \wedge p.gender == 1)$ **then**
      $p.alpha - -137$
      $p.bravo = 541$
      $p.gamma = 11.4$
    else if $(p.gender == 1)$ **then**
      $p.alpha - -256$
      $p.bravo = 615$
      $p.gamma = 11.4$
    else if $(p.age \leq 30 \wedge p.gender == 2)$ **then**
      $p.alpha - -282$
      $p.bravo = 615$
      $p.gamma = 10$
    else if $(p.age \leq 60 \wedge p.gender == 2)$ **then**
      $p.alpha - -11.6$
      $p.bravo = 502$
      $p.gamma = 8.18$
    else if $(p.gender == 2)$ **then**
      $p.alpha - 10.72$
      $p.bravo = 421$
      $p.gamma = 8.52$
    return $(p.alpha, p.bravo, p.gamma)$

Note. This algorithm sets the parameters for calculating EI based on age and gender, using data from [102].

Algorithm 24 Calculate Loss

```
procedure CALCULATELOSS(𝒪, 𝒞)
    double Loss = 0
    for all year x ∈ 𝒪 do
        for all i ∈ 𝒪_x do
            double a = 2 * (𝒪_xi − 𝒞_xi)^2
            Loss+ = a
            if (i mod 2 == 0) then
                double b = 𝒪_xi − 𝒪_x(i+1)
                if (b < 0) then
                    Loss+ = b^2
    return Loss/40
```

Appendix C

Outputs

C.1 Sensitivity Analysis

Plus 5%

% Variation: -0.22%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.033	-0.021	0.072	0.156	0.023	0.028	-0.005	-0.016	-0.054	-0.019	0.034	-0.020
0.091	-0.114	0.049	0.048	0.045	0.114	0.007	0.004	-0.036	0.001	0.016	0.007
0.000	0.000	0.029	0.000	0.061	0.046	-0.022	0.003	-0.011	0.020	0.042	-0.002
0.000	0.000	0.041	0.048	0.059	0.094	-0.040	0.048	-0.011	0.070	-0.001	-0.043
0.000	0.000	0.037	0.027	0.027	0.131	0.005	0.045	-0.006	-0.001	0.048	0.046
0.000	0.000	0.042	0.139	0.044	0.054	0.019	0.005	-0.003	0.000	0.026	0.026
0.000	0.000	0.026	0.025	0.068	0.135	0.030	0.030	0.002	-0.033	0.002	0.032
0.000	0.000	0.036	0.022	0.111	0.077	0.054	0.003	0.014	-0.025	0.001	0.016
0.000	0.000	-0.008	-0.004	0.105	0.087	0.034	-0.016	0.002	0.000	-0.003	-0.015
0.000	0.000	-0.010	0.012	0.095	0.075	0.006	-0.028	0.017	-0.060	-0.081	-0.057

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.114	-0.068	0.026	0.016	0.039	0.000	-0.007	0.030	-0.005	0.010	0.146	0.007
-0.021	0.014	0.032	0.095	0.005	0.010	0.001	0.128	-0.062	0.001	0.167	0.057
0.000	0.000	-0.033	-0.021	0.056	-0.036	0.026	-0.001	-0.065	-0.019	0.233	0.154
0.000	0.000	-0.005	-0.012	0.000	-0.014	0.014	0.001	-0.028	0.020	0.159	0.095
0.000	0.000	-0.029	-0.027	0.049	0.031	0.017	-0.051	-0.051	0.000	0.103	0.053
0.000	0.000	-0.058	-0.022	0.036	-0.004	-0.007	-0.040	0.028	0.031	0.121	0.112
0.000	0.000	-0.050	-0.200	0.079	-0.013	0.050	-0.004	-0.005	0.002	0.121	0.041
0.000	0.000	-0.072	-0.100	0.022	-0.015	-0.020	-0.050	0.046	0.008	0.127	0.082
0.000	0.000	-0.053	-0.067	0.027	-0.004	0.019	-0.025	-0.007	0.037	0.040	0.086
0.000	0.000	-0.053	0.044	0.017	-0.071	0.048	-0.006	-0.041	0.021	0.028	0.072

Minus 5%

% Variation: -1.36%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.044	0.102	0.025	0.129	0.027	0.069	-0.006	0.017	0.005	-0.064	-0.004	-0.055
0.100	0.014	0.030	0.040	0.056	0.111	0.020	0.028	0.017	-0.077	-0.013	-0.025
0.000	0.000	0.033	0.014	0.069	0.089	0.009	0.021	0.054	-0.018	0.004	-0.013
0.000	0.000	0.040	0.069	0.068	0.091	-0.006	0.073	0.015	0.007	0.006	-0.065
0.000	0.000	0.046	0.028	0.021	0.088	0.002	0.079	-0.019	-0.002	0.038	0.017
0.000	0.000	0.052	0.123	0.039	0.009	0.021	0.043	0.009	-0.006	0.061	-0.022
0.000	0.000	0.004	-0.069	0.039	0.131	0.039	0.029	0.011	-0.013	0.070	0.008
0.000	0.000	0.009	0.000	0.031	0.089	0.035	0.037	-0.015	-0.007	0.004	-0.032
0.000	0.000	0.015	-0.004	0.043	0.056	0.017	-0.016	-0.041	0.054	0.013	-0.049
0.000	0.000	0.009	-0.012	0.065	0.040	-0.002	-0.007	0.008	-0.006	0.006	-0.067

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.009	0.000	0.077	-0.012	0.006	-0.013	0.052	-0.001	0.023	-0.037	0.085	0.007
0.006	0.023	0.086	0.112	-0.043	0.021	0.024	-0.009	-0.040	-0.091	0.105	0.166
0.000	0.000	0.030	-0.012	-0.023	-0.058	0.059	0.001	-0.081	-0.058	0.137	0.166
0.000	0.000	0.051	-0.007	-0.024	-0.055	-0.012	-0.013	-0.032	-0.078	0.115	-0.001
0.000	0.000	0.044	-0.006	0.024	0.008	-0.041	-0.127	-0.060	-0.031	0.022	0.018
0.000	0.000	0.036	0.002	0.010	-0.004	0.018	-0.069	-0.019	0.008	0.051	0.102
0.000	0.000	0.031	-0.026	0.070	-0.006	0.013	-0.004	-0.045	0.007	0.037	-0.054
0.000	0.000	0.045	0.005	0.015	0.021	-0.013	-0.117	0.001	-0.028	0.023	-0.077
0.000	0.000	0.046	0.067	-0.004	-0.016	-0.027	-0.048	0.006	-0.006	-0.011	0.025
0.000	0.000	0.040	0.077	0.043	0.019	0.018	-0.056	-0.025	0.019	-0.009	0.000

TABLE C.1: BMI Adjustment - BMI Influence on Network

Plus 5%

% Variation: -0.32%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.065	-0.030	0.029	0.052	0.005	0.062	-0.034	-0.001	0.030	-0.013	0.007	-0.060
0.047	-0.120	-0.016	-0.001	-0.043	0.101	-0.030	-0.021	0.026	-0.007	-0.001	0.060
-0.008	-0.008	-0.017	-0.035	0.016	0.072	-0.049	-0.010	0.044	-0.012	0.011	-0.016
-0.011	-0.004	-0.010	-0.009	0.033	0.113	-0.071	0.019	0.009	0.041	0.037	-0.032
-0.013	-0.015	-0.021	0.000	-0.004	0.069	-0.025	0.008	-0.008	0.059	0.012	0.041
-0.013	-0.010	-0.011	0.099	-0.013	0.003	-0.035	-0.032	-0.041	-0.073	-0.015	0.011
-0.013	-0.005	0.002	-0.063	-0.012	0.092	-0.036	-0.028	-0.024	-0.126	0.014	-0.022
-0.013	-0.009	-0.016	-0.030	-0.007	0.022	-0.035	-0.033	-0.008	-0.050	-0.055	-0.022
-0.014	-0.014	-0.055	-0.059	-0.010	-0.024	-0.053	-0.076	-0.016	0.007	-0.051	-0.070
-0.014	-0.022	-0.078	-0.052	-0.027	-0.051	-0.048	-0.071	-0.018	-0.040	-0.122	-0.123

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.086	0.025	0.080	0.013	-0.020	-0.040	0.033	0.024	-0.073	-0.005	0.109	0.157
-0.005	-0.005	0.140	0.111	-0.069	-0.030	-0.001	0.055	-0.085	0.036	0.128	0.216
-0.011	-0.009	0.028	-0.047	-0.057	-0.139	0.071	0.128	-0.090	-0.017	0.212	0.204
-0.015	-0.010	0.056	-0.038	-0.048	-0.022	0.015	0.012	-0.063	0.011	0.210	0.131
-0.020	-0.017	0.020	-0.041	-0.008	0.035	-0.010	0.003	-0.072	0.008	0.062	0.073
-0.019	-0.029	0.021	-0.048	-0.034	-0.032	-0.038	-0.029	0.002	0.121	0.100	0.132
-0.018	-0.018	-0.026	-0.040	0.005	-0.073	-0.044	-0.044	-0.031	0.069	0.046	0.042
-0.017	-0.026	-0.036	-0.053	-0.095	-0.042	-0.076	-0.068	-0.005	0.024	-0.002	-0.091
-0.016	-0.021	-0.006	-0.008	-0.065	-0.061	-0.111	-0.056	-0.028	0.009	-0.079	-0.069
-0.016	-0.018	-0.026	0.019	-0.053	-0.096	-0.080	-0.028	-0.063	0.004	-0.134	-0.110

Minus 5%

% Variation: -1.22%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.003	0.004	0.073	0.083	0.030	0.027	-0.014	0.010	-0.026	-0.012	0.012	-0.049
0.119	0.020	0.038	0.038	0.040	0.114	0.024	0.054	0.016	0.014	-0.005	0.018
0.009	0.009	0.067	-0.009	0.066	0.061	0.021	0.047	0.019	0.015	0.017	0.014
0.012	0.005	0.079	0.049	0.103	0.143	0.019	0.098	-0.004	-0.028	0.012	0.013
0.014	0.016	0.066	0.056	0.057	0.137	0.062	0.121	0.012	-0.030	0.071	0.103
0.014	0.011	0.084	0.086	0.038	0.041	0.079	0.060	0.007	-0.041	0.043	0.083
0.014	0.005	0.034	0.052	0.086	0.146	0.096	0.105	0.004	-0.111	0.037	0.079
0.015	0.010	0.048	0.050	0.135	0.064	0.111	0.127	0.042	-0.013	0.038	0.058
0.015	0.015	0.059	0.058	0.147	0.134	0.092	0.046	0.036	0.041	0.075	0.069
0.016	0.024	0.044	0.041	0.136	0.099	0.102	0.078	0.063	0.023	0.028	0.062

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.053	0.009	0.089	0.032	0.022	-0.074	0.031	0.039	0.051	0.010	0.065	0.011
0.035	0.034	0.094	0.094	-0.015	-0.029	0.045	0.279	0.003	0.050	0.122	0.116
0.012	0.010	0.037	0.015	0.001	-0.095	0.067	0.181	0.028	-0.017	0.179	-0.034
0.017	0.010	0.065	0.026	0.023	-0.028	0.055	0.084	0.040	0.073	0.166	0.035
0.022	0.018	0.038	0.031	0.107	0.071	0.059	0.015	0.036	0.084	0.173	0.126
0.021	0.032	0.066	0.048	0.108	-0.026	0.095	0.042	0.055	0.185	0.218	0.241
0.020	0.020	0.078	-0.003	0.114	0.044	0.112	0.076	0.026	0.081	0.221	0.151
0.018	0.029	0.087	0.057	0.064	0.046	0.104	0.032	0.089	0.081	0.255	0.154
0.018	0.023	0.101	0.093	0.074	0.022	0.099	0.042	0.050	0.101	0.177	0.127
0.018	0.019	0.052	0.130	0.116	0.074	0.163	0.042	0.049	0.108	0.197	0.176

TABLE C.2: BMI Factor - Environmental Influence

Plus 5%

% Variation: -1.81%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.028	-0.013	0.037	0.113	0.004	0.021	-0.009	-0.001	-0.087	-0.127	0.016	-0.062
0.062	0.000	0.063	0.059	-0.018	0.090	0.022	-0.026	-0.062	-0.119	-0.002	-0.060
0.000	0.000	0.046	0.032	0.004	0.043	-0.007	0.005	-0.050	-0.079	0.035	-0.009
0.000	0.000	0.053	0.030	0.014	0.088	-0.002	0.033	-0.059	-0.063	0.093	-0.065
0.000	0.000	0.071	0.028	-0.026	0.088	0.027	0.051	-0.053	-0.066	0.049	0.021
0.000	0.000	0.069	0.139	-0.037	-0.012	0.036	-0.006	-0.061	-0.155	0.013	-0.013
0.000	0.000	0.039	0.014	0.009	0.155	0.029	0.030	-0.022	-0.152	0.042	-0.025
0.000	0.000	0.050	0.022	0.015	-0.008	0.061	0.032	-0.045	-0.127	0.019	-0.029
0.000	0.000	0.044	0.005	0.035	0.000	0.040	0.037	-0.028	-0.049	0.020	-0.124
0.000	0.000	0.048	0.034	0.048	-0.023	0.032	-0.003	-0.012	-0.101	-0.005	-0.056

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.009	0.000	0.042	0.011	0.025	-0.101	-0.022	-0.004	0.035	-0.010	0.068	0.007
-0.036	0.014	0.085	0.146	-0.011	-0.096	-0.041	0.070	-0.050	-0.064	0.123	0.000
0.000	0.000	0.029	-0.001	-0.018	-0.121	0.018	-0.041	-0.036	-0.035	0.156	-0.020
0.000	0.000	0.061	0.006	-0.010	-0.055	0.006	-0.017	-0.036	-0.085	0.133	-0.043
0.000	0.000	0.050	-0.008	0.068	-0.002	-0.017	-0.138	-0.043	0.000	0.119	0.052
0.000	0.000	0.027	-0.009	0.058	-0.055	-0.008	-0.124	0.010	-0.013	0.138	0.084
0.000	0.000	0.024	-0.059	0.040	-0.058	0.034	-0.054	0.010	-0.022	0.092	0.000
0.000	0.000	0.045	-0.023	-0.051	-0.012	-0.017	-0.117	0.001	-0.038	0.128	-0.032
0.000	0.000	0.058	0.014	-0.024	-0.038	-0.018	-0.061	-0.039	-0.074	-0.020	0.003
0.000	0.000	0.031	0.045	0.012	-0.020	0.008	-0.074	-0.059	-0.058	-0.002	-0.033

Minus 5%

% Variation: -1.36%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.037	0.102	0.060	0.092	-0.002	0.012	0.012	0.037	-0.043	0.006	0.040	-0.055
0.129	0.007	0.032	0.045	-0.010	0.069	0.013	0.028	-0.041	0.001	0.020	-0.060
0.000	0.000	0.049	0.071	0.021	0.032	0.004	-0.001	-0.023	0.036	0.112	-0.006
0.000	0.000	0.060	0.076	0.016	0.060	-0.022	0.051	-0.046	0.061	0.019	0.002
0.000	0.000	0.070	0.028	0.001	0.050	0.007	0.055	-0.029	0.001	0.077	0.036
0.000	0.000	0.064	0.128	0.021	0.025	0.004	0.030	-0.043	-0.004	0.085	0.092
0.000	0.000	0.033	-0.023	0.049	0.137	0.004	0.048	-0.008	-0.075	0.106	0.051
0.000	0.000	0.059	0.048	0.094	0.110	0.009	0.013	-0.010	-0.057	0.078	0.109
0.000	0.000	0.049	0.066	0.055	0.056	0.015	-0.019	-0.011	0.043	0.023	0.071
0.000	0.000	0.043	0.034	0.054	0.077	0.005	-0.002	0.023	0.010	-0.005	0.002

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.036	0.001	0.041	-0.012	0.004	-0.038	-0.037	0.035	0.040	-0.001	0.124	0.007
-0.024	0.014	0.098	0.078	-0.018	-0.027	-0.027	0.113	0.013	0.070	0.127	0.050
0.000	0.000	0.063	0.001	-0.027	-0.065	0.009	-0.002	0.012	0.006	0.202	-0.007
0.000	0.000	0.063	0.003	0.002	-0.055	-0.026	-0.013	0.031	-0.016	0.151	0.003
0.000	0.000	0.050	-0.006	0.050	-0.001	-0.018	-0.072	-0.052	-0.059	0.123	0.080
0.000	0.000	0.072	0.000	0.080	-0.035	0.007	0.003	0.015	0.026	0.127	0.173
0.000	0.000	0.023	-0.095	0.080	-0.022	0.033	-0.014	-0.035	0.000	0.121	0.073
0.000	0.000	0.032	0.000	-0.014	-0.018	-0.005	-0.088	0.005	-0.055	0.120	0.130
0.000	0.000	0.045	0.032	0.038	-0.037	-0.050	-0.042	-0.044	0.000	0.056	0.123
0.000	0.000	0.011	0.059	0.050	-0.017	-0.021	-0.065	-0.059	-0.002	0.063	0.071

TABLE C.3: Diet Time

Plus 5%

% Variation: -1.60%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.020	0.047	0.058	0.123	0.003	0.033	-0.065	0.015	-0.104	-0.127	0.019	0.001
0.068	0.000	0.040	0.026	-0.019	0.117	-0.021	0.048	-0.098	-0.090	-0.027	-0.049
0.000	0.000	0.006	-0.036	-0.006	0.079	-0.029	0.035	-0.084	-0.050	0.041	-0.002
0.000	0.000	0.036	0.020	0.021	0.092	-0.036	0.053	-0.106	-0.056	-0.041	-0.068
0.000	0.000	0.041	0.028	0.016	0.088	0.036	0.087	-0.119	-0.123	-0.008	0.003
0.000	0.000	0.006	0.128	-0.014	0.031	0.063	0.040	-0.096	-0.117	-0.074	-0.023
0.000	0.000	-0.002	0.034	0.050	0.136	0.041	0.028	-0.074	-0.156	-0.069	-0.002
0.000	0.000	0.019	0.038	0.040	0.080	0.047	0.057	-0.056	-0.107	-0.102	-0.040
0.000	0.000	0.004	0.006	0.067	0.036	0.035	0.039	-0.068	-0.017	-0.104	-0.091
0.000	0.000	0.014	0.008	0.045	0.030	0.005	0.003	-0.031	-0.004	-0.069	-0.108

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.125	0.022	0.033	0.000	0.062	-0.010	0.019	0.030	0.063	0.010	0.030	0.006
0.029	0.045	0.076	0.066	0.016	0.044	0.027	0.175	0.033	0.001	0.120	0.146
0.000	0.000	0.030	-0.051	-0.002	0.006	0.031	0.027	-0.041	-0.040	0.120	0.060
0.000	0.000	0.055	-0.031	0.005	0.018	0.021	-0.013	0.019	-0.075	0.105	0.000
0.000	0.000	0.031	-0.047	0.079	0.093	0.029	-0.052	-0.011	-0.027	0.070	0.114
0.000	0.000	0.050	-0.001	0.055	-0.058	0.030	0.004	0.039	0.078	0.089	0.170
0.000	0.000	0.053	-0.026	0.080	-0.007	0.030	0.011	0.019	0.012	0.103	0.109
0.000	0.000	0.037	0.005	0.033	0.011	0.020	-0.047	0.046	0.033	0.109	0.055
0.000	0.000	0.073	0.067	0.035	-0.003	0.019	0.015	0.031	0.067	0.033	0.021
0.000	0.000	0.051	0.059	0.082	0.003	0.058	0.051	-0.003	0.085	0.018	0.053

Minus 5%

% Variation: -0.78%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.143	0.069	0.026	0.076	-0.034	0.030	-0.032	-0.036	-0.032	-0.064	0.009	0.003
0.161	0.014	0.003	0.026	-0.023	0.069	0.016	-0.051	-0.036	-0.071	0.036	0.011
0.000	0.000	0.042	0.027	-0.004	0.043	-0.020	-0.004	-0.012	-0.030	0.076	0.000
0.000	0.000	0.059	0.070	0.003	0.117	-0.034	0.017	-0.034	-0.029	-0.004	0.001
0.000	0.000	0.049	0.028	-0.007	0.002	-0.013	0.019	-0.022	-0.040	0.023	0.051
0.000	0.000	0.053	0.081	-0.007	-0.004	0.013	-0.016	-0.009	-0.014	0.069	0.041
0.000	0.000	0.041	0.014	-0.008	0.031	0.024	-0.004	-0.021	-0.108	0.052	0.026
0.000	0.000	0.065	0.022	0.025	-0.003	0.019	0.017	0.016	-0.069	0.003	-0.032
0.000	0.000	0.053	0.010	0.002	-0.040	0.014	-0.027	0.009	0.016	-0.010	-0.013
0.000	0.000	0.084	0.034	0.006	0.012	0.009	-0.006	0.020	-0.011	-0.050	-0.081

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.100	-0.068	0.074	0.016	-0.026	-0.074	-0.006	-0.011	-0.066	-0.046	0.046	-0.099
0.023	0.000	0.066	0.102	-0.043	-0.096	-0.021	-0.010	-0.095	-0.011	0.048	-0.049
0.000	0.000	0.004	0.003	-0.028	-0.135	-0.014	-0.001	-0.068	-0.029	0.073	-0.120
0.000	0.000	0.023	0.067	-0.016	-0.058	-0.025	-0.018	-0.049	-0.072	0.077	-0.229
0.000	0.000	-0.016	0.000	0.032	0.073	-0.044	-0.122	-0.070	-0.043	0.056	-0.093
0.000	0.000	0.000	-0.008	0.001	-0.061	-0.042	-0.094	-0.021	0.028	0.015	-0.022
0.000	0.000	-0.002	-0.116	0.051	-0.013	-0.014	-0.054	-0.039	0.005	-0.003	-0.065
0.000	0.000	-0.018	0.000	0.006	0.001	-0.056	-0.098	0.018	-0.043	0.006	-0.061
0.000	0.000	-0.002	0.013	0.023	-0.034	-0.058	-0.037	-0.019	-0.022	-0.054	-0.059
0.000	0.000	-0.002	-0.050	0.025	0.007	-0.034	-0.062	-0.037	-0.036	-0.087	-0.071

TABLE C.4: Educational Level

Plus 1

% Variation: -1.02%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.059	0.027	0.062	0.122	0.029	0.069	-0.017	0.011	-0.040	-0.064	-0.006	-0.055
0.172	0.014	0.074	0.046	0.045	0.201	-0.007	0.002	0.022	-0.034	-0.001	0.010
0.000	0.000	0.063	0.022	0.068	0.162	-0.026	0.000	0.059	-0.009	0.071	0.001
0.000	0.000	0.078	0.043	0.080	0.173	-0.046	0.017	0.022	0.006	-0.020	0.006
0.000	0.000	0.093	0.027	0.038	0.145	-0.006	0.043	0.008	0.084	0.109	0.036
0.000	0.000	0.081	0.038	0.103	0.108	0.010	-0.009	-0.011	-0.041	0.005	0.078
0.000	0.000	0.061	0.006	0.121	0.231	-0.007	0.014	0.011	-0.002	-0.013	0.001
0.000	0.000	0.065	0.017	0.121	0.166	0.002	-0.010	0.017	0.000	0.003	-0.008
0.000	0.000	0.061	0.000	0.094	0.095	-0.018	-0.024	0.009	0.076	0.007	-0.065
0.000	0.000	0.044	0.012	0.083	0.077	-0.021	-0.021	0.032	0.015	-0.046	-0.066

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.055	-0.018	-0.013	0.000	-0.028	-0.018	-0.014	0.031	-0.002	0.021	0.010	0.006
-0.026	0.023	-0.011	0.040	-0.079	-0.054	0.019	0.188	-0.076	0.034	0.081	0.044
0.000	0.000	-0.046	-0.047	-0.059	-0.114	0.051	0.002	-0.033	0.058	0.109	0.035
0.000	0.000	-0.022	-0.015	-0.099	0.012	0.047	0.007	-0.030	0.039	0.128	0.000
0.000	0.000	-0.050	-0.011	0.018	0.027	0.018	-0.110	-0.064	0.045	0.123	0.109
0.000	0.000	-0.091	-0.008	0.043	-0.017	-0.007	-0.059	0.024	0.121	0.099	0.211
0.000	0.000	-0.066	-0.113	0.074	-0.013	0.002	-0.003	0.014	0.082	0.081	0.038
0.000	0.000	-0.070	0.000	-0.033	-0.025	-0.009	-0.025	0.041	0.018	0.108	0.055
0.000	0.000	-0.018	0.053	0.018	-0.035	-0.034	0.006	0.050	0.020	0.031	0.048
0.000	0.000	-0.042	0.044	-0.012	-0.016	0.024	-0.002	0.009	0.075	0.021	-0.009

Minus 1

% Variation: -1.19%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.010	0.077	0.124	0.158	0.018	0.021	-0.003	0.028	-0.059	-0.018	0.057	0.006
0.117	0.007	0.105	0.050	0.012	0.102	0.008	0.029	-0.066	-0.007	0.136	0.076
0.000	0.000	0.081	0.055	0.063	0.047	-0.004	0.023	-0.026	0.002	0.152	0.007
0.000	0.000	0.117	0.044	0.056	0.109	-0.023	0.051	-0.034	0.053	0.064	-0.033
0.000	0.000	0.112	0.058	0.037	0.102	-0.011	0.046	-0.037	0.067	0.087	0.021
0.000	0.000	0.119	0.190	0.027	0.022	-0.016	0.016	-0.044	-0.015	0.024	-0.011
0.000	0.000	0.122	0.053	0.110	0.168	-0.012	0.018	-0.014	-0.092	0.023	-0.001
0.000	0.000	0.117	0.042	0.107	0.102	0.000	0.008	-0.016	-0.029	-0.005	-0.017
0.000	0.000	0.098	0.066	0.125	0.080	-0.018	0.014	-0.042	-0.042	-0.038	-0.035
0.000	0.000	0.064	0.034	0.121	0.090	-0.012	0.017	-0.014	-0.009	-0.050	-0.094

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.060	-0.018	0.002	-0.006	0.001	-0.001	0.027	0.067	-0.026	-0.053	0.038	0.002
0.014	0.000	0.021	0.093	-0.005	0.008	0.030	0.245	-0.078	-0.082	0.076	0.033
0.000	0.000	-0.013	-0.025	0.002	-0.071	0.012	0.003	-0.002	-0.049	0.155	0.035
0.000	0.000	0.012	-0.015	0.016	-0.018	-0.007	-0.023	0.016	-0.072	0.158	0.009
0.000	0.000	-0.007	-0.012	0.087	0.113	-0.035	-0.100	-0.021	0.020	0.135	0.110
0.000	0.000	-0.030	-0.022	0.026	-0.002	-0.004	-0.046	0.050	0.060	0.199	0.194
0.000	0.000	-0.040	-0.103	0.046	0.010	0.067	-0.016	0.017	0.005	0.163	0.076
0.000	0.000	-0.036	0.005	-0.002	0.039	0.002	-0.009	0.070	0.011	0.168	-0.029
0.000	0.000	-0.014	0.067	0.000	0.005	-0.040	0.000	0.089	0.041	0.052	-0.007
0.000	0.000	-0.009	0.052	0.034	0.012	-0.019	0.004	0.011	0.079	0.041	0.000

TABLE C.5: Lag

Plus 1

% Variation: -0.48%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.032	0.003	-0.027	0.096	0.024	0.020	0.020	0.015	-0.009	-0.064	-0.037	-0.064
0.116	0.004	-0.007	0.009	0.031	0.075	0.030	0.027	0.014	0.003	-0.050	-0.006
0.000	0.000	0.002	-0.001	0.074	0.042	0.006	0.016	0.017	0.002	-0.028	-0.007
0.000	0.000	0.006	0.015	0.060	0.070	-0.007	0.033	-0.011	-0.002	-0.040	-0.065
0.000	0.000	-0.018	0.027	0.012	0.000	0.040	0.055	-0.051	-0.001	0.026	0.052
0.000	0.000	0.008	0.137	0.021	0.019	0.024	-0.006	0.001	0.000	0.002	-0.013
0.000	0.000	-0.007	-0.023	0.059	0.039	0.011	0.010	-0.007	-0.033	0.033	-0.003
0.000	0.000	0.016	0.022	0.026	0.007	0.008	0.007	0.023	-0.057	0.027	0.024
0.000	0.000	-0.020	-0.028	0.014	0.009	-0.009	-0.024	0.035	0.052	-0.004	-0.067
0.000	0.000	-0.002	0.000	0.009	0.035	-0.044	-0.037	0.042	0.012	-0.051	-0.101

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.048	-0.109	-0.014	-0.051	0.015	0.003	-0.027	-0.028	-0.004	-0.123	0.075	0.007
-0.014	-0.018	0.044	0.059	-0.020	-0.013	-0.047	-0.098	-0.044	-0.099	0.133	0.112
0.000	0.000	0.007	-0.012	0.042	-0.042	-0.026	-0.051	-0.074	-0.131	0.137	0.161
0.000	0.000	0.018	-0.006	0.044	-0.003	-0.080	-0.105	-0.064	-0.080	0.105	0.036
0.000	0.000	0.009	-0.011	0.065	0.000	-0.070	-0.180	-0.092	-0.137	0.109	0.104
0.000	0.000	0.026	0.000	0.081	-0.061	-0.090	-0.145	-0.020	-0.024	0.092	0.199
0.000	0.000	0.018	0.000	0.113	-0.009	-0.060	-0.093	-0.017	-0.047	0.061	0.057
0.000	0.000	-0.018	-0.011	0.032	0.005	-0.063	-0.135	0.020	-0.050	0.110	-0.024
0.000	0.000	-0.022	-0.023	0.040	-0.003	-0.060	-0.122	-0.031	-0.039	0.006	0.005
0.000	0.000	-0.013	-0.050	0.052	-0.009	-0.006	-0.081	-0.061	-0.069	-0.018	-0.005

Minus 1

% Variation: -1.34%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.032	-0.012	0.041	0.158	0.008	0.068	-0.025	-0.005	-0.036	-0.012	0.008	-0.061
0.122	0.014	0.007	0.044	0.021	0.164	-0.013	0.005	-0.003	0.000	0.048	0.005
0.000	0.000	0.044	0.000	0.048	0.054	-0.027	0.021	0.013	0.039	0.039	0.001
0.000	0.000	0.067	0.015	0.049	0.101	-0.037	0.034	0.009	0.007	-0.050	-0.064
0.000	0.000	0.034	0.028	0.020	0.088	0.018	0.055	0.003	-0.067	0.024	0.036
0.000	0.000	0.041	0.133	0.018	0.008	0.027	0.014	0.006	-0.045	0.063	0.004
0.000	0.000	0.021	0.051	0.055	0.137	0.041	0.057	0.009	-0.089	0.069	0.030
0.000	0.000	0.014	0.045	0.061	0.069	0.048	0.062	0.018	-0.038	0.015	-0.029
0.000	0.000	0.017	0.022	0.043	0.061	0.028	0.027	-0.004	0.019	0.038	0.011
0.000	0.000	-0.001	0.034	0.072	0.069	0.019	0.018	0.026	-0.012	-0.006	0.017

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.053	0.062	0.017	0.001	-0.014	-0.061	-0.036	0.064	0.008	-0.011	0.105	0.007
0.033	0.018	0.053	0.160	-0.036	-0.106	-0.033	0.078	0.000	-0.044	0.132	0.122
0.000	0.000	0.019	0.003	-0.064	-0.163	0.012	0.180	-0.001	-0.033	0.156	0.131
0.000	0.000	0.039	0.020	-0.089	-0.110	-0.019	0.052	0.032	-0.060	0.102	0.002
0.000	0.000	0.054	0.031	-0.021	-0.075	0.008	-0.038	-0.012	0.010	0.050	0.073
0.000	0.000	0.062	0.010	-0.032	-0.129	0.036	-0.028	-0.002	0.072	0.074	0.172
0.000	0.000	0.082	0.018	0.015	-0.093	0.021	0.003	-0.021	0.000	0.077	0.077
0.000	0.000	0.127	0.021	-0.097	-0.080	-0.036	-0.049	0.020	0.001	0.055	-0.033
0.000	0.000	0.085	0.081	-0.074	-0.049	-0.063	0.012	-0.002	0.020	-0.004	0.002
0.000	0.000	0.066	0.183	-0.024	-0.016	-0.013	0.050	-0.036	0.012	-0.005	-0.015

TABLE C.6: Memory

Plus 5%

% Variation: -0.41%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.053	-0.021	0.114	0.126	0.018	0.044	-0.016	0.024	-0.047	-0.018	0.010	0.003
0.089	-0.101	0.091	0.032	0.019	0.117	0.033	0.064	-0.012	-0.034	-0.007	0.009
0.000	0.000	0.090	0.009	0.068	0.065	0.004	0.022	0.004	-0.015	-0.007	-0.013
0.000	0.000	0.122	0.074	0.050	0.127	0.003	0.049	-0.026	0.008	-0.002	-0.061
0.000	0.000	0.104	0.028	0.008	0.105	0.039	0.124	-0.042	-0.073	0.059	0.040
0.000	0.000	0.110	0.155	-0.006	0.025	0.053	0.046	-0.042	-0.075	0.043	0.060
0.000	0.000	0.069	0.014	0.009	0.151	0.044	0.113	0.000	-0.082	0.040	0.019
0.000	0.000	0.038	0.000	0.037	0.085	0.049	0.067	-0.012	-0.025	-0.005	-0.022
0.000	0.000	0.033	-0.004	0.070	0.087	0.054	0.046	-0.004	0.044	-0.002	0.000
0.000	0.000	0.012	0.006	0.114	0.083	0.032	0.035	0.015	-0.005	-0.018	-0.082

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.001	0.011	0.037	0.016	-0.060	-0.047	0.003	0.000	0.055	0.038	0.083	0.102
0.030	0.045	0.078	0.053	-0.072	-0.070	-0.018	0.089	0.021	0.000	0.124	0.170
0.000	0.000	0.007	-0.003	-0.097	-0.117	0.036	0.051	-0.016	-0.054	0.185	0.177
0.000	0.000	0.046	0.001	-0.079	-0.056	0.017	0.007	-0.013	-0.036	0.201	0.113
0.000	0.000	0.063	-0.011	-0.018	-0.048	-0.003	-0.097	-0.011	0.021	0.157	0.174
0.000	0.000	0.050	-0.005	-0.004	-0.111	0.019	-0.071	0.045	0.073	0.180	0.271
0.000	0.000	-0.011	-0.051	0.025	-0.017	0.042	-0.010	-0.008	0.005	0.158	0.147
0.000	0.000	0.019	0.005	-0.064	-0.052	-0.033	-0.076	0.024	0.005	0.116	0.088
0.000	0.000	0.033	0.067	-0.043	-0.038	-0.049	0.013	0.013	-0.019	0.055	0.119
0.000	0.000	0.003	0.078	0.014	-0.025	-0.025	-0.009	-0.040	0.016	0.026	0.091

Minus 5%

% Variation: -0.78%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.152	0.134	0.063	0.079	-0.001	0.008	0.004	0.003	-0.081	-0.064	0.048	0.030
0.181	0.007	0.064	0.021	0.017	0.070	0.015	0.007	-0.053	-0.058	0.055	0.069
0.000	0.000	0.079	0.022	0.042	0.040	0.011	0.000	-0.038	-0.001	0.084	0.001
0.000	0.000	0.103	0.055	0.029	0.104	-0.027	0.030	-0.067	-0.048	0.082	-0.047
0.000	0.000	0.093	0.028	0.008	0.098	-0.027	0.006	-0.033	-0.040	0.125	0.078
0.000	0.000	0.090	0.140	0.024	0.025	-0.013	0.010	-0.032	-0.093	0.083	0.085
0.000	0.000	0.059	0.002	0.038	0.125	-0.013	0.000	-0.006	-0.072	0.023	0.029
0.000	0.000	0.079	0.034	0.023	0.005	-0.017	0.003	-0.010	-0.062	-0.033	0.024
0.000	0.000	0.057	0.010	0.061	-0.028	-0.015	-0.028	-0.020	-0.027	-0.086	0.005
0.000	0.000	0.065	0.034	0.060	0.013	-0.023	-0.028	0.008	-0.036	-0.026	-0.010

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.025	0.011	0.054	0.010	-0.023	-0.026	0.018	0.056	-0.018	-0.045	0.074	0.006
0.040	0.014	0.110	0.082	-0.067	0.008	0.020	0.252	-0.135	-0.023	0.163	0.204
0.000	0.000	0.057	-0.011	-0.056	-0.064	0.051	0.076	-0.094	-0.063	0.199	0.090
0.000	0.000	0.063	0.001	-0.054	-0.001	-0.001	0.012	-0.051	-0.072	0.166	0.033
0.000	0.000	0.072	0.000	-0.043	0.054	0.037	-0.055	-0.084	-0.086	0.103	0.098
0.000	0.000	0.076	0.002	-0.052	-0.086	0.005	-0.049	0.019	-0.013	0.100	0.195
0.000	0.000	0.030	0.000	0.032	-0.015	-0.011	-0.016	0.019	0.012	0.093	0.017
0.000	0.000	0.052	0.005	-0.063	-0.041	-0.034	-0.090	0.073	0.006	0.081	-0.012
0.000	0.000	0.032	0.040	-0.027	-0.038	-0.060	-0.030	0.029	-0.020	0.007	0.012
0.000	0.000	0.012	0.045	0.026	0.012	-0.034	-0.074	-0.050	-0.024	0.001	0.048

TABLE C.7: Norms

Plus 5%

% Variation: 0.10%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.036	0.027	0.018	0.108	0.040	0.058	0.002	-0.004	-0.019	-0.009	0.056	-0.019
0.135	0.000	0.006	0.040	0.018	0.119	0.000	-0.005	-0.019	-0.058	0.039	0.065
0.000	0.000	0.014	-0.009	0.072	0.065	-0.022	0.023	0.012	-0.005	0.087	-0.004
0.000	0.000	0.049	0.015	0.064	0.127	-0.012	0.045	0.000	-0.037	0.025	-0.016
0.000	0.000	0.031	0.002	0.044	0.102	0.027	0.048	0.021	0.001	0.049	0.065
0.000	0.000	0.035	0.013	0.042	0.050	0.052	0.015	0.003	0.058	0.034	0.079
0.000	0.000	-0.011	-0.033	0.089	0.184	0.029	0.012	0.019	-0.072	0.049	0.044
0.000	0.000	-0.002	0.000	0.082	0.090	0.019	0.023	0.020	-0.042	0.008	0.038
0.000	0.000	-0.030	-0.015	0.065	0.023	0.019	-0.005	0.009	-0.003	0.012	0.050
0.000	0.000	-0.019	0.012	0.053	0.022	0.009	-0.017	0.020	-0.001	-0.054	-0.033

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.049	-0.018	0.009	-0.109	-0.002	-0.032	-0.066	-0.002	0.073	0.010	0.046	0.003
-0.044	-0.030	0.037	0.055	-0.040	-0.013	-0.055	-0.037	0.032	0.001	0.062	0.053
0.000	0.000	-0.034	-0.019	-0.003	-0.045	-0.008	-0.024	0.001	0.006	0.100	-0.039
0.000	0.000	-0.021	-0.013	-0.037	-0.057	-0.043	-0.058	0.014	0.001	0.102	-0.029
0.000	0.000	-0.012	0.000	0.026	-0.001	-0.030	-0.171	-0.042	0.051	0.049	0.008
0.000	0.000	-0.022	-0.002	0.009	-0.108	-0.043	-0.111	0.034	0.120	0.073	0.072
0.000	0.000	-0.016	-0.185	0.026	-0.031	-0.037	-0.079	-0.011	0.019	0.031	0.018
0.000	0.000	0.005	-0.057	-0.060	-0.037	-0.053	-0.112	-0.017	0.001	0.056	-0.049
0.000	0.000	-0.003	0.000	-0.019	-0.038	-0.034	-0.098	-0.019	-0.022	-0.011	-0.003
0.000	0.000	-0.043	-0.027	0.009	-0.040	-0.029	-0.102	-0.062	-0.014	-0.039	0.052

Minus 5%

% Variation: -0.81%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.045	0.027	0.055	0.123	-0.031	0.000	0.011	0.011	-0.084	-0.064	0.139	0.065
0.074	-0.101	0.038	0.044	-0.005	0.066	0.019	-0.003	-0.062	-0.060	0.038	0.068
0.000	0.000	0.053	0.021	0.018	0.044	0.039	0.021	-0.046	-0.056	0.077	0.003
0.000	0.000	0.065	0.046	0.022	0.051	0.024	0.070	-0.087	-0.071	0.041	-0.062
0.000	0.000	0.060	0.028	0.024	0.099	0.014	0.044	-0.074	-0.117	0.062	0.054
0.000	0.000	0.063	0.144	-0.018	0.021	0.033	0.029	-0.063	-0.171	0.064	0.072
0.000	0.000	0.059	0.032	0.060	0.172	0.020	0.037	-0.036	-0.177	0.106	0.030
0.000	0.000	0.038	0.011	0.044	0.066	0.023	0.048	-0.058	-0.132	0.032	0.010
0.000	0.000	0.032	0.066	0.079	0.032	-0.003	0.009	-0.028	-0.056	0.015	0.045
0.000	0.000	0.045	0.023	0.049	0.034	-0.006	0.006	-0.027	-0.075	-0.058	0.004

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.026	0.011	-0.030	0.000	-0.004	-0.047	0.018	0.047	0.019	0.003	0.106	0.092
0.053	0.040	-0.013	0.051	-0.075	-0.013	0.009	0.089	-0.022	0.001	0.085	0.178
0.000	0.000	-0.020	-0.001	-0.083	-0.118	0.055	0.113	-0.018	0.006	0.166	0.135
0.000	0.000	-0.015	0.020	-0.069	-0.065	-0.039	0.032	0.013	0.000	0.111	0.036
0.000	0.000	0.031	0.000	0.016	0.000	0.025	-0.068	0.017	0.020	0.118	0.115
0.000	0.000	0.014	0.002	0.027	-0.051	0.014	-0.016	0.069	0.038	0.120	0.225
0.000	0.000	0.009	-0.095	0.058	-0.011	0.015	-0.022	0.040	0.012	0.074	0.112
0.000	0.000	0.050	0.005	-0.028	-0.013	-0.035	-0.093	0.071	0.032	0.061	0.075
0.000	0.000	0.039	0.040	-0.014	0.001	-0.025	-0.055	0.046	0.010	-0.035	0.014
0.000	0.000	0.025	0.086	-0.017	0.002	-0.015	-0.040	-0.001	0.017	-0.023	0.034

TABLE C.8: Perceived Behavioural Control

Plus %5

% Variation: -1.10%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.018	-0.024	0.023	0.123	0.002	0.020	-0.018	0.083	-0.098	-0.162	0.010	0.026
0.076	0.000	0.021	0.045	0.007	0.068	0.011	0.062	-0.044	-0.090	0.059	0.001
0.000	0.000	0.013	0.000	0.032	0.083	-0.001	0.040	-0.036	-0.032	-0.010	-0.006
0.000	0.000	0.035	0.014	0.035	0.124	-0.026	0.056	-0.060	-0.023	-0.052	-0.063
0.000	0.000	0.027	0.015	0.043	0.125	0.024	0.071	-0.062	-0.069	0.023	0.055
0.000	0.000	0.050	0.083	0.017	0.090	0.028	0.056	-0.072	-0.097	0.002	0.044
0.000	0.000	0.016	-0.062	0.032	0.155	0.030	0.128	-0.078	-0.099	0.010	0.028
0.000	0.000	0.022	0.000	0.072	0.046	0.031	0.078	-0.050	-0.110	-0.022	-0.008
0.000	0.000	-0.009	-0.015	0.087	0.027	0.029	0.061	-0.025	0.029	-0.054	-0.042
0.000	0.000	-0.027	-0.012	0.051	0.073	0.018	0.029	-0.012	0.006	-0.099	-0.088

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.012	-0.018	0.137	0.116	-0.018	-0.013	0.016	0.015	-0.002	0.013	0.029	-0.024
-0.050	0.014	0.140	0.194	-0.017	0.024	0.011	0.076	-0.064	-0.034	0.031	0.001
0.000	0.000	0.092	0.112	-0.023	0.020	0.040	0.001	-0.060	-0.058	0.116	0.038
0.000	0.000	0.106	0.173	-0.052	0.012	0.019	0.008	0.000	-0.072	0.093	0.001
0.000	0.000	0.087	0.038	0.026	0.058	-0.014	-0.029	-0.027	-0.057	0.106	0.105
0.000	0.000	0.070	0.086	-0.018	-0.004	0.029	0.017	0.017	-0.013	0.165	0.199
0.000	0.000	0.064	0.000	0.061	-0.002	0.006	0.021	-0.015	-0.084	0.087	0.129
0.000	0.000	0.076	0.005	0.013	0.031	0.013	-0.038	0.046	-0.048	0.070	0.036
0.000	0.000	0.080	0.067	0.024	0.020	0.016	0.019	-0.001	-0.030	0.019	0.048
0.000	0.000	0.045	0.082	0.061	0.023	0.013	-0.056	0.025	0.053	0.001	0.069

Plus %5

% Variation: -1.10%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.018	-0.024	0.023	0.123	0.002	0.020	-0.018	0.083	-0.098	-0.162	0.010	0.026
0.076	0.000	0.021	0.045	0.007	0.068	0.011	0.062	-0.044	-0.090	0.059	0.001
0.000	0.000	0.013	0.000	0.032	0.083	-0.001	0.040	-0.036	-0.032	-0.010	-0.006
0.000	0.000	0.035	0.014	0.035	0.124	-0.026	0.056	-0.060	-0.023	-0.052	-0.063
0.000	0.000	0.027	0.015	0.043	0.125	0.024	0.071	-0.062	-0.069	0.023	0.055
0.000	0.000	0.050	0.083	0.017	0.090	0.028	0.056	-0.072	-0.097	0.002	0.044
0.000	0.000	0.016	-0.062	0.032	0.155	0.030	0.128	-0.078	-0.099	0.010	0.028
0.000	0.000	0.022	0.000	0.072	0.046	0.031	0.078	-0.050	-0.110	-0.022	-0.008
0.000	0.000	-0.009	-0.015	0.087	0.027	0.029	0.061	-0.025	0.029	-0.054	-0.042
0.000	0.000	-0.027	-0.012	0.051	0.073	0.018	0.029	-0.012	0.006	-0.099	-0.088

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.012	-0.018	0.137	0.116	-0.018	-0.013	0.016	0.015	-0.002	0.013	0.029	-0.024
-0.050	0.014	0.140	0.194	-0.017	0.024	0.011	0.076	-0.064	-0.034	0.031	0.001
0.000	0.000	0.092	0.112	-0.023	0.020	0.040	0.001	-0.060	-0.058	0.116	0.038
0.000	0.000	0.106	0.173	-0.052	0.012	0.019	0.008	0.000	-0.072	0.093	0.001
0.000	0.000	0.087	0.038	0.026	0.058	-0.014	-0.029	-0.027	-0.057	0.106	0.105
0.000	0.000	0.070	0.086	-0.018	-0.004	0.029	0.017	0.017	-0.013	0.165	0.199
0.000	0.000	0.064	0.000	0.061	-0.002	0.006	0.021	-0.015	-0.084	0.087	0.129
0.000	0.000	0.076	0.005	0.013	0.031	0.013	-0.038	0.046	-0.048	0.070	0.036
0.000	0.000	0.080	0.067	0.024	0.020	0.016	0.019	-0.001	-0.030	0.019	0.048
0.000	0.000	0.045	0.082	0.061	0.023	0.013	-0.056	0.025	0.053	0.001	0.069

TABLE C.9: Range

Plus 5%

% Variation: -0.81%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.008	-0.012	0.054	0.138	-0.036	0.030	0.016	0.055	-0.068	-0.034	-0.052	-0.055
0.090	0.007	-0.019	0.015	-0.035	0.109	0.025	0.080	-0.030	-0.071	0.009	-0.060
0.000	0.000	0.010	-0.038	0.000	0.047	0.011	0.062	-0.030	-0.011	-0.032	-0.009
0.000	0.000	0.018	0.006	0.005	0.117	0.002	0.072	-0.016	-0.008	-0.052	-0.065
0.000	0.000	-0.002	0.002	-0.034	0.010	0.017	0.055	0.003	-0.016	0.019	0.021
0.000	0.000	0.011	0.068	-0.040	0.018	0.010	0.005	-0.020	-0.063	0.014	0.013
0.000	0.000	0.015	-0.033	0.025	0.108	-0.017	0.042	-0.029	-0.081	0.035	-0.005
0.000	0.000	0.007	0.017	0.006	0.012	-0.002	0.059	0.001	-0.008	-0.016	-0.032
0.000	0.000	-0.006	-0.003	0.037	0.024	-0.004	0.025	-0.006	0.060	-0.050	-0.056
0.000	0.000	0.012	0.000	0.027	-0.006	-0.020	-0.022	0.042	0.032	-0.074	-0.053

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.046	0.000	0.085	0.016	0.027	0.047	0.004	0.031	0.044	0.132	0.068	0.002
-0.014	-0.008	0.102	0.160	-0.051	0.048	0.043	0.160	-0.021	0.017	0.148	0.135
0.000	0.000	0.027	0.018	-0.064	-0.007	0.042	0.002	-0.033	-0.019	0.174	0.177
0.000	0.000	0.029	0.035	-0.052	0.015	0.020	0.016	-0.012	-0.001	0.123	0.065
0.000	0.000	0.034	0.017	-0.009	0.060	-0.020	-0.021	-0.006	0.006	0.163	0.132
0.000	0.000	0.048	0.010	0.000	0.039	0.022	-0.004	0.057	0.058	0.131	0.271
0.000	0.000	0.002	-0.067	0.026	0.049	0.032	0.006	0.013	0.002	0.095	0.109
0.000	0.000	0.005	-0.034	-0.015	0.060	-0.007	-0.048	0.068	-0.019	0.110	0.018
0.000	0.000	0.005	0.000	0.039	0.030	-0.016	-0.015	-0.004	-0.019	0.000	0.048
0.000	0.000	-0.001	0.080	0.012	0.019	0.060	0.013	-0.037	-0.046	0.073	0.071

Minus 5%

%V ariation: -1.39%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.057	0.082	0.047	0.123	0.044	0.073	-0.029	-0.005	-0.064	-0.012	0.072	0.059
0.107	0.007	0.055	0.045	0.039	0.113	0.000	0.028	-0.009	-0.035	0.034	0.024
0.000	0.000	0.049	0.060	0.057	0.041	-0.002	0.042	-0.034	-0.024	0.063	0.001
0.000	0.000	0.084	0.076	0.055	0.086	-0.006	0.067	-0.046	-0.055	0.077	-0.061
0.000	0.000	0.073	0.052	0.012	0.102	0.042	0.049	-0.089	-0.081	0.103	0.035
0.000	0.000	0.077	0.147	0.002	0.004	0.058	0.012	-0.100	-0.095	0.080	0.029
0.000	0.000	0.050	0.053	0.044	0.186	0.016	0.018	-0.089	-0.173	0.083	0.019
0.000	0.000	0.032	0.040	0.062	0.116	0.036	0.027	-0.076	-0.056	0.090	0.007
0.000	0.000	0.015	0.038	0.072	0.056	-0.003	-0.013	-0.042	0.000	0.068	0.010
0.000	0.000	0.014	0.034	0.082	0.074	0.025	0.008	-0.031	-0.005	0.037	-0.067

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.051	0.000	0.082	0.078	-0.053	-0.010	-0.012	-0.010	0.034	0.007	0.031	0.006
-0.021	0.014	0.098	0.128	-0.058	-0.050	-0.032	0.070	-0.006	0.036	0.074	0.098
0.000	0.000	0.044	-0.001	-0.066	-0.111	0.009	0.040	-0.017	0.006	0.158	0.064
0.000	0.000	0.065	0.008	-0.039	-0.053	-0.028	0.012	-0.029	0.021	0.121	0.035
0.000	0.000	0.052	0.007	-0.010	-0.005	-0.017	-0.019	-0.002	-0.027	0.079	0.093
0.000	0.000	0.036	0.017	-0.023	-0.030	-0.013	0.031	0.018	0.067	0.087	0.166
0.000	0.000	0.006	-0.066	0.032	-0.009	0.013	0.019	-0.008	0.001	0.070	0.075
0.000	0.000	-0.052	0.000	0.008	0.018	-0.038	-0.027	0.041	0.001	0.107	0.012
0.000	0.000	-0.022	0.014	0.029	0.005	-0.049	0.006	0.019	0.008	0.014	0.012
0.000	0.000	-0.027	0.059	0.083	0.066	0.006	-0.007	-0.051	-0.006	0.018	0.071

TABLE C.10: Satisficing Number

Plus 5%.

% Variation: -1.03%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.022	0.000	0.063	0.106	-0.022	-0.007	-0.008	0.015	-0.056	-0.057	-0.010	-0.055
0.025	0.007	0.046	0.040	0.001	0.066	0.037	0.017	-0.040	-0.059	-0.014	-0.002
0.000	0.000	0.062	-0.023	0.013	0.058	-0.003	0.024	-0.020	-0.048	0.055	-0.002
0.000	0.000	0.085	0.014	0.030	0.110	-0.003	0.067	-0.060	-0.045	0.044	0.050
0.000	0.000	0.052	0.027	0.037	0.089	0.024	0.088	-0.066	-0.099	0.110	0.121
0.000	0.000	0.059	0.057	0.049	0.065	0.017	0.043	-0.047	-0.091	0.072	0.080
0.000	0.000	0.033	-0.025	0.063	0.178	0.023	0.069	-0.014	-0.069	0.018	0.004
0.000	0.000	0.012	0.000	0.071	0.098	0.042	0.074	-0.041	-0.074	-0.008	-0.010
0.000	0.000	0.009	-0.003	0.047	0.091	0.043	0.063	-0.062	-0.044	-0.055	-0.015
0.000	0.000	0.027	-0.012	0.063	0.064	0.027	0.047	-0.047	-0.062	-0.033	-0.074

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.028	0.000	0.061	0.000	0.026	0.014	0.007	0.014	-0.048	-0.009	0.149	0.099
0.023	0.000	0.055	0.055	-0.002	-0.012	-0.008	0.076	-0.078	-0.006	0.139	0.142
0.000	0.000	0.026	-0.012	0.010	-0.023	0.042	0.001	-0.068	-0.016	0.246	0.295
0.000	0.000	0.044	-0.008	0.000	0.012	-0.015	0.007	-0.054	-0.090	0.285	0.279
0.000	0.000	0.017	-0.011	0.066	-0.004	0.002	-0.047	-0.036	0.000	0.237	0.301
0.000	0.000	0.004	-0.010	0.058	0.038	-0.003	0.004	0.016	0.004	0.321	0.438
0.000	0.000	0.019	-0.110	-0.002	-0.005	0.031	0.016	-0.013	-0.045	0.229	0.232
0.000	0.000	0.053	0.003	-0.055	-0.005	-0.043	-0.056	-0.013	-0.051	0.250	0.149
0.000	0.000	0.066	0.014	-0.066	-0.033	-0.046	0.013	-0.014	-0.017	0.041	0.098
0.000	0.000	0.051	0.077	-0.040	-0.005	0.010	0.017	-0.034	-0.044	0.018	0.071

Minus 5%

% Variation: -2.71%

M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.061	0.186	0.035	0.058	-0.070	-0.030	0.044	0.102	-0.039	-0.064	-0.004	-0.064
0.131	0.000	0.049	0.044	-0.081	-0.014	0.028	0.112	-0.024	-0.058	-0.027	0.031
0.000	0.000	0.054	0.036	-0.044	-0.058	0.010	0.076	0.013	-0.027	-0.007	-0.003
0.000	0.000	0.075	0.044	-0.020	-0.004	-0.017	0.072	-0.033	-0.041	-0.060	-0.066
0.000	0.000	0.090	0.043	-0.082	-0.172	0.032	0.102	-0.054	-0.005	-0.037	0.003
0.000	0.000	0.078	0.164	-0.070	-0.039	0.038	0.046	-0.053	-0.054	-0.009	-0.026
0.000	0.000	0.062	0.056	-0.024	0.050	0.000	0.024	-0.039	-0.042	-0.003	-0.025
0.000	0.000	0.057	0.068	0.005	0.013	-0.015	0.003	-0.025	0.000	-0.006	-0.035
0.000	0.000	0.045	0.087	0.044	0.009	-0.028	-0.020	0.007	0.083	-0.059	-0.113
0.000	0.000	0.031	0.046	0.034	0.012	-0.039	-0.027	0.011	0.034	-0.091	-0.127

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.005	0.022	0.043	0.010	-0.019	-0.010	0.013	0.031	-0.012	0.003	0.024	-0.021
0.047	0.020	0.107	0.166	-0.013	-0.015	-0.038	-0.037	-0.068	-0.004	0.075	-0.020
0.000	0.000	0.042	-0.002	-0.002	-0.042	0.027	0.003	-0.058	-0.058	0.148	-0.045
0.000	0.000	0.077	-0.006	0.003	0.012	-0.015	-0.063	-0.020	0.000	0.094	-0.054
0.000	0.000	0.075	-0.005	-0.004	0.000	-0.045	-0.124	-0.036	0.020	0.002	0.040
0.000	0.000	0.048	-0.010	0.020	-0.004	-0.027	-0.084	0.010	0.081	0.005	0.072
0.000	0.000	0.033	-0.103	0.046	-0.009	0.017	0.000	0.006	0.029	0.000	-0.094
0.000	0.000	0.024	0.005	-0.007	0.020	-0.049	-0.049	0.020	-0.007	0.053	-0.057
0.000	0.000	0.053	0.022	0.022	-0.007	-0.064	-0.029	0.027	0.003	-0.049	0.048
0.000	0.000	0.044	0.106	0.029	-0.006	-0.030	-0.062	-0.029	0.038	-0.020	0.062

TABLE C.11: Trigger

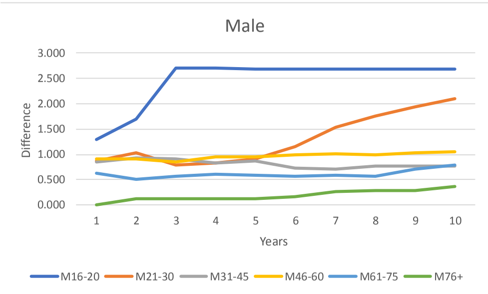
C.2 Scenarios

Scenario 1 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.635	21.337	25.070	24.200	27.593	26.741	28.523	27.619	28.858	28.226	28.009	28.009
2015	22.723	21.017	24.809	23.784	27.425	26.500	28.823	27.918	29.053	28.537	28.112	27.993
2016	23.722	21.017	24.314	23.529	27.506	26.600	28.851	27.993	29.173	28.610	28.375	28.263
2017	23.787	21.089	24.448	23.622	27.425	26.599	29.016	28.072	29.216	28.612	28.615	28.494
2018	23.852	21.161	24.289	23.382	27.496	26.630	29.065	28.123	29.441	28.850	28.727	28.595
2019	23.852	21.161	24.301	23.141	27.353	26.623	29.136	28.154	29.736	29.179	28.870	28.700
2020	23.852	21.161	24.162	22.636	27.307	26.604	29.237	28.220	29.812	29.223	29.135	28.867
2021	23.852	21.161	24.155	22.396	27.272	26.502	29.337	28.340	29.883	29.314	29.253	28.976
2022	23.852	21.161	24.336	22.396	27.111	26.342	29.366	28.328	30.010	29.296	29.578	29.288
2023	23.852	21.161	24.304	22.202	27.042	26.282	29.492	28.435	30.121	29.330	29.868	29.497

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.717	22.709	25.498	24.050	26.873	25.598	28.551	26.991	28.730	28.022	28.280	27.693
23.723	22.760	25.337	23.995	27.059	25.757	28.578	27.052	28.835	28.077	28.672	28.188
22.603	22.014	25.239	23.941	27.248	25.808	28.722	26.990	28.997	28.235	28.936	28.513
22.693	22.115	25.506	24.183	27.063	25.718	28.939	27.126	29.240	28.486	29.049	28.559
22.783	22.217	25.482	24.261	27.209	25.812	29.122	27.524	29.240	28.429	29.314	28.682
22.783	22.217	25.422	24.351	27.159	25.751	29.189	27.452	29.454	28.755	29.414	28.844
22.783	22.217	25.356	24.170	27.216	25.765	29.343	27.629	29.551	28.766	29.534	28.939
22.783	22.217	25.104	23.950	27.345	25.874	29.447	27.756	29.640	28.834	29.754	29.123
22.783	22.217	24.966	23.700	27.559	26.129	29.424	27.566	29.814	28.939	29.887	29.142
22.783	22.217	24.700	23.580	27.692	26.124	29.410	27.665	29.960	28.850	30.210	29.448

Mean v's Median Differences:



M16-20	M21-30	M31-45	M46-60	M61-75	M76+
1.298	0.870	0.852	0.903	0.632	0.000
1.706	1.025	0.925	0.904	0.515	0.119
2.705	0.786	0.905	0.858	0.563	0.112
2.698	0.826	0.826	0.945	0.604	0.121
2.692	0.908	0.866	0.942	0.591	0.132
2.692	1.160	0.730	0.982	0.556	0.169
2.692	1.527	0.702	1.017	0.589	0.268
2.692	1.759	0.770	0.996	0.568	0.278
2.692	1.940	0.769	1.038	0.714	0.290
2.692	2.102	0.760	1.057	0.791	0.371

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
1.008	1.448	1.275	1.560	0.708	0.586
0.963	1.342	1.302	1.525	0.757	0.485
0.589	1.298	1.440	1.732	0.762	0.423
0.578	1.323	1.345	1.813	0.754	0.490
0.566	1.221	1.397	1.597	0.810	0.632
0.566	1.071	1.408	1.737	0.700	0.570
0.566	1.186	1.450	1.714	0.785	0.595
0.566	1.154	1.471	1.691	0.806	0.631
0.566	1.266	1.430	1.858	0.875	0.745
0.566	1.120	1.568	1.745	1.110	0.762

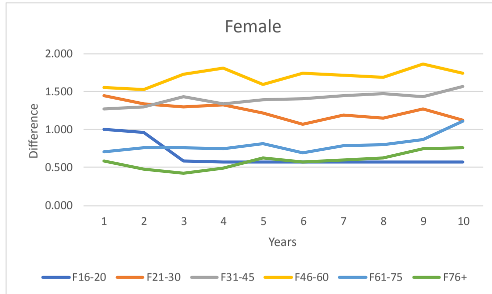


TABLE C.12: Scenario 1 - Raw Forecast

Scenario 1: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.992	21.828	26.071	25.336	27.572	26.316	28.869	27.842	28.916	28.485	27.408	27.507
2015	23.397	22.572	25.413	24.651	26.831	25.536	28.608	27.627	28.624	28.328	26.956	27.213
2016	24.250	21.652	25.797	25.187	27.342	26.124	28.988	27.721	28.917	28.345	27.272	27.329
2017	24.977	22.488	25.921	25.341	27.291	26.008	29.113	27.910	29.201	28.562	27.568	27.563
2018	24.577	20.880	25.295	24.307	27.199	26.089	29.238	28.056	29.137	28.349	27.634	27.654
2019	24.139	20.663	25.764	24.819	27.337	26.608	29.071	27.917	29.603	28.792	27.795	27.847
2020	24.631	21.261	25.549	24.441	27.066	26.470	29.153	27.906	29.342	28.535	27.894	28.114
2021	24.298	20.439	26.409	24.659	27.602	26.929	29.235	27.941	29.599	28.671	28.346	28.396
2022	24.205	20.683	26.264	24.578	27.421	26.684	29.307	27.694	29.566	28.524	28.573	28.553
2023	25.335	21.121	26.392	24.531	27.847	26.974	29.213	27.697	29.476	28.363	28.466	28.217

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	24.436	23.452	26.483	24.859	26.813	25.678	28.266	26.732	28.627	28.290	28.374	28.114
2015	25.668	24.136	25.545	24.178	27.127	26.085	28.247	26.633	28.957	28.662	27.965	27.725
2016	24.776	23.169	25.804	24.504	27.218	25.724	28.346	26.581	28.742	28.395	28.251	28.018
2017	24.166	23.163	25.822	24.592	26.697	25.514	28.277	26.539	29.035	28.408	28.152	28.072
2018	23.362	22.646	26.240	25.043	26.631	25.159	28.407	27.007	28.666	28.226	27.856	27.303
2019	23.258	22.643	25.867	24.747	27.073	25.648	28.553	27.067	28.865	28.329	28.064	27.645
2020	23.687	23.752	26.481	24.899	26.998	25.409	28.784	27.180	28.996	28.020	27.861	27.681
2021	22.851	22.625	26.431	24.996	27.468	26.075	28.822	27.250	29.057	28.393	28.362	27.812
2022	23.585	22.754	25.933	24.273	27.775	26.220	28.724	27.029	28.703	27.744	28.584	28.253
2023	23.191	22.842	26.225	24.488	27.570	25.880	28.601	26.968	28.491	27.265	28.115	27.716

Adjusted Scenario 1 v's Comparator Data(2013):

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.400	-0.493	0.177	0.145	-0.008	-0.605	0.128	-0.339	0.367	0.501	-0.399	0.039
2015	0.005	0.251	-0.481	-0.541	-0.749	-1.384	-0.132	-0.554	0.075	0.344	-0.850	-0.256
2016	0.858	-0.668	-0.097	-0.005	-0.238	-0.797	0.247	-0.460	0.367	0.360	-0.534	-0.140
2017	1.585	0.168	0.027	0.149	-0.289	-0.913	0.372	-0.272	0.652	0.577	-0.238	0.094
2018	1.185	-1.440	-0.600	-0.885	-0.381	-0.832	0.497	-0.125	0.587	0.365	-0.172	0.186
2019	0.748	-1.658	-0.130	-0.373	-0.243	-0.313	0.330	-0.264	1.053	0.808	-0.011	0.378
2020	1.240	-1.059	-0.345	-0.751	-0.514	-0.451	0.412	-0.275	0.792	0.550	0.088	0.645
2021	0.906	-1.882	0.514	-0.533	0.021	0.008	0.495	-0.241	1.049	0.686	0.540	0.927
2022	0.813	-1.637	0.370	-0.614	-0.160	-0.237	0.566	-0.488	1.016	0.540	0.767	1.085
2023	1.943	-1.200	0.497	-0.661	0.266	0.053	0.472	-0.484	0.926	0.378	0.659	0.748

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.629	0.716	0.630	0.170	-0.326	-0.043	0.445	0.048	0.378	1.140	1.136	1.075
2015	1.861	1.401	-0.309	-0.511	-0.012	0.365	0.426	-0.051	0.708	1.513	0.727	0.686
2016	0.969	0.433	-0.049	-0.185	0.079	0.004	0.524	-0.103	0.493	1.245	1.014	0.979
2017	0.359	0.427	-0.032	-0.097	-0.442	-0.206	0.456	-0.145	0.786	1.258	0.914	1.033
2018	-0.445	-0.090	0.387	0.355	-0.508	-0.562	0.586	0.323	0.417	1.076	0.619	0.264
2019	-0.549	-0.093	0.013	0.058	-0.066	-0.072	0.732	0.383	0.616	1.180	0.826	0.606
2020	-0.121	1.016	0.628	0.210	-0.142	-0.311	0.963	0.495	0.747	0.871	0.623	0.642
2021	-0.956	-0.111	0.578	0.307	0.329	0.355	1.001	0.565	0.808	1.243	1.125	0.773
2022	-0.222	0.018	0.080	-0.416	0.636	0.500	0.903	0.344	0.454	0.595	1.347	1.214
2023	-0.616	0.107	0.372	-0.201	0.431	0.160	0.780	0.284	0.242	0.115	0.878	0.677

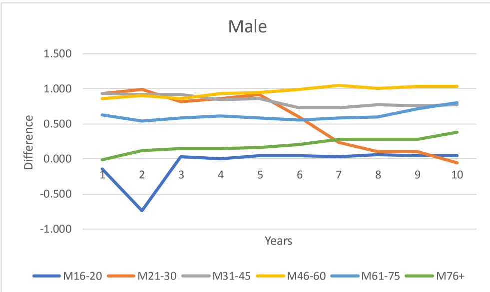
TABLE C.13: Scenario 1 - Adjusted Forecast and Comparison

Scenario 2 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.502	22.639	25.065	24.134	27.559	26.620	28.505	27.647	28.855	28.220	28.021	28.039
2015	22.445	23.179	24.766	23.770	27.403	26.482	28.816	27.918	29.029	28.490	28.132	28.011
2016	23.411	23.373	24.283	23.463	27.522	26.604	28.858	28.003	29.155	28.568	28.411	28.257
2017	23.519	23.513	24.406	23.540	27.418	26.574	29.030	28.102	29.189	28.575	28.633	28.488
2018	23.626	23.583	24.283	23.367	27.469	26.615	29.076	28.124	29.408	28.830	28.740	28.577
2019	23.719	23.676	24.242	23.648	27.337	26.604	29.148	28.154	29.705	29.145	28.891	28.689
2020	23.803	23.778	24.057	23.817	27.309	26.577	29.256	28.211	29.789	29.204	29.133	28.848
2021	23.892	23.837	24.044	23.940	27.270	26.499	29.349	28.342	29.871	29.274	29.237	28.958
2022	23.973	23.924	24.172	24.073	27.090	26.327	29.362	28.332	29.985	29.273	29.539	29.256
2023	24.093	24.051	24.134	24.187	27.017	26.248	29.484	28.452	30.117	29.316	29.847	29.472

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.899	23.200	25.486	23.987	26.875	25.628	28.577	27.062	28.701	28.022	28.212	27.582
24.084	23.284	25.326	23.982	27.051	25.780	28.611	27.094	28.784	28.076	28.644	28.122
23.440	23.446	25.250	23.928	27.262	25.861	28.730	27.043	28.972	28.233	28.885	28.408
23.584	23.560	25.494	24.144	27.057	25.713	28.978	27.139	29.202	28.468	28.975	28.322
23.714	23.683	25.437	24.236	27.197	25.794	29.145	27.558	29.208	28.408	29.263	28.607
23.812	23.803	25.479	24.195	27.142	25.718	29.217	27.486	29.448	28.752	29.368	28.801
23.909	23.924	25.446	24.129	27.211	25.719	29.363	27.653	29.548	28.761	29.490	28.924
24.021	24.033	25.228	24.243	27.340	25.827	29.453	27.765	29.629	28.810	29.699	29.099
24.143	24.139	25.220	24.381	27.532	26.082	29.435	27.586	29.792	28.932	29.868	29.167
24.230	24.235	25.069	24.482	27.689	26.085	29.430	27.665	29.937	28.919	30.171	29.435

Mean v's Median Differences:



M16-20	M21-30	M31-45	M46-60	M61-75	M76+
-0.137	0.931	0.939	0.858	0.635	-0.018
-0.734	0.996	0.921	0.898	0.539	0.121
0.038	0.820	0.918	0.856	0.587	0.154
0.005	0.866	0.844	0.929	0.615	0.144
0.043	0.917	0.854	0.952	0.578	0.163
0.044	0.594	0.733	0.993	0.560	0.202
0.026	0.241	0.732	1.045	0.585	0.285
0.054	0.105	0.771	1.007	0.597	0.279
0.049	0.099	0.763	1.030	0.712	0.282
0.042	-0.054	0.769	1.032	0.801	0.375

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
0.699	1.499	1.248	1.514	0.679	0.630
0.801	1.344	1.270	1.517	0.708	0.522
-0.006	1.322	1.401	1.687	0.739	0.477
0.025	1.350	1.344	1.839	0.734	0.653
0.032	1.201	1.403	1.588	0.800	0.656
0.009	1.284	1.424	1.732	0.696	0.566
-0.015	1.317	1.492	1.710	0.787	0.566
-0.013	0.986	1.513	1.688	0.819	0.600
0.005	0.838	1.450	1.849	0.860	0.701
-0.004	0.587	1.604	1.765	1.019	0.736

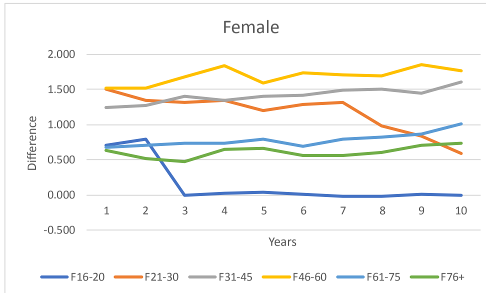


TABLE C.14: Scenario 2 - Raw Forecast

Scenario 2: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.859	23.130	26.066	25.270	27.538	26.195	28.852	27.870	28.913	28.479	27.420	27.537
2015	23.119	24.734	25.371	24.637	26.809	25.519	28.601	27.627	28.600	28.280	26.976	27.231
2016	23.938	24.008	25.766	25.121	27.358	26.127	28.995	27.731	28.899	28.303	27.308	27.323
2017	24.708	24.913	25.880	25.259	27.284	25.982	29.127	27.939	29.175	28.524	27.586	27.557
2018	24.350	23.303	25.289	24.292	27.172	26.074	29.249	28.057	29.104	28.329	27.648	27.636
2019	24.006	23.178	25.705	25.326	27.322	26.590	29.082	27.917	29.572	28.758	27.816	27.835
2020	24.582	23.878	25.444	25.622	27.068	26.443	29.172	27.897	29.319	28.515	27.892	28.095
2021	24.337	23.115	26.298	26.202	27.600	26.925	29.248	27.943	29.588	28.631	28.330	28.378
2022	24.325	23.446	26.101	26.255	27.400	26.669	29.304	27.698	29.542	28.502	28.534	28.521
2023	25.575	24.010	26.222	26.516	27.821	26.940	29.204	27.714	29.471	28.349	28.445	28.191

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	24.618	23.942	26.471	24.796	26.816	25.707	28.292	26.804	28.598	28.289	28.306	28.003
2015	26.029	24.660	25.533	24.165	27.119	26.108	28.281	26.675	28.906	28.661	27.937	27.659
2016	25.613	24.600	25.815	24.490	27.232	25.777	28.353	26.634	28.717	28.393	28.201	27.914
2017	25.057	24.607	25.810	24.553	26.691	25.509	28.316	26.552	28.996	28.390	28.078	27.835
2018	24.294	24.112	26.196	25.019	26.620	25.141	28.431	27.041	28.634	28.204	27.805	27.227
2019	24.288	24.230	25.924	24.591	27.057	25.616	28.581	27.100	28.859	28.327	28.017	27.603
2020	24.813	25.460	26.572	24.858	26.993	25.362	28.804	27.203	28.993	28.015	27.817	27.665
2021	24.089	24.441	26.556	25.289	27.463	26.028	28.827	27.259	29.046	28.369	28.307	27.789
2022	24.946	24.676	26.186	24.954	27.748	26.173	28.735	27.049	28.681	27.737	28.566	28.278
2023	24.639	24.861	26.594	25.390	27.567	25.841	28.622	26.968	28.468	27.334	28.076	27.704

Scenario 2 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.133	1.302	-0.006	-0.066	-0.034	-0.121	-0.017	0.028	-0.003	-0.006	0.012	0.030
2015	-0.278	2.162	-0.043	-0.014	-0.022	-0.018	-0.007	-0.001	-0.024	-0.048	0.020	0.018
2016	-0.311	2.356	-0.031	-0.066	0.016	0.003	0.007	0.010	-0.018	-0.041	0.036	-0.006
2017	-0.269	2.424	-0.042	-0.082	-0.007	-0.025	0.014	0.030	-0.027	-0.037	0.018	-0.006
2018	-0.227	2.422	-0.006	-0.015	-0.027	-0.015	0.011	0.001	-0.033	-0.021	0.013	-0.018
2019	-0.133	2.515	-0.060	0.507	-0.015	-0.019	0.011	0.000	-0.031	-0.034	0.021	-0.012
2020	-0.049	2.617	-0.105	1.181	0.003	-0.027	0.019	-0.010	-0.023	-0.019	-0.002	-0.019
2021	0.039	2.676	-0.111	1.543	-0.002	-0.003	0.012	0.002	-0.011	-0.040	-0.016	-0.017
2022	0.121	2.763	-0.164	1.677	-0.021	-0.015	-0.004	0.005	-0.024	-0.022	-0.040	-0.032
2023	0.240	2.890	-0.170	1.986	-0.026	-0.034	-0.008	0.017	-0.004	-0.014	-0.021	-0.025

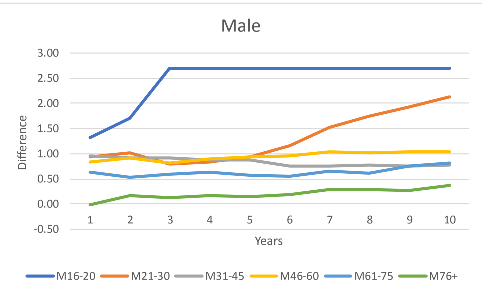
	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.182	0.491	-0.012	-0.063	0.002	0.030	0.025	0.071	-0.029	0.000	-0.068	-0.111
2015	0.361	0.524	-0.011	-0.013	-0.009	0.023	0.034	0.042	-0.051	-0.001	-0.028	-0.066
2016	0.837	1.432	0.011	-0.013	0.014	0.053	0.008	0.053	-0.025	-0.002	-0.051	-0.105
2017	0.891	1.444	-0.011	-0.038	-0.006	-0.005	0.039	0.013	-0.038	-0.018	-0.074	-0.237
2018	0.932	1.466	-0.044	-0.025	-0.012	-0.018	0.024	0.033	-0.032	-0.022	-0.051	-0.076
2019	1.030	1.587	0.057	-0.156	-0.016	-0.032	0.028	0.033	-0.006	-0.003	-0.046	-0.042
2020	1.126	1.707	0.090	-0.040	-0.004	-0.046	0.020	0.024	-0.003	-0.005	-0.044	-0.015
2021	1.238	1.817	0.125	0.293	-0.005	-0.046	0.005	0.009	-0.011	-0.024	-0.055	-0.024
2022	1.361	1.922	0.253	0.681	-0.027	-0.047	0.011	0.020	-0.022	-0.007	-0.019	0.025
2023	1.448	2.018	0.369	0.903	-0.003	-0.039	0.021	0.000	-0.023	0.069	-0.039	-0.013

TABLE C.15: Scenario 2 - Adjusted Forecast and Comparison

Scenario 3 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.66	21.34	25.07	24.13	27.57	26.62	28.55	27.72	28.87	28.24	28.03	28.03
2015	22.72	21.02	24.78	23.77	27.41	26.50	28.84	27.92	29.05	28.52	28.17	27.99
2016	23.72	21.02	24.32	23.52	27.50	26.58	28.88	28.06	29.19	28.59	28.41	28.29
2017	23.79	21.09	24.46	23.62	27.42	26.54	29.04	28.15	29.20	28.57	28.66	28.49
2018	23.85	21.16	24.32	23.38	27.49	26.62	29.10	28.16	29.44	28.86	28.74	28.59
2019	23.85	21.16	24.28	23.13	27.36	26.61	29.17	28.20	29.73	29.17	28.89	28.71
2020	23.85	21.16	24.13	22.61	27.32	26.57	29.29	28.25	29.81	29.16	29.16	28.87
2021	23.85	21.15	24.14	22.38	27.27	26.50	29.37	28.36	29.88	29.28	29.26	28.98
2022	23.85	21.15	24.32	22.39	27.10	26.35	29.38	28.34	30.02	29.27	29.55	29.27
2023	23.85	21.15	24.30	22.17	27.03	26.26	29.49	28.45	30.13	29.31	29.85	29.49

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.74	22.63	25.46	24.04	26.86	25.60	28.53	27.02	28.69	28.02	28.26	27.68
23.72	22.71	25.31	24.00	27.05	25.73	28.58	27.05	28.81	28.11	28.67	28.14
22.60	22.01	25.24	23.95	27.24	25.82	28.71	26.99	28.98	28.26	28.89	28.46
22.69	22.11	25.48	24.14	27.05	25.71	28.94	27.13	29.21	28.50	29.00	28.47
22.78	22.21	25.45	24.25	27.17	25.78	29.12	27.46	29.22	28.43	29.29	28.63
22.78	22.21	25.42	24.34	27.12	25.69	29.17	27.47	29.44	28.71	29.40	28.79
22.78	22.21	25.37	24.14	27.19	25.75	29.31	27.62	29.56	28.75	29.49	28.89
22.78	22.21	25.12	23.88	27.34	25.85	29.41	27.74	29.63	28.80	29.72	29.08
22.78	22.21	24.96	23.68	27.55	26.10	29.39	27.55	29.78	28.89	29.86	29.13
22.78	22.21	24.69	23.56	27.68	26.07	29.39	27.65	29.95	28.73	30.18	29.39



Mean v's Median Differences:

M16-20	M21-30	M31-45	M46-60	M61-75	M76+
1.33	0.94	0.95	0.83	0.63	-0.01
1.70	1.01	0.91	0.91	0.53	0.17
2.71	0.80	0.92	0.81	0.60	0.12
2.70	0.83	0.88	0.90	0.63	0.17
2.69	0.95	0.87	0.94	0.58	0.16
2.69	1.16	0.74	0.97	0.55	0.18
2.69	1.51	0.75	1.04	0.64	0.29
2.69	1.75	0.77	1.01	0.61	0.28
2.70	1.93	0.75	1.04	0.75	0.28
2.70	2.13	0.77	1.04	0.82	0.37

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
1.11	1.42	1.27	1.52	0.67	0.58
1.01	1.31	1.31	1.52	0.70	0.53
0.59	1.28	1.42	1.72	0.72	0.43
0.58	1.34	1.33	1.81	0.71	0.54
0.57	1.20	1.40	1.66	0.79	0.66
0.57	1.08	1.43	1.71	0.73	0.62
0.57	1.23	1.44	1.69	0.81	0.60
0.57	1.23	1.49	1.66	0.83	0.63
0.57	1.28	1.45	1.84	0.88	0.74
0.57	1.13	1.62	1.74	1.21	0.79

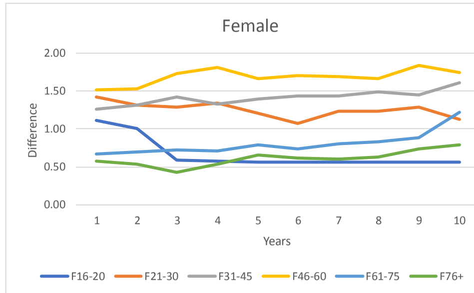


TABLE C.16: Scenario 3 - Raw Forecast

Scenario 3: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	23.02	21.83	26.07	25.27	27.55	26.20	28.89	27.94	28.93	28.50	27.43	27.53
2015	23.39	22.57	25.39	24.64	26.82	25.54	28.62	27.63	28.62	28.31	27.01	27.21
2016	24.25	21.65	25.81	25.18	27.34	26.10	29.01	27.79	28.93	28.32	27.31	27.35
2017	24.98	22.49	25.93	25.34	27.29	25.95	29.14	27.98	29.18	28.52	27.62	27.56
2018	24.58	20.88	25.33	24.30	27.19	26.08	29.27	28.09	29.14	28.36	27.65	27.64
2019	24.14	20.66	25.75	24.80	27.34	26.60	29.10	27.97	29.59	28.79	27.82	27.86
2020	24.63	21.26	25.51	24.42	27.08	26.44	29.20	27.93	29.34	28.47	27.92	28.12
2021	24.29	20.43	26.39	24.65	27.60	26.93	29.26	27.96	29.60	28.63	28.35	28.40
2022	24.20	20.68	26.25	24.57	27.41	26.69	29.32	27.71	29.57	28.50	28.55	28.54
2023	25.33	21.11	26.39	24.50	27.83	26.95	29.21	27.71	29.48	28.34	28.45	28.20

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
24.46	23.37	26.44	24.85	26.80	25.68	28.25	26.76	28.59	28.29	28.36	28.10
25.66	24.09	25.52	24.18	27.11	26.06	28.25	26.63	28.93	28.69	27.96	27.67
24.77	23.17	25.80	24.52	27.21	25.74	28.33	26.58	28.72	28.42	28.20	27.96
24.16	23.16	25.80	24.55	26.68	25.51	28.28	26.54	29.01	28.43	28.11	27.98
23.36	22.64	26.21	25.03	26.60	25.12	28.41	26.95	28.64	28.22	27.84	27.25
23.25	22.64	25.87	24.74	27.04	25.59	28.54	27.08	28.85	28.28	28.05	27.59
23.68	23.75	26.50	24.87	26.97	25.39	28.76	27.17	29.00	28.00	27.82	27.63
22.85	22.62	26.44	24.93	27.46	26.05	28.78	27.24	29.04	28.36	28.33	27.77
23.58	22.75	25.93	24.25	27.77	26.19	28.69	27.01	28.67	27.70	28.56	28.24
23.18	22.83	26.22	24.47	27.56	25.82	28.58	26.95	28.48	27.15	28.08	27.66

Scenario 3 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.03	0.00	0.00	-0.07	-0.02	-0.12	0.02	0.10	0.01	0.01	0.02	0.02
2015	0.00	0.00	-0.03	-0.01	-0.01	0.00	0.01	0.00	0.00	-0.01	0.05	0.00
2016	0.00	0.00	0.01	-0.01	0.00	-0.02	0.02	0.07	0.01	-0.02	0.04	0.02
2017	0.00	0.00	0.01	0.00	0.00	-0.05	0.03	0.08	-0.02	-0.04	0.05	0.00
2018	0.00	0.00	0.03	-0.01	0.00	-0.01	0.04	0.04	0.00	0.01	0.01	-0.01
2019	0.00	0.00	-0.02	-0.01	0.00	-0.01	0.03	0.05	-0.01	0.00	0.02	0.01
2020	0.00	-0.01	-0.04	-0.02	0.01	-0.03	0.05	0.03	-0.01	-0.06	0.03	0.00
2021	0.00	-0.01	-0.02	-0.01	0.00	0.00	0.03	0.01	0.00	-0.04	0.01	0.00
2022	0.00	-0.01	-0.02	-0.01	-0.01	0.00	0.01	0.02	0.01	-0.03	-0.03	-0.02
2023	0.00	-0.01	-0.01	-0.03	-0.02	-0.03	0.00	0.02	0.01	-0.02	-0.01	-0.01

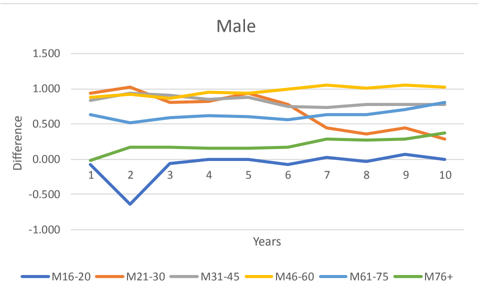
F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.02	-0.08	-0.04	-0.01	-0.01	0.00	-0.02	0.03	-0.04	0.00	-0.02	-0.01
0.00	-0.05	-0.03	0.00	-0.01	-0.02	0.00	0.00	-0.03	0.03	0.00	-0.05
0.00	0.00	0.00	0.01	-0.01	0.01	-0.01	0.00	-0.02	0.02	-0.05	-0.05
0.00	0.00	-0.03	-0.04	-0.02	-0.01	0.00	0.00	-0.03	0.02	-0.05	-0.09
0.00	0.00	-0.03	-0.01	-0.04	-0.03	0.00	-0.06	-0.02	0.00	-0.02	-0.05
0.00	-0.01	0.00	-0.01	-0.03	-0.06	-0.02	0.02	-0.02	-0.05	-0.01	-0.06
0.00	-0.01	0.02	-0.03	-0.03	-0.02	-0.03	-0.01	0.01	-0.02	-0.04	-0.05
-0.01	-0.01	0.01	-0.07	-0.01	-0.02	-0.04	-0.01	-0.01	-0.04	-0.04	-0.04
-0.01	-0.01	-0.01	-0.02	-0.01	-0.03	-0.04	-0.02	-0.04	-0.05	-0.02	-0.02
-0.01	-0.01	-0.01	-0.02	-0.01	-0.06	-0.02	-0.01	-0.01	-0.12	-0.03	-0.06

TABLE C.17: Scenario 3- Adjusted Forecast and Comparison

Scenario 4 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.489	22.567	25.091	24.155	27.570	26.738	28.520	27.647	28.857	28.226	28.025	28.042
2015	22.276	22.920	24.823	23.798	27.434	26.499	28.832	27.917	29.047	28.526	28.178	28.008
2016	23.012	23.071	24.331	23.523	27.518	26.609	28.855	27.995	29.159	28.567	28.417	28.244
2017	22.978	22.976	24.450	23.633	27.446	26.597	29.024	28.070	29.198	28.574	28.648	28.493
2018	23.035	23.038	24.320	23.383	27.509	26.629	29.093	28.154	29.430	28.824	28.730	28.577
2019	22.922	22.995	24.279	23.495	27.374	26.623	29.162	28.162	29.713	29.147	28.871	28.698
2020	22.807	22.786	24.041	23.589	27.335	26.599	29.274	28.229	29.788	29.161	29.153	28.860
2021	22.688	22.718	23.960	23.599	27.298	26.521	29.369	28.355	29.853	29.226	29.246	28.968
2022	22.474	22.411	24.005	23.554	27.138	26.362	29.386	28.339	29.978	29.269	29.539	29.247
2023	22.427	22.433	23.900	23.619	27.056	26.273	29.481	28.452	30.107	29.294	29.854	29.482

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.798	23.034	25.512	24.051	26.863	25.602	28.569	27.039	28.744	28.055	28.280	27.688
23.897	22.996	25.365	24.001	27.036	25.750	28.602	27.093	28.855	28.159	28.671	28.147
23.011	22.947	25.250	23.967	27.248	25.822	28.722	27.043	29.043	28.277	28.908	28.410
23.096	23.082	25.504	24.172	27.063	25.709	28.962	27.149	29.259	28.561	28.998	28.432
23.082	23.018	25.480	24.257	27.204	25.791	29.124	27.537	29.277	28.482	29.271	28.607
22.989	22.974	25.481	24.013	27.172	25.742	29.173	27.490	29.493	28.832	29.399	28.826
22.889	22.895	25.404	23.971	27.225	25.754	29.336	27.666	29.588	28.784	29.516	28.939
22.807	22.748	25.069	23.912	27.358	25.862	29.435	27.767	29.651	28.799	29.724	29.100
22.802	22.808	25.001	23.939	27.568	26.128	29.430	27.572	29.826	28.912	29.865	29.131
22.676	22.689	24.707	23.825	27.684	26.107	29.423	27.667	29.966	28.744	30.196	29.402



Mean v's Median Differences:

M16-20	M21-30	M31-45	M46-60	M61-75	M76+
-0.078	0.935	0.832	0.873	0.631	-0.017
-0.644	1.025	0.935	0.915	0.521	0.170
-0.059	0.809	0.909	0.861	0.592	0.172
0.002	0.817	0.849	0.955	0.624	0.155
-0.003	0.937	0.880	0.939	0.606	0.153
-0.072	0.784	0.751	0.999	0.566	0.173
0.021	0.451	0.735	1.045	0.628	0.292
-0.030	0.361	0.777	1.014	0.626	0.278
0.063	0.450	0.775	1.047	0.709	0.292
-0.006	0.281	0.783	1.028	0.813	0.372

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
0.764	1.460	1.260	1.530	0.689	0.592
0.902	1.364	1.285	1.510	0.695	0.523
0.064	1.284	1.425	1.679	0.766	0.498
0.014	1.332	1.353	1.813	0.699	0.566
0.064	1.223	1.413	1.586	0.795	0.664
0.015	1.468	1.430	1.683	0.661	0.574
-0.006	1.433	1.471	1.670	0.804	0.576
0.059	1.157	1.496	1.668	0.853	0.624
-0.006	1.062	1.439	1.859	0.914	0.734
-0.013	0.882	1.577	1.756	1.222	0.795

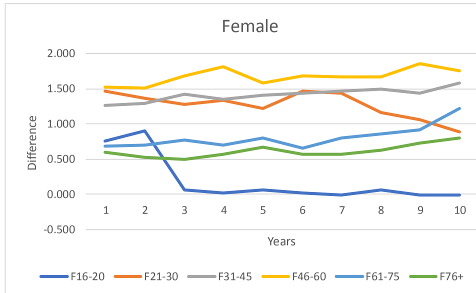


TABLE C.18: Scenario 4 - Raw Forecast

Scenario 4: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.845	23.058	26.092	25.291	27.549	26.313	28.867	27.870	28.915	28.485	27.423	27.541
2015	22.949	24.475	25.428	24.665	26.840	25.536	28.617	27.626	28.619	28.317	27.022	27.228
2016	23.539	23.706	25.814	25.181	27.355	26.132	28.992	27.723	28.904	28.302	27.314	27.311
2017	24.168	24.376	25.924	25.351	27.312	26.005	29.121	27.908	29.183	28.524	27.601	27.562
2018	23.760	22.758	25.326	24.308	27.213	26.087	29.266	28.088	29.126	28.324	27.638	27.636
2019	23.209	22.496	25.742	25.173	27.359	26.609	29.096	27.925	29.580	28.759	27.796	27.845
2020	23.586	22.886	25.428	25.395	27.094	26.465	29.190	27.915	29.318	28.472	27.911	28.107
2021	23.133	21.996	26.214	25.862	27.628	26.948	29.268	27.956	29.569	28.583	28.339	28.388
2022	22.826	21.934	25.933	25.736	27.447	26.704	29.328	27.705	29.535	28.498	28.534	28.512
2023	23.909	22.393	25.988	25.948	27.861	26.965	29.201	27.714	29.461	28.327	28.452	28.202

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	24.517	23.777	26.497	24.861	26.803	25.682	28.284	26.780	28.641	28.322	28.374	28.109
2015	25.842	24.372	25.573	24.184	27.104	26.078	28.272	26.673	28.977	28.745	27.964	27.684
2016	25.185	24.102	25.815	24.529	27.217	25.738	28.346	26.634	28.787	28.436	28.224	27.915
2017	24.569	24.130	25.820	24.581	26.697	25.505	28.300	26.562	29.054	28.483	28.101	27.945
2018	23.662	23.447	26.238	25.039	26.626	25.138	28.409	27.020	28.703	28.279	27.813	27.227
2019	23.464	23.400	25.925	24.408	27.086	25.640	28.537	27.104	28.904	28.407	28.049	27.627
2020	23.793	24.431	26.529	24.700	27.007	25.397	28.777	27.216	29.033	28.038	27.842	27.681
2021	22.875	23.156	26.396	24.958	27.481	26.063	28.809	27.260	29.068	28.358	28.333	27.790
2022	23.604	23.345	25.968	24.512	27.784	26.219	28.731	27.035	28.715	27.717	28.562	28.242
2023	23.084	23.315	26.232	24.733	27.562	25.863	28.615	26.970	28.497	27.159	28.101	27.670

Scenario 4 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.146	1.230	0.021	-0.045	-0.023	-0.003	-0.003	0.028	-0.001	0.000	0.016	0.033
2015	-0.448	1.903	0.014	0.014	0.009	-0.001	0.009	-0.001	-0.006	-0.011	0.066	0.015
2016	-0.710	2.054	0.017	-0.006	0.013	0.009	0.005	0.002	-0.014	-0.043	0.042	-0.018
2017	-0.809	1.887	0.002	0.011	0.021	-0.003	0.008	-0.002	-0.018	-0.038	0.033	-0.001
2018	-0.817	1.877	0.031	0.001	0.013	-0.001	0.028	0.032	-0.011	-0.026	0.003	-0.018
2019	-0.930	1.834	-0.023	0.354	0.022	0.000	0.025	0.008	-0.023	-0.033	0.001	-0.002
2020	-1.045	1.625	-0.122	0.954	0.028	-0.005	0.037	0.009	-0.024	-0.063	0.017	-0.007
2021	-1.164	1.557	-0.195	1.203	0.026	0.019	0.033	0.015	-0.030	-0.088	-0.007	-0.008
2022	-1.379	1.250	-0.331	1.158	0.027	0.020	0.021	0.012	-0.031	-0.027	-0.039	-0.042
2023	-1.425	1.272	-0.404	1.417	0.014	-0.009	-0.012	0.017	-0.014	-0.036	-0.014	-0.015

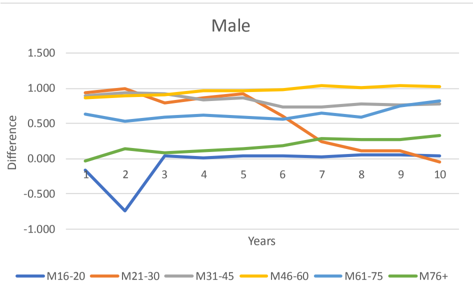
	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.081	0.325	0.014	0.001	-0.010	0.004	0.017	0.048	0.014	0.032	0.000	-0.006
2015	0.174	0.236	0.028	0.006	-0.024	-0.007	0.025	0.040	0.020	0.082	-0.001	-0.040
2016	0.409	0.933	0.011	0.026	0.000	0.014	0.000	0.053	0.045	0.041	-0.028	-0.103
2017	0.403	0.967	-0.002	-0.011	-0.001	-0.009	0.023	0.023	0.019	0.075	-0.051	-0.127
2018	0.299	0.801	-0.002	-0.004	-0.005	-0.021	0.002	0.013	0.038	0.053	-0.043	-0.076
2019	0.206	0.757	0.059	-0.338	0.013	-0.008	-0.016	0.037	0.039	0.078	-0.015	-0.018
2020	0.106	0.678	0.048	-0.199	0.009	-0.011	-0.007	0.036	0.037	0.018	-0.019	0.000
2021	0.024	0.531	-0.035	-0.038	0.013	-0.011	-0.013	0.011	0.011	-0.036	-0.029	-0.023
2022	0.019	0.591	0.035	0.239	0.008	-0.001	0.006	0.006	0.012	-0.028	-0.022	-0.011
2023	-0.107	0.473	0.007	0.245	-0.008	-0.017	0.014	0.002	0.006	-0.105	-0.014	-0.046

TABLE C.19: Scenario 4- Adjusted Forecast and Comparison

Scenario 5 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.478	22.647	25.076	24.135	27.593	26.695	28.527	27.668	28.872	28.239	28.015	28.039
2015	22.440	23.180	24.769	23.774	27.451	26.511	28.818	27.918	29.075	28.542	28.163	28.019
2016	23.412	23.374	24.288	23.492	27.543	26.623	28.861	27.951	29.193	28.605	28.358	28.280
2017	23.520	23.515	24.421	23.560	27.439	26.611	29.031	28.064	29.220	28.596	28.606	28.492
2018	23.628	23.587	24.292	23.376	27.486	26.627	29.105	28.144	29.434	28.843	28.723	28.581
2019	23.722	23.678	24.258	23.656	27.352	26.618	29.170	28.188	29.736	29.179	28.886	28.708
2020	23.807	23.781	24.070	23.825	27.316	26.580	29.272	28.235	29.811	29.168	29.153	28.864
2021	23.896	23.843	24.064	23.951	27.281	26.503	29.367	28.363	29.874	29.286	29.254	28.975
2022	23.977	23.929	24.197	24.084	27.122	26.361	29.381	28.350	30.033	29.280	29.548	29.274
2023	24.098	24.058	24.157	24.205	27.045	26.271	29.490	28.468	30.153	29.339	29.828	29.496

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.848	23.158	25.495	24.039	26.889	25.623	28.596	27.026	28.717	28.023	28.294	27.779
24.035	23.284	25.337	23.976	27.076	25.807	28.623	27.091	28.830	28.110	28.689	28.231
23.441	23.447	25.235	23.930	27.283	25.861	28.721	27.054	29.023	28.248	28.909	28.479
23.586	23.562	25.479	24.096	27.085	25.738	28.971	27.136	29.233	28.519	29.003	28.496
23.717	23.687	25.418	24.219	27.244	25.823	29.142	27.525	29.246	28.434	29.287	28.666
23.816	23.805	25.403	24.145	27.184	25.750	29.201	27.474	29.466	28.819	29.370	28.843
23.913	23.928	25.412	24.136	27.263	25.765	29.367	27.651	29.563	28.755	29.501	28.973
24.025	24.039	25.215	24.254	27.363	25.844	29.457	27.765	29.657	28.819	29.739	29.143
24.149	24.145	25.188	24.395	27.563	26.110	29.431	27.603	29.831	28.908	29.883	29.235
24.237	24.242	25.047	24.495	27.681	26.118	29.458	27.676	29.996	28.842	30.202	29.472



Mean v's Median Differences:

M16-20	M21-30	M31-45	M46-60	M61-75	M76+
-0.169	0.941	0.899	0.859	0.633	-0.025
-0.740	0.994	0.940	0.899	0.533	0.143
0.038	0.797	0.920	0.910	0.588	0.078
0.005	0.861	0.828	0.967	0.624	0.114
0.041	0.916	0.859	0.961	0.591	0.142
0.044	0.602	0.734	0.982	0.556	0.179
0.026	0.244	0.735	1.038	0.643	0.289
0.052	0.113	0.778	1.004	0.587	0.279
0.048	0.112	0.761	1.031	0.753	0.274
0.040	-0.047	0.774	1.022	0.814	0.332

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
0.690	1.456	1.266	1.571	0.693	0.515
0.751	1.361	1.269	1.531	0.720	0.457
-0.006	1.305	1.422	1.666	0.774	0.431
0.024	1.384	1.347	1.835	0.714	0.507
0.030	1.199	1.420	1.617	0.812	0.621
0.011	1.258	1.434	1.727	0.647	0.527
-0.015	1.276	1.498	1.716	0.808	0.528
-0.013	0.961	1.519	1.692	0.838	0.596
0.005	0.794	1.453	1.828	0.923	0.649
-0.005	0.552	1.563	1.782	1.155	0.729

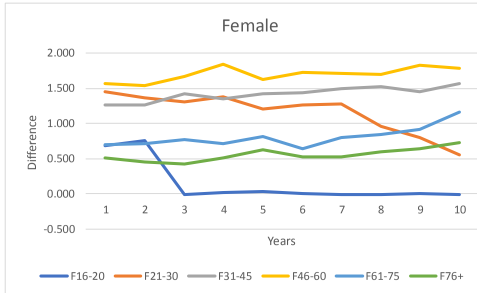


TABLE C.20: Scenario 5 - Raw Forecast

Scenario 5: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.835	23.138	26.077	25.271	27.573	26.270	28.873	27.891	28.930	28.498	27.413	27.538
2015	23.113	24.734	25.373	24.641	26.857	25.547	28.603	27.627	28.646	28.332	27.007	27.239
2016	23.939	24.009	25.771	25.150	27.379	26.147	28.998	27.680	28.938	28.340	27.255	27.346
2017	24.710	24.914	25.895	25.279	27.305	26.019	29.127	27.902	29.206	28.546	27.560	27.562
2018	24.353	23.306	25.297	24.301	27.189	26.085	29.277	28.077	29.130	28.342	27.630	27.640
2019	24.009	23.180	25.721	25.334	27.337	26.604	29.104	27.951	29.603	28.792	27.811	27.854
2020	24.586	23.881	25.456	25.631	27.075	26.446	29.188	27.921	29.341	28.479	27.912	28.111
2021	24.341	23.121	26.318	26.214	27.611	26.930	29.266	27.964	29.590	28.643	28.347	28.395
2022	24.329	23.451	26.125	26.266	27.432	26.703	29.323	27.716	29.590	28.508	28.543	28.539
2023	25.580	24.017	26.245	26.534	27.849	26.963	29.210	27.729	29.507	28.372	28.425	28.216

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
24.567	23.901	26.481	24.848	26.829	25.702	28.312	26.767	28.614	28.291	28.388	28.200
25.979	24.661	25.545	24.159	27.145	26.135	28.292	26.672	28.952	28.695	27.981	27.768
25.614	24.602	25.800	24.492	27.253	25.776	28.344	26.646	28.768	28.408	28.225	27.984
25.059	24.610	25.796	24.505	26.719	25.534	28.309	26.549	29.028	28.441	28.106	28.009
24.297	24.116	26.176	25.001	26.666	25.170	28.428	27.008	28.672	28.231	27.829	27.287
24.291	24.231	25.847	24.541	27.098	25.648	28.565	27.089	28.876	28.394	28.020	27.644
24.817	25.463	26.537	24.866	27.045	25.409	28.808	27.201	29.008	28.009	27.827	27.714
24.094	24.447	26.542	25.300	27.486	26.045	28.832	27.259	29.073	28.378	28.348	27.832
24.952	24.682	26.155	24.967	27.780	26.201	28.731	27.066	28.720	27.713	28.581	28.345
24.645	24.867	26.572	25.403	27.559	25.874	28.650	26.980	28.527	27.257	28.107	27.741

Scenario 5 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.157	1.310	0.005	-0.065	0.000	-0.046	0.004	0.049	0.014	0.013	0.005	0.030
2015	-0.284	2.163	-0.040	-0.010	0.026	0.011	-0.005	0.000	0.022	0.004	0.051	0.026
2016	-0.310	2.356	-0.026	-0.037	0.037	0.023	0.010	-0.041	0.020	-0.004	-0.017	0.017
2017	-0.267	2.426	-0.027	-0.062	0.014	0.012	0.014	-0.008	0.004	-0.016	-0.008	-0.001
2018	-0.224	2.426	0.003	-0.005	-0.010	-0.003	0.040	0.021	-0.007	-0.007	-0.004	-0.015
2019	-0.130	2.517	-0.043	0.515	-0.001	-0.005	0.034	0.034	0.000	0.000	0.016	0.007
2020	-0.046	2.620	-0.093	1.190	0.009	-0.024	0.035	0.014	-0.001	-0.055	0.018	-0.003
2021	0.043	2.682	-0.091	1.555	0.009	0.001	0.030	0.023	-0.009	-0.028	0.001	0.000
2022	0.124	2.768	-0.139	1.688	0.011	0.019	0.015	0.022	0.024	-0.016	-0.030	-0.014
2023	0.245	2.897	-0.146	2.003	0.002	-0.011	-0.003	0.032	0.031	0.009	-0.040	-0.001

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.131	0.449	-0.003	-0.011	0.016	0.025	0.045	0.035	-0.013	0.001	0.014	0.086
0.312	0.524	0.000	-0.019	0.017	0.050	0.045	0.039	-0.004	0.033	0.017	0.044
0.838	1.433	-0.004	-0.011	0.035	0.052	-0.001	0.065	0.025	0.013	-0.027	-0.034
0.893	1.447	-0.026	-0.087	0.022	0.020	0.032	0.010	-0.007	0.033	-0.046	-0.063
0.934	1.470	-0.064	-0.042	0.035	0.011	0.021	0.001	0.007	0.005	-0.027	-0.016
1.033	1.588	-0.019	-0.206	0.025	0.000	0.012	0.022	0.012	0.065	-0.044	-0.001
1.130	1.711	0.056	-0.033	0.047	0.000	0.023	0.021	0.012	-0.011	-0.034	0.034
1.243	1.822	0.111	0.304	0.018	-0.030	0.010	0.009	0.017	-0.015	-0.015	0.020
1.367	1.928	0.222	0.694	0.004	-0.019	0.007	0.037	0.017	-0.031	-0.004	0.093
1.455	2.025	0.347	0.916	-0.011	-0.006	0.048	0.012	0.036	-0.008	-0.008	0.024

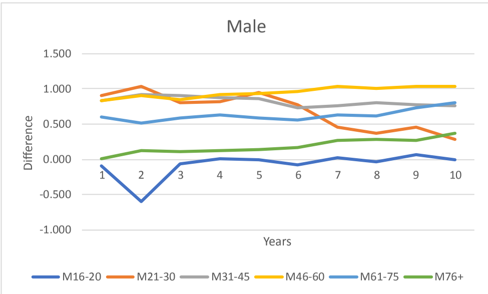
TABLE C.21: Scenario 5- Adjusted Forecast and Comparison

Scenario 6 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.469	22.567	25.096	24.184	27.579	26.738	28.527	27.690	28.830	28.226	28.003	27.998
2015	22.316	22.921	24.808	23.774	27.426	26.500	28.823	27.920	29.023	28.502	28.133	28.009
2016	23.013	23.072	24.327	23.524	27.524	26.611	28.859	28.008	29.158	28.569	28.372	28.265
2017	22.980	22.978	24.458	23.643	27.435	26.562	29.040	28.119	29.201	28.576	28.608	28.482
2018	23.038	23.040	24.342	23.386	27.485	26.626	29.099	28.163	29.410	28.828	28.707	28.569
2019	22.925	22.998	24.279	23.497	27.352	26.615	29.188	28.223	29.705	29.142	28.859	28.698
2020	22.810	22.789	24.046	23.596	27.335	26.578	29.282	28.248	29.798	29.165	29.148	28.879
2021	22.692	22.723	23.975	23.599	27.311	26.510	29.372	28.363	29.884	29.264	29.260	28.981
2022	22.478	22.414	24.015	23.562	27.152	26.373	29.393	28.360	29.999	29.268	29.559	29.293
2023	22.432	22.440	23.905	23.628	27.060	26.303	29.503	28.463	30.118	29.311	29.870	29.496

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.873	23.034	25.501	24.014	26.856	25.592	28.542	27.009	28.733	28.028	28.264	27.688
23.959	22.997	25.349	23.981	27.025	25.733	28.584	27.078	28.807	28.133	28.708	28.159
23.012	22.949	25.286	23.941	27.244	25.803	28.697	26.992	28.983	28.268	28.924	28.454
23.098	23.084	25.518	24.174	27.068	25.716	28.951	27.152	29.204	28.525	29.029	28.473
23.085	23.021	25.454	24.249	27.219	25.783	29.113	27.527	29.235	28.486	29.320	28.611
22.992	22.976	25.480	24.019	27.154	25.701	29.165	27.471	29.489	28.814	29.438	28.843
22.893	22.898	25.456	23.999	27.240	25.739	29.324	27.644	29.582	28.780	29.558	28.953
22.812	22.754	25.138	23.922	27.361	25.812	29.437	27.757	29.639	28.811	29.778	29.113
22.808	22.814	25.059	23.968	27.572	26.113	29.415	27.564	29.794	28.931	29.902	29.233
22.683	22.695	24.742	23.860	27.715	26.109	29.418	27.670	29.947	28.840	30.248	29.472

Mean v's Median Differences:



M16-20	M21-30	M31-45	M46-60	M61-75	M76+
-0.098	0.912	0.840	0.838	0.604	0.005
-0.605	1.034	0.926	0.903	0.521	0.124
-0.059	0.803	0.913	0.851	0.589	0.108
0.002	0.815	0.874	0.921	0.626	0.126
-0.002	0.956	0.859	0.936	0.582	0.139
-0.073	0.782	0.737	0.965	0.563	0.160
0.021	0.450	0.757	1.035	0.632	0.269
-0.030	0.376	0.801	1.009	0.620	0.279
0.063	0.452	0.779	1.032	0.730	0.266
-0.008	0.277	0.757	1.040	0.808	0.374

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
0.839	1.487	1.264	1.533	0.705	0.575
0.963	1.368	1.292	1.506	0.674	0.549
0.064	1.345	1.440	1.705	0.715	0.470
0.014	1.344	1.352	1.799	0.679	0.556
0.064	1.205	1.436	1.587	0.749	0.709
0.016	1.461	1.453	1.694	0.675	0.596
-0.006	1.457	1.500	1.680	0.802	0.605
0.058	1.216	1.549	1.680	0.827	0.665
-0.006	1.092	1.460	1.851	0.863	0.669
-0.012	0.882	1.606	1.747	1.107	0.776

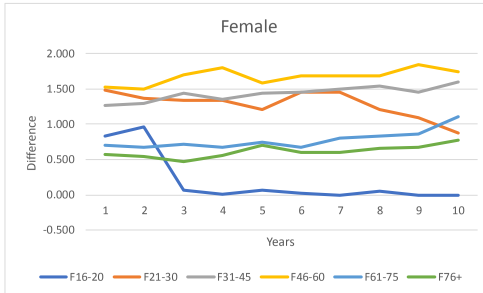


TABLE C.22: Scenario 6 - Raw Forecast

Scenario 6: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.825	23.058	26.097	25.321	27.558	26.313	28.874	27.913	28.888	28.485	27.401	27.496
2015	22.989	24.475	25.412	24.641	26.832	25.536	28.609	27.629	28.595	28.293	26.977	27.229
2016	23.540	23.707	25.810	25.183	27.360	26.135	28.996	27.737	28.902	28.304	27.269	27.331
2017	24.169	24.377	25.932	25.362	27.302	25.970	29.136	27.957	29.187	28.525	27.561	27.551
2018	23.762	22.759	25.347	24.311	27.188	26.084	29.272	28.096	29.106	28.328	27.615	27.628
2019	23.212	22.499	25.742	25.175	27.337	26.600	29.122	27.986	29.572	28.754	27.784	27.845
2020	23.589	22.889	25.433	25.401	27.094	26.443	29.198	27.934	29.328	28.477	27.907	28.126
2021	23.137	22.000	26.229	25.862	27.641	26.937	29.271	27.963	29.601	28.621	28.353	28.401
2022	22.830	21.937	25.943	25.744	27.462	26.715	29.334	27.726	29.555	28.497	28.554	28.558
2023	23.914	22.400	25.993	25.957	27.865	26.995	29.223	27.724	29.473	28.343	28.468	28.216

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	24.592	23.777	26.487	24.824	26.796	25.671	28.257	26.750	28.630	28.296	28.357	28.109
2015	25.904	24.373	25.557	24.164	27.093	26.061	28.253	26.659	28.929	28.718	28.000	27.696
2016	25.186	24.103	25.851	24.504	27.213	25.719	28.321	26.583	28.728	28.428	28.239	27.959
2017	24.571	24.132	25.834	24.583	26.702	25.512	28.289	26.565	28.998	28.447	28.132	27.986
2018	23.664	23.450	26.213	25.032	26.641	25.130	28.399	27.010	28.661	28.283	27.862	27.231
2019	23.468	23.403	25.924	24.415	27.069	25.598	28.529	27.085	28.899	28.389	28.088	27.644
2020	23.797	24.434	26.582	24.729	27.021	25.383	28.765	27.194	29.026	28.034	27.885	27.694
2021	22.880	23.162	26.466	24.968	27.484	26.012	28.811	27.250	29.055	28.370	28.387	27.802
2022	23.610	23.351	26.026	24.541	27.788	26.204	28.715	27.027	28.683	27.736	28.599	28.344
2023	23.091	23.321	26.267	24.769	27.593	25.865	28.609	26.974	28.478	27.255	28.153	27.741

Scenario 6 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.166	1.230	0.026	-0.016	-0.014	-0.003	0.004	0.070	-0.028	0.000	-0.006	-0.011
2015	-0.407	1.904	-0.001	-0.010	0.001	0.000	0.001	0.002	-0.030	-0.035	0.021	0.016
2016	-0.709	2.055	0.012	-0.004	0.018	0.011	0.008	0.016	-0.015	-0.040	-0.003	0.002
2017	-0.808	1.889	0.011	0.021	0.010	-0.038	0.024	0.047	-0.015	-0.036	-0.007	-0.012
2018	-0.815	1.879	0.053	0.004	-0.011	-0.004	0.035	0.040	-0.031	-0.022	-0.019	-0.026
2019	-0.927	1.837	-0.023	0.356	-0.001	-0.008	0.052	0.069	-0.031	-0.038	-0.011	-0.002
2020	-1.042	1.628	-0.117	0.960	0.028	-0.027	0.045	0.027	-0.015	-0.058	0.013	0.012
2021	-1.160	1.562	-0.180	1.203	0.039	0.008	0.036	0.023	0.002	-0.050	0.007	0.005
2022	-1.375	1.253	-0.321	1.166	0.041	0.030	0.027	0.033	-0.011	-0.027	-0.020	0.004
2023	-1.421	1.279	-0.399	1.426	0.018	0.021	0.010	0.027	-0.003	-0.019	0.002	-0.001

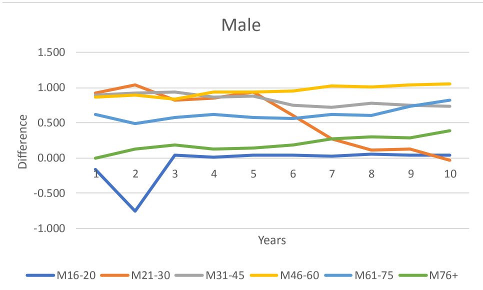
	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.156	0.325	0.003	-0.036	-0.018	-0.006	-0.009	0.017	0.003	0.006	-0.016	-0.005
2015	0.236	0.237	0.012	-0.014	-0.034	-0.024	0.006	0.026	-0.028	0.056	0.035	-0.029
2016	0.410	0.935	0.047	0.000	-0.004	-0.005	-0.025	0.002	-0.014	0.033	-0.012	-0.059
2017	0.405	0.969	0.012	-0.009	0.004	-0.002	0.012	0.026	-0.036	0.038	-0.020	-0.086
2018	0.302	0.804	-0.027	-0.012	0.010	-0.029	-0.008	0.002	-0.004	0.057	0.006	-0.071
2019	0.210	0.760	0.058	-0.332	-0.004	-0.050	-0.025	0.019	0.034	0.059	0.024	-0.001
2020	0.110	0.682	0.101	-0.170	0.024	-0.026	-0.019	0.014	0.030	0.013	0.024	0.014
2021	0.029	0.537	0.035	-0.028	0.016	-0.062	-0.010	0.001	-0.002	-0.023	0.024	-0.010
2022	0.025	0.597	0.093	0.268	0.013	-0.016	-0.009	-0.002	-0.020	-0.008	0.015	0.091
2023	-0.100	0.478	0.042	0.281	0.023	-0.015	0.008	0.006	-0.013	-0.010	0.038	0.024

TABLE C.23: Scenario 6- Adjusted Forecast and Comparison

Scenario 7 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.508	22.675	25.099	24.185	27.576	26.686	28.508	27.641	28.876	28.259	28.001	27.996
2015	22.420	23.180	24.835	23.795	27.439	26.516	28.809	27.919	29.070	28.575	28.118	27.994
2016	23.413	23.374	24.346	23.523	27.539	26.602	28.845	28.006	29.183	28.605	28.398	28.219
2017	23.522	23.517	24.462	23.615	27.451	26.581	29.016	28.076	29.220	28.607	28.627	28.493
2018	23.631	23.591	24.316	23.387	27.514	26.630	29.077	28.147	29.449	28.881	28.730	28.588
2019	23.725	23.681	24.269	23.661	27.370	26.628	29.121	28.173	29.750	29.182	28.891	28.709
2020	23.810	23.785	24.096	23.831	27.331	26.613	29.242	28.214	29.838	29.213	29.153	28.882
2021	23.899	23.850	24.075	23.959	27.303	26.524	29.351	28.348	29.902	29.293	29.282	28.987
2022	23.981	23.934	24.226	24.093	27.131	26.383	29.377	28.345	30.017	29.286	29.597	29.313
2023	24.102	24.065	24.189	24.217	27.056	26.315	29.505	28.459	30.154	29.327	29.904	29.511

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.824	23.158	25.457	24.050	26.846	25.589	28.541	26.950	28.739	28.039	28.275	27.729
24.022	23.285	25.308	23.980	27.033	25.734	28.594	27.054	28.818	28.146	28.684	28.164
23.442	23.448	25.188	23.932	27.222	25.784	28.714	27.033	29.004	28.275	28.928	28.451
23.588	23.565	25.435	24.100	27.033	25.698	28.965	27.166	29.227	28.522	29.034	28.496
23.720	23.691	25.427	24.224	27.194	25.800	29.137	27.526	29.241	28.434	29.316	28.661
23.819	23.807	25.424	24.206	27.159	25.755	29.180	27.480	29.459	28.709	29.416	28.848
23.917	23.932	25.390	24.143	27.218	25.777	29.325	27.655	29.565	28.757	29.550	28.942
24.030	24.044	25.226	24.252	27.367	25.854	29.452	27.772	29.633	28.819	29.752	29.111
24.155	24.155	25.192	24.408	27.570	26.138	29.446	27.579	29.822	28.932	29.909	29.180
24.244	24.248	25.043	24.510	27.708	26.141	29.442	27.689	29.961	28.766	30.222	29.487



Mean v's Median Differences:

M16-20	M21-30	M31-45	M46-60	M61-75	M76+
-0.167	0.915	0.890	0.867	0.616	0.005
-0.761	1.040	0.923	0.890	0.495	0.124
0.039	0.823	0.938	0.839	0.578	0.179
0.005	0.847	0.871	0.941	0.612	0.134
0.040	0.929	0.885	0.930	0.568	0.142
0.044	0.608	0.741	0.949	0.568	0.182
0.026	0.265	0.718	1.028	0.626	0.271
0.050	0.116	0.779	1.004	0.609	0.295
0.046	0.133	0.748	1.032	0.730	0.283
0.037	-0.028	0.741	1.045	0.827	0.392

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
0.666	1.407	1.257	1.591	0.700	0.547
0.737	1.328	1.300	1.540	0.672	0.521
-0.007	1.256	1.438	1.681	0.729	0.477
0.023	1.335	1.335	1.798	0.705	0.538
0.028	1.203	1.394	1.612	0.807	0.655
0.011	1.218	1.403	1.700	0.750	0.568
-0.015	1.246	1.441	1.670	0.808	0.608
-0.014	0.974	1.513	1.680	0.814	0.641
0.000	0.784	1.432	1.867	0.890	0.729
-0.004	0.533	1.568	1.753	1.194	0.735

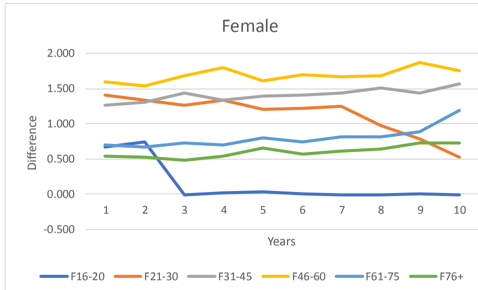


TABLE C.24: Scenario 7 - Raw Forecast

Scenario 7: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.864	23.166	26.100	25.321	27.555	26.261	28.854	27.864	28.934	28.518	27.399	27.495
2015	23.093	24.735	25.439	24.661	26.845	25.552	28.594	27.628	28.642	28.366	26.962	27.214
2016	23.940	24.009	25.829	25.181	27.375	26.125	28.982	27.735	28.927	28.340	27.295	27.285
2017	24.711	24.916	25.936	25.334	27.317	25.989	29.113	27.914	29.205	28.557	27.580	27.562
2018	24.355	23.310	25.322	24.312	27.218	26.088	29.250	28.080	29.145	28.380	27.637	27.647
2019	24.012	23.182	25.731	25.339	27.354	26.614	29.056	27.935	29.617	28.795	27.816	27.855
2020	24.589	23.885	25.483	25.636	27.090	26.478	29.158	27.900	29.368	28.524	27.912	28.129
2021	24.344	23.127	26.329	26.221	27.633	26.951	29.250	27.948	29.618	28.650	28.375	28.407
2022	24.333	23.457	26.154	26.275	27.440	26.725	29.319	27.711	29.573	28.515	28.592	28.578
2023	25.584	24.024	26.277	26.546	27.860	27.007	29.225	27.721	29.508	28.360	28.501	28.231

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	24.543	23.901	26.442	24.859	26.787	25.669	28.256	26.692	28.636	28.306	28.369	28.149
2015	25.967	24.661	25.516	24.163	27.102	26.062	28.263	26.635	28.940	28.731	27.977	27.701
2016	25.615	24.603	25.752	24.494	27.191	25.700	28.337	26.624	28.749	28.434	28.243	27.956
2017	25.061	24.613	25.751	24.509	26.667	25.494	28.303	26.579	29.021	28.444	28.136	28.009
2018	24.299	24.120	26.185	25.006	26.616	25.147	28.423	27.009	28.667	28.231	27.859	27.281
2019	24.294	24.234	25.868	24.602	27.073	25.653	28.544	27.094	28.869	28.284	28.066	27.649
2020	24.820	25.467	26.515	24.873	27.000	25.420	28.766	27.205	29.010	28.011	27.877	27.683
2021	24.098	24.452	26.553	25.298	27.490	26.055	28.826	27.265	29.050	28.378	28.361	27.801
2022	24.958	24.692	26.159	24.980	27.786	26.229	28.746	27.042	28.711	27.737	28.607	28.291
2023	24.652	24.874	26.568	25.418	27.586	25.896	28.634	26.992	28.492	27.181	28.127	27.755

Scenario 7 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.127	1.338	0.029	-0.015	-0.018	-0.056	-0.015	0.022	0.017	0.033	-0.008	-0.013
2015	-0.304	2.163	0.026	0.011	0.014	0.016	-0.014	0.001	0.017	0.038	0.006	0.001
2016	-0.309	2.357	0.031	-0.006	0.034	0.001	-0.006	0.013	0.010	-0.005	0.023	-0.044
2017	-0.266	2.428	0.014	-0.007	0.026	-0.019	0.000	0.004	0.003	-0.005	0.012	-0.001
2018	-0.222	2.430	0.027	0.005	0.018	-0.001	0.012	0.024	0.008	0.030	0.003	-0.007
2019	-0.127	2.520	-0.033	0.520	0.017	0.005	-0.015	0.018	0.014	0.003	0.021	0.008
2020	-0.042	2.624	-0.066	1.195	0.024	0.008	0.005	-0.007	0.026	-0.011	0.018	0.015
2021	0.047	2.689	-0.080	1.563	0.031	0.022	0.015	0.007	0.019	-0.021	0.029	0.011
2022	0.128	2.773	-0.110	1.697	0.020	0.041	0.012	0.017	0.007	-0.009	0.018	0.025
2023	0.250	2.904	-0.114	2.015	0.013	0.033	0.012	0.024	0.033	-0.003	0.036	0.014

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.106	0.449	-0.041	0.000	-0.027	-0.009	-0.010	-0.041	0.009	0.017	-0.004	0.035
2015	0.299	0.525	-0.028	-0.015	-0.026	-0.023	0.016	0.002	-0.017	0.068	0.012	-0.024
2016	0.839	1.434	-0.051	-0.010	-0.027	-0.024	-0.008	0.043	0.006	0.039	-0.008	-0.062
2017	0.895	1.450	-0.071	-0.083	-0.030	-0.020	0.026	0.040	-0.013	0.036	-0.016	-0.063
2018	0.937	1.474	-0.055	-0.038	-0.015	-0.012	0.016	0.001	0.002	0.005	0.002	-0.022
2019	1.036	1.591	0.002	-0.145	0.000	0.005	-0.009	0.027	0.004	-0.046	0.002	0.004
2020	1.134	1.715	0.034	-0.026	0.002	0.011	-0.019	0.026	0.014	-0.009	0.016	0.003
2021	1.247	1.827	0.122	0.302	0.022	-0.020	0.004	0.016	-0.007	-0.015	-0.002	-0.011
2022	1.373	1.938	0.226	0.708	0.011	0.008	0.022	0.013	0.008	-0.007	0.022	0.038
2023	1.461	2.031	0.343	0.930	0.016	0.017	0.032	0.024	0.001	-0.083	0.012	0.039

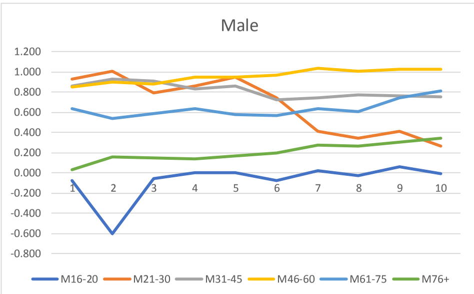
TABLE C.25: Scenario 7- Adjusted Forecast and Comparison

Scenario 8 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.696	21.337	25.042	24.134	27.575	26.672	28.520	27.641	28.869	28.251	28.040	27.984
2015	22.753	21.018	24.764	23.775	27.428	26.511	28.828	27.912	29.056	28.542	28.159	28.017
2016	23.723	21.018	24.318	23.542	27.508	26.597	28.869	27.998	29.175	28.613	28.429	28.242
2017	23.789	21.091	24.454	23.666	27.413	26.574	29.039	28.075	29.216	28.622	28.677	28.499
2018	23.855	21.164	24.328	23.391	27.491	26.623	29.091	28.127	29.428	28.829	28.766	28.592
2019	23.855	21.166	24.335	23.152	27.348	26.628	29.162	28.162	29.721	29.122	28.923	28.718
2020	23.856	21.167	24.187	22.627	27.323	26.578	29.260	28.223	29.800	29.186	29.148	28.884
2021	23.856	21.168	24.202	22.404	27.282	26.506	29.353	28.342	29.888	29.295	29.270	28.994
2022	23.857	21.169	24.367	22.407	27.119	26.323	29.392	28.345	30.003	29.272	29.597	29.324
2023	23.857	21.170	24.341	22.215	27.038	26.263	29.509	28.460	30.141	29.332	29.893	29.512

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.660	22.509	25.480	24.011	26.879	25.623	28.586	27.101	28.723	28.029	28.263	27.754
23.704	22.635	25.286	23.971	27.080	25.781	28.628	27.121	28.811	28.080	28.689	28.274
22.604	22.016	25.195	23.935	27.269	25.838	28.742	27.136	29.002	28.229	28.946	28.507
22.695	22.118	25.447	24.099	27.086	25.713	28.989	27.167	29.226	28.503	29.021	28.530
22.786	22.221	25.387	24.194	27.226	25.812	29.144	27.608	29.255	28.434	29.282	28.686
22.786	22.222	25.384	24.270	27.163	25.755	29.206	27.497	29.466	28.726	29.417	28.881
22.787	22.223	25.342	24.133	27.228	25.749	29.347	27.699	29.592	28.767	29.542	28.962
22.788	22.224	25.098	23.847	27.360	25.856	29.448	27.789	29.666	28.818	29.761	29.161
22.789	22.225	24.936	23.682	27.580	26.128	29.455	27.611	29.855	28.966	29.925	29.243
22.790	22.226	24.687	23.570	27.712	26.131	29.477	27.697	30.016	28.949	30.226	29.487

Mean v's Median Differences:



M16-20	M21-30	M31-45	M46-60	M61-75	M76+
1.359	0.909	0.903	0.879	0.618	0.055
1.735	0.989	0.917	0.917	0.513	0.142
2.705	0.776	0.911	0.871	0.561	0.187
2.698	0.788	0.839	0.964	0.595	0.178
2.690	0.938	0.868	0.964	0.599	0.174
2.690	1.183	0.720	1.000	0.599	0.205
2.689	1.560	0.745	1.037	0.614	0.264
2.688	1.798	0.776	1.011	0.593	0.276
2.688	1.959	0.796	1.047	0.731	0.273
2.687	2.126	0.775	1.050	0.810	0.381

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
1.152	1.469	1.256	1.485	0.694	0.508
1.069	1.316	1.299	1.507	0.732	0.414
0.588	1.261	1.431	1.607	0.774	0.439
0.577	1.348	1.373	1.822	0.723	0.491
0.565	1.193	1.414	1.536	0.821	0.595
0.565	1.114	1.407	1.709	0.740	0.537
0.564	1.209	1.479	1.648	0.824	0.580
0.564	1.252	1.504	1.659	0.848	0.600
0.564	1.254	1.452	1.845	0.888	0.682
0.564	1.117	1.581	1.780	1.067	0.739

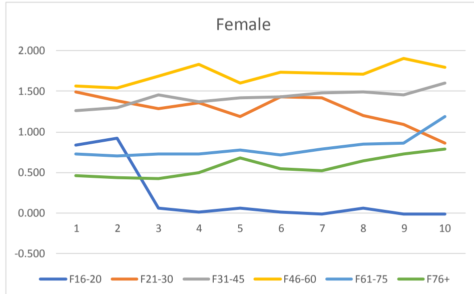


TABLE C.26: Scenario 8 - Raw Forecast

Scenario 8: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	23.053	21.828	26.043	25.270	27.554	26.247	28.866	27.864	28.927	28.510	27.438	27.483
2015	23.427	22.572	25.368	24.641	26.834	25.547	28.614	27.620	28.627	28.333	27.003	27.237
2016	24.250	21.653	25.801	25.200	27.344	26.120	29.006	27.727	28.919	28.348	27.326	27.308
2017	24.978	22.490	25.928	25.385	27.279	25.982	29.136	27.913	29.202	28.572	27.631	27.569
2018	24.579	20.884	25.334	24.316	27.194	26.081	29.264	28.061	29.124	28.328	27.674	27.651
2019	24.142	20.667	25.798	24.830	27.333	26.614	29.096	27.925	29.588	28.735	27.848	27.865
2020	24.635	21.267	25.574	24.432	27.082	26.443	29.176	27.909	29.330	28.497	27.906	28.131
2021	24.301	20.445	26.456	24.667	27.612	26.932	29.252	27.942	29.605	28.652	28.362	28.414
2022	24.209	20.691	26.295	24.589	27.429	26.665	29.333	27.711	29.560	28.501	28.592	28.589
2023	25.339	21.130	26.429	24.544	27.843	26.955	29.230	27.721	29.496	28.365	28.490	28.232

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
24.379	23.251	26.465	24.820	26.819	25.703	28.301	26.842	28.620	28.296	28.357	28.175
25.649	24.012	25.494	24.153	27.148	26.109	28.298	26.702	28.933	28.665	27.982	27.811
24.777	23.170	25.760	24.497	27.239	25.754	28.366	26.727	28.747	28.388	28.262	28.013
24.168	23.165	25.763	24.508	26.720	25.509	28.328	26.580	29.021	28.426	28.124	28.043
23.365	22.650	26.146	24.977	26.648	25.159	28.429	27.091	28.680	28.231	27.824	27.306
23.262	22.648	25.829	24.666	27.077	25.653	28.570	27.111	28.877	28.301	28.067	27.682
23.691	23.758	26.467	24.862	27.010	25.393	28.788	27.249	29.036	28.021	27.869	27.703
22.856	22.632	26.426	24.893	27.483	26.057	28.822	27.283	29.083	28.377	28.369	27.851
23.592	22.762	25.903	24.254	27.796	26.219	28.756	27.074	28.744	27.771	28.623	28.354
23.198	22.852	26.212	24.478	27.589	25.887	28.669	27.001	28.548	27.364	28.131	27.756

Scenario 8 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.061	0.000	-0.028	-0.067	-0.018	-0.070	-0.003	0.022	0.010	0.024	0.031	-0.025
2015	0.030	0.001	-0.045	-0.009	0.003	0.011	0.006	-0.007	0.003	0.005	0.047	0.024
2016	0.001	0.001	0.004	0.013	0.002	-0.003	0.019	0.006	0.002	0.004	0.054	-0.021
2017	0.001	0.002	0.006	0.044	-0.012	-0.026	0.023	0.004	0.000	0.010	0.063	0.006
2018	0.002	0.003	0.039	0.009	-0.005	-0.007	0.026	0.004	-0.013	-0.022	0.039	-0.003
2019	0.003	0.005	0.034	0.011	-0.004	0.005	0.026	0.008	-0.015	-0.057	0.053	0.018
2020	0.003	0.006	0.025	-0.009	0.016	-0.027	0.022	0.003	-0.012	-0.037	0.013	0.017
2021	0.004	0.007	0.047	0.008	0.010	0.003	0.016	0.001	0.006	-0.019	0.017	0.018
2022	0.004	0.008	0.031	0.011	0.008	-0.020	0.026	0.017	-0.006	-0.023	0.019	0.036
2023	0.005	0.009	0.038	0.013	-0.004	-0.019	0.017	0.024	0.020	0.002	0.024	0.015

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
-0.057	-0.200	-0.018	-0.039	0.006	0.025	0.035	0.110	-0.007	0.006	-0.017	0.061
-0.019	-0.125	-0.050	-0.024	0.020	0.024	0.051	0.069	-0.024	0.002	0.017	0.087
0.001	0.002	-0.043	-0.007	0.021	0.030	0.020	0.146	0.005	-0.007	0.010	-0.006
0.002	0.003	-0.059	-0.084	0.023	-0.005	0.051	0.041	-0.014	0.017	-0.029	-0.029
0.003	0.004	-0.094	-0.067	0.017	0.000	0.022	0.083	0.015	0.005	-0.032	0.004
0.004	0.005	-0.038	-0.081	0.004	0.005	0.017	0.044	0.012	-0.028	0.003	0.037
0.005	0.006	-0.014	-0.037	0.012	-0.016	0.004	0.069	0.040	0.001	0.008	0.022
0.005	0.007	-0.005	-0.103	0.015	-0.018	0.001	0.033	0.026	-0.016	0.007	0.038
0.006	0.008	-0.030	-0.018	0.021	-0.001	0.031	0.045	0.041	0.027	0.038	0.101
0.007	0.009	-0.013	-0.009	0.019	0.007	0.067	0.032	0.056	0.099	0.016	0.039

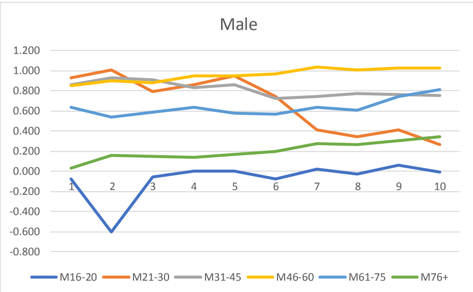
TABLE C.27: Scenario 8- Adjusted Forecast and Comparison

Scenario 9 Forecast:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.454	22.527	25.062	24.134	27.568	26.707	28.516	27.662	28.851	28.220	28.037	28.009
2015	22.322	22.921	24.786	23.775	27.430	26.501	28.819	27.919	29.051	28.509	28.149	27.994
2016	23.014	23.073	24.314	23.517	27.523	26.614	28.848	27.970	29.154	28.570	28.393	28.245
2017	22.981	22.980	24.454	23.598	27.431	26.604	29.022	28.075	29.214	28.576	28.622	28.484
2018	23.040	23.041	24.324	23.376	27.488	26.631	29.079	28.135	29.433	28.855	28.739	28.575
2019	22.928	23.001	24.242	23.501	27.351	26.628	29.157	28.193	29.718	29.148	28.898	28.704
2020	22.814	22.792	24.018	23.602	27.333	26.586	29.270	28.238	29.809	29.173	29.142	28.867
2021	22.696	22.727	23.947	23.603	27.298	26.524	29.362	28.354	29.899	29.293	29.247	28.982
2022	22.482	22.417	23.983	23.571	27.145	26.383	29.380	28.354	30.020	29.272	29.563	29.255
2023	22.437	22.447	23.906	23.640	27.052	26.301	29.486	28.460	30.159	29.349	29.846	29.500

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
23.837	23.000	25.539	24.050	26.872	25.611	28.582	27.016	28.713	27.983	28.280	27.820
23.919	22.997	25.379	23.999	27.064	25.759	28.636	27.094	28.799	28.095	28.716	28.275
23.013	22.950	25.274	23.983	27.262	25.806	28.749	27.056	28.973	28.237	28.970	28.549
23.100	23.086	25.506	24.150	27.086	25.719	28.987	27.155	29.202	28.475	29.059	28.563
23.087	23.023	25.451	24.254	27.249	25.828	29.154	27.548	29.210	28.434	29.363	28.686
22.996	22.979	25.464	24.026	27.182	25.755	29.210	27.477	29.455	28.736	29.461	28.914
22.897	22.903	25.430	24.007	27.245	25.763	29.362	27.642	29.548	28.757	29.568	29.039
22.817	22.759	25.133	23.932	27.382	25.885	29.471	27.765	29.630	28.785	29.802	29.161
22.814	22.819	25.070	23.980	27.589	26.136	29.460	27.557	29.803	28.944	29.972	29.246
22.690	22.701	24.735	23.875	27.728	26.131	29.485	27.683	29.957	28.766	30.284	29.490

Mean v's Median Differences:



M16-20	M21-30	M31-45	M46-60	M61-75	M76+
-0.073	0.928	0.861	0.854	0.632	0.028
-0.600	1.011	0.929	0.900	0.542	0.155
-0.059	0.797	0.909	0.878	0.584	0.148
0.001	0.856	0.828	0.947	0.638	0.138
-0.001	0.948	0.857	0.943	0.578	0.164
-0.073	0.742	0.722	0.964	0.570	0.194
0.022	0.416	0.747	1.033	0.636	0.274
-0.031	0.344	0.774	1.009	0.606	0.265
0.064	0.412	0.762	1.026	0.748	0.309
-0.011	0.266	0.752	1.026	0.810	0.346

F16-20	F21-30	F31-45	F46-60	F61-75	F76+
0.836	1.489	1.261	1.566	0.730	0.460
0.922	1.379	1.305	1.542	0.704	0.441
0.063	1.292	1.457	1.693	0.735	0.421
0.013	1.356	1.367	1.832	0.728	0.496
0.064	1.196	1.421	1.606	0.776	0.677
0.016	1.438	1.427	1.733	0.718	0.546
-0.007	1.423	1.482	1.720	0.791	0.528
0.057	1.201	1.497	1.707	0.845	0.641
-0.005	1.089	1.453	1.903	0.858	0.726
-0.011	0.860	1.597	1.803	1.191	0.795

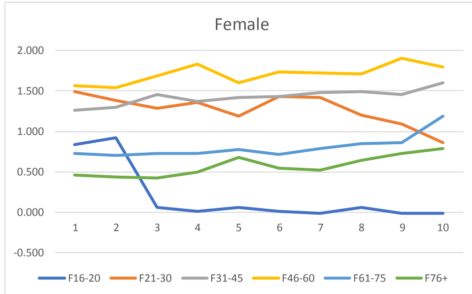


TABLE C.28: Scenario 9 - Raw Forecast

Scenario 9: Adjusted Forecast

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	22.811	23.018	26.063	25.270	27.547	26.282	28.862	27.885	28.909	28.479	27.435	27.508
2015	22.995	24.476	25.391	24.641	26.836	25.537	28.605	27.628	28.623	28.299	26.993	27.214
2016	23.541	23.708	25.797	25.176	27.359	26.137	28.985	27.698	28.898	28.305	27.290	27.311
2017	24.171	24.379	25.928	25.317	27.297	26.012	29.118	27.913	29.199	28.526	27.575	27.553
2018	23.765	22.760	25.329	24.301	27.191	26.089	29.251	28.068	29.129	28.354	27.646	27.634
2019	23.215	22.503	25.705	25.179	27.335	26.614	29.091	27.956	29.585	28.761	27.823	27.851
2020	23.593	22.892	25.405	25.407	27.092	26.451	29.186	27.924	29.339	28.484	27.900	28.114
2021	23.141	22.005	26.201	25.865	27.627	26.951	29.261	27.954	29.616	28.650	28.340	28.402
2022	22.834	21.940	25.911	25.753	27.455	26.725	29.321	27.720	29.577	28.501	28.558	28.520
2023	23.919	22.407	25.994	25.969	27.857	26.993	29.207	27.722	29.513	28.382	28.443	28.219

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
24.556	23.743	26.525	24.859	26.812	25.690	28.297	26.758	28.610	28.250	28.374	28.241
25.864	24.374	25.586	24.182	27.132	26.086	28.306	26.675	28.921	28.680	28.009	27.812
25.187	24.105	25.839	24.545	27.232	25.721	28.372	26.648	28.717	28.397	28.285	28.054
24.573	24.134	25.822	24.559	26.721	25.515	28.325	26.568	28.997	28.397	28.162	28.076
23.667	23.453	26.209	25.036	26.672	25.174	28.440	27.031	28.636	28.231	27.906	27.306
23.471	23.405	25.908	24.422	27.097	25.653	28.574	27.091	28.865	28.311	28.110	27.715
23.801	24.439	26.555	24.736	27.027	25.406	28.803	27.192	28.992	28.011	27.894	27.781
22.885	23.167	26.460	24.978	27.505	26.086	28.846	27.258	29.046	28.344	28.411	27.851
23.617	23.357	26.036	24.553	27.805	26.227	28.760	27.020	28.692	27.749	28.670	28.357
23.098	23.326	26.260	24.783	27.606	25.887	28.677	26.986	28.488	27.181	28.190	27.758

Scenario 9 v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.181	1.190	-0.008	-0.066	-0.026	-0.034	-0.007	0.043	-0.007	-0.007	0.027	0.000
2015	-0.402	1.904	-0.023	-0.009	0.005	0.001	-0.003	0.001	-0.002	-0.029	0.037	0.001
2016	-0.709	2.056	0.000	-0.011	0.017	0.013	-0.003	-0.023	-0.019	-0.040	0.018	-0.018
2017	-0.806	1.891	0.007	-0.024	0.006	0.004	0.006	0.003	-0.003	-0.036	0.007	-0.010
2018	-0.812	1.880	0.035	-0.006	-0.008	0.000	0.014	0.012	-0.008	0.005	0.012	-0.021
2019	-0.924	1.840	-0.059	0.360	-0.002	0.005	0.021	0.039	-0.018	-0.031	0.028	0.004
2020	-1.039	1.631	-0.144	0.966	0.026	-0.019	0.033	0.017	-0.003	-0.051	0.007	0.000
2021	-1.156	1.566	-0.208	1.206	0.026	0.022	0.025	0.013	0.017	-0.021	-0.006	0.006
2022	-1.371	1.256	-0.353	1.175	0.034	0.040	0.014	0.026	0.011	-0.023	-0.015	-0.034
2023	-1.416	1.286	-0.398	1.438	0.010	0.019	-0.006	0.025	0.038	0.019	-0.022	0.003

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.119	0.291	0.042	0.000	-0.002	0.013	0.031	0.025	-0.017	-0.039	0.001	0.127
0.196	0.237	0.042	0.004	0.005	0.001	0.059	0.042	-0.036	0.017	0.044	0.087
0.411	0.936	0.035	0.041	0.014	-0.003	0.027	0.067	-0.025	0.002	0.034	0.036
0.407	0.971	0.000	-0.033	0.023	0.001	0.048	0.029	-0.038	-0.011	0.010	0.003
0.305	0.807	-0.031	-0.007	0.040	0.016	0.033	0.024	-0.029	0.005	0.050	0.004
0.213	0.762	0.042	-0.325	0.024	0.005	0.021	0.024	0.000	-0.018	0.047	0.071
0.114	0.687	0.074	-0.162	0.030	-0.002	0.019	0.013	-0.004	-0.010	0.033	0.100
0.034	0.543	0.029	-0.018	0.037	0.011	0.024	0.009	-0.011	-0.049	0.048	0.039
0.031	0.603	0.103	0.280	0.030	0.007	0.036	-0.009	-0.011	0.005	0.085	0.104
-0.093	0.484	0.035	0.295	0.036	0.007	0.076	0.018	-0.003	-0.083	0.075	0.042

TABLE C.29: Scenario 9- Adjusted Forecast and Comparison

C.3 Extreme Scenarios

Rising New Agent BMI v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.022	1.372	0.014	-0.016	-0.032	-0.120	0.001	0.060	-0.006	-0.007	0.069	0.037
2015	-0.013	2.535	-0.034	-0.001	-0.021	-0.017	-0.016	0.004	-0.016	-0.011	0.090	0.083
2016	0.288	3.081	-0.007	-0.016	0.014	-0.005	-0.018	0.016	0.008	-0.020	0.060	0.035
2017	0.543	3.211	0.015	-0.026	0.003	-0.039	-0.018	0.015	0.002	-0.036	0.051	0.006
2018	0.662	3.336	0.021	-0.005	-0.015	-0.004	0.011	0.038	-0.015	-0.002	0.029	0.000
2019	1.066	3.744	-0.007	0.690	0.012	-0.005	-0.009	0.029	-0.027	-0.033	0.045	0.017
2020	1.449	4.051	0.041	1.422	0.011	0.002	0.018	0.013	-0.023	-0.058	0.037	0.011
2021	1.849	4.473	0.079	1.860	0.021	0.011	0.024	0.016	-0.012	-0.025	0.038	0.009
2022	2.374	5.109	0.223	2.410	0.007	0.030	-0.005	0.008	-0.011	-0.006	0.011	-0.011
2023	2.745	5.478	0.299	2.805	0.006	-0.001	-0.012	0.018	-0.004	-0.004	0.007	0.000

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.186	0.748	0.050	0.000	-0.020	0.000	-0.012	-0.042	-0.001	0.000	-0.030	-0.071
0.607	0.986	0.026	0.000	-0.027	-0.006	0.014	0.015	-0.007	0.032	-0.010	-0.068
1.480	2.064	0.004	-0.011	-0.028	0.003	0.000	0.023	0.012	0.000	-0.040	-0.103
1.626	2.194	-0.011	-0.003	-0.041	-0.023	0.027	0.009	-0.017	-0.015	-0.079	-0.203
1.883	2.479	-0.020	-0.015	-0.025	-0.029	0.009	-0.075	-0.011	-0.015	-0.027	-0.071
2.269	2.848	0.114	0.018	-0.043	-0.002	0.004	0.021	0.013	-0.039	-0.025	-0.001
2.661	3.267	0.226	0.404	-0.022	-0.061	0.005	0.018	0.014	-0.011	-0.002	-0.008
3.065	3.637	0.367	0.929	0.000	-0.062	-0.009	0.009	0.022	-0.026	-0.010	-0.019
3.381	3.848	0.585	1.362	0.006	-0.046	-0.019	0.015	0.026	-0.104	0.032	0.043
3.789	4.332	0.935	1.830	0.012	-0.015	0.008	0.008	0.016	-0.099	-0.001	0.024

Falling New Agent BMI v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.136	1.262	-0.013	-0.067	-0.012	-0.003	0.032	0.070	-0.007	-0.007	0.011	0.030
2015	-0.342	2.003	-0.041	-0.013	0.007	0.015	0.018	0.004	-0.028	-0.059	0.028	0.016
2016	-0.510	2.191	-0.018	-0.029	0.026	0.007	0.009	0.048	-0.004	-0.055	0.012	0.004
2017	-0.537	2.167	-0.015	-0.020	0.004	-0.012	0.028	0.046	-0.005	-0.038	0.012	-0.003
2018	-0.520	2.178	-0.003	-0.005	-0.002	-0.004	0.031	0.040	-0.009	-0.022	0.011	-0.004
2019	-0.529	2.158	-0.064	0.379	-0.010	-0.005	0.033	0.050	-0.009	-0.043	0.006	0.008
2020	-0.544	2.145	-0.113	1.050	-0.001	-0.021	0.022	0.010	0.002	-0.052	0.012	0.011
2021	-0.559	2.139	-0.147	1.363	0.004	0.005	0.019	0.012	-0.001	-0.023	-0.002	0.005
2022	-0.625	2.071	-0.244	1.485	-0.001	0.015	0.021	0.016	-0.005	-0.025	-0.004	-0.001
2023	-0.588	2.117	-0.274	1.711	-0.011	-0.010	0.003	0.025	0.001	0.003	0.007	-0.002

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.195	0.383	-0.022	0.000	0.022	0.025	0.012	0.006	-0.042	-0.011	-0.002	-0.005
0.302	0.412	-0.025	-0.038	-0.001	0.024	0.030	0.031	-0.054	0.000	-0.002	-0.039
0.624	1.175	0.002	-0.008	-0.002	0.028	-0.008	0.055	-0.018	0.000	-0.038	-0.004
0.649	1.222	-0.028	-0.035	-0.014	-0.002	0.033	0.031	-0.036	-0.037	-0.047	-0.066
0.618	1.101	-0.073	-0.022	0.006	-0.008	0.015	0.004	-0.030	-0.031	-0.031	-0.071
0.621	1.182	0.003	-0.234	-0.028	-0.050	0.017	0.022	-0.004	-0.046	-0.020	-0.001
0.620	1.226	0.039	-0.140	-0.016	-0.036	0.016	0.011	-0.031	-0.019	-0.020	0.013
0.636	1.191	0.044	0.136	-0.008	-0.031	0.020	0.008	-0.029	-0.030	-0.003	-0.017
0.696	1.261	0.138	0.500	-0.007	-0.016	0.022	0.014	-0.034	-0.089	-0.004	0.042
0.677	1.268	0.213	0.596	0.018	-0.006	0.033	0.003	-0.017	-0.102	-0.011	0.024

TABLE C.30: BMI - Extreme Rates of Change

Rising APCCC v's Scenario 1

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.034	0.000	0.014	-0.040	-0.054	-0.034	0.004	0.022	0.023	0.013	-0.017	-0.044
2015	0.017	0.001	-0.016	-0.009	-0.021	0.001	0.004	0.003	0.020	-0.008	0.011	0.013
2016	0.002	0.002	0.007	-0.001	0.004	0.019	0.001	0.008	0.028	-0.010	0.012	0.003
2017	0.003	0.004	0.006	0.004	-0.012	-0.018	0.006	0.023	0.016	-0.023	0.010	0.001
2018	0.004	0.007	0.037	0.001	-0.021	0.003	0.015	0.043	0.030	-0.010	-0.006	-0.012
2019	0.005	0.009	0.035	0.004	0.008	0.007	-0.001	0.043	0.039	-0.002	0.010	0.005
2020	0.006	0.011	-0.015	-0.008	0.013	-0.001	0.009	0.021	0.045	-0.009	-0.025	0.000
2021	0.007	0.014	0.017	0.014	0.021	0.044	0.011	0.032	0.056	-0.054	-0.010	0.017
2022	0.008	0.016	-0.006	0.022	0.014	0.051	0.006	0.037	0.031	0.008	-0.011	-0.008
2023	0.010	0.018	0.023	0.015	0.008	0.037	-0.018	0.037	0.051	0.029	0.019	0.018

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.019	-0.034	-0.049	-0.032	-0.021	0.001	0.032	0.031	-0.002	0.007	-0.031	-0.005
2015	0.083	0.163	-0.045	-0.010	-0.009	-0.001	0.032	0.043	0.007	0.035	0.008	-0.024
2016	0.002	0.003	0.007	0.003	0.005	0.038	0.008	0.047	0.008	0.030	-0.004	-0.049
2017	0.004	0.006	-0.013	-0.033	-0.005	0.004	0.026	0.031	0.000	0.038	-0.018	-0.061
2018	0.006	0.008	-0.027	-0.028	-0.010	-0.014	0.038	0.034	0.023	0.019	0.043	-0.013
2019	0.008	0.010	0.027	-0.045	-0.017	0.000	0.012	0.052	0.030	0.060	0.052	0.010
2020	0.009	0.012	0.036	-0.013	0.014	-0.014	0.014	0.034	0.043	0.024	0.019	0.007
2021	0.011	0.015	0.052	0.017	0.029	-0.024	0.021	0.034	0.035	0.006	0.036	-0.003
2022	0.013	0.017	0.070	0.017	0.036	-0.021	0.024	0.036	0.027	0.018	0.046	0.110
2023	0.014	0.019	0.080	0.066	0.027	-0.003	0.049	0.045	0.046	0.119	0.040	0.054

Falling APCCC v's Scenario 1

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.045	0.000	-0.008	-0.066	0.008	-0.004	0.007	0.070	-0.014	0.007	0.012	0.030
2015	-0.031	-0.001	-0.025	-0.014	0.013	0.000	0.019	0.006	-0.024	-0.010	0.026	0.012
2016	-0.002	-0.002	-0.016	-0.022	0.032	0.007	0.014	0.070	-0.015	-0.022	0.038	0.010
2017	-0.003	-0.004	-0.019	-0.066	0.017	-0.004	0.026	0.066	-0.007	-0.020	-0.018	-0.001
2018	-0.004	-0.007	0.012	-0.011	-0.012	-0.011	0.024	0.036	0.005	0.018	0.002	-0.013
2019	-0.005	-0.009	-0.024	-0.013	-0.005	-0.011	0.010	0.014	-0.005	-0.010	0.010	-0.004
2020	-0.006	-0.011	-0.039	-0.041	0.001	-0.009	-0.001	-0.001	-0.013	-0.019	0.009	-0.005
2021	-0.007	-0.014	-0.024	-0.018	0.005	-0.012	0.000	0.005	0.011	-0.020	0.003	-0.015
2022	-0.008	-0.016	-0.037	-0.022	-0.026	0.001	0.002	0.007	0.002	-0.022	-0.042	-0.052
2023	-0.010	-0.018	-0.034	-0.012	-0.039	-0.040	-0.023	0.010	-0.003	-0.011	-0.022	-0.021

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.083	-0.128	-0.044	-0.011	-0.015	0.029	-0.004	-0.053	0.012	0.047	-0.011	-0.071
2015	0.083	0.045	-0.044	-0.004	-0.033	0.007	0.024	-0.009	-0.010	0.104	-0.010	-0.079
2016	-0.002	-0.003	-0.006	0.042	-0.011	0.023	0.002	-0.008	-0.009	0.035	-0.035	-0.079
2017	-0.004	-0.006	0.000	-0.017	-0.025	-0.029	0.026	-0.004	-0.039	0.027	-0.037	-0.094
2018	-0.006	-0.008	-0.024	-0.017	-0.010	-0.033	0.028	0.016	-0.017	-0.006	-0.026	-0.062
2019	-0.008	-0.010	0.028	-0.005	-0.012	-0.043	-0.014	0.010	-0.004	-0.007	0.001	-0.013
2020	-0.009	-0.012	0.038	-0.008	0.004	-0.040	-0.022	0.004	-0.016	-0.027	-0.013	0.011
2021	-0.011	-0.015	0.022	-0.022	0.009	-0.036	-0.020	-0.015	-0.033	-0.045	-0.021	-0.011
2022	-0.013	-0.017	0.012	-0.019	-0.002	-0.034	-0.022	0.002	-0.042	-0.045	-0.003	0.080
2023	-0.014	-0.019	0.034	-0.035	-0.038	-0.046	-0.017	-0.026	-0.021	-0.074	-0.041	0.013

TABLE C.31: APCCC - Extreme Rates of Change

C.4 Cross-Comparisons

Scenario 7 v's Scenario 4:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.019	0.108	0.009	0.029	0.005	-0.053	-0.012	-0.006	0.019	0.033	-0.024	-0.046
2015	0.144	0.260	0.011	-0.003	0.005	0.016	-0.023	0.002	0.023	0.049	-0.061	-0.014
2016	0.401	0.303	0.014	0.000	0.021	-0.008	-0.010	0.011	0.024	0.038	-0.019	-0.025
2017	0.543	0.540	0.012	-0.017	0.005	-0.016	-0.008	0.006	0.022	0.033	-0.021	0.000
2018	0.595	0.553	-0.004	0.004	0.005	0.001	-0.016	-0.007	0.018	0.056	0.000	0.011
2019	0.803	0.686	-0.010	0.166	-0.005	0.005	-0.040	0.010	0.037	0.035	0.020	0.010
2020	1.003	0.998	0.055	0.242	-0.004	0.013	-0.032	-0.015	0.050	0.052	0.000	0.022
2021	1.211	1.132	0.115	0.359	0.005	0.003	-0.018	-0.008	0.049	0.067	0.036	0.019
2022	1.507	1.523	0.221	0.539	-0.007	0.020	-0.009	0.006	0.038	0.017	0.058	0.067
2023	1.675	1.632	0.290	0.598	0.000	0.042	0.024	0.007	0.047	0.033	0.050	0.029

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.025	0.124	-0.055	-0.001	-0.016	-0.013	-0.027	-0.089	-0.005	-0.016	-0.004	0.041
2015	0.125	0.289	-0.056	-0.021	-0.002	-0.016	-0.008	-0.038	-0.037	-0.014	0.013	0.016
2016	0.430	0.501	-0.063	-0.035	-0.026	-0.038	-0.009	-0.010	-0.039	-0.002	0.020	0.040
2017	0.492	0.483	-0.069	-0.072	-0.030	-0.011	0.003	0.017	-0.032	-0.039	0.036	0.064
2018	0.638	0.673	-0.053	-0.033	-0.010	0.009	0.014	-0.012	-0.036	-0.048	0.046	0.054
2019	0.830	0.834	-0.057	0.193	-0.014	0.013	0.007	-0.010	-0.035	-0.123	0.017	0.022
2020	1.028	1.037	-0.014	0.173	-0.007	0.023	-0.011	-0.011	-0.023	-0.027	0.035	0.003
2021	1.223	1.296	0.157	0.340	0.009	-0.008	0.017	0.005	-0.018	0.020	0.028	0.011
2022	1.354	1.347	0.191	0.468	0.002	0.009	0.015	0.007	-0.004	0.020	0.044	0.049
2023	1.568	1.559	0.336	0.685	0.024	0.033	0.019	0.022	-0.005	0.022	0.026	0.085

Scenario 9 v's Scenario 2:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.048	-0.112	-0.003	0.000	0.009	0.087	0.010	0.015	-0.004	-0.001	0.015	-0.030
2015	-0.123	-0.258	0.020	0.005	0.027	0.019	0.004	0.001	0.022	0.019	0.017	-0.018
2016	-0.397	-0.300	0.031	0.054	0.001	0.010	-0.011	-0.033	-0.002	0.002	-0.018	-0.012
2017	-0.538	-0.534	0.048	0.058	0.013	0.030	-0.008	-0.026	0.024	0.001	-0.011	-0.005
2018	-0.586	-0.542	0.041	0.009	0.019	0.015	0.002	0.011	0.025	0.025	-0.001	-0.003
2019	-0.791	-0.675	0.000	-0.147	0.013	0.024	0.009	0.039	0.013	0.003	0.008	0.016
2020	-0.990	-0.985	-0.039	-0.214	0.023	0.008	0.015	0.027	0.020	-0.031	0.009	0.020
2021	-1.195	-1.110	-0.098	-0.337	0.028	0.025	0.013	0.011	0.028	0.020	0.009	0.024
2022	-1.491	-1.506	-0.189	-0.502	0.055	0.055	0.018	0.021	0.035	-0.001	0.025	-0.002
2023	-1.656	-1.603	-0.228	-0.547	0.035	0.053	0.002	0.008	0.042	0.033	-0.002	0.028

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.062	-0.199	0.053	0.063	-0.004	-0.017	0.005	-0.046	0.012	-0.039	0.069	0.238
2015	-0.165	-0.286	0.053	0.017	0.013	-0.022	0.025	0.000	0.015	0.019	0.072	0.153
2016	-0.426	-0.495	0.024	0.055	0.000	-0.055	0.019	0.014	0.000	0.004	0.084	0.140
2017	-0.485	-0.473	0.012	0.006	0.029	0.007	0.009	0.016	0.001	0.007	0.084	0.241
2018	-0.627	-0.659	0.013	0.018	0.052	0.033	0.009	-0.010	0.002	0.027	0.101	0.079
2019	-0.817	-0.824	-0.015	-0.169	0.040	0.037	-0.007	-0.009	0.006	-0.016	0.093	0.113
2020	-1.012	-1.021	-0.016	-0.122	0.034	0.044	-0.001	-0.011	-0.001	-0.005	0.078	0.115
2021	-1.204	-1.274	-0.096	-0.311	0.042	0.058	0.019	-0.001	0.000	-0.025	0.103	0.062
2022	-1.329	-1.319	-0.150	-0.401	0.057	0.054	0.025	-0.029	0.011	0.012	0.104	0.079
2023	-1.541	-1.534	-0.334	-0.607	0.039	0.046	0.055	0.018	0.020	-0.152	0.114	0.055

TABLE C.32: Cross Scenario Comparisons - 1

Scenario 5 v's Scenario 6:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.010	0.080	-0.021	-0.050	0.015	-0.043	-0.001	-0.022	0.042	0.013	0.012	0.041
2015	0.124	0.259	-0.039	0.000	0.025	0.011	-0.006	-0.002	0.052	0.040	0.030	0.011
2016	0.399	0.302	-0.038	-0.032	0.019	0.012	0.002	-0.057	0.035	0.036	-0.015	0.015
2017	0.540	0.537	-0.037	-0.082	0.004	0.050	-0.009	-0.055	0.019	0.021	-0.001	0.011
2018	0.590	0.547	-0.050	-0.009	0.001	0.001	0.005	-0.020	0.024	0.014	0.015	0.012
2019	0.797	0.680	-0.021	0.159	0.000	0.003	-0.018	-0.035	0.031	0.038	0.028	0.009
2020	0.996	0.992	0.024	0.229	-0.019	0.002	-0.010	-0.013	0.014	0.003	0.005	-0.015
2021	1.203	1.121	0.089	0.352	-0.030	-0.007	-0.005	0.000	-0.010	0.022	-0.006	-0.006
2022	1.499	1.515	0.182	0.522	-0.030	-0.011	-0.012	-0.010	0.035	0.011	-0.011	-0.019
2023	1.666	1.618	0.253	0.577	-0.016	-0.032	-0.013	0.005	0.034	0.028	-0.042	0.000

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.025	0.124	-0.006	0.025	0.033	0.031	0.054	0.017	-0.016	-0.005	0.031	0.091
2015	0.076	0.288	-0.012	-0.005	0.052	0.074	0.039	0.013	0.024	-0.023	-0.019	0.072
2016	0.428	0.498	-0.051	-0.011	0.039	0.057	0.024	0.063	0.040	-0.020	-0.014	0.025
2017	0.489	0.478	-0.039	-0.078	0.017	0.022	0.020	-0.016	0.029	-0.006	-0.026	0.023
2018	0.632	0.666	-0.037	-0.031	0.025	0.040	0.029	-0.001	0.011	-0.052	-0.033	0.055
2019	0.823	0.828	-0.077	0.126	0.029	0.049	0.037	0.003	-0.023	0.005	-0.068	0.000
2020	1.020	1.029	-0.044	0.137	0.024	0.026	0.043	0.007	-0.019	-0.024	-0.058	0.020
2021	1.214	1.285	0.077	0.332	0.002	0.032	0.020	0.008	0.018	0.008	-0.039	0.030
2022	1.342	1.331	0.129	0.427	-0.009	-0.002	0.016	0.039	0.037	-0.023	-0.019	0.002
2023	1.554	1.547	0.305	0.635	-0.034	0.009	0.040	0.006	0.049	0.002	-0.047	0.000

Scenario 8 v's Scenario 3:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.032	0.000	-0.031	0.000	0.006	0.050	-0.026	-0.075	0.000	0.011	0.012	-0.049
2015	0.034	0.001	-0.018	0.001	0.015	0.012	-0.008	-0.010	0.006	0.018	-0.007	0.025
2016	0.002	0.002	-0.006	0.019	0.004	0.016	-0.006	-0.065	-0.011	0.024	0.019	-0.044
2017	0.003	0.004	-0.003	0.043	-0.010	0.029	-0.003	-0.071	0.017	0.050	0.015	0.007
2018	0.004	0.007	0.007	0.014	-0.001	0.005	-0.009	-0.032	-0.013	-0.036	0.025	0.007
2019	0.005	0.009	0.050	0.025	-0.009	0.016	-0.008	-0.042	-0.007	-0.053	0.032	0.007
2020	0.006	0.011	0.060	0.014	0.005	0.007	-0.026	-0.023	-0.006	0.026	-0.016	0.013
2021	0.007	0.014	0.066	0.022	0.010	0.007	-0.013	-0.014	0.004	0.020	0.008	0.016
2022	0.008	0.016	0.050	0.022	0.019	-0.024	0.013	0.002	-0.013	0.006	0.045	0.051
2023	0.010	0.018	0.043	0.042	0.011	0.006	0.018	0.007	0.014	0.025	0.039	0.027

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.079	-0.120	0.024	-0.028	0.015	0.025	0.052	0.083	0.028	0.007	0.001	0.072
2015	-0.016	-0.077	-0.022	-0.027	0.034	0.047	0.053	0.069	0.006	-0.030	0.021	0.137
2016	0.002	0.003	-0.042	-0.020	0.027	0.016	0.032	0.150	0.023	-0.031	0.059	0.049
2017	0.004	0.006	-0.034	-0.046	0.040	0.000	0.046	0.038	0.015	0.000	0.018	0.064
2018	0.006	0.008	-0.061	-0.054	0.053	0.035	0.020	0.144	0.039	0.007	-0.011	0.052
2019	0.008	0.010	-0.037	-0.075	0.038	0.061	0.033	0.029	0.028	0.017	0.014	0.095
2020	0.009	0.012	-0.030	-0.010	0.040	0.003	0.033	0.076	0.034	0.018	0.050	0.075
2021	0.011	0.015	-0.019	-0.036	0.022	0.006	0.039	0.045	0.039	0.021	0.044	0.078
2022	0.013	0.017	-0.025	0.005	0.029	0.028	0.069	0.062	0.078	0.074	0.061	0.117
2023	0.014	0.019	-0.004	0.008	0.028	0.063	0.085	0.046	0.071	0.217	0.049	0.098

TABLE C.33: Cross Comparisons - 2

C.5 Counterfactuals

Scenario 1 (no network effect) v's Scenario 1:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.013	0.000	0.009	-0.008	-0.009	-0.024	0.004	0.014	-0.014	-0.009	0.022	0.030
2015	0.001	0.000	-0.005	-0.010	-0.004	-0.010	-0.013	-0.004	-0.008	-0.067	0.020	0.016
2016	0.000	0.000	0.008	-0.013	0.037	0.011	-0.006	-0.002	-0.013	-0.063	0.038	0.000
2017	0.000	0.000	-0.003	-0.026	0.021	-0.024	-0.009	0.002	-0.002	-0.049	0.017	0.003
2018	0.000	0.000	0.025	-0.012	0.010	-0.001	0.017	0.016	-0.010	-0.026	0.008	-0.005
2019	0.000	0.000	-0.013	-0.020	0.016	-0.008	0.024	0.013	-0.018	-0.047	0.046	0.007
2020	0.000	0.000	-0.007	-0.038	0.028	-0.026	0.019	0.014	-0.001	-0.059	0.022	0.011
2021	0.000	0.000	0.002	0.000	0.035	0.029	0.021	0.014	-0.010	-0.092	-0.011	0.008
2022	0.000	0.000	-0.003	0.000	0.022	0.031	0.015	0.020	-0.012	-0.060	-0.014	0.006
2023	0.000	0.000	0.009	0.000	-0.001	-0.011	0.011	0.018	-0.003	-0.036	-0.006	0.000

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.069	-0.127	-0.017	-0.005	0.006	0.025	-0.001	-0.039	-0.020	0.000	-0.036	-0.005
0.114	0.162	-0.021	-0.014	-0.008	0.024	0.009	0.018	-0.005	0.033	-0.039	-0.009
0.000	0.000	-0.006	-0.011	-0.001	0.091	-0.029	-0.002	0.013	0.032	-0.052	-0.004
0.000	0.000	-0.030	-0.080	0.011	0.043	0.002	0.008	0.000	0.033	-0.066	-0.066
0.000	0.000	-0.070	-0.070	0.014	0.007	0.006	0.006	-0.005	0.001	-0.029	-0.021
0.000	0.000	-0.032	-0.087	0.016	-0.001	-0.001	0.019	-0.001	-0.008	-0.032	-0.001
0.000	0.000	-0.021	-0.045	0.016	-0.026	-0.011	-0.003	0.029	0.001	-0.024	0.034
0.000	0.000	0.003	-0.073	0.016	0.000	0.010	-0.002	0.012	-0.046	-0.019	0.001
0.000	0.000	0.013	-0.027	0.022	-0.016	-0.012	-0.009	-0.010	-0.035	0.016	0.097
0.000	0.000	0.060	0.006	-0.002	-0.011	0.025	0.007	0.015	-0.100	-0.005	0.022

TABLE C.34: Comparison - No Network Effect Scenario 1.

Scenario 4 (no network effect) v's Scenario 4:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.037	0.080	-0.037	-0.022	0.009	-0.035	-0.016	-0.024	0.004	0.033	0.015	-0.003
2015	0.191	0.259	-0.022	-0.025	-0.012	0.000	-0.012	-0.001	0.002	0.014	-0.021	0.016
2016	0.399	0.302	0.004	-0.008	-0.001	-0.011	-0.009	-0.004	0.033	0.029	-0.026	0.014
2017	0.540	0.537	0.012	-0.074	-0.033	-0.026	-0.012	0.003	0.025	0.004	0.020	-0.001
2018	0.591	0.545	-0.004	-0.010	-0.024	-0.005	-0.042	-0.036	-0.001	0.040	0.018	0.009
2019	0.797	0.681	-0.008	0.157	-0.022	-0.011	-0.026	-0.002	0.023	0.006	0.019	0.002
2020	0.996	0.991	0.067	0.229	-0.025	0.000	-0.060	-0.016	0.032	0.028	-0.006	0.011
2021	1.203	1.119	0.115	0.344	-0.022	0.000	-0.052	-0.020	0.033	0.060	0.018	0.008
2022	1.499	1.512	0.210	0.519	-0.015	0.001	-0.026	-0.017	0.027	-0.001	0.048	0.019
2023	1.666	1.618	0.275	0.573	-0.023	0.016	0.001	-0.014	0.012	0.019	0.011	0.002

F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.114	0.124	-0.047	-0.030	-0.044	-0.005	0.002	0.034	-0.021	-0.033	-0.022	0.000
0.217	0.288	-0.058	-0.038	-0.015	0.000	-0.009	-0.010	-0.034	-0.015	0.007	0.028
0.428	0.498	-0.008	-0.040	-0.060	-0.021	-0.019	-0.052	-0.043	-0.021	-0.008	0.049
0.489	0.477	-0.023	-0.079	-0.024	-0.017	-0.013	-0.011	-0.044	-0.045	-0.005	0.058
0.632	0.665	-0.021	-0.046	-0.023	-0.014	0.006	0.021	-0.061	-0.055	-0.005	0.055
0.823	0.830	-0.007	0.132	-0.027	0.000	0.012	-0.008	-0.038	-0.122	0.007	0.014
1.020	1.029	0.059	0.158	-0.023	0.000	-0.002	-0.034	-0.047	-0.036	-0.001	-0.015
1.214	1.285	0.193	0.326	-0.030	-0.015	-0.020	-0.016	-0.042	0.001	0.001	0.028
1.342	1.331	0.260	0.440	-0.037	-0.042	-0.031	-0.022	-0.041	0.006	0.022	0.092
1.554	1.546	0.383	0.657	-0.031	-0.020	-0.054	-0.010	-0.010	0.168	-0.006	0.066

TABLE C.35: Comparison - No Network Effect Scenario 4.

Scenario 7(no network effect) v's Scenario 7:

	M16-20		M21-30		M31-45		M46-60		M61-75		M76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	-0.087	-0.107	-0.040	-0.027	0.005	0.126	0.024	0.033	-0.030	-0.033	0.002	-0.041
2015	-0.127	-0.259	-0.073	-0.020	-0.013	-0.015	0.034	0.000	-0.046	-0.080	0.006	0.000
2016	-0.399	-0.301	-0.063	-0.029	-0.003	0.000	0.030	0.002	-0.042	-0.046	0.004	0.017
2017	-0.540	-0.537	-0.035	-0.046	-0.024	-0.002	0.036	0.031	-0.020	-0.039	0.019	0.003
2018	-0.590	-0.550	-0.033	-0.006	-0.031	0.000	0.037	-0.018	-0.030	-0.065	-0.013	-0.018
2019	-0.797	-0.680	-0.048	-0.160	-0.026	-0.006	0.053	0.016	-0.022	-0.053	-0.011	0.000
2020	-0.996	-0.992	-0.079	-0.229	-0.020	0.001	0.062	0.022	-0.011	-0.039	0.008	-0.020
2021	-1.203	-1.123	-0.127	-0.356	-0.024	0.000	0.038	0.024	-0.012	-0.049	-0.004	-0.005
2022	-1.499	-1.517	-0.224	-0.522	-0.012	-0.009	0.031	0.009	-0.019	-0.032	-0.009	-0.016
2023	-1.665	-1.617	-0.297	-0.577	-0.018	-0.037	0.019	0.004	-0.015	-0.017	-0.016	-0.011

	F16-20		F21-30		F31-45		F46-60		F61-75		F76+	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
2014	0.012	-0.116	-0.025	-0.065	0.002	-0.034	0.035	0.068	-0.005	-0.016	0.009	0.037
2015	-0.060	-0.287	-0.010	-0.006	-0.005	-0.018	-0.001	0.030	0.013	-0.025	0.012	0.104
2016	-0.428	-0.498	0.006	0.000	0.025	-0.006	0.004	-0.041	0.019	-0.011	0.005	0.060
2017	-0.488	-0.479	0.028	0.021	0.009	-0.004	-0.002	-0.018	0.014	0.007	-0.011	-0.001
2018	-0.632	-0.668	0.013	0.017	-0.007	-0.012	0.000	-0.032	-0.010	0.000	-0.010	0.000
2019	-0.823	-0.828	0.021	-0.180	-0.014	-0.072	0.020	-0.001	0.008	0.087	0.016	0.001
2020	-1.020	-1.028	0.019	-0.136	-0.020	-0.036	0.027	0.003	0.000	0.009	-0.009	0.018
2021	-1.213	-1.285	-0.137	-0.320	-0.024	-0.023	0.024	-0.002	-0.013	-0.001	-0.011	0.019
2022	-1.341	-1.336	-0.162	-0.430	-0.031	-0.046	0.013	0.005	-0.014	0.001	0.017	-0.003
2023	-1.554	-1.547	-0.298	-0.635	-0.055	-0.042	0.038	0.001	-0.020	0.085	0.020	0.000

TABLE C.36: Comparison - No Network Effect Scenario 7.

Appendix D

Model Simplification using Graphical Methods

D.1 Introduction

A critical resource in understanding how different factors impact on the spread of obesity was the obesity atlas [27]. However it's complexity raised a number of issues, especially in the early stages of the project when the focus was on identifying causality and feedback loops. [29] describes two algorithms that treat causal loop diagrams as directed graphs and then use graphical methods to analyse them. In the original work Matlab was used to implement the algorithms, in our iteration we took advantage of some of the new functionality in Python to deliver them and also add some functionality.

D.2 Purpose

The first algorithm is focused on understanding causality, it generates a 'pruned' graph in which all the predecessors to a specified node are identified, but any repetitions are removed ('pruned'). The output is structured in the form of a tree diagram, so that causality with regard to the specified node is clearly visible.

The second algorithm identifies loops within the causal loop diagram, and maps any overlaps between those loops.

D.3 Methodology

Both the pruning and loop algorithms, take as their starting point a weighted adjacency matrix and a dictionary of labels for each node. The matrix describes a weakly connected directed graph (causal loop diagram), the weights are limited to +1 or -1. The former indicate a positive correlation between the originating node and its dependent, conversely -1 indicates a negative correlation.

D.3.1 Pruning Algorithm

The algorithm can be described as follows:

1. Initially the algorithm iterates through each value in the the matrix (0, 1 or -1) squaring each item within it, turning it into an un-weighted adjacency matrix (0,1).
2. This matrix is then transposed.
3. The transposed matrix is then subjected to Breadth First Search starting at the node of interest.

   (A Breadth First Search algorithm starts at an identified node and returns the shortest path to every other node in the network that can be reached from the identified node, in a transposed matrix this becomes the shortest path from every predecessor node, or when structured as a tree diagram, the pruned graph suggested in [29]).
4. The resulting graph is then exported in a suitable format.

D.3.2 Loop Algorithm

This algorithm takes the original weighted adjacency matrix and the dictionary of descriptors, and uses functions from the Networkx package for Python. The output is exported in a number of csv files.

It replicates the work in [29] in that it identifies the number of loops in the given diagram (using the nx.simple-cycles function from the Networkx package. It then describes them in terms of Length (no. of original nodes), average Eccentricity, Components (the individual nodes and their descriptors), subsets (a subset exists where a larger node contains all the nodes, but not edges of a smaller loop) and identifying where loops share one or more components, creating the structures that in the system dynamics paradigm are often classified as system archetypes.

It extends the work by also classifying the loops as balancing or reinforcing, this is achieved by multiplying together the edge weightings for each edge in the loop (recall that each edge is given a value of $+1$ or -1 according to whether the correlation between the nodes it links is positive or negative), thus a result of $+1$ indicates a reinforcing loop and a result of -1 a balancing loop.

After testing on small examples the initial runs on the full causal loop diagram [27], caused an extremely long run which was eventually halted after 8 hours. Having checked the code thoroughly, the obvious hypothesis was that the run time was a consequence of a very high number of loops. A literature search failed to find any theory that might predict the expected number of loops, so a brief investigation was carried out.

D.3.3 Loop Occurrence in Causal Loop Diagrams

A causal loop diagram could be viewed as a weakly connected, directed, random graph. To facilitate the investigation an 'experimentation space' was defined, comprising a range of randomly generated directed Erdös-Renyi graphs containing from 10 to 30 nodes comprising a single component, with edge:node ratios from 1.5 to 2.5 (rising in increments of 0.1).

To explore the space, the original python programme was re-purposed to simulate large numbers of appropriate graphs with specified numbers of nodes and edges, (the latter specified by the edge:node ratio). The programme then counted the number of loops (simple cycles) in each graph.

Initial experimentation made it clear that the actual number of loops for any given combination of nodes and edge ratio varied considerably, and that their distribution was not symmetrical exhibiting a long 'tail'. For this reason the median (rather than mean) value from 1,000 iterations was used to describe the output from each node/edge:node combination .

This was then repeated for each combination in the experimentation space (231 in total).The output was plotted onto surface charts to describe the distribution of the median values over the experimentation space. The expected number of loops for 30 nodes and an Edge:Node ratio of 2.5 was 1,656, so the number of loops displayed was truncated at 50, to give a more detailed view of the main area of interest (from a practical perspective it was assumed that 50 was the maximum number of loops a researcher might be interested in analysing), see Figure D.1.:

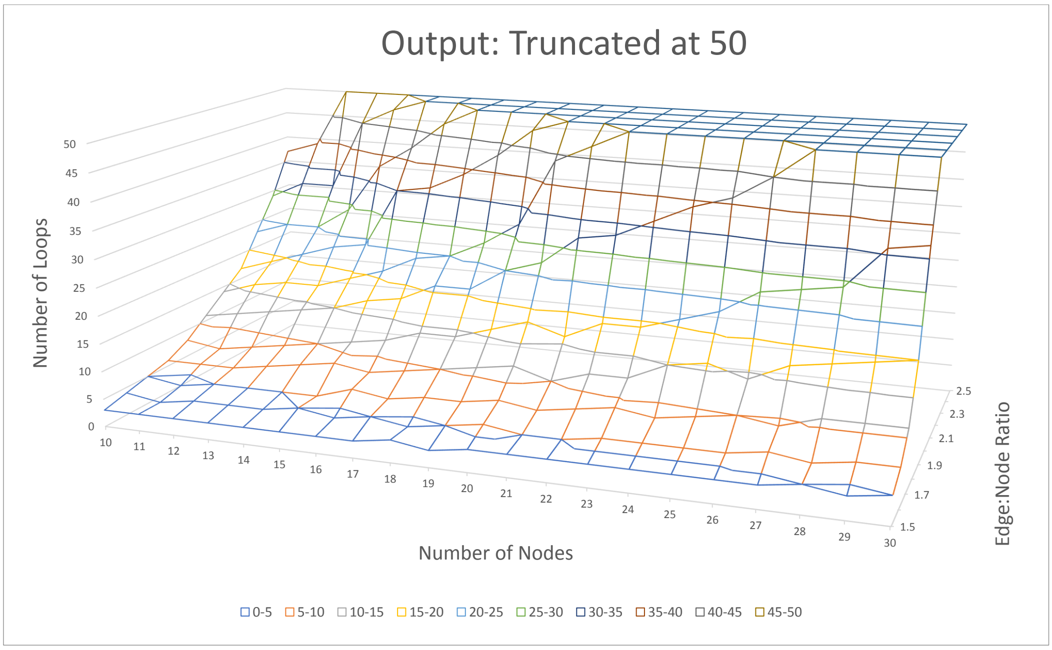


FIGURE D.1: Expected number of Loops

The programme was also used to generate histograms for a range of points across the experimentation space, to better understand the relative distributions (these were each based on 10,000 samples).

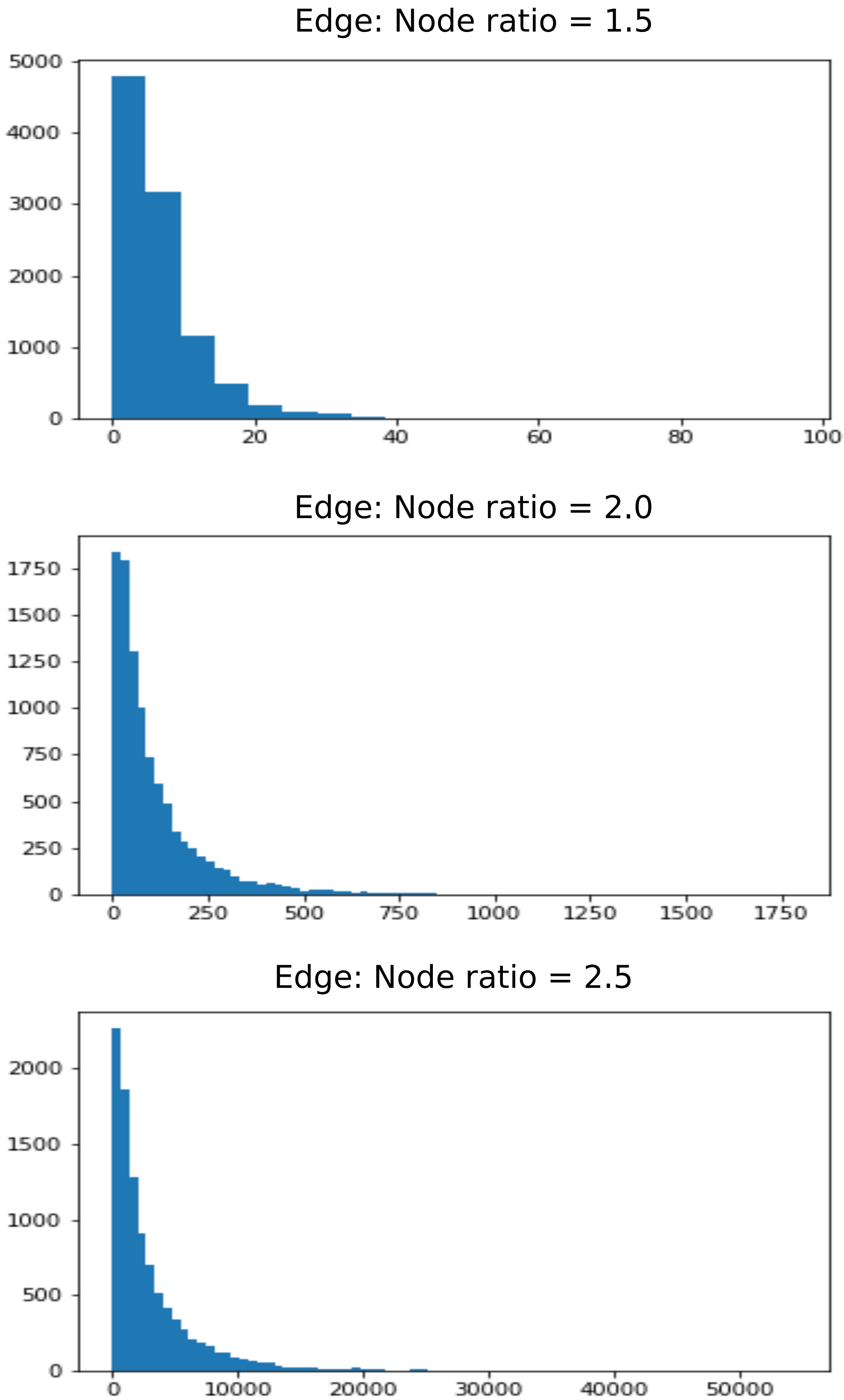


FIGURE D.2: Loop Distribution for 30 Nodes at Various Edge:Node Ratios.

Finally the output was used to estimate the expected number of loops in the original causal loop map [27], this was achieved by extrapolating node numbers from the the Edge:Node ratio within the experiment space ratio closest to that of the original map. This contained 107 nodes and 294 edges (2.74), so the 2.5 ratio was used. An exponential distribution provided a good fit to the data (R^2 value of 0.9966) and gave:

$$No.of Loops = 27.37e^{0.1956 Nodes}$$

Thus a conservative estimate for the expected number of loops in the original map [27] is $3.4 * 10^{10}$.

D.4 Example Output

After experimentation with a number of approaches, the thematic clusters proposed by the authors were used to divide the graph into four sub-graphs based on the "key levers" ([27, Map. 19]), and to exclude the core loop (containing the nested loops at the centre of the diagram). These were:

- Force of Dietary Habit - Food Consumption and Food Production
- Physical Activity - Physical Activity Environment and Individual Physical Activity
- Psychological Ambivalence - Social Psychology and Individual Psychology
- Degree of Primary Appetite Control - Physiology

Each of these sub graphs was examined separately and subjected to both the pruning and loop analysis functions.

An example of the final output for the pruning algorithm (the Psychological Ambivalence lever) is shown in Figure D.3.

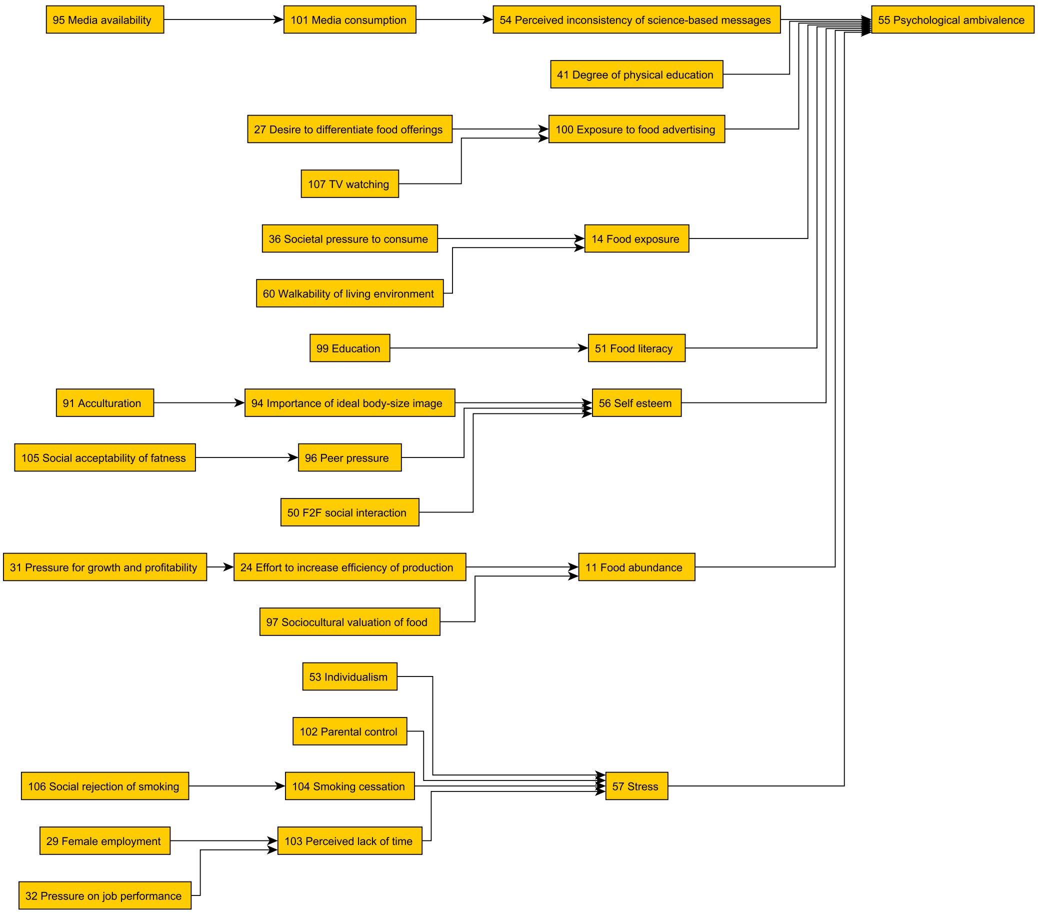


FIGURE D.3: Psychological Ambivalence (Pruned).

It was derived from Figure D.4

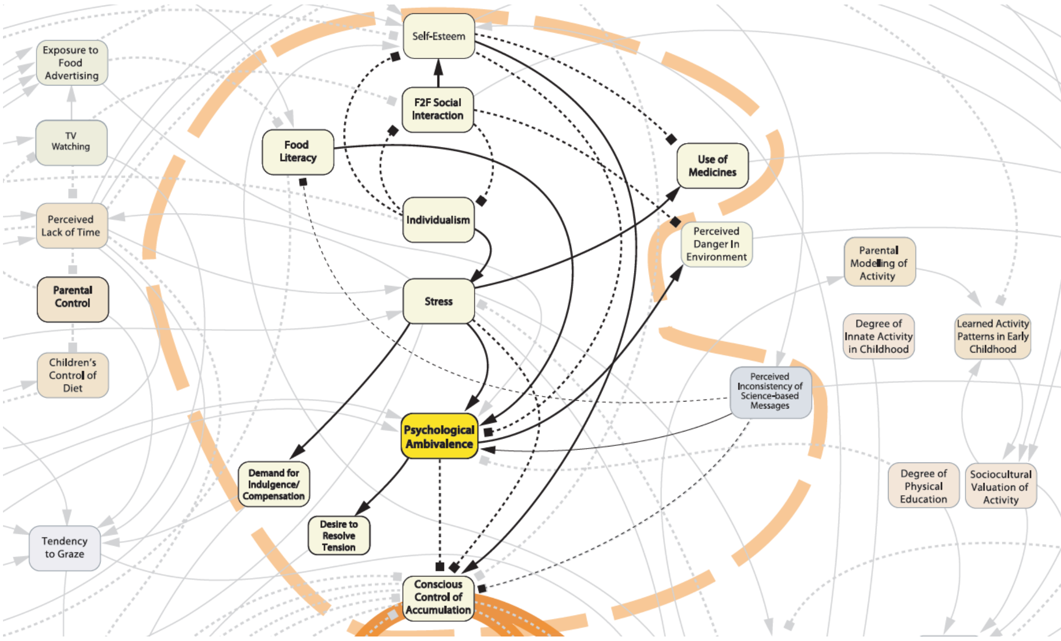


FIGURE D.4: Psychological Ambivalence (Original).

Separately the loop analysis function was applied separately to each of the four sub-graphs, again an example of the output is given below.
Looking at Psychological Ambivalence in more detail the loops are:

- Loop 0. Reinforcing: F2F social interaction, Individualism.
- Loop 1. Reinforcing: Stress, Perceived lack of time.
- Loop 2. Balancing: Stress, Perceived lack of time, Parental control.
- Loop 3. Balancing: TV watching, Perceived lack of time, Parental control.
- Loop 4. Reinforcing: F2F social interaction, Individualism, Stress, Perceived lack of time, Parental control, TV watching.

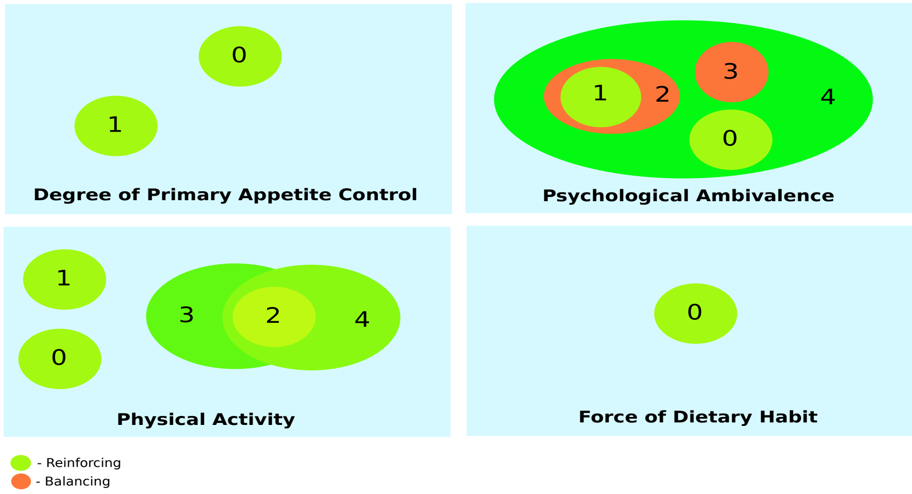


FIGURE D.5: Systemic View of Loops

The diagram shown in Figure D.5 is a systematic attempt to describe the interrelationship of the loops:

- Where two loops intersect as in loop 3 and loop 4 in the 'Physical Activity' quadrant they share common nodes, this is an indicator of the possibility of system archetypes as described in System Dynamics theory.
- Where a loop is wholly contained within another as loop 0 is contained in loop 4 in the 'Psychological Ambivalence' quadrant, then the nodes of the former are a subset of the latter.
- Balancing and Reinforcing loops are coloured shades of red or green respectively.

The two sets of information; the pruned sub-graphs and the loops as they relate to each of those sub graphs were then considered with a view to developing a simplified model based on the expert view encapsulated within the broader causal loop diagram.

Appendix E

Social Networks

E.1 Overview

[118, p. 12]defines a social network as a set of nodes with a description of the relationships between them. In such a network the nodes and mapping represent individuals or groups of individuals (incl. organisations) and the relationships that tie them together into a broader network.

He goes on to suggest that the motivations for individuals to form social networks are threefold:

- Safety - getting support and help from other network members
- Effectance - making connections to make progress and improve a current situation
- Status - improving social standing typically within pyramid social structures, based on occupational or socio-economic models

The relationships can represent one or more of a range of functions including; explicit exchange of information, influence and persuasion (conscious or unconscious), friendship, membership of common groups/organisations, commercial dependencies, family ties and so on.

[118] also suggests that there are three main mechanisms for fulfilling these functions, through such a network:

- Contact involving influence, persuasion or coercion.
- Contact that involves some kind of emulation.
- Contact that involves adoption or emulation at one or more removes (something that has been read or seen).

There are also cognitive limits to the number of relationships an individual can deal with in the context of such a network with one figure suggesting 300 as an average number, although there seems to be considerable variation in this figure. Notwithstanding the differences in mean figures quoted by researchers (perhaps an issue of different cultural and social norms), there is also the distribution of this figure, which doesn't follow a Normal distribution but instead is skewed with a long tail indicating a very few individuals with very high numbers of connections [118, p. 133].

It is these relationships that create the complexity and variation between individuals and between and within different sections of the network, that in turn account for the huge variations in behaviour of different networks in apparently similar circumstances.

From a social science perspective there are three basic types of network:

- Egocentric describes the relationships connected to a single individual/node (ego), the other individuals or nodes are typically known as 'alters. Examples might include an individuals network of friends, or the organisations that do 'business' with a specific organisation. To be more than a simple list the network must also describe any relationships between the alters as well as their relationship to the ego.
- Sociocentric map the relationships between a group of nodes within a set of defined boundaries, the members of a club, children in a class room, a workplace and so on.
- Open-system also map the relationships between a group of nodes, but here the boundaries are not clear, membership of a particular social group, relationships between states and so on.

Two types of structure are also frequently used in network analysis:

- Dyads; in a directed graph the relationship between two nodes is described as a dyad. The classification of dyads and their occurrence forms a significant part of the social network research process, with significance placed on the comparison of different types of dyad in a network with the statistical likelihood of their being formed by chance.
- Triads (sometimes called triples); in a directed network the relationship between three nodes is called a triad, there are 16 different triad classifications ranging from the situation where there are no links between any of the three to that where there are reciprocal relationships between each

of the three [118, p. 24]. The analysis of triad types and quantities within a network and comparison with their chance occurrence can provide a range of insights into the functioning of that network (see Transitivity and Clustering below).

E.2 Features

Social networks exhibit a number of features not common in other types of network, these include:

- Relationship Direction; in the mapping of the relationship between nodes there are a number of options, both in terms of direction and scale. For scale the options are binary (there is a relationship or there isn't), or scalar providing a relative or absolute measure of the 'strength' of the flow. Similarly the options in terms of direction comprise: directional (implying a direction of influence or flow of information), symmetric (suggesting flow in both directions) and non-directional where the relationship is viewed as a 'neutral' conduit for flow. The various combinations of direction and scale provide a range of options for exploring the behaviour of networks, but in practise availability of information about the network is often the limiting factor.
- Node Attributes; the nature of the two processes used to map networks tends to focus attention on the relationships or connections between the nodes, but in the investigation of social networks the nodes themselves will have attributes that will be of interest to the researcher and which will also need to be represented.
- Homophily; the tendency of individuals with common attributes to form connections, these attributes can range from class, race, gender, ethnicity and nationality to marital status, education, occupation and common values and behaviours. The concept is interesting since its manifestation is ambiguous, in some circumstances the common attributes create the context for the connections but equally on occasion a common connection can promote the formation of common attributes. In practise there are probably three processes at work [118, p. 20]:
  - People with similar attributes are drawn to each other.
  - If there is an existing relationship then they are likely to influence each other creating similarity in their mutual attributes.

  - Individuals can end up in the same 'space', and once they are there, the 'space' itself influences them to become more similar in their attributes

  Homophily is of particular interest since it would seem that it plays a role within networks in a health context, significantly improving the adoption of new health behaviours. Centola [119]

- Integration and Consolidation; a network exhibits a high degree of integration if there are a lot of connections between heterogeneous individuals within the network. A network is highly consolidated if there are a lot of connections between agents with similar traits. Generally if there is a high level of consolidation then the level of integration is likely to be low. Centola [120] explores the hypothesis that reducing levels of consolidation promotes diffusion of values and behaviours across social networks, concluding that reduction of consolidation works until a threshold value, at which point further reduction hinders that spread.

- Propinquity; the tendency of individuals who occupy the same 'space' (at the same time), to form relationships. The 'space' can be a physical location (a neighbourhood), or describe a social 'space' such as attendance at a common recreational or business function or event.

- Multiplexity; in a social network individuals may have a number of different relationships, they might be family members who share a workplace, and have a common recreational interest. Each of these implies a potentially different relationship between the same individuals depending on the context [118, p. 202]. This 'layering' of relationships is known as Multiplexity. The example above describes role multiplexity, there is also a concept of content multiplexity whereby as a consequence of a given role a relationship on a number of levels is expected.

- Memory or History; in social networks once a connections is made between two individuals even if it is subsequently broken, it's probability of being restored is higher that would otherwise be the case, this concept is often referred to as memory.

E.3 Network Metrics

A critical tool in the analysis of social networks is Isomorphism, this occurs when the pattern of vertex and edge incidences in two graphs are the same.

Whilst two social networks are unlikely to be exactly Isomorphic, almost all social analysis techniques rely on being able to compare large or very large networks and assign some measure of isomorphism (similarity or difference) to them. The structural measures used to do this include...

- No. of nodes.
- No. of edges.
- Average Node Degree.
- Transitivity; linked to the concept of triad analysis described previously, this is defined by the ratio of closed triads or triplets within the network to the total number of triplets within the network, closed or open (an open triplet is defined as three nodes joined by two edges).
- Average Clustering Co-efficient; the average of the local clustering coeficients for each node, local clustering co-efficient is defined as the number of existing connections amongst a nodes neighbours compared to the total number possible.
- Assortativity; Correlation of node degree between adjacent nodes.
- Components; the number of separate sub-graphs within the network.
- Average Shortest Path Length; provides an average of the shortest path length between each pair of connected nodes.
- Diameter; the longest of the shortest paths between any two vertices in the component.
- The minimum/maximum degree distribution of the nodes.
- Density; this is defined as the number of actual ties in a network divided by the number of possible connections. High density is associated with the quick transmission of ideas, high levels of mutual visibility (often but not always associated with high levels of mutual support). Given the limit to the number of connections and individual can maintain smaller networks tend to be denser than larger ones [118, p. 29].
- Centrality; this is a simple concept but potentially complex in application, in an undirected network the degree or centrality of a node is defined as the number of connections to that node, the higher the number by comparison to other nodes the more 'central' it is perceived to be [118,

p. 31]. In a directed network each node has an in-degree that represents the number of of connections into the node and an out-degree that represents the number of connections from that node. The nature of the nodes that connect into the node in question can also have an impact on centrality if they too have high in-degrees.

- Distance; the (geodesic) distance between two nodes is defined as the shortest distance between two nodes via the edges between the intervening nodes, average distances for a network (averaging the distance between each node and every other node in the network) can be obtained giving an indication of compactness [118, p. 32]. In a similar vein the first order zone represent the direct connections to a node, in networks involving personal connections this is sometimes referred to as the interpersonal environment. The second order zone represents their direct connections and the third order zone represents in turn the connections to the second order zone (all in relation to the original node).
- Segmentation; in practise any large network based on social connections is unlikely to be uniform in its structure and there will be regions where the connections are denser and areas where they are less dense in effect forming clusters or segments that are more cohesive surrounded by regions that are less densely connected. Whilst easy to describe this feature has proven hard to define in a useful fashion with a rigorous mathematical approach. The algorithms used invariable create separate or hierarchical clusters with no lateral overlap a clear and common feature of social networks. Having said that they are still very useful for statistical analysis [118, p. 47].
- Structural similarity and structural equivalence; are alternative approaches used to create sub-sets or clusters from within a network, in the first clusters are formed from nodes with similar patterns of connections, in a manufacturing environment this might result in line managers being clustered together. Using structural equivalence, nodes with the same patterns and connections are clustered together, so in this case line managers from the same section would be clustered together. [118, p. 49].
- Weak Ties; an important concept in understanding networks is that of weak ties or acquaintances not part of the immediate cluster of 'friends' (strong ties)linking to other more distant parts of the network. They provide a useful function in connecting the network as a whole and speeding the diffusion of ideas or information across the whole network [118,

p. 31]. In a similar vein [71] suggests that there is competitive advantage to be obtained within less dense networks, by acting as a strategic link between different parts of network, filling what he terms as 'structural holes'.

- Core/Periphery; the simplest form of network segmentation is represented by the concept of a 'core' group and a 'peripheral' group, where the 'core' is densely interconnected and a 'peripheral' group who perhaps only have one connection to a member of the 'core'. This an asymmetric model with diffusion coming from a self-contained core with little or no contribution from the periphery. There are a number of variations depending on whether the network is directed or symmetrical, these are typically represented using block models

[47] argues that social networks differ from most other networks in 2 significant ways, and supports this by contrasting them with a range of technological and biological examples:

- Clustering or transitivity; an individuals connections are much more likely to be inter-connected than would be the case in a random network.
- Social Networks exhibit assortative mixing or positive correlations between the degrees of adjacent nodes, contrary to the case in random networks.

They conclude that this is caused primarily by the formation of 'communities' of various sizes within the network.

E.4 Intervention Strategies in Social Networks

[117] identifies four network intervention strategies in the context of health behaviours:

- Individual; network data is used to identify champions or opinion leaders, either through 'nominations' from other network members or algorithms based on one of the models of 'centrality'; closeness (shortest paths to all other nodes), betweenness ('gate-keeping' on shortest path between nodes). There are some caveats, central nodes can often link to the same people, leaders are often invested in the 'status quo', if speed of dissemination is important then it might be best to target 'bridging' nodes between clusters. This approach was explored in a counterfactual simulation by El-Sayed, Seemann, Scarborough, and Galea [121] with equivocal results, two scenarios were tested with structured interventions and with random interventions. In the first scenario (prevention of obesity in 10% of the population) the targeted intervention out performed the random intervention, but at only with a very high level of 'contagion'. In the second scenario (treatment of obesity in 10% of the population) the random outperformed the targeted, suggesting that at the very least such initiatives should be thoroughly tested before implementation.

- Segmentation; instead of looking at individuals this looks at key groups/cliques (membership of a group is exclusive, clique memberships may overlap), perhaps carrying out a specific role within the community, or fulfilling certain criteria. Again there are algorithms for identifying groups and cliques within networks. A common structure in organisations is core-periphery, mobilising such a network with limited resources is best achieved by focusing the resources on the core.

- Induction; this seeks to create peer to peer interactions in order to diffuse information and/or cascade behaviour through the network. An example of this approach is Respondent Driven Sampling (RDS) in which an initial population (seeds) are identified to go out to their individual networks and promote the behaviour or spread the change. Network outreach is similar except that the seeds recruit their network to take part in an intervention together as one group. Outreach is usually perceived as more effective than RDS because of the additional group dynamics involved [117].

- Modification; the options here include adding/subtracting nodes and links and re-wiring existing links. An example of adding nodes is the use of change agents introduced into the community to promote and facilitate the desired change, node deletion is an effective strategy for degrading networks effectiveness (criminal and terrorist organisations) and for slowing the spread of disease. Re-wiring is a strategy often used to improve efficiency/performance in relation to the networks objectives, many of the social network experiments conducted on social media focus on rewiring strategies [119], [122] and Centola and Rijt [93].

With the exception of the modification approach used in the context of social media, there seems to be very little literature describing examples of the application of these methodologies in a healthcare setting. Suggesting that there is scope for more research in this area.

E.5 Social Network Analysis Approaches

Social network analysis involves in essence the use of network metrics to link network topography to observed effects, testing the hypothesis that structure ’a’ has some causative effect on effect ’b’, or vice versa. The topology is described using the features and metrics described earlier, and as previously mentioned the concept of Isomorphism (comparing graph topologies) is key.

There are two broad approaches to testing hypotheses in social network analysis (Borgatti, Everett, and Johnson [123]). In the first permutation tests are used in conjunction with techniques like regression (the work described in [5] is an example of this approach).

The Quadratic Assignment Procedure (QAP) is also an example of this approach. QAP allows the testing of hypotheses against observed results and the development of statistical distributions to assess their significance by comparing data from the network under consideration with data obtained by generating multiple random permutations of that network using statistical analysis. This approach can be used to examine relationships between networks, or between network relationships (dyads) and attributes (categorical or continuous), diffusion and homphily can be included [123].

The second approach develops statistical models to examine the distribution and evolution of links in a network. Two common approaches are the use of Exponential Random Graphs (ERG) and Stochastic Actor Based Models (SAB). A relevant example of this approach is [72], which explores the utility of high levels of clustering v’s weak ties in enabling the rapid diffusion of health behaviours through an online network. Again the benchmark measures of clustering are obtained by using multiple randomly generated ERG’s, these are then used to test whether specific patterns and configurations within the network occur at a higher frequency than might be expected in a randomly organised network. The type of patterns examined will depend on the hypothesis being tested Robbins, G. Pattison [124].

Stochastic Actor Based (SAB) models suggested by Snijders, Bunt, and Steglich [125] and [123] use panel data to model the likelihood of agents (actors) making or breaking ties based on an evaluation/utility function set up by the researcher to test the hypothesis. The underlying paradigm is a continuous time

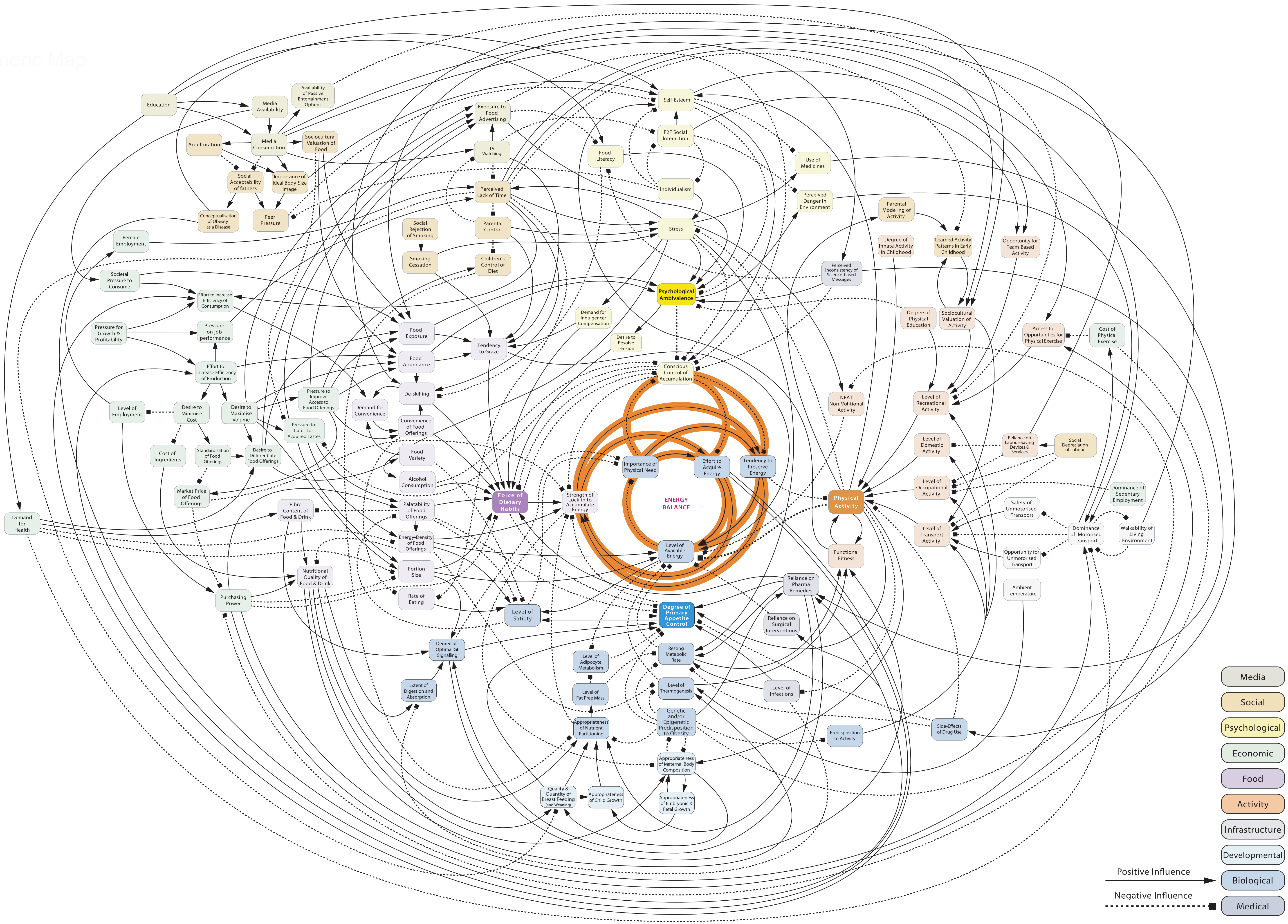
Markov model.

With the exception of SAB models, the predominant theme in all of these approaches is analysis of the existing structure (assuming structure as cause), SAB is different treating structure as effect. Perhaps explaining its popularity as a tool for forecasting use of page links for commercial web based applications.

Appendix F

Foresight Tackling Obesities Future Choices - Obesity System Atlas

Vandenbroek, P. Goossens, J. Clemens [27]



Education
Media Availability
Availability of Passive Entertainment Options
Acculturation
Media Consumption
Sociocultural Valuation of Food
Social Acceptability of fatness
Importance of Ideal Body-Size Image
Conceptualisation of Obesity as a Disease
Peer Pressure
Exposure to Food Advertising
TV Watching
Perceived Lack of Time
Social Rejection of Smoking
Smoking Cessation
Parental Control
Children's Control of Diet
Self-Esteem
F2F Social Interaction
Food Literacy
Individualism
Use of Medicines
Perceived Danger In Environment
Stress
Parental Modelling of Activity
Degree of Innate Activity in Childhood
Learned Activity Patterns in Early Childhood
Opportunity for Team-Based Activity
Perceived Inconsistency of Science-based Messages
Psychological Ambivalence
Female Employment
Societal Pressure to Consume
Effort to Increase Efficiency of Consumption
Pressure for Growth & Profitability
Pressure on job performance
Effort to Increase Efficiency of Production
Food Exposure
Tendency to Graze
Food Abundance
Demand for Indulgence/ Compensation
Desire to Resolve Tension
Degree of Physical Education
Sociocultural Valuation of Activity
Access to Opportunities for Physical Exercise
Cost of Physical Exercise
Conscious Control of Accumulation
NEAT Non-Volitional Activity
Level of Recreational Activity
Level of Employment
Desire to Minimise Cost
Desire to Maximise Volume
Pressure to Improve Access to Food Offerings
De-skilling
Demand for Convenience
Convenience of Food Offerings
Pressure to Cater for Acquired Tastes
Cost of Ingredients
Standardisation of Food Offerings
Desire to Differentiate Food Offerings
Food Variety
Alcohol Consumption
Importance of Physical Need
Effort to Acquire Energy
Tendency to Preserve Energy
Level of Domestic Activity
Reliance on Labour-Saving Devices & Services
Social Depreciation of Labour
Level of Occupational Activity
Dominance of Sedentary Employment
Market Price of Food Offerings
Demand for Health
Fibre Content of Food & Drink
Palatability of Food Offerings
Force of Dietary Habits
Strength of Lock-in to Accumulate Energy
ENERGY BALANCE
Physical Activity
Safety of Unmotorised Transport
Level of Transport Activity
Dominance of Motorised Transport
Walkability of Living Environment
Opportunity for Unmotorised Transport
Energy-Density of Food Offerings
Level of Available Energy
Functional Fitness
Nutritional Quality of Food & Drink
Portion Size
Reliance on Pharma Remedies
Ambient Temperature
Purchasing Power
Rate of Eating
Level of Satiety
Degree of Primary Appetite Control
Reliance on Surgical Interventions
Degree of Optimal GI Signalling
Level of Adipocyte Metabolism
Resting Metabolic Rate
Extent of Digestion and Absorption
Level of Fat-Free Mass
Level of Thermogenesis
Level of Infections
Appropriateness of Nutrient Partitioning
Genetic and/or Epigenetic Predisposition to Obesity
Predisposition to Activity
Side-Effects of Drug Use
Appropriateness of Maternal Body Composition
Quality & Quantity of Breast Feeding (and Weaning)
Appropriateness of Child Growth
Appropriateness of Embryonic & Fetal Growth
Media
Social
Psychological
Economic
Food
Activity
Infrastructure
Developmental
Biological
Medical
Positive Influence
Negative Influence

Bibliography

[1] S. Mendis, S. Davis, and B. Norrving, "Organizational Update: The World Health Organization Global Status Report on Noncommunicable Diseases 2014; One More Landmark Step in the Combat Against Stroke and Vascular Disease", *Stroke*, vol. 46, no. 5, pp. 121–122, 2015.

[2] *WHO | Noncommunicable Diseases Progress Monitor 2017*. [Online]. Available: `https://www.who.int/nmh/publications/ncd-progress-monitor-2017/en/` (visited on 11/05/2019).

[3] D. T. Finegood, T. D. N. Merth, and H. Rutter, "Implications of the foresight obesity system map for solutions to childhood obesity.", *Obesity (Silver Spring, Md.)*, vol. 18, pp. 13–16, 2010.

[4] L. N. Allen and A. B. Feigl, "Reframing non-communicable diseases as socially transmitted conditions.", *The Lancet. Global health*, vol. 5, no. 7, pp. 644–646, 2017.

[5] N. A. Christakis and J. H. Fowler, "The spread of obesity in a large social network over 32 years.", *The New England journal of medicine*, vol. 357, no. 4, pp. 370–9, 2007.

[6] R. J. Brooks, "Modelling the increase in BMI and obesity in England since the early 1990s", *Unpublished work*, 2016.

[7] T Dawber, G Meadors, F Moore, and Jr., "Epidemiological approaches to heart disease: the Framingham Study.", *American journal of public health and the nation's health*, vol. 41, no. 3, pp. 279–81, 1951.

[8] J. D. Klein, "The National Longitudinal Study on Adolescent Health", *JAMA*, vol. 278, no. 10, p. 864, 1997.

[9] E. Cohen-Cole and J. M. Fletcher, "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic", *Journal of Health Economics*, vol. 27, no. 5, pp. 1382–1387, 2008.

[10] J. H. Fowler and N. A. Christakis, "Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais", *Journal of Health Economics*, vol. 27, no. 5, pp. 1400–1405, 2008.

[11] N. A. Christakis and J. H. Fowler, "Social contagion theory: Examining dynamic social networks and humanbehavior", *Statistics in Medicine*, vol. 32, no. 4, pp. 556–577, 2013.

[12] T. J. VanderWeele, "Inference for influence over multiple degrees of separation on a social network", *Statistics in Medicine*, vol. 32, no. 4, pp. 591–596, 2013.

[13] D. J. Hruschka, A. A. Brewis, A. Wutich, and B. Morin, "Shared norms and their explanation for the social clustering of obesity", *American Journal of Public Health*, vol. 101, no. 1, pp. 295–300, 2011.

[14] H. B. Shakya, N. A. Christakis, and J. H. Fowler, "Self-comparisons as motivators for healthy behavior", *Obesity*, vol. 23, no. 12, pp. 2477–2484, 2015.

[15] P. J. H. Hulshof, N. Kortbeek, R. J. Boucherie, *et al.*, "Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS", *Health Systems*, vol. 1, no. 2, pp. 129–175, 2012.

[16] S. C. Brailsford, P. R. Harper, B Patel, and M Pitt, "An analysis of the academic literature on simulation and modelling in health care", *Journal of Simulation*, vol. 3, pp. 130–140, 2009.

[17] N. Mustafee and J. H. Powell, "From hybrid simulation to hybrid systems modelling", in *Proceedings - Winter Simulation Conference*, vol. 2018-Decem, Institute of Electrical and Electronics Engineers Inc., 2019, pp. 1430–1439, ISBN: 9781538665725. DOI: 10.1109/WSC.2018.8632528.

[18] R. A. Nianogo and O. A. Arah, "Agent-based modeling of noncommunicable diseases: a systematic review.", *American journal of public health*, vol. 105, no. 3, pp. 20–31, 2015.

[19] T. Monks, C. S. M. Currie, B. S. Onggo, *et al.*, "Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines", *Journal of Simulation*, pp. 1–13, 2018.

[20] J. Bentham, M. Di Cesare, V. Bilano, *et al.*, "Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults", *The Lancet*, vol. 390, no. 10113, pp. 2627–2642, 2017, ISSN: 1474547X. DOI: 10.1016/S0140-6736(17)32129-3.

[21] M. M. Finucane, G. A. Stevens, M. J. Cowan, *et al.*, "National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960

country-years and 91 million participants”, *The Lancet*, vol. 377, no. 9765, pp. 557–567, 2011.

[22] C. Baker, *UK Parliament, Commons Library briefing - Obesity Statistics*. [Online]. Available: https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN03336 (visited on 10/24/2019).

[23] E. A. Finkelstein, O. A. Khavjou, H. Thompson, *et al.*, “Obesity and Severe Obesity Forecasts Through 2030”, *American Journal of Preventive Medecine*, vol. 42, no. 6, pp. 563–570, 2012.

[24] J. Kearney, “Food consumption trends and drivers”, *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1554, pp. 2793–2807, 2010.

[25] B. M. Popkin, “Technology, transport, globalization and the nutrition transition food policy”, *Food Policy*, vol. 31, no. 6, pp. 554–569, 2006.

[26] D. F. Gleich, L.-H. Lim, and Y. Yu, “Multilinear PageRank”, pp. 1–33, 2014.

[27] M Vandenbroek, P. Goossens, J. Clemens, *Foresight Tackling Obesities: Future Choices — Obesity System Atlas*, 2007. [Online]. Available: https://www.gov.uk/government/publications/reducing-obesity-obesity-system-map (visited on 10/25/2019).

[28] M. C. IP Vandenbroeck, J Goossens, *Foresight tackling obesities: future choices-building the obesity system map*, 2007. [Online]. Available: https://www.gov.uk/government/publications/reducing-obesity-obesity-system-map (visited on 10/25/2019).

[29] R. Oliva, “Model structure analysis through graph theory: partition heuristics and feedback structure decomposition”, *System Dynamics Review*, vol. 20, no. 4, pp. 313–336, 2004.

[30] E. Ostman, J. Britton, M. Jonsson, Ed., *Treating and Preventing Obesity*. Wiley, 2004.

[31] G. Gerlach, S. Herpertz, and S. Loeber, “Personality traits and obesity: A systematic review”, *Obesity Reviews*, vol. 16, no. 1, pp. 32–63, 2015.

[32] P. Kopelman, S. A. Jebb, and B. Butland, “Executive summary: Foresight ’tackling obesities: Future choices’ project”, *Obesity Reviews*, vol. 8, no. SUPPL. 1, 2007.

[33] F. C. Pampel, P. M. Krueger, and J. T. Denney, “Socioeconomic Disparities in Health Behaviors”, *Annual Review of Sociology*, vol. 36, no. 1, pp. 349–370, 2010.

[34] D. M. Cutler and A. Lleras-Muney, “Understanding differences in health behaviors by education.”, *Journal of health economics*, vol. 29, no. 1, pp. 1–28, 2010.

[35] I. Ajzen, “The theory of planned behavior”, *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 1991.

[36] S. S. Lim, T. Vos, A. D. Flaxman, *et al.*, “A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010.”, *Lancet*, vol. 380, no. 9859, pp. 2224–60, 2012.

[37] WHO, *Ten threats to global health in 2019*. [Online]. Available: `https://www.who.int/emergencies/ten-threats-to-global-health-in-2019` (visited on 10/25/2019).

[38] M. L. Specchia, M. A. Veneziano, C. Cadeddu, *et al.*, “Economic impact of adult obesity on health systems: A systematic review”, *European Journal of Public Health*, vol. 25, no. 2, pp. 255–262, 2015.

[39] J. Linton and L. James, *Handbook of Obesity Intervention for the Lifespan*. Springer New York, 2009. DOI: `10.1007/978-0-387-78305-5`.

[40] NICE, *Obesity: identification, assessment and management*, eng, 2014. [Online]. Available: `https://www.nice.org.uk/guidance/cg189/chapter/1-recommendations{\#}surgical-interventions` (visited on 02/29/2016).

[41] M. Manyema, L. J. Veerman, L. Chola, *et al.*, “The potential impact of a 20% tax on sugar-sweetened beverages on obesity in South African adults: A mathematical model”, *PLoS ONE*, vol. 9, no. 8, 2014.

[42] A. D. M. Briggs, O. T. Mytton, A. Kehlbacher, *et al.*, “Overall and income specific effect on prevalence of overweight and obesity of 20% sugar sweetened drink tax in UK: econometric and comparative risk assessment modelling study”, *BMJ*, vol. 347, pp. 6189–6189, 2013.

[43] D. Thomas, S. K. Das, J. A. Levine, *et al.*, “New fat free mass - fat mass model for use in physiological energy balance equations.”, *Nutrition & metabolism*, vol. 7, p. 39, 2010.

[44] J. A. Ello-Martin, J. H. Ledikwe, and B. J. Rolls, “The influence of food portion size and energy density on energy intake: implications for weight management.”, *The American journal of clinical nutrition*, vol. 82, pp. 236–241, 2005.

[45] T. Cruwys, K. E. Bevelander, and R. C. J. Hermans, "Social modeling of eating: A review of when and why social influence affects food intake and choice", *Appetite*, vol. 86, pp. 3–18, 2015.

[46] G. Bernstein and K. O'Brien, "Stochastic Agent-Based Simulations of Social Networks", in *Proceedings of 46th Annual Simulation Symposium*, 2013, pp. 33–40.

[47] M. E. Newman and J. Park, "Why social networks are different from other types of networks", *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 68, no. 3, p. 8, 2003.

[48] E. Kiesling, M. Günther, C. Stummer, and L. M. Wakolbinger, "Agent-based simulation of innovation diffusion: A review", *Central European Journal of Operations Research*, vol. 20, no. 2, pp. 183–230, 2012.

[49] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks.", *Nature*, vol. 393, no. 6684, pp. 440–2, 1998.

[50] E. Z. Erbach-Schoenberg, S. Bullock, and S. Brailsford, "A Model of Spatially Constrained Social Network Dynamics", *Social Science Computer Review*, vol. 32, no. 3, pp. 373–392, 2013.

[51] N. Mustafee, K. Katsaliaki, and S. J. Taylor, "Profiling Literature in Healthcare Simulation", *Simulation*, vol. 86, no. 8-9, pp. 543–558, 2010.

[52] H Xue, L Slivka, T Igusa, *et al.*, "Applications of systems modelling in obesity research", *Obesity Reviews*, vol. 19, pp. 1293–1308, 2018.

[53] T. K. Abdel-Hamid, "Exercise and Diet in Obesity Treatment: An Integrative System Dynamics Perspective", *Medicine & Science in Sports & Exercise*, vol. 35, no. 3, pp. 400–413, 2003.

[54] J. B. Homer and G. B. Hirsch, "System dynamics modeling for public health: Background and opportunities", *American Journal of Public Health*, vol. 96, no. 3, pp. 452–458, 2006.

[55] N. Z. Abidin, M. Mamat, B. Dangerfield, *et al.*, "Combating obesity through healthy eating behavior: A call for system dynamics optimization", *PLoS ONE*, vol. 9, no. 12, pp. 1–17, 2014.

[56] S. Liu, N. Osgood, Q. Gao, *et al.*, "Systems simulation model for assessing the sustainability and synergistic impacts of sugar-sweetened beverages tax and revenue recycling on childhood obesity prevention", *Journal of the Operational Research Society*, vol. 67, no. 5, pp. 708–721, 2016.

[57] R. Beheshti, T. Igusa, and J. Jones-Smith, "Simulated Models Suggest That Price per Calorie Is the Dominant Price Metric That Low-Income

Individuals Use for Food Decision Making.", *The Journal of nutrition*, vol. 146, no. 11, pp. 2304–2311, 2016.

[58] J. Zhang, D. A. Shoham, E. Tesdahl, and S. B. Gesell, "Network interventions on physical activity in an afterschool program: An agent-based social network study", *American Journal of Public Health*, vol. 105, pp. 236–243, 2015.

[59] Y. Yang and A. Diez-Roux, "Using an agent-based model to simulate children's active travel to school", *International Journal of Behavioral Nutrition and Physical Activity*, vol. 10, 2013.

[60] J. E. Hurwitz, J. A. Lee, K. K. Lopiano, *et al.*, "A flexible simulation platform to quantify and manage emergency department crowding.", *BMC medical informatics and decision making*, vol. 14, no. 1, p. 50, 2014, ISSN: 1472-6947. DOI: 10.1186/1472-6947-14-50.

[61] M. Rauner, S. Brailsford, and S. Flessa, "Use of discrete-event simulation to evaluate strategies for the prevention of mother-to-child transmission of HIV in developing countries", *Journal of the Operational Research Society*, vol. 56, no. 2, 2005. DOI: 10.1057/palgrave.jors.2601884.

[62] S. C. Brailsford, P. R. Harper, and J. Sykes, "Incorporating human behaviour in simulation models of screening for breast cancer", *European Journal of Operational Research*, vol. 219, no. 3, pp. 491–507, 2012, ISSN: 03772217. DOI: 10.1016/j.ejor.2011.10.041.

[63] I. T. Vieira, R. C. H. Cheng, P. R. Harper, and V. de Senna, "Small world network models of the dynamics of HIV infection", *Annals of Operations Research*, vol. 178, no. 1, pp. 173–200, 2010, ISSN: 02545330. DOI: 10.1007/s10479-009-0571-y.

[64] C. Swinerd and K. R. Mcnaught, "Design classes for hybrid simulations involving agent-based and system dynamics models", *Simulation Modelling Practise and Theory*, vol. 25, pp. 118–133, 2012.

[65] S. C. Brailsford, "System dynamics: What's in it for healthcare simulation modelers", *Proceedings - Winter Simulation Conference*, no. 1968, pp. 1478–1483, 2008.

[66] M. Kunc, P. Harper, and K. Katsikopoulos, "A review of implementation of behavioural aspects in the application of OR in healthcare", *Journal of the Operational Research Society*, pp. 1–18, 2018.

[67] R. A. Hammond and J. T. Ornstein, "A model of social influence on body mass index", *Annals of the New York Academy of Sciences*, vol. 1331, no. 1, pp. 34–42, 2014.

[68] J. R. Busemeyer and A. Diederich, "Survey of decision field theory", *Mathematical social sciences*, vol. 43, no. 3, pp. 345–370, 2002.

[69] D. J. Higham, P. Grindrod, A. V. Mantzaris, *et al.*, "Anticipating Activity in Social Media Spikes", 2014.

[70] X. Fang, P. Jen-Hwa Hu, Z. Li, and W. Tsai, "Predicting Adoption Probabilities in Social Networks", *Information Systems Research*, vol. 24, no. 1, pp. 128–145, 2013.

[71] R. S. Burt, *Structural Holes*. Harvard University Press, 1995.

[72] D. M. Centola, "The Spread of Behavior in an Online Social Network Experiment", *Science*, vol. 1194, no. September, pp. 1194–1198, 2010.

[73] D. Centola, "Social media and the science of health behavior", *Circulation*, vol. 127, no. 21, pp. 2135–2144, 2013.

[74] R. West and J. Brown, *Theory of Addiction*. John Wiley & Sons, 2013.

[75] R. H. Thaler and C. R. Sunstein, *Nudge : improving decisions about health, wealth and happiness*. Penguin Books, 2009.

[76] "The foundations of statistics. By Leonard J. Savage, John Wiley & Sons, Inc., 1954, 294 pp", *Naval Research Logistics Quarterly*, vol. 1, no. 3, 1954.

[77] G. S. Becker and K. M. Murphy, "A Theory of Rational Addiction", *Journal of Political Economy*, vol. 96, no. 4, p. 675, 1988.

[78] K. Glanz and B. K. Rimer, "Theory at a Glance: A Guide for Health Promotion Practice", *Health (San Francisco)*, p. 52, 2005.

[79] I. M. Rosenstock, "Historical Origins of the Health Belief Model", *Health Education & Behavior*, vol. 2, no. 4, pp. 328–335, 1974.

[80] C. Diclemente, "Stages and Processes of Self-Change of Smoking-Toward An Integrative Model of Change", *Journal of Consulting and Clinical Psychology*, 1983.

[81] V. H. Taylor, C. M. Curtis, and C. Davis, "The obesity epidemic: the role of addiction.", *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*, vol. 182, no. 4, pp. 327–8, 2010.

[82] A. N. Gearhardt, W. R. Corbin, and K. D. Brownell, "Preliminary validation of the Yale Food Addiction Scale.", *Appetite*, vol. 52, no. 2, pp. 430–6, 2009.

[83] C. Davis, C. Curtis, R. D. Levitan, *et al.*, "Evidence that 'food addiction' is a valid phenotype of obesity", *Appetite*, vol. 57, no. 3, pp. 711–717, 2011.

[84] G. T. Wilson, "Eating disorders, obesity and addiction", *European Eating Disorders Review*, vol. 18, no. 5, pp. 341–351, 2010.

[85] H Ziauddeen, I. S. Farooqi, and P. C. Fletcher, "Obesity and the brain: how convincing is the addiction model?", *Nature Reviews. Neuroscience*, vol. 13, no. 4, pp. 279–286, 2012.

[86] M. Baddeley, *Behavioural Economics and Finance*. Routledge, 2013.

[87] N. T. Hendy and N. Montargot, "Understanding Academic dishonesty among business school students in France using the theory of planned behavior", *The International Journal of Management Education*, vol. 17, no. 1, pp. 85–93, 2019.

[88] F. Khan, W. Ahmed, and A. Najmi, "Understanding consumers' behavior intentions towards dealing with the plastic waste: Perspective of a developing country", *Resources, Conservation and Recycling*, vol. 142, pp. 49–58, 2019.

[89] K. Kwok, K. Li, S. Lee, *et al.*, "Multi-centre study on cultural dimensions and perceived attitudes of nurses towards influenza vaccination uptake", *Journal of Hospital Infection*, 2018.

[90] J. P. Guidry, K. E. Carlyle, J. G. LaRose, *et al.*, "Using the Health Belief Model to Analyze Instagram Posts about Zika for Public Health Communications", *Emerging Infectious Diseases*, vol. 25, no. 1, pp. 179–180, 2019.

[91] R. E. Cushing, K. L. Braun, and S. Alden, "A Qualitative Study Exploring Yoga in Veterans with PTSD Symptoms", *International Journal of Yoga Therapy*, vol. 28, no. 1, pp. 63–70, 2018.

[92] J. Shida, K. Kuwana, and K. Takahashi, "Behavioral intention to prevent cervical cancer and related factors among female high school students in Japan", *Japan Journal of Nursing Science*, vol. 15, no. 4, pp. 375–388, 2018.

[93] D. Centola and A. van de Rijt, "Choosing your network: Social preferences in an online health community.", *Social science & medicine (1982)*, vol. 125, pp. 1–13, 2014.

[94] G. Forbes, "Lean body mass-body fat interrelationships in humans", *Nutrition Reviews*, vol. 45, no. 8, pp. 225–231, 1987.

[95] D. M. Thomas, C. K. Martin, S. Heymsfield, *et al.*, "A simple model predicting individual weight change in humans", *Journal of Biological Dynamics*, vol. 5, no. 6, pp. 579–599, 2011.

[96] D. M. Thomas, A. E. Ivanescu, C. K. Martin, *et al.*, "Predicting successful long term weight loss from short term weight loss outcomes : new insights from a dynamic energy balance model", *American Journal of Clinical Nutrition*, vol. 101, no. 3, pp. 449–454, 2015.

[97] K. D. Hall, "Body fat and fat-free mass inter-relationships: Forbes's theory revisited", *British Journal of Nutrition*, vol. 97, no. 06, p. 1059, 2007.

[98] K. D. Hall, "What is the required energy deficit per unit weight loss?", *International Journal of Obesity*, vol. 32, no. 3, pp. 573–576, 2008.

[99] K. D. Hall, D Chandramohan, C. C. Chow, *et al.*, "Quantification of the effect of energy imbalance on bodyweight", *Lancet*, vol. 378, pp. 826–37, 2011.

[100] K. R. Westerterp, J. H. H. L. M. Donkers, E. W. H. M. Fredrix, *et al.*, "Energy intake, physical activity and body weight: a simulation model", *British Journal of Nutrition*, vol. 73, no. 03, p. 337, 1995.

[101] Henry C J, "Basal metabolic rate studies in humans: measurement and development of new equations", *Public Health Nutr.*, vol. 8, no. 7A, pp. 1133–1152, 2005.

[102] Scientific Advisory Committee on Nutrition, *Dietary Reference Values for Energy*, 2011. [Online]. Available: https://www.gov.uk/government/publications/sacn-dietary-reference-values-for-energy (visited on 11/05/2019).

[103] S. Meeuwsen, G. Horgan, and M. Elia, "The relationship between BMI and percent body fat, measured by bioelectrical impedance, in a large adult sample is curvilinear and influenced by age and sex", *Clinical Nutrition*, vol. 29, no. 5, pp. 560–566, 2010.

[104] D. M. Thomas, C. K. Martin, L. M. Redman, *et al.*, "Effect of dietary adherence on the body weight plateau: A mathematical model incorporating intermittent compliance with energy intake prescription", *American Journal of Clinical Nutrition*, vol. 100, no. 3, pp. 787–795, 2014.

[105] C. C. Chow and K. D. Hall, "The Dynamics of Human Body Weight Change", *PLoS Comput Biol*, vol. 4, no. 310, 2008.

[106] S. Robinson, *Simulation : the practice of model development and use*. MacMillan, 2014, p. 367, ISBN: 9781137328021.

[107] S. F. Railsback and V. Grimm, *Agent-based and individual-based modeling : a practical introduction*. Princeton University Press, 2012, p. 329.

[108] M. Pidd, *Computer simulation in management science*. Wiley, 2004.

[109] J. C. Spall, *Introduction to Stochastic Search and Optimization*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2003.

[110] V. Grimm and S. F. Railsback, "Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, pp. 298–310, 2012.

[111] J. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 817–823, 1998.

[112] HSE, *Health Survey England 2012*, 2012. [Online]. Available: `http://www.hscic.gov.uk/catalogue/PUB13218` (visited on 01/09/2018).

[113] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[114] D. Shoham, L. Tong, P. Lamberson, *et al.*, "An actor-based model of social network influence on adolescent body size, screen time, and playing sports", *PLoS ONE*, vol. 7, no. 6, 2012.

[115] J. Daw, R. Margolis, and A. M. Verdery, "Siblings, friends, course-mates, club-mates: How adolescent health behavior homophily varies by race, class, gender, and health status", *Social Science and Medicine*, vol. 125, pp. 32–39, 2015.

[116] R. M. Viner, S. Kinra, D. Nicholls, *et al.*, "Burden of child and adolescent obesity on health services in England", *Archives of Disease in Childhood*, 2017.

[117] T. W. Valente, "Network Interventions", *Science*, vol. 337, no. 6090, pp. 49–53, 2012.

[118] C. Kadushin, *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford University Press, USA, 2012.

[119] D. Centola, "An Experimental Study of Homophily in the Adoption of Health Behavior", *Science*, vol. 334, no. 6060, pp. 1269–1272, 2011.

[120] D Centola, "The Social Origins of Networks and Diffusion1", *American Journal of Sociology*, vol. 120, no. 5, pp. 1295–1338, 2015.

[121] A. El-Sayed, L. Seemann, P. Scarborough, and S. Galea, "Are network-based interventions a useful antiobesity strategy? An application of simulation models for causal inference in epidemiology", *American Journal of Epidemiology*, vol. 178, no. 2, 2013.

[122] R. M. Bond, C. J. Fariss, J. J. Jones, *et al.*, "A 61-million-person experiment in social influence and political mobilization", *Nature*, vol. 489, no. 7415, pp. 295–298, 2012.

[123] S. P. Borgatti, M. G. Everett, and J. C. J. C. Johnson, *Analyzing social networks*. SAGE Publications, 2013.

[124] P. Robbins, G. Pattison, "Logit models and logistic regression for social networks: III. valued relations", *Psychometrika*, vol. 64, no. 3, pp. 371–394, 1999.

[125] T. A. B. Snijders, G. G. van de Bunt, and C. E. G. Steglich, "Introduction to stochastic actor-based models for network dynamics", *Social Networks*, vol. 32, no. 1, pp. 44–60, 2010.