Measuring the Obesogenic Environment of Childhood Obesity

Kimberley Lonsdale Procter

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

The University of Leeds

School of Medicine

November 2007

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others

Academic Acknowledgements

The development of the java code for two of the three SimObesity algorithms was carried out by a team, which has included the author, Dianna Smith and Kirk Harland. The author's contribution to the work has been the choice and coding of the mathematical algorithms to use in SimObesity alongside Dianna Smith, with the ultimate decisions for SimObesity resting with the author. Mr. Kirk Harland provided the bulk of the java code (i.e. all the non-algorithm code) for the programme. This collaboration was to develop two static deterministic re-weighting spatial microsimulation models, SimObesity and SimHealth.

Survey data from the UK Data Archive (UKDA) was used in this research. Specifically the Health Survey for England 2002 provided by the National Centre for Social Research and University College London. And the Expenditure and Food Survey 2005 provided by the Office for National Statistics and Department for Environment, Food and Rural Affairs. Both datasets are managed by the University of Essex.

The 2001 Census statistics used in this thesis are Crown Copyright produced by the Office for National Statistics (ONS). Licensed for academic use by the Economic and Social Research Council (ESRC) and the Joint Information Systems Committee (JISC) Census Programme, which funded free access to the data for researchers in the UK. The data service providers is the Census Dissemination Unit through the Manchester Information and Associated Services (MIMAS) of Manchester Computing, University of Manchester, through an interface called CASWEB. Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

The boundary data used in this thesis are Crown Copyright produced by the Office for National Statistics (ONS). Licensed for academic use by the Economic and Social Research Council (ESRC) with the support of the Joint Information Systems Committee (JISC), which funded free access to the data for researchers in the UK. The data service providers are the Census Geography Data Unit (UK Boundary Outline and Reference Database for Education and Research Study: UKBORDERS) via Edinburgh University Data Library (EDINA).

The 2001 Census Output Area, Super Output Area and Ward Boundaries are Crown copyright 2003 where Crown copyright material is reproduced with the permission of the Controller of HMS.

The Postcode Directories were also used in order to locate and convert postcode locations to Census geographies, such as super output areas. These directories provided by the Census Geography Data Unit (UK Boundary Outline and Reference Database for Education and Research Study: UKBORDERS) via Edinburgh University Data Library (EDINA). The directories' copyright is © Crown Copyright 2006. Source: National Statistics / Ordnance Survey. Extracts are Crown Copyright and may only be reproduced by permission.

Acknowledgements

This research would not have been possible without the help, support and advice of many people, for which I am enormously grateful.

First and foremost, I would like to thank my supervisors: Prof. Janet Cade, Prof. Graham Clarke and Dr. Joan Ransley. They have consistently been there for me, offering guidance in the direction of my work, supporting my training requests, and providing feedback on papers and my thesis. This support has been invaluable to me and facilitated the completion of my research project.

I would also like to acknowledge the funding that was provided by the Economic and Social Research Council (ESRC) and Medical Research Council (MRC) as a joint award. Without this funding my PhD would not have been possible.

This PhD was supported by the collaboration of many partners, in particular in relation to the provision of height and weight data of children (with appropriate ethics approvals). The Leeds Primary Care Trusts (PCTs) provided access to the records of routinely collected data for children born since 1995 that was held on the National Child Health Computer System (NCHCS). Additionally the "Trends" project team assisted in providing me access to their study's height and weight data collected in 2004-2006, as did the "RADs" project team for their data collected in 2004-2006. This PhD would have been impossible without the use of this data and I would like to take this opportunity to thank everyone involved in facilitating my access to it.

There are also many people who have provided assistance on particular aspects of my work: Dr. Dimitris Ballas in relation to spatial microsimulation modelling; Prof. Mary Rudolf for childhood obesity expertise; Dr. Myles Gould for suggesting I use multi-level modelling (MLM), pointing me towards the MLWin courses in the University of Bristol, and wading through some MLM outputs with me; Dr. Richard Feltbower for statistical support, particularly MLM; Prof. Stuart Fotheringham and Dr. Martin Charlton for their expertise and advice on interpreting geographically weighted regression outputs; Dr. Dan Vickers regarding the Census Open Area Classification system; Dr. Charatdao Kongmuang for showing me her model and sharing ideas for formatting maps; Dr. Alison Heppenstall for discussions about clusters and hot spots; Anna Clough for actually being able to get the ArcGIS software to work every time it failed. They generously gave me their time and expertise, which was immensely valuable to me, and for which I will always be grateful.

Other people have helped me, directly and indirectly, and it is hard to acknowledge everyone individually. Thanks to the Nutritional Epidemiology Group for their general support - many information discussions gave me ideas to use in my research. Also I would like to thank the members of the Centre for Spatial Analysis and Policy (CSAP) and Multi-Agent Systems and Simulation (MASS) research clusters in the School of Geography for their support and advice.

Last but not least, I would like to thank my family, for their help, support and encouragement throughout the last three years. Somebody better tell Mark there is no need to go away climbing without me for long periods of time now that I have finished.

Abstract

Obesity prevalence has accelerated over the last two decades and is predicted to continue to rise, bringing with it increased morbidity and mortality as well as rising dramatically health care costs. Obesogenic environments are one of the explanations for the rising prevalence. Accordingly this thesis investigates the obesogenic environment factors, as well as obesogenic behaviour factors, associated with the increased prevalence of childhood obesity, using familiar geographical techniques in novel ways. These results were then applied to develop a targeted childhood obesity prevention policy for Leeds to reduce the risk of childhood obesity in different populations.

In this ecological study body mass index in Leeds for children aged 3 to 13 years old was examined to measure variations in childhood obesity. Spatial microsimulation modelling was utilised to give synthetic individual estimates of obesogenic covariates (e.g. obesogenic environment variables such as socio-economic characteristics and perceived social capital; individuals' behavioural variables such as dietary variables and physical activity levels) at the micro level. Additionally two demographic indices based on the 2001 Census were employed. The relationship between childhood obesity and the obesogenic covariates were considered at the home and school level using a combination of spatial statistical techniques.

Spatial microsimulation modelling was shown to be a robust method to estimate obesogenic covariates at the micro-level. In the design of a spatial microsimulation model using a deterministic re-weighting algorithm, the input variables must be strongly correlated with the output variables to be able to accurately simulate micro-area estimates. Also this thesis has highlighted that there is considerable advantage to analysing health data at a small scale, otherwise micro-level differences are simply "averaged" away and missed.

As well as showing that individuals' behaviours are important in determining risk of childhood obesity, this study adds to the increasing evidence of the existence of "obesogenic environments": features of the local environment in Leeds may affect childhood obesity by changing health behaviours. There was significant variation in childhood obesity across Leeds, with "hot spots" in both deprived and affluent areas. Further, relationships between obesogenic covariates and childhood obesity were not uniform across Leeds, highlighting the need for tailored, multifaceted public health policies that are based on locally relevant evidence.

Table of Contents

Academic Acknowledgements.....i
Acknowledgements.....ii
Abstract.....iii
Table of Contents.....iv
List of Figures.....vii
List of Tables.....ix
Abbreviations.....xi

Chapter 1: Introduction.....1
1.1 Introduction.....1
1.2 Hypothesis.....2
1.3 Aims and objectives of the research.....2
1.4 Thesis structure.....3

Chapter 2: The Aetiology of Childhood Obesity: A Review.....8
2.1 Introduction.....8
2.2 Impact of childhood obesity on health.....9
2.3 Definition and measurement of obesity.....10
2.3.1 British classification system.....12
2.3.2 International classification system (IOTF).....13
2.4 Current trends in childhood obesity.....17
2.5 Determinants of health behaviour.....20
2.6 Genetics of the environment?.....22
2.7 Aetiology of primary childhood obesity.....23
2.7.1 Physical (in)activity levels.....23
2.7.2 Diet.....25
2.7.3 Obesogenic environments.....28
2.8 Conclusion.....43

Chapter 3: The Geography of Obesity.....45
3.1 Introduction.....45
3.2 The geography of health.....45
3.2.1 Compositional or contextual?.....46
3.2.2 Social capital.....47
3.2.3 Scale of analysis.....48
3.3 The geography of obesity.....48
3.4 Discussion.....51

Chapter 4: Spatial variations in childhood obesity in Leeds.....53
4.1 Introduction.....53
4.2 Background.....54
4.2.1 Use of routinely collected data to monitor childhood obesity.....54
4.2.2 Spatial analysis of health data.....55
4.3 Methodology.....59
4.3.1 Study area.....59
4.3.2 Sources of data.....60
4.3.3 Analysis of data.....64
4.4 Results.....68
4.4.1 Routine data coverage.....69
4.4.2 Global analysis.....71
4.4.3 Spatial analysis.....77
4.5 Discussion.....84
4.5.1 Data collection.....84
4.5.2 Global analysis.....87

4.5.3 Local analysis.....88
4.5.4 Limitations.....90
4.5.5 Conclusion.....92

Chapter 5: Measuring the school impact on child obesity 94
5.1 Introduction.....94
5.2 Background.....95
5.3 Methodology.....97
5.3.1 Measurement of pupils' growth.....97
5.3.2 Analysis of data.....97
5.3.3 Development of model.....98
5.4 Results.....100
5.5 Discussion.....105

Chapter 6: A spatial microsimulation model of obesogenic environments and behaviours in Leeds: SimObesity 108
6.1 Introduction.....108
6.2 Small area population estimation methods.....109
6.3 SimObesity Model Specification.....111
6.3.1 Choice of obesogenic variables.....113
6.3.2 Choice of constraint variables.....116
6.3.3 Structure of SimObesity.....130
6.3.4 Input data.....132
6.3.5 Algorithm Methodology.....136
6.3.6 Output data.....148
6.4 Validation of synthetic micro-data 151
6.5 Discussion.....163
6.51 Validation / Calibration.....163
6.5.2 Methodology limitations.....166
6.5.3 Uses of SimObesity.....168
6.5.4 Conclusion.....170

Chapter 7: Micro-level analysis of childhood obesity, diet, physical activity, residential socio-economic and social capital variables: where are the obesogenic environments in Leeds?.....171
7.1 Introduction.....171
7.2 Background.....172
7.2.1 Obesogenic Environments.....172
7.2.2 Geographically Weighted Regression 173
7.3 Methodology.....176
7.3.1 Sources of data.....176
7.3.2 Analysis of data.....177
7.4 Results 179
7.4.1 Descriptive analysis.....179
7.4.2 Global analyses.....184
7.4.3 Geographically weighted regression analyses.....186
7.5 Discussion.....196
7.5.1 Global analyses.....196
7.5.2 Local analyses.....200
7.5.3 Relationship with income.....201
7.5.4 Limitations.....202
7.5.5 Conclusion.....203

Chapter 8: The development of a childhood obesity prevention policy for Leeds using an ANGELO-style framework.....205
8.1 Introduction.....205
8.2 Background.....206

8.2.1 Interventions to prevent obesity in children.....206
8.2.2 The ANGELO model.....210
8.2.3 Current obesity prevention policy in Leeds.....212
8.3 Methodology.....216
8.4 Results.....218
8.4.1 ANGELO-style framework to prevent obesity.....218
8.4.2 Identify three case study locations.....218
8.4.3 Obesity variance across each area.....225
8.4.4 Childhood obesity and deprivation / Open Area Classification super groups.....227
8.4.5 Childhood obesity and the simulated covariates.....227
8.4.6 Tailoring the framework to prevent obesity in Leeds.....233
8.4.7 Comparison to Leeds policy.....235
8.5 Discussion.....238

Chapter 9: Conclusion.....242
9.1 Introduction.....242
9.2 Summary of research findings.....242
9.3 Limitations of the research.....246
9.4 Possibilities for future research.....247
9.5 Conclusion.....248

References.....249

Appendix A: Diagram of causes of obesity.....283
Appendix B: SimObesity: structure of java code.....284
Appendix C: SimObesity: population files.....285
Appendix D: SimObesity: constraint files.....286
Appendix E: SimObesity: summary files.....288
Appendix F: SimObesiy:micro_config files.....289
Appendix G: Definitions and categorisation of the simulated obesogenic variables.....290

List of Figures

Figure 1.1 Structure of thesis chapters.....4
Figure 2.1 Growth charts for the British dataset.....15
Figure 2.2 Growth charts for international (IOTF) dataset.....15
Figure 2.3 Diagram of the global increase in the prevalence of childhood obesity.....18
Figure 2.4 Worldwide prevalence of overweight in children.....18
Figure 2.5 Worldwide prevalence of overweight and obesity.....19
Figure 2.6 Prevalence of obesity amongst UK children.....19
Figure 2.7 Ecological model of the determinants of health behaviour.....20
Figure 4.1 Map of study area.....60
Figure 4.2 Proportion of children classified as Acceptable, Overweight or Obese by year of measurement.....72
Figure 4.3 Proportion of children classified as Acceptable, Overweight or Obese by age group.....74
Figure 4.4 Scatter plot of deprivation score against mean BMI SDS for each SOA.....76
Figure 4.5 Map of proportion of obese children by SOA.....78
Figure 4.6 Map of proportion of obese children by ward.....78
Figure 4.7 Map of proportion of overweight and obese children by SOA.....78
Figure 4.8 Map of proportion of overweight and obese children by ward.....78
Figure 4.9 Map of mean BMI SDS by SOA.....78
Figure 4.10 Map of mean BMI SDS by ward.....78
Figure 4.11 Map of Index of Multiple Deprivation by SOA.....79
Figure 4.12 Map of OAC Super Groups by OA.....79
Figure 4.13 Map of relative risk of obesity, with hot spots highlighted.....81
Figure 4.14 Map of relative risk of overweight and obesity, with hot spots highlighted.....81
Figure 4.15 Map of mean BMI SDS, with hot spots highlighted.....81
Figure 4.16 Map amalgamating all "obesity" significant hot spots.....81
Figure 4.17 2 level multi-level model, with child at level 1 and SOA at level 2.....82
Figure 4.18 Plot of multi-level model's residuals, with 95% confidence intervals.....82
Figure 4.19 Map of hot and cold spots of obesity, after adjusting for deprivation and OAC super groups.....83
Figure 5.1 The steps involved in data analysis in the development of the methodology.....99
Figure 5.2 Geographic location of the schools, showing schools' mean BMI SDS and deprivation in Leeds.....104
Figure 5.3 Hot and cold spots of obesity in schools in Leeds (as per Value Added Model)....104
Figure 6.1 Venn diagram of obesogenic covariates.....111
Figure 6.2 Scatterplot of simulated and actual values for home ownership, using tenure as an input variable.....118
Figure 6.3 Scatterplot of simulated and actual values for home ownership, with tenure not an input variable.....118
Figure 6.4 Structure of SimObesity.....131
Figure 6.5 Formula to undertake the input variable adjustment.....135
Figure 6.6 Formula for first step of the reweighting algorithm.....140
Figure 6.7 Exert of java code corresponding to the first step of the reweighting algorithm....140
Figure 6.8 Formula for second step of the reweighting algorithm.....140
Figure 6.9 Exert of java code corresponding to the second step of the reweighting Algorithm.....140
Figure 6.10 Exert of java code corresponding to the first step of the integisation Algorithm.....146
Figure 6.11 Exert of java code corresponding to the second step of the integisation Algorithm.....147
Figure 6.12 Map of distribution of perceived low social capital in Leeds.....150
Figure 6.13 Map of distribution of populations with low income, no car, rent home and low food expenditure.....150
Figure 6.14 Map of distribution of obesogenic behaviours in children.....150

Figure 6.15 Map of distribution of people with low neighbourhood safety, low income, in deprived areas.....151
Figure 6.16 Bar charts of standarised absolute error for HSE simulation at ward level.....156
Figure 6.17 Bar charts of standarised absolute error for EFS simulation at ward level.....157
Figure 6.18 Scatterplots for each pair (simulated and actual) of constraint variables from HSE simulation.....158
Figure 6.19 Scatterplots for each pair of non-constraint variables from HSE simulation.....159
Figure 6.20 Scatterplots for each pair of constraint variables from EFS simulation.....160
Figure 6.21 Simulated against actual proportion of population who are male for each ward (HSE simulation).....161
Figure 6.22 Simulated against actual proportion of population in large households for each ward (EFS simulation).....161
Figure 7.1 Graphic illustration problems aggregating spatial data.....174
Figure 7.2 Key equations for a Geographic Weighted R regression model.....176
Figure 7.3 Map of childhood obesity across Leeds.....180
Figure 7.4 Maps of obesogenic covariates – 1.....181
Figure 7.5 Maps of obesogenic covariates – 2.....182
Figure 7.6 Maps of obesogenic covariates – 3.....183
Figure 7.7 Maps of parameter coefficient for key global covariates 189
Figure 7.8 Map of Leeds showing the location of the wards.....192
Figure 7.9 Map of areas where key local and global determinants of childhood obesity differ.....193
Figure 7.10 Maps of income, car availability, SEG and deprivation score across Leeds.....195
Figure 8.1 The ANGELO framework.....212
Figure 8.2 Map of location of the three case study wards.....224
Figure 8.3 Map of distribution of perceived low social capital in Leeds.....224
Figure 8.4 Map of less affluent populations.....224
Figure 8.5 Map of obesogenic behaviours in children.....224
Figure 8.6 Map of high deprivation and low neighbourhood safety.....224
Figure 8.7 Map of obesity in Leeds, presented as percentage obese and mean BMI SDS.....226
Figure 8.8 Map of 'hot spot' primary schools in case study wards.....226
Figure 8.9 Map of relationship between obesity and covariates in Wetherby.....230
Figure 8.10 Map of relationship between obesity and covariates in City and Holbeck.....231
Figure 8.11 Map of relationship between obesity and covariates in Morley South.....232

List of Tables

Table 1.1 Thesis outline.....3
Table 2.1 Age and gender specific BMI cut-offs to define overweight and obesity 16
Table 4.1 Details of number of children measured through routinely collected data.....69
Table 4.2 Details of number of children measured through the Trends and RADs studies...69
Table 4.3 Coverage of routinely collected data.....70
Table 4.4 Summary of routinely collected data by year of measurement.....71
Table 4.5 Summary of all three BMI datasets by age group.....71
Table 4.6 Summary of logistic regression for obese children over time.....73
Table 4.7 Summary of logistic regression for obese boys or girls over time.....73
Table 4.8 Summary of key statistics from logistic regression for boys or girls over time.....73
Table 4.9 Summary of logistic regression for obese children by age group.....75
Table 4.10 Summary of logistic regression for obese boys or girls by age group.....75
Table 4.11 Summary of key statistics from logistic regression for boys or girls by age group.....75
Table 4.12 Summary of the seven OAC super groups.....75
Table 4.13 Summary of the number of children measured by ward, SOA and OA.....77
Table 4.14 Summary of characteristics of the hot and cold spots of obesity that are evident after adjusting for deprivation and OAC super groups.....83
Table 5.1 Data on 35 schools, ranking the schools by the three methods.....102
Table 6.1 Details of obesogenic variables available in UK national surveys.....114
Table 6.2 Example showing why exact age is required to calculate if a child is obese or not.....121
Table 6.3 Details of participants in Health Survey for England 2002 excluded from study dataset.....123
Table 6.4 Comparison of Health Survey for England 2002 to 2001 Census data for Leeds.....123
Table 6.5 Results of stepwise logistic regression analysis for HSE dataset using all possible constraints.....124
Table 6.6 Results of entry logistic regression analysis using the optimal combination of constraints.....125
Table 6.7 Comparison of Expenditure and Food Survey 2005 to 2001 Census data for Leeds.....127
Table 6.8 Summary of key statistics for logistic regression analysis for EFS dataset.....129
Table 6.9 Details of output variables in EFS dataset that have correlations with the constraint variables.....129
Table 6.10 Summary of variables included within SimObesity.....131
Table 6.11 Example of the structure of a population file.....132
Table 6.12 Details of the source tables for the 2001 Census data.....133
Table 6.13 Example of the structure of a constraint file.....134
Table 6.14 Example of input adjustment calculation.....135
Table 6.15 Example of the structure of a summary file.....135
Table 6.16 Example of a micro_config files (used to read files in and out of SimObesity)..136
Table 6.17 Constraint tables used in worked example of the reweighting algorithm.....142
Table 6.18 Population table used in worked example of the reweighting algorithm.....142
Table 6.19 Summary tables used in worked examples of the reweighting algorithm.....142
Table 6.20 Results from first step of reweighting algorithm for the first constraint.....143
Table 6.21 Results from second step p of reweighting algorithm for the first constraint.....143
Table 6.22 Results from first step of reweighting algorithm for the second constraint.....144
Table 6.23 Results from the second step of reweighting algorithm for the second Constraint.....144
Table 6.24 Illustration of the integisation algorithm.....145
Table 6.25 Example of SimObesity output following the two algorithms in java code.....148
Table 6.26 Example of final SimObesity output following the conversion to aggregate Data.....148
Table 6.27 Comparison of simulated population (from HSE) to 2001 Census data.....153

Table 6.28 Comparison of simulated population (from EFS) to2001 Census data.....153
Table 6.29 Results of equal variance 2-tailed t-test for simulated (HSE) against actual (census) data.....162
Table 6.30 Results of equal variance 2-tailed t-test for simulated (EFS) against actual (census) data.....162
Table 7.1 Example of problems of using a global model to predict non-stationary Processes.....175
Table 7.2 Summary statistics for the obesogenic variables for whole of Leeds.....180
Table 7.3 Correlations between covariates and percentage of obese children in each SOA.....185
Table 7.4 Correlations between the obesogenic variables and deprivation / income.....186
Table 7.5 Summary of results of global and local GWR analyses.....187
Table 7.6 Summary of parameter estimates for each short-listed covariate and obesity.....188
Table 7.7 Summary of variability in parameter estimates by ward.....191
Table 7.8 Summary by ward of the important local determinants of childhood obesity.....194
Table 8.1 Summary of the key interventions to prevent childhood obesity.....207
Table 8.2 Outline of Department of Health priority initiatives to tackle obesity.....213
Table 8.3 Summary of the Leeds LPSA targets that may impact obesity prevalence.....213
Table 8.4 ANGELO-style framework for prevention of childhood obesity in Leeds.....219
Table 8.5 Leeds wards ranked by deprivation score.....223
Table 8.6 Details of the OAC super group for each of the case study wards.....223
Table 8.7 Obesity data for the case study wards.....225
Table 8.8 Summary of the OAC super groups by mean BMI SDS.....227
Table 8.9 Summary of variables associated with higher/lower obesity in each case study area.....229

Abbreviations

ABM	Agent based modelling
ANGELO	ANalysis Grid for Environments Linked to Obesity
ArcGIS	GIS software package
BIA	Bioimpedance analysis
BMI	Body mass index
BMI SDS	Body mass index standard deviation score
CDC	Centres for Disease Control and Prevention (USA)
CH	City and Holbeck ward
CT	Computer tomography
DEFRA	Department for Environment, Food and Rural Affairs
DH	Department of Health
DXA	Dual energy X-ray absorptiometry
EFS	Expenditure and Food Survey, 2005
FLeXscan	Flexible scan statistic
GIS	Geographic Information System
GWR	Geographically weighted regression
HSE	Health Survey for England, 2002
IMD	Index of Multiple Deprivation (relating to Index of Deprivation, 2004)
IOTF	International obesity task force
LAA	Local Area Agreement
LPSA	Local Public Service Agreements
MRI	Magnetic resonance imaging
MAUP	Modifiable Areal Unit Problem
MLM	Multi-level modelling
MLwiN	Multi level modelling software package
MS	Morley South ward
NCHCS	National Child Health Computer System
NHANES	National Health and Nutrition Examination Survey (USA)
NSMC	National Social Marketing Centre
OA	Census output area
OAC	National Statistics 2001 Open Area Classification of Output Areas
PCT	Primary Care Trust
PE	Physical education
RCT	randomised controlled trials
RADS	Rugby League and Athletics Development Scheme
SAE	Standarised absolute error
SaTScan	Spatial Scan Statistic
SEG	Socio-economic group
SES	Socio-economic status
SimObesity	A spatial microsimulation model for the obesogenic environment
SOA	Census low super output area
TAE	Total absolute error
TV	Television
W	Wetherby ward
W/H	Weight for height indices
WHO	World Health Organisation
WinBUGS	A statistical software package (part of the BUGS project – Bayesian inference Using Gibbs Sampling)

Chapter 1: Introduction

1.1 Introduction
1.2 Hypothesis
1.3 Aims and objectives of the research
1.4 Thesis Structure

1.1 Introduction

"Obesity [is] abnormal or excessive fat accumulation that may impair health"

(World Health Organisation, 2006)

"Through most of human history, weight gain and fat storage have been viewed as signs of health and prosperity. In times of hard labour and frequent food shortages, securing an adequate energy intake to meet requirements has been the major nutritional concern. Today, however, as standards of living continue to rise, weight gain and obesity are posing a growing threat to health in countries all over the world".

(World Health Organisation, 2004a)

Obesity in children is a rapidly growing problem and is associated with a number of co-morbidities in childhood and with increased risk of adult disease, particularly cardiovascular disease, hypertension and type 2 diabetes. Reducing childhood obesity and health inequalities are at the centre of the UK government's health policy. Halting growth in childhood obesity is a prime objective (Department of Health, 2004) (reviewed by Foster & Buttriss, 2005). There is a debate around the reasons for this increasing prevalence and obesogenic environments are one of the explanations. However, whilst the importance of the environment in controlling obesity is widely acknowledged, environmental strategies to prevent obesity remain relatively unexplored. Obesity is notoriously difficult to correct after becoming established plus obese children are more likely to become obese adults. Accordingly it follows that prevention of obesity in children is key.

This study brings together the disciplines of medicine, nutrition, geography and public health to investigate the environmental factors associated with the increased prevalence of childhood obesity, using familiar geographical techniques in novel ways to further our understanding of childhood obesity. It also seeks to develop methods to facilitate the prevention of obesity in children.

In this ecological study body mass index (BMI) in Leeds for children aged 3 to 13 years old was examined to measure variations in childhood obesity. These data are derived from the data collected routinely by the Primary Care Trusts and from two studies, "Trends" and "RADs", that are being conducted in Leeds. Spatial microsimulation modelling was utilised to give synthetic individual estimates of obesogenic covariates (e.g. obesogenic environment variables

such as socio-economic characteristics and perceived social capital; individual obesogenic behaviour variables such as dietary variables and physical activity levels) at the micro level, as otherwise these data are only available in national surveys. This is cheaper, quicker and more detailed than conducting a survey in Leeds. Additionally two demographic indices based on the 2001 Census were employed: the Index of Deprivation 2004 and the Census Open Area Classification system. The relationships between childhood obesity and the obesogenic covariates were considered at the home and school level using a combination of spatial statistical techniques (including, geographically weighted regression, multi-level modelling, and spatial scan statistics). These results are applied to develop a framework for a targeted childhood obesity prevention policy for Leeds to reduce the risk of childhood obesity in different populations, which is compared and contrasted to the actual policies in Leeds.

1.2 Hypothesis

Obesogenic covariates for the locality will be related to patterns of obesity in children.

1.3 Aims & objectives of the research

The principal aim of this thesis is to investigate the micro-level variability in childhood obesity and obesogenic environments/behaviours. More specifically:

- To describe, measure and map childhood obesity and obesogenic covariates across Leeds.
- To identify relationships between obesogenic covariates and childhood obesity, highlighting those factors with the strongest associations and considering the variation in these relationships across Leeds.
- To identify populations 'at risk' of childhood obesity, both at the home and school levels, defined either spatially and/or by any of the covariates analysed.
- To develop targeted environmental interventions and health policies to reduce childhood obesity in Leeds.

In order to achieve these aims, the following research objectives were formulated:

1 To review the aetiology and geography of childhood obesity, and interventions to prevent childhood obesity.
2 To evaluate small area population estimation methods and spatial analysis techniques.
3 To describe obesogenic environments/behaviours and the "ANGELO" model.
4 To examine current obesity prevention policy in Leeds.
5 To identify and obtain the available data for childhood obesity in Leeds.
6 To investigate and determine sources of data for obesogenic covariates.

7 To explore the temporal and micro-level spatial variations in childhood obesity in Leeds at residential level, using spatial analysis techniques, and identify "hot" and "cold" spots of childhood obesity.

8 To explore the variation in childhood obesity at the school level, using spatial analysis techniques, and identify the schools that are under or over performing in respect to their levels of pupils' obesity.

9 Build and validate SimObesity, a static spatial microsimulation model for obesogenic covariates.

10 Use SimObesity to estimate obesogenic covariates at the individual level, and aggregate to Census low super output area, in Leeds.

11 To study the relationship between childhood obesity and obesogenic covariates in Leeds, using spatial analysis techniques, to identify those covariates with the strongest associations with childhood obesity and also to determine how the relationships vary at the micro-level.

12 To apply the knowledge gained in objectives 1-11 to develop a framework for childhood obesity prevention policy, and tailor it for different population groups.

13 Compare and contrast the theoretical childhood obesity prevention policy with the actual obesity policy in Leeds.

14 To evaluate the success of the research and propose the possibilities for future work.

1.4 Thesis Structure

In order to accomplish the research objectives set out above, this thesis is structured into nine chapters as outlined in Table 1.1. Each chapter relates to one or more research objective, and each research objective may be covered in more than one chapter as knowledge builds from one chapter to the next. The way in which each chapter is related to others is described in Figure 1.1.

Chapter		Objective(s)
Chapter 2:	The Aetiology of Childhood Obesity: A Review	1
Chapter 3:	The Geography of Obesity	1
Chapter 4:	Spatial variations in childhood obesity in Leeds	2, 5, 7
Chapter 5:	Measuring the school impact on child obesity	1, 2, 8
Chapter 6:	A spatial microsimulation model of obesogenic environments and behaviours in Leeds: SimObesity	2, 6, 9, 10
Chapter 7:	Micro-level analysis of childhood obesity, diet, physical activity, residential socio-economic and social capital variables: where are the obesogenic environments in Leeds?	2, 3, 11
Chapter 8:	The development of a childhood obesity prevention policy for Leeds using an ANGELO-style framework	1, 3, 4, 12, 13
Chapter 9:	Conclusions	14

Table 1.1. Thesis outline

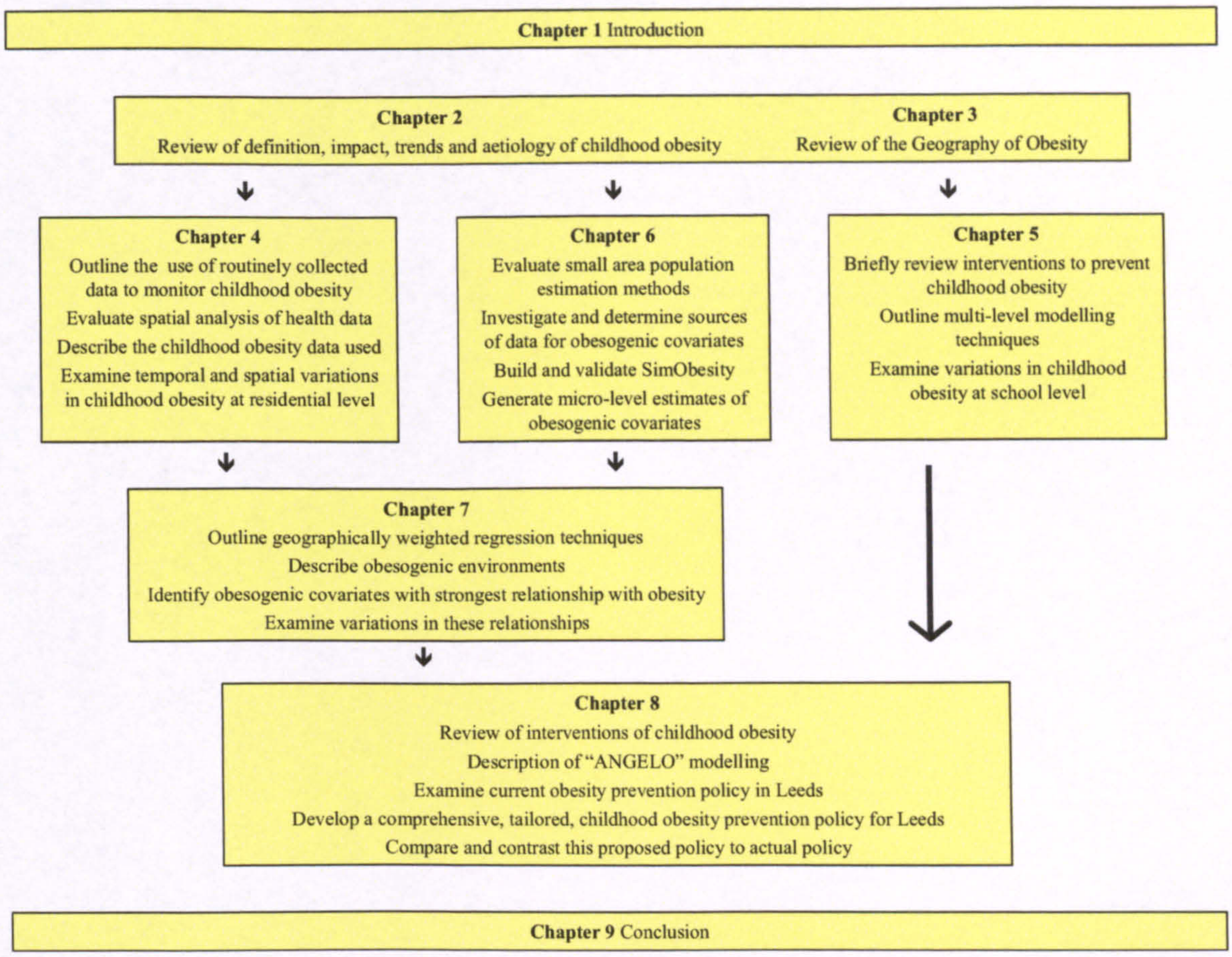


Figure 1.1. Structure of thesis chapters

Chapter 2: The Aetiology of Childhood Obesity: A Review

This chapter reviews the literature to examine the implications of obesity in children, from both the perspective of the increased health risk to the individual and the high economic cost of its treatment. Obesity is then defined, the methods to measure childhood obesity examined, and recent worldwide trends explored. The determinants of health behaviour are outlined to facilitate understanding of why the aetiology of obesity is so complex. The question of whether genetics or the environment are leading the rapid rise in prevalence of obesity is also investigated. The paper then considers the evidence base for the multi-factorial aetiology of childhood obesity under three key headings: physical activity, diet and obesogenic environments. Finally the chapter draws conclusions for different causes of childhood obesity, in particular considering the importance of the obesogenic environment.

Chapter 3: The Geography of Obesity

Following a definition for the broad concept of the geography of health, this chapter reviews selective areas of health geography that can be applied to obesity, in particular examining the composition and contextual effects debate, followed by an explanation of the concept of social capital and health. The issue of spatial scale is also addressed. There is a brief overview of the geography of obesity literature, focusing on obesity research regarding food access, green space availability, neighbourhood characteristics, and degree of urbanization, as well as examining the effect of multiple environmental determinants on obesity. Finally there is a discussion about how this type of analysis can add value to obesity research and prevention initiatives.

Chapter 4: Spatial variations in childhood obesity in Leeds

There are two key foci to this chapter. First, in relation to whether it is feasible to use routinely collected data to monitor trends in prevalence of childhood obesity in the UK. Secondly, a micro-level spatial analysis of childhood obesity in Leeds (using the routinely collected data, together with data gathered from two studies in Leeds). By way of background, this chapter outlines the use of routinely collected data to monitor childhood obesity, and then the use of Geographic Information Systems (GIS) and spatial statistics to analyse health data together with an exploration of the problems associated with this type of analysis. Then the methodology is described, starting with a description of the sources of data used, followed by an explanation of the analysis undertaken. The results of the analysis of the routinely collected data are presented first, followed by the examination of the obesity data. The routinely collected height and weight data and survey data on older children from two studies in Leeds (children aged 3-6 years, 9 years, 11 years and 13 years) are used to describe childhood obesity in Leeds, from both a global (whole of Leeds) and micro-level (spatial) perspective, together with a consideration of the demographics across Leeds. This will identify "hot" and "cold" spots at the residential micro-level, with consideration of how these populations differ from each other. Finally the discussion section puts the results into context of the existing body of work and an indication of the insights that these results give to the understanding of the whole area is given.

Chapter 5: Measuring the school impact on child obesity

This chapter explores the impact that schools have on their pupils' obesity, identifying those where targeted input is most needed. It provides some background information regarding obesity and the obesogenic environment, briefly overviewing interventions to prevent childhood obesity (more detail is given in Chapter 8) and the role that schools may be able to play to facilitate the prevention of obesity in children. The methodology covers the description of the childhood obesity data used and the three steps in the modelling process that was developed. There is an explanation and appraisal of multi-level modelling as a technique. The results show that there was variation between the schools in terms of their levels of obesity, although only a

small proportion of this variation was accounted for by residential deprivation score and ethnicity of the child. The results of the three different rankings of the schools are also explained. This modelling process enabled the identification of the schools whose levels differed from that expected given the socio-demographic make up of the pupils attending, highlighting the "highest" and "lowest" performing schools, and suggested that there may be a significant school effect.

Chapter 6: A spatial microsimulation model of obesogenic environments and behaviours in Leeds: SimObesity

SimObesity is a spatial microsimulation model that is designed to synthesise obesogenic covariates at the small-area level in Leeds in order that the relationship between actual obesity data and these synthetic obesogenic covariates (obesogenic environment variables and individuals' behaviour variables) data can be analysed at the micro level. This is a novel application of spatial microsimulation modelling. This chapter explains how the SimObesity model was built and validated. It firstly explores the alternatives for estimating small area populations. Then it describes the SimObesity model specification, including the choice of variables to use in the model, and how the model is structured and works. This includes a summary of the investigation and determination of the available sources of data for obesogenic covariates, plus a detailed explanation of the algorithms used. The model output is the creation of synthetic individuals in Leeds whose characteristics match as closely as possible the characteristics of the actual individuals living in Leeds as shown in the 2001 Census. SimObesity combines individual micro-data from the Health Survey for England 2002 (HSE), and separately from the Expenditure and Food Survey 2005 (EFS), with census statistics for lower Super Output Areas (SOAs) to create synthetic micro-data estimates for SOAs in Leeds. The resulting, synthesised, micro dataset includes all the attributes from both the survey and the census datasets, thereby creating estimates of obesogenic variables at the micro-level in Leeds. The validation of the synthetic micro-data is deliberated, with a concluding discussion of the robustness of the methodology to estimate small area level obesogenic covariate data.

Chapter 7: Micro-level analysis of childhood obesity, diet, physical activity, residential socio-economic and social capital variables: where are the obesogenic environments in Leeds?

This chapter describes global and local analyses of the relationship between childhood obesity and many potential obesogenic variables - including obesogenic environment variables such as deprivation, urbanisation, socio-economic group, income, access to local amenities and perceived neighbourhood safety; as well as individuals' obesogenic behaviour variables such as fruit and vegetable consumption, household expenditure on food generally and school meals, and physical activity behaviours. The covariate data were all synthesised at the individual/household level using spatial microsimulation (SimObesity). This builds on the work

undertaken in chapters 4 and 6. This chapter provides a definition and description of obesogenic environments, as well as the rationale for using geographically weighted regression techniques. Then, using spatial analysis techniques, it considers whether various neighbourhood features or individuals' behaviours, such as perceived supermarket accessibility or fruit and vegetable consumption (respectively), contribute to a higher risk for childhood obesity. Accordingly these analyses identify the covariates with the strongest relationships with obesity, as well as highlighting the micro-level variation in these relationships across the study area. This highlights "at-risk" populations, defined either spatially and/or by any of the covariates analysed. It seeks to demonstrate the importance of analysis at the micro-level in order to provide health planners with additional information with which to tailor interventions and health policies to prevent childhood obesity.

Chapter 8: The development of a childhood obesity prevention policy for Leeds using an ANGELO-style framework

This chapter seeks to pull together the work undertaken in previous chapters into a practical application. There is a review of the interventions to prevent childhood obesity, together with a description of the ANGELO framework to prevent obesity in children, and an investigation into the current obesity prevention policy in Leeds. Then, having identified the key obesogenic factors at the micro-level, three case studies of different neighbourhoods in Leeds are studied in depth – the most affluent, the most deprived, and a middle-of-the-road ward. An ANGELO-style framework was used to develop and prioritise potential targeted interventions and health policies to facilitate the prevention of childhood obesity in these three areas, using factors that are amenable to measurement, intervention, and change. These suggestions are compared and contrasted with actual local policy in Leeds with a discussion regarding what could, and perhaps, should be altered to facilitate the prevention of childhood obesity.

Chapter 9: Conclusions

This last chapter concludes the study by summarising the findings of the research, considering whether the aims and objectives have been met, and discussing the limitations of this work. It then looks to possibilities for future work, before finishing with the key messages from this thesis.

Chapter 2: The Aetiology of Childhood Obesity: A Review

2.1 Introduction
2.2 Impact of childhood obesity on health
2.3 Definition and measurement of obesity
2.3.1 British classification system
2.3.2 International classification system (IOTF)
2.4 Current trends in childhood obesity
2.5 Determinants of health behaviour
2.6 Genetics of the environment?
2.7 Aetiology of primary childhood obesity
2.7.1 Physical (in)activity levels
2.7.2 Diet
2.7.3 Obesogenic environments
2.8 Conclusion

2.1 Introduction

This review is focusing on the increasing prevalence of childhood primary obesity, a condition caused by chronic positive energy imbalance due to excess energy intake and/or insufficient energy expenditure. (as opposed to the rare instances of Secondary Obesity, which can occur due to endocrine problems (e.g. Cushing's syndrome, hypothyroidism) and genetic abnormalities (e.g. Down's syndrome, Prader-Willi syndrome) (Flodmark et al, 2004)). What has changed in society and behavioural patterns in recent years to warrant the rapid rise in the prevalence of childhood obesity that is currently evident? Why are children consuming too many calories and/or not taking enough exercise?

This chapter starts by briefly considering why we should be concerned about obesity in children: looking at the impact of childhood obesity, both in terms of the child's health and the strain it imposes on the health system. These factors facilitate an understanding of why prevention of obesity in children is so important. It then examines the definition of obesity and how childhood obesity is measured. The chapter then looks at the current increasing trends in the prevalence of childhood obesity. Next the determinants of health behaviour are explored, moving on to consider the question of whether genetics or the environment are leading the rapid rise in prevalence of obesity. Next, the different causes of childhood obesity and the rationale for these increasing trends are investigated, in order to understand what is changing in society and our children's behaviour that is causing the positive energy balance leading to obesity. Finally the chapter draws conclusions about the evidence base for different causes of childhood obesity, in particular considering the importance of the obesogenic environment.

2.2 Impact of childhood obesity on health

Obesity in children, and adults, is a rapidly growing problem in the UK and worldwide and has been increasing at accelerating rates in more recent years. Childhood obesity is associated with a number of co-morbidities in childhood and with increased risk of adult disease, particularly cardiovascular disease, hypertension and type 2 diabetes. Obese children tend to be more isolated and have lower self-esteem than their peers (Strauss, 2000).

Reducing childhood obesity and health inequalities are at the centre of the UK government's health policy. The government's "Choosing Health" white paper on improving public health in England (Department of Health, 2004) (reviewed by Foster & Buttriss, 2005) outlines a number of actions to tackle key current public health issues. Specifically, six key priorities for action have been identified, with children's health, particularly childhood obesity, being a major focus. Halting growth in childhood obesity is their prime objective. This objective was very recently updated by the government to "by 2020, we aim to reduce the proportion of overweight and obese children to 2000 levels" (DH, 2008)

One of the steps towards achieving this is the development of a national social marketing strategy. Health related social marketing is "the systematic application of marketing concepts and techniques, to achieve specific behavioural goals, to improve health and reduce inequalities" (NSMC, 2006). Importantly this process addresses short, medium and long term issues, recognising that encouraging healthy choices and associated behavioural change is a complex process, requiring more than merely increased public awareness of health issues. This shows how government public health policy is moving away from considering disease groupings in isolation, towards a population approach that considers the determinants of health – which is why obesity has suddenly risen up the agenda.

Obesity related diseases account for a substantial proportion of costs of health care resources worldwide (WHO, 2004a). The Select Committee Report on Obesity (2004) estimated that the total cost of treating obesity in the UK was £3.3-3.7 billion in 2002 and will increase to £7 billion by 2020. The Foresight report (2007) has now increased this estimate to £45.5 billion by 2050. As well as being expensive, the treatment of obesity is difficult and time consuming; whilst it can be effective, weight regain is common. Obesity is notoriously difficult to correct after becoming established plus obese children are more likely to become obese adults (Guo et al, 1994; Freedman et al, 2002), with all the corresponding health and social disadvantages. Whilst recognising that treatment of obesity is also an important approach that needs to be addressed concurrent with prevention approaches, prevention of obesity is likely to be more cost effective than treatment (WHO, 2004a). Without a focus on prevention, the unavoidable high

cost of managing the obesity epidemic will almost certainly be too expensive for many countries. Accordingly it follows that prevention of obesity in children is key.

The importance of the environment in controlling obesity is widely acknowledged. A WHO report (2003) states that major social and environmental changes to make healthier choices more accessible and preferable are required to prevent obesity. The strength of an environmental approach is that significant population benefits can result from even fairly small effects if a large number of people are exposed to that environment (Swinburn & Egger, 2002).

Yet whilst prevention of childhood obesity is the only viable, enduring, cost effective, solution, effective methods for it remain elusive. Furthermore strategies to influence obesogenic environments remain relatively unexplored. In order to be able to develop powerful population level interventions and public health policies to prevent childhood obesity, we need to fully understand its aetiology and those environments that are most amenable to measurable change, which is what this review seeks to consider.

2.3 Definition and measurement of obesity

The definition of obesity is not merely an excess of body fat, but where that excess is sufficient to be detrimental to health (WHO, 2006).

So how should that level of 'excess fatness detrimental to health' (obesity) be measured? There are four alternatives. (1) To measure the increased risk of adverse health consequences. It is difficult to use measures of increased risk of adverse health consequences to measure childhood obesity because excess fatness has a gradual negative impact on health (Neovius et al, 2004), which in turn depends upon the extent of excess fatness and the duration (WHO, 2004a). Accordingly, obesity-related morbidity in adults is clearer than in children. (2) To measure body fat. There are many methods available to estimate body fatness, including dual energy X-ray absorptiometry (DXA), underwater weighing, magnetic resonance imaging (MRI), bioimpedance analysis (BIA), computer tomography (CT) (Flodmark et al, 2004; Freedman et al, 2004). However, whilst they are accurate, they are expensive and time consuming, making them impractical to use on a routine basis (and so for epidemiological studies). Their use is normally limited to research purposes in clinical settings. (3) Other anthropometric measurements may be used to define childhood obesity, such as skinfold thickness and waist circumference. However cut off points are not universally agreed. Furthermore, assumptions required for these methods can be invalid during growth and maturation leading to inaccurate measurements in children (Freedman et al, 2004). Also they can be very impractical to undertake on infants and very young children. (4) Given the limitations with the above methods

of estimating obesity, weight for height indices are generally used to estimate body fatness, and in particular, Body Mass Index (BMI). BMI more closely correlates with other measures of adiposity (e.g. skinfold thickness) than other weight for height indices, namely W/H^3 and W/H^p (Michielutte et al, 1984; Lazarus et al, 1996; Frontini et al, 2001; Freedman et al, 2004).

BMI is a useful proxy measurement of body fat at a population perspective. It has been validated (by the IOTF - Dietz & Bellizzi, 1999) against other, more direct, measures of body fatness (such as DXA and BIA) and its use to measure obesity in children is generally accepted (Bellizzi & Dietz, 1999; Dietz & Bellizzi, 1999). It is safe, easy and inexpensive to undertake. It is calculated as a person's weight in kilograms divided by the square of their height in metres. The main limitation of BMI is its association with height and the exponential relationship between weight and height, so it may be (more) inaccurate for individuals at height extremes (Lara-Esqueda et al, 2004). In particular, it tends to under estimate obesity for taller children. Also BMI varies with sex, race and maturation (Daniels et al, 1997; Neovius et al, 2004), which complicate the choice of appropriate cut-off points for children. During childhood BMI shows significant variations (see Figure 2.1, below), which is the reason why age and gender specific reference standards must be used. Further, the time course of these variations will fluctuate due to differences in maturation patterns (Guo et al, 1998).

There are ethnic differences in the use of BMI to predict increased morbidity and mortality (WHO Expert Consultation, 2004). In adult Caucasians the risk of increased ill health due to obesity commences at (and above) a BMI of 25 kg m^{-2}. However in Asians this risk commences at (and above) a BMI of around 23 kg m^{-2} (although this varies from 22-25 kg m^{-2}). Whilst the cut-offs are not as clear cut for children, it is likely that this trend is true for child populations, although this difference in adults may be due to differences in height between ethnic groups (Lara-Esqueda et al, 2004).

In adults (in the West) there is consensus that overweight is defined as BMI between 25-30 kg m^{-2} and obesity as BMI $\geq$ 30 kg m^{-2} (WHO, 2004a). These figures are based, albeit crudely, on known risk ratios for these different levels of BMI. No such data exist for children. In children it is difficult to determine at what BMI increased risk of health consequences occurs – due to maturation differences, age, gender, the time lag between excess fatness and disease, etc. So, unlike for adults, there is no universal consensus of the cut off points to use for children.

Notwithstanding this, there are two accepted ways to determine obesity in children. One classification system uses international reference data and the other uses national reference data. The debate seems to continue about which dataset should be used when, particularly, for some reason, in the UK (Reilly (2002) provides a good review). In summary, clinicians should use

the relevant national dataset, and researchers who wish to make international comparisons should use the IOTF definition. If in doubt, use both. It is important to use an international dataset when making secular or cross-cultural comparisons between large epidemiological datasets for children. The IOTF dataset is the gold standard international reference to compare worldwide populations (albeit limited to those in Western countries) and it combines data from 6 different countries (GB, Netherlands, Brazil, US, HK, Singapore). However when working with data for children from only one country, then that countries reference data (if available) can be used. The British reference dataset was produced using data on nearly 15,000 children from the combination of several UK surveys undertaken between 1980-1990 (Cole et al, 1995). Other examples of national datasets include NHANES and CDC, which are both US based national datasets and so not accurate for UK children. However, in practice, even if the children in the data set under examination are only from one country, the IOTF dataset should be used (probably as well as the national definition) to enable the work to be compared with other international work.

2.3.1 British reference system

When using the British reference system, BMI Z-scores are used to compare an individual or specified population against the British reference population, which is then used to identify the numbers of obese children. A Z-score is based on the difference between the observed value and the median reference value of a population, standardized against the standard deviation of the reference population, and a BMI Z-score is calculated as follows:

$$\frac{\text{(observed value) - (median reference value of a population)}}{\text{standard deviation of reference population}}$$

Therefore for a normal Gaussian distribution, 50% of children would be expected to have a z-score greater than 0, 16% of children would be expected to have a Z score of greater than +1, 2% of children would be expected to have a Z score greater than +2, assuming the sampled population is identical to the reference population. That is, by definition 2% of children would be expected to have a BMI above the 98th centile value.

Figure 2.1 shows the growth charts for the British dataset. This clearly shows how BMI varies with age (there is a different graph for boys and girls as BMI varies with gender as well as age). BMI typically rises during the first months after birth, falls after the first year and rises again around the fifth or sixth year of life - this second rise is referred to as 'the adiposity rebound'. A given value of BMI therefore needs to be evaluated against age- and gender-specific reference values.

The British dataset conventionally uses the 91st centile as the cut off for overweight and the 98th centile as the cut off for obesity (Rudolf et al, 2005a), although some authors do not stick to this (for example, using the 85th and 95th centiles as the cut off for overweight and obesity, respectively, which will obviously give a larger number of obese children, thus increasing the strength of any associations with risk, or protective, factors). This is probably because the 95th centile as a cut-off for obesity has been recommended in the past (Barlow & Dietz, 1998) and this is the cut-off for other national datasets (Rudolf et al, 2005a). So when comparing studies, even if authors are using the same reference population the actual cut off point used should also be considered.

The use of a BMI Z-score has advantages, permitting comparison between populations, the depiction of temporal trends and a comprehensive statistical description of an individual or a population. Also BMI-for-age charts can be useful in a clinical setting as a child can be described as being above or below a particular centile (e.g. the 91st centile).

However this method is not without its difficulties. Firstly in choosing the appropriate reference population. The data are derived from a single reference population and classifying the individual as overweight or obese assumes they are comparable to that reference population. If you consider how much the demographics of the UK population are changing, that is increasing numbers of ethnic minorities, this might not be a valid assumption. Plus clinicians may erroneously assume the centiles represent an ideal population, when the data may come from a reference population with a high prevalence of obesity (e.g. USA reference data). Other problems include the fact that the Z-score methodology requires a degree of statistical skills and/or the appropriate software in order to do the calculation. Most importantly perhaps the cut-offs to define overweight and obesity are based only on statistical convenience (e.g. BMI Z-score > 2) rather than a known health risk.

2.3.2 International classification system (IOTF)

Moving on to look at the IOTF definition, which has a different methodology. The IOTF reference dataset uses growth data from six different countries, namely Brazil, Britain, HK, Singapore, Netherlands and USA. These are shown in Figure 2.2.

As it is not clear at which BMI level adverse health risk factors increase in children, the IOTF definitions do not relate to specified centiles. Instead the IOTF extrapolated the adult cut offs of 25 and 30 kg/m^2 for overweight and obesity, respectively, at age 18, back to childhood to identify the age and gender specific BMI cut offs to define overweight and obesity in children corresponding to the adult cut off points. Each of the six datasets were used to estimate these

centiles, which were then averaged to provide the published cut-off points. These are listed in Table 2.1. Where an age is between the half years given, interpolation between the two values is used. As with the British reference data a given value of BMI still needs to be evaluated against age- and gender- specific reference values.

Whilst the IOTF definition is the 'gold standard' to use in order to facilitate comparisons across borders, it is not perfect. The IOTF cut-offs have been questioned by Voss et al (2005) (running the EarlyBird study in Plymouth) as these child cut-offs are based on adult thresholds, which are based on known health risks for adults, not for children. Results from the EarlyBird study show that these cut-offs in young children are poor indicators of insulin resistance and so of metabolic risk and diabetes. They are concerned that the use of these cut-offs could stigmatise the heavier child unduly. It should be remembered that different national and international cut-offs produce different estimates of childhood obesity for the same child, which is clearly demonstrated by Flodmark et al (2004) and Chinn & Rona (2002). The latter authors compared the IOTF definitions with the British reference data using BMI data from 6000 white children from primary schools across England. The IOTF cut-offs resulted in higher prevalence of obesity in girls and lower in boys (a consequence of averaging curves of different shapes).

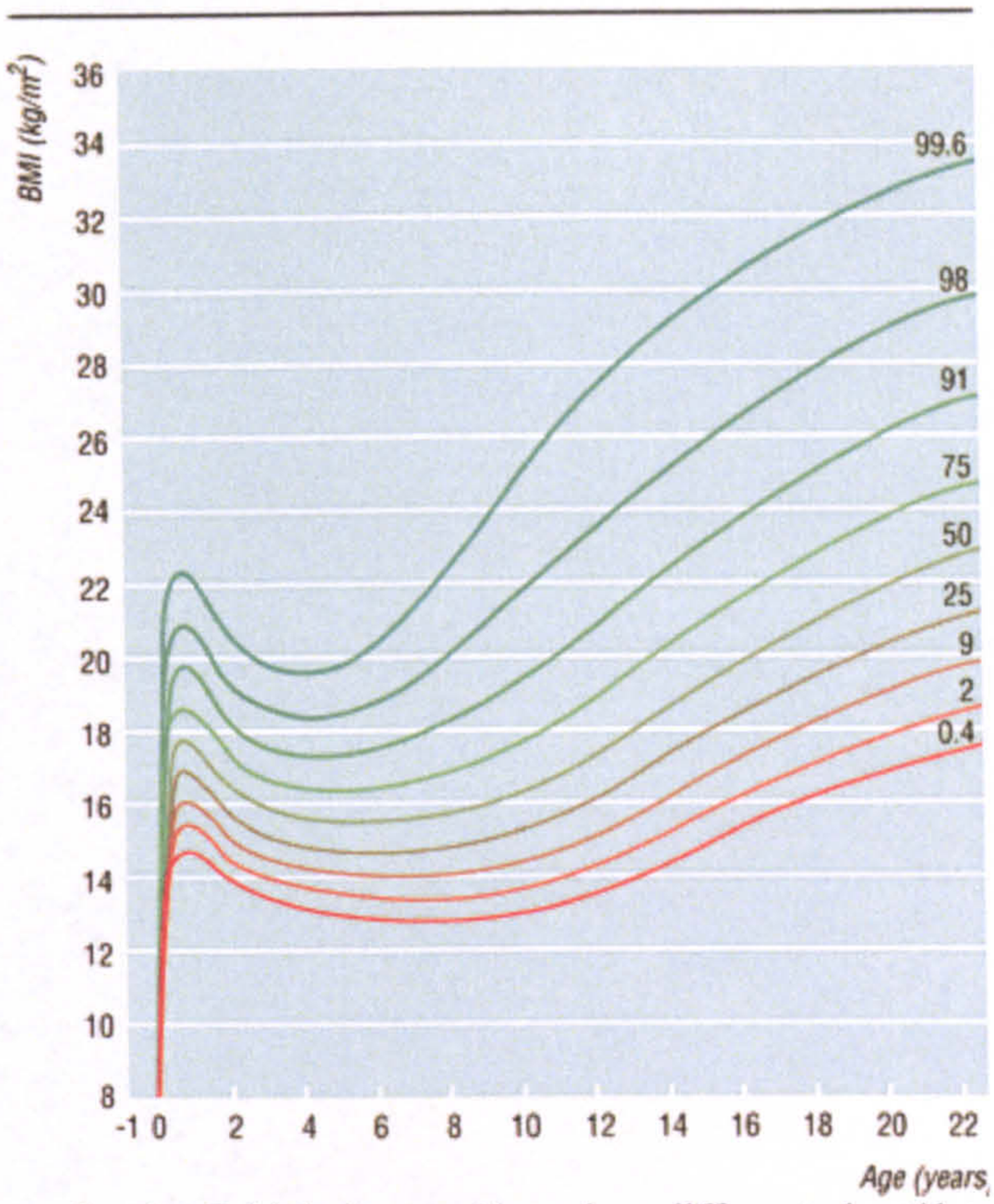


Figure 2.1. Growth charts for the British dataset (there is a different chart/dataset for boys and girls as BMI varies with gender as well as age) (Prentice, 1998). Age is along the x axis and BMI along the y axis. These charts include nine centile curves based on divisions of two thirds of a standard deviation, thus ranging from the 0.4th to 99.6th centile. The centiles were fitted using a least mean squared method, with the BMI distribution being adjusted at various ages for differing amounts of skewness (Cole & Green, 1992). Age and gender specific BMI cut offs define overweight as above the 91st centile and obesity as above the 98th centile (Cole et al, 1995).

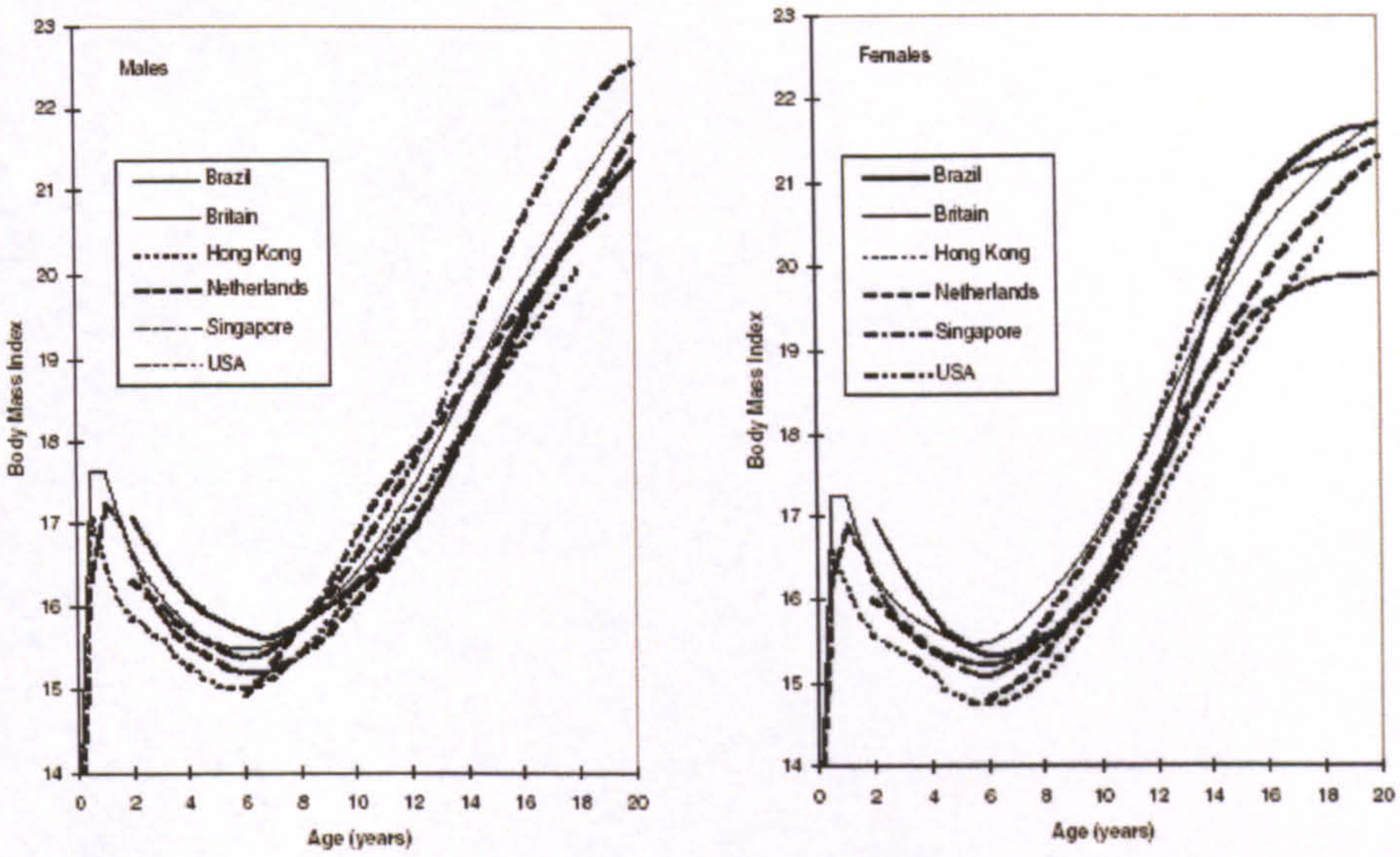


Figure 2.2. BMI plotted against age for six large-scale surveys of children, boys on the left and girls on the right, showing the 'rebound' after age 5 years (Lobstein et al, 2004)

Age (years)	Body mass index 25 kg m^2		Body mass index 30 kg m^2	
	Males	Females	Males	Females
2	18.41	18.02	20.09	19.81
2.5	18.13	17.76	19.80	19.55
3	17.89	17.56	19.57	19.36
3.5	17.69	17.40	19.39	19.23
4	17.55	17.28	19.29	19.15
4.5	17.47	17.19	19.26	19.12
5	17.42	17.15	19.30	19.17
5.5	17.45	17.20	19.47	19.34
6	17.55	17.34	19.78	19.65
6.5	17.71	17.53	20.23	20.08
7	17.92	17.75	20.63	20.51
7.5	18.16	18.03	21.09	21.01
8	18.44	18.35	21.60	21.57
8.5	18.76	18.69	22.17	22.18
9	19.10	19.07	22.77	22.81
9.5	19.46	19.45	23.39	23.46
10	19.84	19.86	24.00	24.11
10.5	20.20	20.29	24.57	24.77
11	20.55	20.74	25.10	25.42
11.5	20.89	21.20	25.58	26.05
12	21.22	21.68	26.02	26.67
12.5	21.56	22.14	26.43	27.24
13	21.91	22.58	26.84	27.76
13.5	22.27	22.98	27.25	28.20
14	22.62	23.34	27.63	28.57
14.5	22.96	23.66	27.98	28.87
15	23.29	23.94	28.30	29.11
15.5	23.60	24.17	28.60	29.29
16	23.90	24.37	28.89	29.43
16.5	24.19	24.54	29.14	29.56
17	24.46	24.70	29.41	29.69
17.5	24.73	24.85	29.70	29.84
18	25	25	30	30

Table 2.1. Age and gender specific BMI cut offs to define overweight and obesity (IOTF) (Cole et al, 2000)

2.4 Current trends in childhood obesity

Obesity is a significant medical, social and economic problem. Its prevalence has escalated over the last two decades, reaching pandemic levels in the developed world and is also increasing across the developing world (Wang & Lobstein, 2006). Worldwide over 22 million children under the age of 5 years were overweight in 2002 (Kumanyika et al, 2002). Figure 2.3 demonstrates how rates of childhood obesity have been increasing worldwide, and not just in the West. Figure 2.4 shows how the prevalence of childhood overweight (not obesity) has been increasing worldwide since the 1970s. All of the countries (USA, Canada, Brazil, China, Australia, Spain and the UK) show an increasing trend, with overweight in the UK rising from about 8% in the mid 1980s to about 20% in 2000. The bar chart in Figure 2.5 is a summary of current prevalence of childhood overweight and obesity in different global regions. This shows how America stands out with the highest rates, but that Europe comes a close second.

Furthermore, in the UK the prevalence of obesity amongst children of all ages is also increasing. In particular, between 1995 and 2002 the prevalence of obesity in children aged 2 to 15 years increased from 10.4% to 16.6% (60% rise) in boys and 11.7% to 16.7% (43% rise) in girls (Sproston & Primatesta, 2002). Figure 2.6 shows UK rates of childhood obesity for different age groups, with increasing trends, particularly for the older age groups. In actual fact, research in the UK, has shown that nine years old is probably a key age to start seeing rises in obesity (Rudolf et al, 2005b), although, as will be shown later, the seeds may have been sown in much earlier childhood.

This epidemic has affected most ethnic groups and children of every socio-economic background, though in disproportionate ways and there is some evidence of a social class gradient in obesity in English children of school age (Kinra et al, 2000). Also data over ten years to 1999 shows that the rise in prevalence of childhood obesity was focused on the French children from lower socio-economic groups, with no change in rates of obesity over the period in children from the higher socio-economic groups (Romon et al, 2005). Furthermore, indicators of morbidity and mortality have long shown significant differences between areas of high and low income and variations in fruit and vegetable consumption between social classes (negative correlation) have also been shown to be highly significant (Billson et al, 1999; Shohaimi et al, 2004). There is also growing evidence of spatial variations in the geography of obesity in the UK, with Scottish children having a higher prevalence of obesity than English children (Chinn & Rona, 2001). On a more local level, it has also been shown that poor access to good quality food retail outlets (within so called food deserts on many highly deprived council housing estates) has led to poor diets amongst residents, especially those with young children (Whelan et al, 2002; Wrigley et al, 2003).

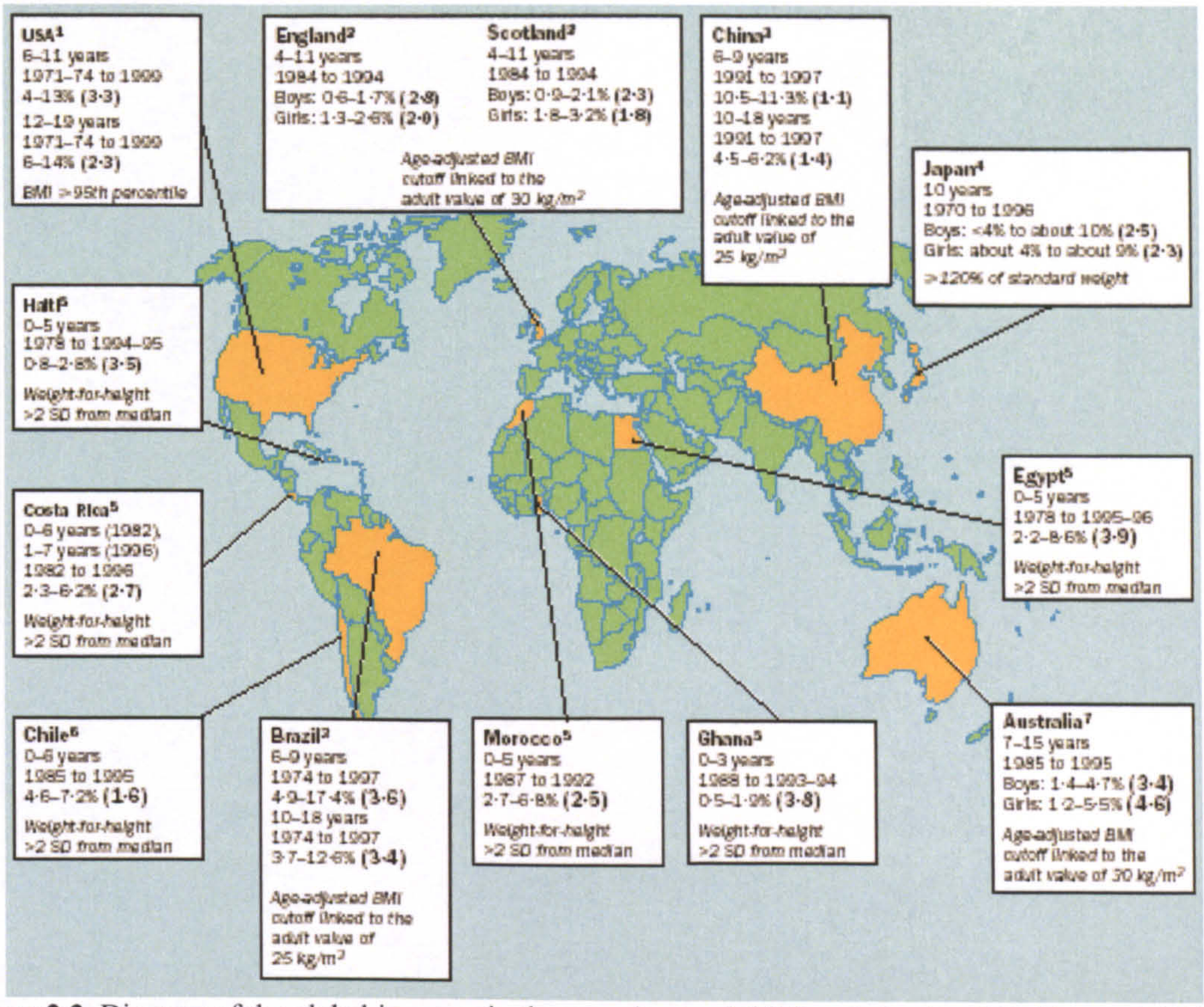


Figure 2.3. Diagram of the global increase in the prevalence of childhood obesity (Ebbeling et al, 2002)

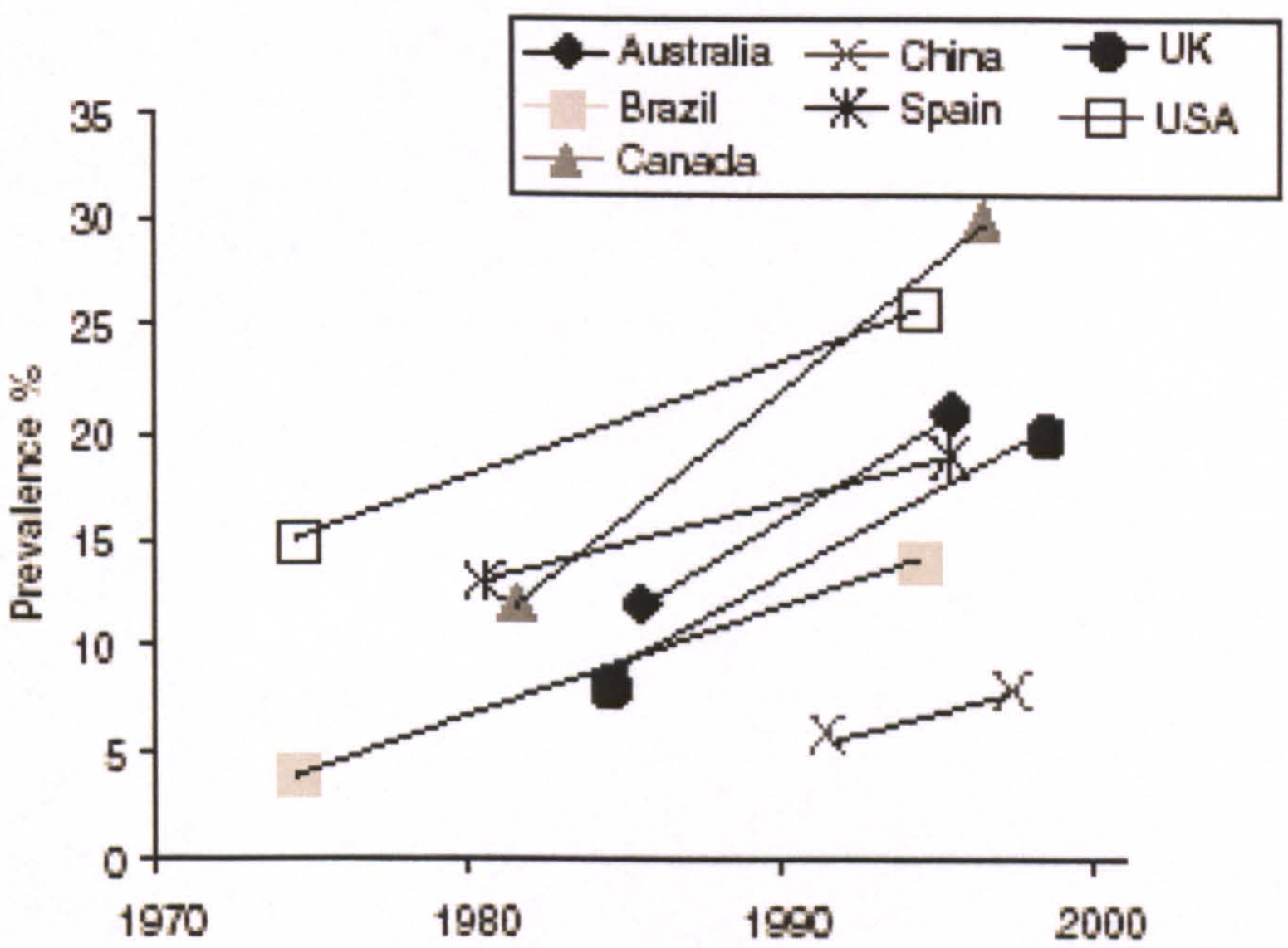


Figure 2.4. Worldwide prevalence of overweight among children (1970 – 2000). Time is along the x axis. The rate of overweight shown on the y axis (overweight defined by IOTF criteria) (Lobstein et al, 2004).

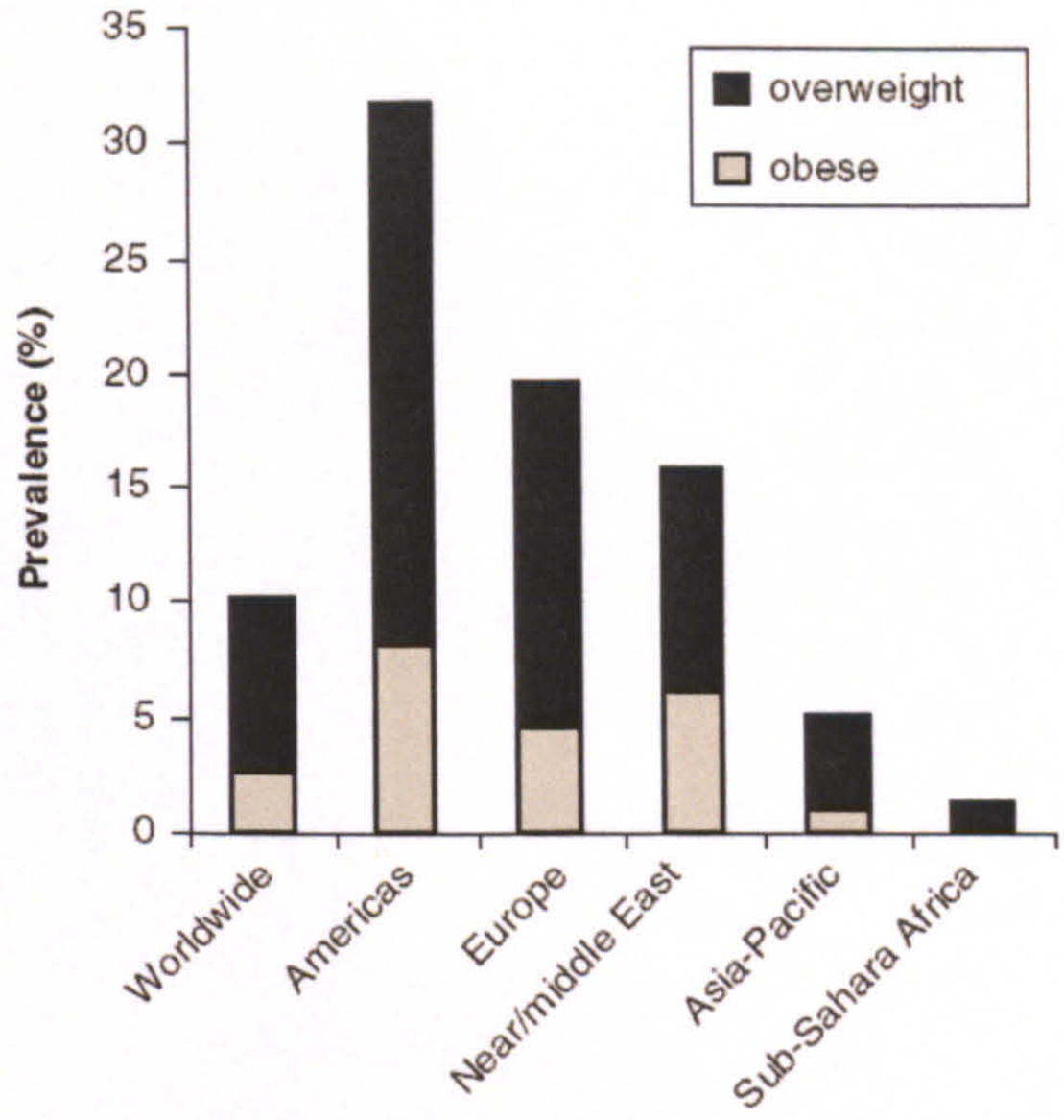


Figure 2.5. Worldwide prevalence of overweight and obesity. The global region is along the x axis and prevalence of overweight and obesity along the y axis (using IOTF definitions) among school aged children (aged 5-17 years old). Each countries data is based on a survey post 1990 (all in different years, so not entirely comparable) (Lobstein et al, 2004).

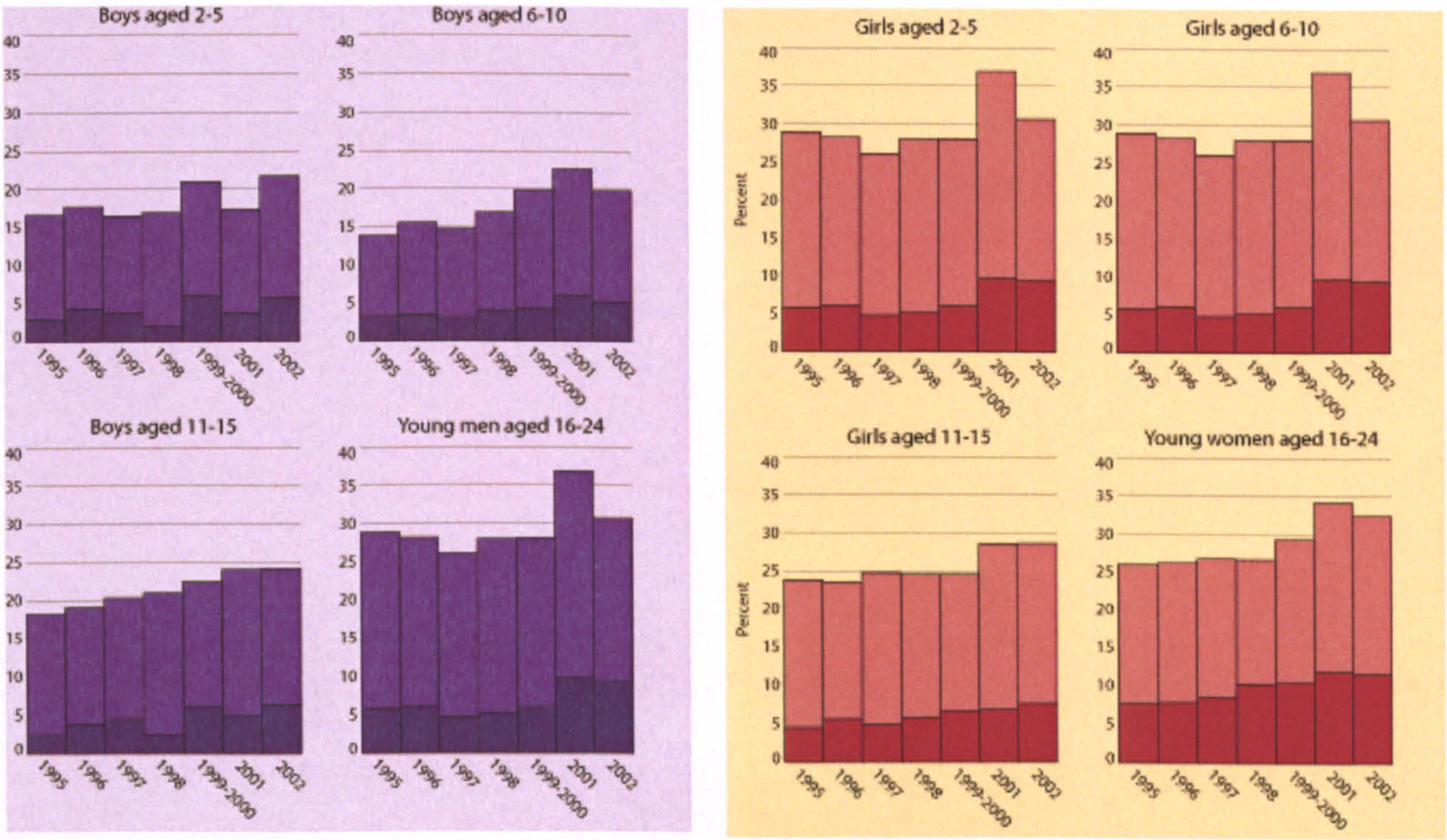


Figure 2.6. Prevalence of obesity (using IOTF definitions) among children in the UK (1995 – 2002). Boys of different age groups are on the left hand side and similarly for girls on the right hand side (Sproston & Primatesta, 2002).

2.5 Determinants of health behaviour

In order to fully understand the complex multi-factorial aetiology of obesity and to identify the role of the broader environmental influences (obesogenic factors) on energy balance, it is beneficial to use an "ecological model" to look at the multiple levels of influence on the determinants of health behaviour. Various models have been proposed, including: Ecological Systems Theory (Davison & Birch, 2001); Epidemiological Triad (Swinburn & Egger, 2002; Egger et al, 2003); Ecological Model (Egger & Swinburn, 1997; Swinburn et al, 1999). All expand the energy balance equation to look at the broader environmental factors and their role in influencing energy balance, in order to facilitate the identification of obesogenic factors and the prevention of obesity. Similarly, Flodmark et al (2004) suggest, without the use of a model, that there are six levels that should each be considered when addressing a preventative programme for childhood obesity.

Each of these models concur that the determinants of obesity sit at many different levels, and agree that successful prevention of obesity needs to work at all of these levels. However, it is how these levels are defined and summarised that varies between the models. For example, the ecology model in Figure 2.7 considers the multiple levels of influence (both within and outside the individual) on the determinants of health behaviour, seeking to address the complex web of factors that impact a person's dietary and physical activity choices. Accordingly this model subdivides the influences on obesity health behaviours into three broad categories: individual factors, social and cultural factors, and the physical environment.

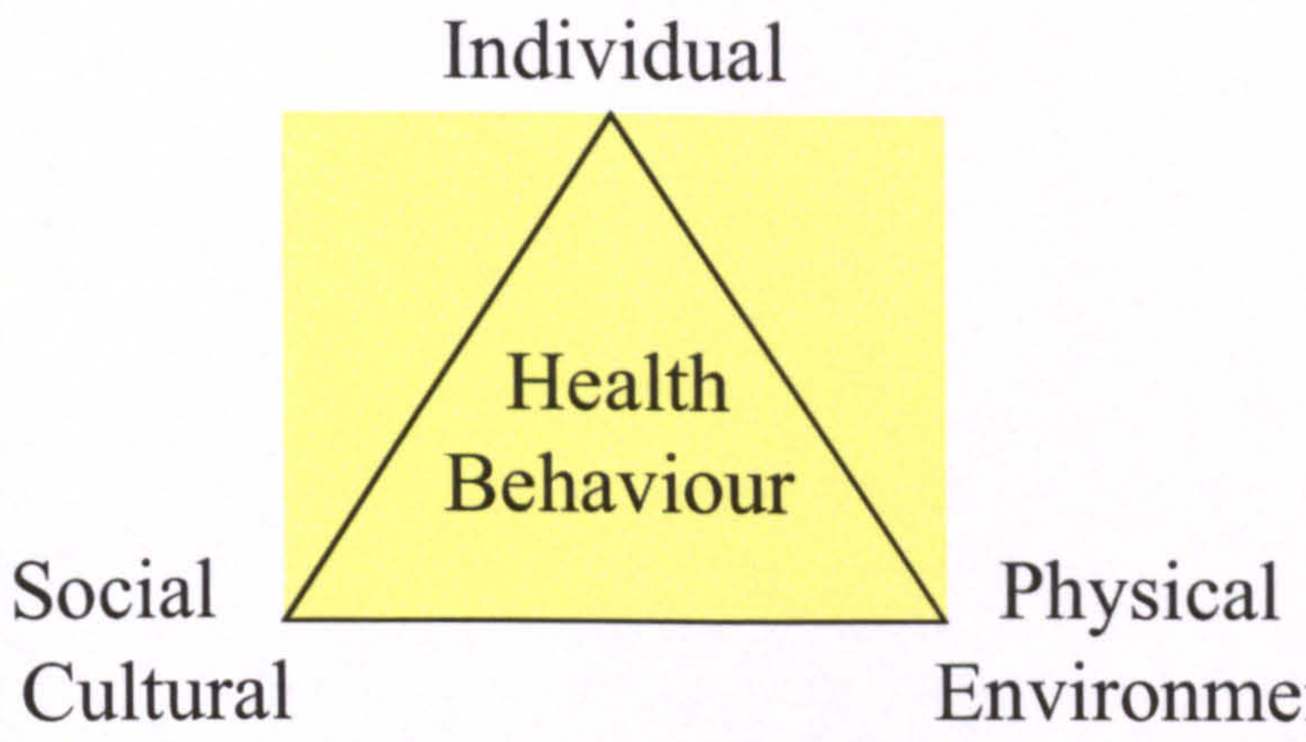


Figure 2.7. The ecological approach states that health behaviour is influenced by more than just individual factors (such as attitudes, beliefs and knowledge). Factors outside the individual (i.e. social and cultural and the physical environment) also impact the choices people make in relation to health behaviour, as does the interaction between the individual and these external factors.

(1) Individual factors (such as the knowledge, attitudes and beliefs of the individual). It is recognised that individual factors in determining health behaviour are important. However the ecology approach is also concerned with factors outside the individual and with the interaction

between the individual and these external factors. That is, obesity results from positive energy balance (when people burn too little energy in energy expenditure relative to the intake of energy from eating). Eating and physical activity patterns are individual characteristics (individuals decide which food to consume and how much exercise to take). However this is only part of the story. The amount and type of food consumed and the amount of exercise undertaken are also determined by factors outside the individual. These external obesogenic factors are encompassed in the next two categories.

(2) Social and Cultural factors including, for example, the impact of the influence and behaviours of friends, family, peers, neighbours, and all rules (whether formal, e.g. laws, regulations, policies, or informal, e.g. institutional rules, including in the home) on the eating and physical activity behaviours of the individual. On a micro level, this encompasses the 'culture' or 'ethos' of a school, home, workplace or neighbourhood. On a macro level, this includes the media's impact on influencing the socio-cultural aspects of food and physical activity, particularly through advertising and marketing.

(3) The Physical Environment category looks at what is available. It includes, for example, food and physical activity choices that may be impacted by climate, geography and crime rates (both perceived and actual), as well as nutrition and exercise expertise, available technology, and food labelling. This category would also encompass financial factors, including both costs and incomes for consumers, money spent on the promotion of healthy lifestyles by health departments, advertising by fast food outlets and government funding of roads, public transport and recreation activities.

This model suggests that health behaviours can be changed by impacting factors other than at the individual level. For example, if crime rates were lower, parents may allow children to play outside more frequently, increasing child physical activity levels and reducing risk of obesity. Consequently interventions to reduce obesity in children would be more effective if targeted at multiple levels of influence of the determinants of health, rather than solely focusing on the child.

It is important to note that the interaction between these categories means that different individuals may be influenced by different environmental factors, or in different ways by the same environmental factors. For example, a person with a high income level may not be influenced by fluctuations in the price of food, yet lower income individuals may be easily influenced by such fluctuations and consume less of the produce, such as fruit and vegetables, when it is higher priced. This makes interpretation very difficult and also means that the

interaction between multiple obesogenic factors needs to be considered rather than just a single obesogenic factor in isolation.

2.6 Genetics or the environment?

The regulation of energy balance and the aetiology of obesity are enormously complex, with numerous genetic, hormonal, neural, metabolic, behavioural, societal, and obesogenic influences (Comuzzie and Allison, 1998; Deckelbaum & Williams, 2001). Although several single-gene mutations have been shown to cause obesity in animal models, the situation in humans is considerably more complex (Comuzzie and Allison, 1998). The most common forms of human obesity arise from the interactions of multiple genes and obesogenic factors (Comuzzie and Allison, 1998; Deckelbaum & Williams, 2001; Liu et al, 2003).

Many studies show a strong genetic link with obesity. That is, an individual is more likely to be obese if he/she has obese relatives. For example adoptees' body mass index (BMI) were more similar to biological parents' BMI (Sorensen et al, 1998) and weight gain in twins showed a genetic factor (Bouchard & Tremblay, 1997). However, the environment has to be, at least partially, responsible for the rapid rise in obesity, as evidenced by the following: (1) The fact that the rise in childhood obesity has been so rapid suggests that environmental factors rather than single gene defects are the primary cause (if the cause was genetic then the increasing prevalence would take longer as it takes time for gene defects to pass between generations). (2) Migrant studies suggest a strong influence of environmental factors on obesity rates, as migrants have higher BMIs than their counterparts still living in the country of origin (McDermott et al, 1998; Popkin & Udry, 1998). Also immigrants' offspring have higher rates of obesity than their parents (Popkin & Udry, 1998) and second generation children have higher obesity rates than first generation children (Popkin, 1998). (3) As developing countries switch to more Western diets (Drewnowski & Popkin, 1997) and reduced physical activity levels (WHO, 2006), the prevalence of obesity in children is rising (Bhave et al, 2004; Wang et al, 2002). Developing countries also show over and under weight children in the same family (Florencio et al, 2001). Both of these instances imply that it is not genetic factors but environmental factors that are influencing levels of obesity.

The difficult question of how much of the variation is explained by each of the genes and the environment has been addressed by Allison et al (2001). This review suggests that about 10% of the population may become overweight even in a leptogenic environment (i.e. an environment that increases the likelihood of being "normal" weight, as opposed to obese) and another 10% would remain slim even in an obesogenic environment. These people have strong genetic predispositions to be obese or slim. The remaining 80% of us possess 'thrifty' genes, which evolved to help us deal with periods of famine and feast and which have not (yet) adapted

to the modern obesogenic world, where energy dense foods are readily available and energy expenditure can be minimal. So, for the majority of us, although we possess the genes to become obese (genetics is the loaded gun), it is the obesogenic environment that is the primary factor causing obesity (the environment pulls the trigger).

2.7 Aetiology of primary childhood obesity

The answers to the questions regarding the causes of the increased prevalence of childhood obesity remain subject to debate, with different authors holding different opinions and studies producing conflicting results. It may be that there isn't a simple or exact answer, particularly as obesity is a condition that develops slowly (so the time lag could mask the causes) plus its cause is likely to be multi-factorial with many confounding factors. Nevertheless, the debate around the reasons for the increasing prevalence of childhood obesity include the following possible explanations: this review will look at how physical activity or inactivity effects obesity, dietary risk factors for obesity, and then at obesogenic environments. This is summarised diagrammatically in Appendix A.

2.7.1 Physical (in)activity levels

There is some evidence, particularly in the US and the UK, of a reduction in habitual energy expenditure in children: reduced walking, cycling, and increased use of cars (DiGuiseppi et al, 1997); increased use of automated transport and technology in the home, with more passive leisure pursuits (WHO, 2004b). A systematic review of studies looking at the relationship between physical activity in children and obesity found roughly half had found no effect and the balance had a negative effect (i.e. increased physical activity levels were protective) (Parsons et al, 1999). A stronger link has been found between lifestyles characterised by lack of physical activity and excessive inactivity (particularly television viewing) with increased risk of obesity (Lowry et al, 2002; Matheson et al, 2004).

It should be noted that physical activity can be measured in a number of ways. Either energy output can be directly measured using calorimetric methods or indicators of energy expenditure (such as the incidence or prevalence of specific physical activities) can be used. Alternatively physical inactivity can be measured as an indicator of low energy expenditure. TV viewing and/or media time (e.g. surfing the web, playing video games, etc) are often used as a proxy for all sedentary leisure activities and so for physical inactivity.

Accordingly many cross sectional and prospective studies have looked at the association between TV viewing and childhood obesity. Some only found a weak association (Robinson et

al, 1993; Maffeis et al, 1998), but most found a positive association (after adjusting for potential confounders, such as maternal overweight, previous overweight, family structure, ethnicity, and socio-economic status) in children all over the world - USA (Dietz & Gortmaker, 1985), Mexico (Hernandez et al, 1999), Native Canadian groups (Hanley et al, 2000), Australia (Wake et al, 2003) and the UK (Reilly et al, 2005).

A prospective study by Gortmaker et al (1996) showed a strong positive dose-response relationship between time watching TV and prevalence of overweight (as measured at the end of the 4 year study). This relationship was found after adjusting for potential confounders, including baseline maternal overweight, previous overweight, family structure, ethnicity, socio-economic status and maternal and child aptitude test scores.

The effect of TV viewing on obesity may be mediated through one or more of the following factors: a reduction in physical activity levels (Wake et al, 2003); an increase in energy intake whilst viewing (particularly snacking on energy dense foods and poor portion control) (Wake et al, 2003; Phillips et al, 2004); a reduction in resting metabolic rate (Klesges et al, 1993;Reilly & McDowell, 2003; Matheson et al, 2004); inappropriate food choices due to TV advertising for foods high in added sugars or fat (Lewis & Hill, 1998; Borzekowski & Robinson, 2001); TV programmes or advertisements may confuse/contradict the message about a healthy lifestyle (Dietz, 2001a).

If increased television viewing does lead to obesity, then factors that increase TV viewing time are important (as this is where interventions should be targeted). A study in America (Wiecha et al, 2001) found the following: a TV in the child's bedroom (raise viewing time by average of 38 minutes per day); additional TV sets in the household (7 minutes per day more per additional set); lack of family dinners (33 minutes per day more viewing); no parental limits on amount of TV watched (29 minutes per day more viewing).

2.7.2 Diet

Increased Energy Intake

It would seem logical that the rise in obesity prevalence might be partly due to increases in energy intake, but paradoxically, in the US at least, while the prevalence of obesity in adolescents has doubled (WHO, 2004a), energy intakes (in adolescents) have apparently decreased (Cavadini et al, 2000). There are, however, concerns about the accuracy of measures relying on reported food intake. Food disappearance data[1] suggest that energy intakes have actually increased while reported food intakes show a decrease (Harnack et al, 2000). Also energy balance is the important factor, so the rise in obesity may be due to energy expenditure decreasing by more than the fall in energy intake.

Eating Patterns

Changes in dietary patterns and eating habits are likely to be factors related to the increased prevalence of childhood obesity.

Snacking is gaining prominence as a potential risk factor for obesity (Takahashi et al, 1999; Berteus Forslund et al, 2005; Jebb, 2005; Sturm, 2005), as is skipping meals. Whilst babies and young children characteristically eat frequently, as children get older this is traditionally (in Western society) replaced by "three square meals a day". However, eating occasions are increasingly becoming less well defined and a "grazing" or snacking culture is permeating our society with "meals" at more frequent or irregular intervals (Jahns et al, 2001) and meals being skipped.

The impact of snacking may be attributed to the types and amounts of foods eaten as well as the frequency of eating. Snacking is often associated with more energy dense foods (and drink) or more total food ingested, particularly outside the home where the types of foods commonly consumed as snacks are often high in fat or high in carbohydrates (sugar and/or starch) (Jebb, 2005). It has been shown body weight is not affected by the frequency of eating - in a lab under isoenergetic conditions. However real life is not isoenergetic. Marmonier et al (2000) demonstrated that snacks (using a 1 MJ afternoon snack) delay the next meal slightly but that the "snacking individual" consumed more total energy over the course of the day. This suggests

---

[1] Food disappearance is equivalent to food available for consumption. It is calculated by adding total food production (plus imports, minus exports) and net losses from processing at the mill level and food fed to animals. These data are a reasonable approximation in all countries of the trends in food consumption at the national level. However, the data do not reflect actual consumption because additional losses in the food chain linking the producers and mills to the consumers are not considered.

that snacking contributes to positive energy balance, over the short term at least. Longer duration studies, which may be more predictive of long-term behaviour, show inconsistent results. Johnstone et al (2000) showed no difference in energy intake between snackers and non-snackers over nine days (using a 4-way cross-over study design with either high protein, high carbohydrate or high fat (isoenergetic) snacks or no snacks – under lab conditions), whereas Blair (1991) showed higher weight loss in subjects who stopped snacking (in "real life" conditions). An observational study of 3 year old children in Japan showed irregular or frequent snacking (no definition provided – subject to individuals' interpretation) was correlated with increased risk of obesity (Takahashi et al, 1999), but a longitudinal study by Phillips et al (2004) with adolescent girls found no relationship between obesity and consumption of energy dense snacks.

However, snacking can be difficult to measure as it is often self reported, which can be highly inaccurate. For example, Barkeling et al (2001) validated self-reported food intake with saliva tests, which showed significant differences in levels of sugary foods consumed between the obese and non-obese groups, yet the food diaries showed no significant differences.

Children who skip breakfast may have a higher risk of subsequent obesity (Wolfe et al, 1994; Siega-Riz et al, 1998). The mechanism is unclear, but it may be due to breakfast consumption being a marker of general good healthy behaviour or being related to decreased fat intake and decreased snacking during the day. Alternatively it may be due to an uneven distribution of energy intake over the course of the day, for example those who do not consume breakfast tend to eat a large amount of food in the evening, and this imbalance could lead to a higher risk of obesity (Thompson et al, 2006).

Also meal times as a family are becoming increasingly uncommon. This has the effect of fewer social controls on eating and opportunities to observe good role models, which can lead to unhealthy eating habits.

Portion sizes of foods/meals are also gaining prominence as a potential risk factor for obesity (Ebbeling et al, 2002). Research has shown that very young children have innate control of appetite and energy balance is achieved, but as children age social and environmental factors take precedence over this biological mechanism (Rolls et al, 2000; McConahy et al, 2002). In light of this and of the increases in standard portion sizes seen both inside and outside the home in recent years (Young & Nestle, 2002), more research is needed to look at the impact of portion size over a long duration (rather than just one meal) and also the factors that influence this and cause the overriding of our natural biological appetite control mechanism.

Psychological factors also play a key role in the development of childhood obesity. Increased depression and boredom in this age group can lead to comfort eating and binge eating, which are associated with increased risk of obesity (Neumark-Sztainer et al, 2007).

Diet composition

Dietary composition may be an important risk factor for obesity. The amount of fat and type of fat may be important in part due to the energy provision of fat. Cross sectional surveys of diet indicate that on average children's intake of fat is close to recommended levels, but that there are big between-children variations in intake levels (Gregory & Lowe, 2000) and they also show that higher fat intakes (as a percentage of energy intake) are associated with higher weight (Tucker et al, 1997; Guillaume et al, 1998; McGloin et al, 2002). Energy density may also be important. A UK-based cross-sectional survey showed high energy dense diets in young children tend to be higher in fat and lower in sugar content than lower energy dense diets (Gibson, 2000), although other longitudinal studies have less clear results (Robertson et al, 1999; Maffeis et al, 1998; Magarey et al, 2001). "Healthy" food intake and fruit and vegetable intake are negatively associated with obesity (WHO, 2003), although potential confounding issues such as deprivation should be taken into account. Refined carbohydrate foods, and particularly those with a high glycaemic index such as sugar-sweetened soft drinks, biscuits and cakes, may be associated with obesity (Livesey, 2005; Nielsen et al, 2005). High glycaemic index foods increase postprandial blood glucose concentration and so could play a part in appetite regulation.

Consumption of unhealthy foods

Another dietary risk factor for obesity, unsurprisingly, is a high consumption of unhealthy foods, and in particular "fast foods" (i.e. foods from fast food outlets) and soft drinks.

The popularity of fast foods has increased over recent years and consumption by children has risen 300% over the last twenty years (St-Onge et al, 2003). It has been shown that when children eat fast food, then that day their energy and fat intake is likely to be higher and fruit and vegetable intake lower than normal (Bowman et al, 2004). Also children who eat fast food frequently "consume more total energy, more energy per gram of food, more total fat, more total carbohydrate, more added sugars, less fibre, less milk, and fewer fruit and vegetables than children who eat fast food infrequently" (Speiser et al, 2005). Accordingly, it is not the consumption of fast food, per se, that leads to obesity (as both lean and obese children consume fast food), but the fact that overweight consumers of fast food are less likely to adjust their daily

energy intake to take account of an energy dense fast food meal than their lean counterparts (Ebbeling et al, 2004).

There has also been a massive increase in the amount of soft drinks consumed. Soft drink intake now accounts for largest single source of non-milk extrinsic sugar intake in young people (Gregory & Lowe, 2000). These fluids tend to replace milk and so calcium intake for adolescents, which is a concern, not least because there is an inverse relationship between calcium intake and adiposity (Heaney et al, 2002). Sugar sweetened soft drinks can lead to increased energy intake as their energy value is often not acknowledged, and the energy intake from solid food not correspondingly reduced. In a study where children were given either a sugar sweetened or aspartame sweetened soft drink with a standardized meal, both groups consumed similar amount of foods, resulting in the sugar sweetened group consuming more energy in total (Wilson, 2000). Furthermore a prospective study by Ludwig et al (2001) has shown consumption of soft drinks is positively associated with obesity in children (over 19 months). Although this observational study cannot prove causality, the regression models did take other dietary and lifestyle differences into account to minimize the impact of confounding on the results, but obviously other unaddressed factors could be at work. Furthermore a longitudinal study over ten years (Phillips et al, 2004) also found an association between soda consumption and BMI (Phillips et al, 2004). A recent cross sectional analysis (O'Connor et al, 2006) appears to contradict these findings, with no association found between total amount of beverage consumed and weight status of the child and whilst higher beverage consumption was associated with total energy intake (positively) it was not related to BMI. However this study considered very young children (2-5 years), which may be too young to see the long-term impact of higher energy intake due to beverages, plus it is limited by its snap shot cross sectional nature.

2.7.3 Obesogenic environments

Obesogenic environments are one of the explanations for the increasing prevalence in obesity. An obesogenic environment considers the combination of factors that influence health behaviour and is one that makes obesity more likely to occur. It is defined as "the sum of influences that the surroundings, opportunities or conditions of life have on promoting obesity in individuals or populations" (Swinburn et al, 1999). Six different obesogenic environments are now considered.

The foetal environment

Birth weight is positively associated with childhood obesity, with an increased risk of obesity for both the heaviest and lightest babies (Fall et al, 1995; Curhan et al, 1996a&b; Parsons et al, 1999), independent of socio-economic status (Barker et al, 1997; Stettler et al, 2002) and gestational age (Sorensen et al, 1997), but may be confounded by maternal weight (Parsons et al, 2001). However other studies suggest that subsequent obesity may actually be independent of foetal growth (birth weight), instead suggesting that unfavourable conditions in the foetal environment are fundamental to the increased risk of subsequent obesity:

Maternal diabetes during pregnancy results in offspring with an increased risk of developing childhood obesity (Whitaker & Dietz, 1998). These infants are likely to be born overweight, revert to normal weight by 12 months, then become overweight / obese as older children (Whitaker & Dietz, 1998; Dabelea et al, 2000). This higher risk of subsequent obesity is independent of birth weight and maternal weight, suggesting that the effect is due to the unfavourable foetal environment.

Maternal smoking during pregnancy is also associated with an increased risk of childhood obesity (Power & Jefferis, 2002). There is a dose-dependent relationship between numbers of cigarettes smoked during pregnancy and extent of childhood overweight/obesity, after accounting for potential confounders (social class, maternal weight and birth weight) (von Kries et al, 1999), which may be due to programming of appetite regulation (Grove et al, 2001; von Kries et al, 2002). There was no association with smoking after pregnancy, suggesting that it is the intrauterine exposure that was fundamental to the increased risk of obesity.

Maternal fatness may promote childhood obesity (Curhan et al, 1996a; Parsons et al, 2001). Furthermore, studies of famine during pregnancy (Ravelli et al, 1999; Biro et al, 2001) again suggest that it is the adverse foetal environment rather than any effect on foetal growth that may be responsible for this relationship with obesity.

The infant environment

There is strong evidence that the environment in early life can determine the risk of subsequent obesity. Contrary to the previous section, Kinra et al (2005) suggest that the critical period when obesity risk is acquired is postnatally, rather than prenatally.

Post natal weight gain (of the infant) is thought to be important in determining risk of subsequent obesity, although the exact pattern of weight gain that is higher risk is controversial.

Rapid weight gain during the first four months increases risk of subsequent obesity (Stettler et al, 2002), as does rapid weight gain during the first twelve months (Reilly et al, 2005), and also children in the highest age standardized weight quarter at age 8 and 18 months (Reilly et al, 2005). Conversely it is suggested that it is the mixture of foetal and infant growth that is important. There is an increased risk of obesity for low birth weight babies who show catch up growth or rapid childhood growth (Ong et al, 2000; Parsons et al, 2001, Reilly et al, 2005).

Further studies suggest it may be the age of adiposity rebound that is crucial. The evidence is strong that the earlier this occurs the higher the risk of subsequent obesity in the child (Parsons et al, 1999, Reilly et al, 2005). However, the mechanism for this relationship is unclear and it is undecided whether the association between early adiposity rebound and subsequent obesity is caused by a biological mechanism or whether it simply reflects a child's predisposition to gain weight easily (as a result of existing genetic or environmental circumstances). It does not appear to be due to high early protein intake (Dorosty et al, 2000).

A systematic review by Baird et al (2005) concluded that the highest risk of subsequent obesity was for infants both at the highest end of the distribution for weight or BMI and those who grow rapidly during infancy. The mechanism for greater fatness earlier in childhood leading to increased risk of subsequent obesity is unclear. It could be because early excessive fatness predicts earlier maturation (at least after 3–4 years of age) (Parsons et al, 1999) and early maturation is associated with increased risk obesity (Power et al, 1997). However adolescents who mature later have higher protein and energy intake as well as higher activity levels, which might be the factors that prevent the obesity rather than the timing of maturation itself (Post & Kemper, 1993).

There is evidence for and against the protective effects of breast feeding. It has shown a dose dependent (better protection with longer duration of breast feeding) reduction in the risk of subsequent childhood obesity (von Kries et al, 1999; Gillman et al, 2001; Armstrong et al, 2002; Dietz, 2001b; Bergmann et al, 2003), although more recent studies have shown no or limited protective effect (Li et al, 2003; Victoria et al, 2003). Furthermore the designs of the studies with protective effects have been called into question (Clifford, 2003). That said a systematic review found that breast feeding had a (small) protective effect against subsequent childhood obesity (Arenz et al, 2004).

The apparent protective effect may be due to confounding variables such as maternal diabetes, maternal BMI, maternal smoking during pregnancy, low birth weight, familial dietary patterns or social class (von Kries et al, 1999; Wadsworth et al, 1999; Hediger et al, 2001; Poulton & Williams, 2001). Alternatively the conflicting results may be due to an interaction between

breast feeding and potential confounding factors. For example, Reilly et al (2005) found that breast feeding amongst non smoking (during pregnancy) women was significantly associated with reduced risk of obesity in the child at age 7 years. This effect was not evident in women who smoked during pregnancy. Before taking this factor into account there was not a significant relationship between breast feeding and obesity.

The mechanism for the proposed protective effect of breast feeding may be due to the timing of weaning, as solid foods increase the energy density of the diet and so could lead to excess energy intake and consequent weight gain. It might also be a factor of the amount of protein in the diet, with bottle feeding and early weaning increasing protein intake (breast milk provides a relatively high amount of energy from fat), which may reduce the age of adiposity rebound and increase the risk of subsequent obesity (Agostoni et al, 2005). Feeding style may also be important to the infant's risk of obesity. A "vigorous" feeding style (Agras et al, 1990), restrictive patterns causing upset to the baby (Wells et al, 1997) and a lack of control over the child's intake (Wardle et al, 2002) have all been associated with subsequent obesity.

Sleep duration (as an infant) has been shown to have a negative independent association with the risk of childhood obesity (Sekine et al, 2002; Agras et al, 2004; Reilly et al, 2005). There are several different possible mechanisms for this effect. It may be due to growth hormone secretion being altered by the duration of sleep or because sleep reduces the child's exposure to obesogenic factors, such as evening food intake or it could be marker for another variable, such as levels of physical activity (more active, more sleep required).

The family environment

It has been shown that family structure, including the family size (Wolfe et al, 1994; Padez et al, 2005), birth order of the child (Wang et al, 2006) as well as whether it is a single or joint parent family (Wolfe et al, 1994) may have an effect on childhood obesity. However relatively few studies have been undertaken and the results are inconsistent (Parsons et al, 1999; Lobstein et al, 2004).

Parent-child interactions, the quality of the home environment and the level of care provided within a family might also be affecting the behaviours related to the risk of obesity. These factors may have more of an impact on the risk of obesity than family structure or deprivation. For example, children with low cognitive stimulation are at increased risk of subsequent obesity (Strauss and Knight, 1999), as are children who suffer parental neglect (Lissau & Sorensen, 1994).

Parenting styles may influence the food and exercise choices of a child. Each member of the family acts as a role model for the child, their behaviour reinforcing and supporting the development of diet and activity behaviours (Davison & Birch, 2002). The family members all share the same environment, which may encourage overeating or a sedentary lifestyle (Lake et al, 1997; Parsons et al, 1999; Hood et al, 2000; Wardle et al, 2001). Dietary and activity behaviours have been shown to "run" in families (Davison & Birch, 2001), primarily due to shared environmental factors rather than genetics (Franks et al, 2005), and parental diet and activity patterns can predict risk of obesity (Davison & Birch, 2002).

Parental BMI (particularly maternal) has a strong positive association with childhood obesity (Maffeis et al, 1998, Strauss & Knight, 1999; Danielzik et al, 2004). This predictor is much stronger with young children (Whitaker et al, 1997), and if both parents are obese (Lake et al, 1997; Wang et al, 2000). This latter increase is systematic - with two lean parents having the leanest children, two obese parents having the fattest children and children of one lean and one obese parent falling in between (Garn et al, 1976, Reilly et al, 2005). This relationship is largely due to lifestyle factors and parents' diet and activity patterns can be used to identify obesogenic or non-obesogenic family clusters, with children in an obesogenic family cluster have a higher risk of obesity (Davison & Birch, 2002).

It is also worth noting that parents of overweight children tend not to recognise that their child has a weight problem (Etelson et al, 2003). However this was a small study with a sample of only 83 parents. Plus the recognition scale, used to determine the parents' perception of how overweight (or otherwise) their child was, tends to produce a normal distribution, whereas the actual BMI percentiles of the children in this sample do not appear to be normally distributed. Accordingly we might expect to see greater differential between the perceived and actual child weights at the heavier end of the scale in this study. Nevertheless a subsequent, larger, longitudinal study (Jeffery et al, 2005) using a five point scale questionnaire to determine parental perception of overweight also found that most of the overweight children (and one third/half of obese girls/boys respectively) were judged by their parents to be of normal weight. These authors suggest that possible reasons for parental low recognition of a child's weight problem may be due to simple denial, an unwillingness to admit that there is a problem or even desensitisation to overweight because this state has become normal.

Ethnicity could also be important. In the West, non-white children are more likely to be obese than white children, however this is largely to do with socio-economic differences, such as parental education and family income (Strauss & Knight, 1999, Strauss & Pollack, 2001; Booth et al, 2001; Ogden et al, 2002; Whincup et al, 2002), although increased prevalence of overweight and obesity after adjustment for socio-economic group has been reported among

Afro-Caribbean and Southern Asian ethnic groups (Saxena et al, 2004) and black girls (Wardle et al, 2006) in the UK. Further, the fact that obesity related diseases (such as type 2 diabetes or high blood pressure) are more common in people from the Indian subcontinent and that the risk of obesity related complications commences at lower BMI for these populations has implications for childhood obesity in these populations – and more research is required into this (Lobstein et al, 2004).

The school environment

The schools' policy (and/or national guidelines) to promote healthy eating might affect obesity levels. That is, the choice of foods available during the school day and the types of foods permitted for classroom events may also impact obesity rates. The availability of vending machines in schools is associated with an obesogenic environment, although not all the evidence supports this view (New and Livingstone, 2003). Children who attend breakfast clubs consume more fat and saturated fat than children who don't attend (Belderson et al, 2003). Children who bring a packed lunch to school consume a less healthy meal than those eating school dinners (Whincup et al, 2005). Externally available foods (i.e. local shops and children being allowed off school premises) may also impact food choices.

School food policies that reduce availability of high fat and high sugar foods are connected with reduced buying of these items (Neumark-Sztainer et al, 2005). However a recent study by Gould et al (2006) of school meals in the UK found that two thirds of schools did not meet the government nutritional guidelines and deprivation was associated with the worst food provision and most unhealthy food choices. This suggests that nutritional standards in isolation do not facilitate healthy eating in schools. Enforcement of the guidelines as well as a pricing policy to encourage healthier food choice (or restrict unhealthy choices) is required to improve the nutrient intake of school children.

Also nutritional and physical education might help to reduce risk of childhood obesity, by promoting healthy eating habits and body image, as well as providing opportunities for regular exercise.

A study in primary schools in Leeds used a population-based approach to implement a health promotion programme to prevent risk factors for obesity (Sahota et al, 2001a). Positive changes were seen in school meals, tuck shops, and playground activities and the implementation of the programme was a success, yet only nominal behavioural changes were seen in the children (Sahota et al, 2001b). A national programme launched in Singapore to promote healthy lifestyles, "Trim and Fit", used similar methods to the Sahota study, as well as giving special

attention to overweight children (Toh et al, 2002). Conversely in Singapore obesity levels have fallen since the commencement of the programme, although this may be due to factors outside of the programme. Similarly school interventions have been run to affect children's activities outside of school. For example, Robinson (1999) ran an intervention aimed at reducing levels of TV viewing which resulted in a positive association between children changes in levels of TV viewing and adiposity. However often children return to baseline after the intervention stops.

It has been shown that primary school children are more active at the weekend than on school days. So although schools are well placed to help tackle childhood obesity, school attendance actually limits levels of physical activity (Metcalf et al, 2002). That said, the level of timetabled physical activity at school does not affect the overall daily amount of activity undertaken by the child, as they compensate out of school (Mallam et al, 2003). Furthermore, although children who walk to primary school expend more energy on that journey than children who are driven, there is no difference between the two groups in total weekly physical activity levels. Again children are compensating elsewhere (Metcalf et al, 2004).

Low achievers at school are more likely to become obese (Guillaume et al, 2002), although it is not clear whether the poor performance leads to obesity or vice versa (Mo-Suwan et al, 1999).

The neighbourhood environment

There are many different aspects of the neighbourhood that may impact levels of obesity in children. For example:

(1) The availability of public transport affects many diet and exercise choices people make, for example with where to do the shopping, what to do with the children, etc. This impact is obviously larger on families without a car.

(2) Food deserts are areas where there is low (or no) access to affordable, healthy food, particularly if the residents don't have access to a car or good public transport links. This may impact on the dietary choices of residents.

(3) It may be that proximity to or access to parks and green spaces has an effect on obesity in children by impacting their physical activity levels (e.g. playing on swings) or diet (e.g. consuming ice creams and sugary drinks), although the little research that has been undertaken in this area tends not to show a relationship (Timperio et al, 2005). It is likely that perceived neighbourhood safety is a more important determinant of childhood obesity, but again the evidence is contrary (Burdette & Whitaker, 2005; Lumeng et al, 2006).

(4) Crime, both perceived and actual, can affect a parent's decision whether to let the child outside to play, as can road safety issues, such as safe road crossings, pavements and the speed of traffic (Timperiol et al, 2005).

Deprivation is commonly associated with obesity, although the relationship is not straightforward, depending the timing of the outcome measure of obesity (that is, whether it is in childhood or adulthood). Also different authors use different measures of deprivation, ranging from a simplistic single indicator of socio-economic status (SES) as a proxy for deprivation to a more sophisticated indicator of deprivation by ranking several different factors.

A thorough review in 1999 (Parsons et al, 1999) found a relationship between low socio-economic status (SES) in childhood and subsequent adulthood obesity, which concurs with subsequent work by Hardy et al (2000) and Okasha et al (2003), both using father's occupation as the indicator of childhood SES. This relationship was also shown more recently and using a more sophisticated indicator of deprivation (a ranking of three different factors - education level, occupation of head of household and current employment status) (Monden et al, 2006). This "SES of origin to subsequent adult obesity" relationship may be due to (1) confounding by parental body size (which insufficient studies have considered – Parsons et al, 1999), and (2) SES acting as a proxy for the effect of multiple adverse childhood circumstances, which are then manifesting as adult obesity in the long term (Power & Parsons, 2000). For example, it has been shown that there is a higher density of fast food outlets in poorer areas, which may (partially) explain the phenomenon (Reidpath et al, 2002).

The 1999 review did not find any relationship between childhood SES and childhood obesity, although conversely several more recent studies have found that children with lower SES / more deprived backgrounds do have an increased risk of childhood obesity. Some used only a single indicator of SES as a proxy for deprivation. For example, household income has been shown to be a significant predictor of childhood obesity (inverse relationship) (Strauss & Knight, 1999; Stamatakis et al, 2005). Similarly using entitlement to free school meals as a proxy for income (Cecil et al, 2005). Cecil found that it was not that these deprived children weighed more than their more affluent peers, in fact the higher BMI was due to shorter height, suggesting possible nutrition related growth restriction in low income families. Also children from families with lower education levels have a higher risk of obesity (Danielzik et al, 2004; Lamerz et al, 2005; Romon et al, 2005). However, this effect could be mediated by confounding factors, such as low income and lower levels of cognitive stimulation (Strauss and Knight, 1999). Other studies have considered multiple SES factors as an index of deprivation. For example, studies using the Townsend Deprivation Score (an index score based on a combination of adult unemployment, household size, and car and home ownership) have shown that children from more deprived

areas have higher risk of obesity (despite lower birth weights) (Kinra et al, 2000; Kinra et al, 2005). However if the deprivation index is based on the electoral ward of the school (rather than the home), no relationship with childhood obesity is present (Dummer et al, 2005).

The increased prevalence of obesity in children from more deprived backgrounds could be due to a multitude of factors: dietary differences are often apparent; no safe play area for the child; lack of opportunity / funds for activities, so TV viewing is the primary leisure activity by default; food deserts (lack of accessible, affordable, healthy (low energy dense) food); constraints on calories per pound, which focuses purchases on energy dense foods.

Also whilst deprivation is commonly associated with obesity, affluence has been less critically considered, yet there may also be a link. Certainly early work in Asia found such an association (Hakeem, 2001; Noor, 2002; Subramanian & Smith, 2006) although this could reflect cultural differences that are not prevalent in Western society (that is, whether fatness or thinness is more highly regarded).

The macro environment

The macro environment relates to those influences on childhood obesity outside of our direct control – namely industry, media and government.

Industry. This aspect of the macro environment largely encompasses all levels of the food industry, from manufacture to retail outlets to eating out. However also included is price and availability of goods that reduce our energy expenditure. It encompasses many different issues, such as hidden fats and/or sugars in prepared foods and greater availability of energy dense foods, increased use of restaurants and fast food outlets, larger portions of food offering better "value" for money, poor labelling of foods, subsidised "bad" foods as loss leaders and expensive "good" foods, more frequent and widespread food purchasing opportunities, and cheap and easy access to labour saving devices / cars.

It has already been discussed that a high dietary fat intake is associated with obesity (Tucker et al, 1997; Guillaume et al, 1998; McGloin et al, 2002) and that high fruit and vegetable consumption is negatively correlated (WHO, 2003). Additionally the rise in soft drink consumption has already been highlighted (Gregory & Lowe, 2000). Furthermore, sugar consumption in general (including sugar, corn sweeteners, honey and other edible syrups, excluding non-caloric sweeteners) has also increased substantially over the last twenty years, largely due to increased high-fructose corn syrup use in beverages, bakery products and processed and prepared foods (Coulston & Johnson, 2002). On top of this, advances in

technology have increased the availability of processed and prepared foods (Cawley, 2006). Accordingly, it follows that consuming a diet composed of a large quantity of processed and prepared foods with high "hidden" fat and sugar content (consumption of which has increased in recent years) may lead, perhaps unwittingly, to increased energy intake and so to obesity. Similarly a diet high in energy dense fast foods will also lead to increased fat consumption and higher energy intake, which may increase risk of obesity as well (Prentice & Jebb, 2003; Bowman et al, 2004; Ebbeling et al, 2004). Accordingly the fact that both eating out generally and fast food consumption have increased in recent years (St-Onge et al, 2003; Tillotson, 2004) should be of concern. Spending on eating in the home is now less than eating out spending (National Statistics, 2006). Furthermore high fast food outlet density in an area is negatively associated with SES, which in turn is considered a social determinant of obesity with, generally, a negative association between obesity and SES (Reidpath et al, 2002).

Another factor contributing to increased energy intake, and thus highlighted as a plausible risk factor for obesity, is larger portion sizes (Hill & Peters, 1998; Ebbeling et al, 2002; Rolls et al, 2004; Diliberti et al, 2004). This factor relates back to industry as the macro environment because most processed and prepared foods have seen rises in the standard portion size over the last twenty years (Young & Nestle, 2002; Young & Nestle, 2003), as have restaurant and fast food outlet portion sizes. The fact that many packaged foods contain multiple (not single) servings and that consumers do not recognize this exacerbates this problem (Pelletier et al, 2004). As well as increasing energy intake on the eating occasion (of the product or at the premises), this may also have a knock on effect of increasing the expected portion size, or that considered appropriate, at a self-serve eating occasion (Geier et al, 2006). This occurrence of "portion distortion" varies for different foods but does show significant differences (mostly increases) in self serve portion sizes over the last two decades (Schwartz & Byrd-Bredbenner, 2006).

The question of food labelling is frequently discussed as a means to facilitate healthy food choices by the consumer. In the UK the Food Standards Agency has proposed a "traffic light" food labelling scheme in this regard. Whilst several retailers have agreed to introduce it on their own products, many other retailers and manufacturers are introducing their own labelling systems, which only serves to add to consumers confusion. Consumers have been shown to change their consumption patterns depending on the information given about the fat content of the food (Roefs & Jansen, 2004), although whether this translates into long term purchasing and consumption patterns remains to be seen. The evidence of providing dietary information about restaurant meals is hampered by the fact few restaurants provide this facility, particularly at point of purchase (Wootan et al, 2006), and it has been shown that consumers largely ignore or do not correctly understand restaurant food labelling (Krukowski et al, 2006).

Poor access to affordable, healthy food is considered to be a contributory factor to poor diet and obesity. Whilst the price of food in real terms has reduced, this is largely for "unhealthy" (energy dense, high fat, high sugar) foods (Cawley, 2006) and it has been shown that these "food deserts" do exist (for example, Clarke et al, 2002; Wrigley, 2002; Whelan et al, 2002). Improving access to food can increase fruit and vegetable intake, which also suggests that limited access to healthy, affordable food does affect the diet consumed (Wrigley et al, 2002). Also Sturm & Datar (2005) found that higher fruit and vegetable prices were positively correlated with change in BMI. Yet all the evidence does not agree, as some authors have not found a positive relationship between amount of fruit and vegetables consumed and food deserts (Pearson et al, 2005; Winkler et al, 2006), although the evidence does seem to be stronger in the USA (Cummins & Macintyre, 2006). A clear way to increase healthy choices over unhealthy choices is to provide an economic incentive, for example, healthy food subsidies and unhealthy food taxes. It was shown that young people do respond to this, with price rises reducing purchases of a particular food, and substitution between healthy and unhealthy foods occurring as prices rise or fall depending on the amount of disposable income (Epstein et al, 2006).

Obesity is also promoted by industry in the macro environment by the more frequent and widespread food purchasing and consuming opportunities that currently exist. An extensive range of tasty, reasonably priced foods are accessible almost ubiquitously (Hill & Peters, 1998). On the other side of the coin, increased access to labour saving devices and use of cars has reduced levels of physical activity (WHO, 2004b), which is further impacted by less habitual energy expenditure (DiGuiseppi et al, 1997; WHO, 2004b).

Given the food industries role in encouraging, or at least facilitating, obesogenic behaviour, accordingly they also have a role in preventing obesity. They could reduce the availability of high fat, high sugar, and energy dense foods. However, realistically this is not going to happen, as it would be too directly damaging to profits. More pragmatically, food companies could make more (in quantity) healthy and, importantly, cheap products available (rather than making these products a high profit margin alternative). Clear food labelling would also help. Finally, a more indirect role could be taken, with encouraging consumers to select healthy produce and to collaborate in research to increase our understanding about food and health (Dwyer & Ouyang, 2000).

Media. There is a broad and strong impact of the media, both negative as well as positive. An example of the positive impact the media can have is the recent success of a UK TV chef in bringing the public's and government's attention to the poor diet given to children in schools, leading to changes in awareness of the issue as well as real changes. Food TV advertising,

especially that aimed at children, is a classic example of a negative impact of the media particularly as this is often for unhealthy foods (Linn, 2004; Neville et al, 2005), which can lead to unhealthy food choices (Lewis & Hill, 1998; Borzekowski & Robinson, 2001). Additionally advertisers often concentrate on building brand loyalty and "creating lifelong customers rather than generating immediate sales" (Connor, 2006), which means any resulting unhealthy food choices will be enduring. Furthermore TV programmes and advertisements may confuse or contradict the message about a healthy lifestyle (Dietz, 2001a). Advertisers refute the claim that they contribute to the obesity problem, stating that they cannot compel people into buying goods (Hoek & Gendall, 2006). However that response is illogical - if the adverts are not successful why would advertisers go to huge lengths and expense to build brands and advertise products? Excessive amounts of money are spent on advertising (especially when compared to governmental budgets for healthy food promotion). Also a review of the ecological evidences showed that there is a significant relationship between TV advertising and prevalence of overweight children (Lobstein & Dibb, 2005). All in all this suggests that in this day and age where obesity is a growing problem, TV advertising aimed at children should be limited, which probably needs to occur at a governmental level as voluntary codes are largely unsuccessful.

Government. It is all well and good saying that diet and exercise are down to individual choice. But this approach is not working as demonstrated by the rising prevalence of obesity. Furthermore, in relation to children, their cognitive ability is not sufficiently developed to enable them to take the future consequences of their actions into account when evaluating what to do. Whilst it can be argued that parents therefore have a role in deciding what foods and how much exercise their children should take, arguably there is also a role for government to help children (and their parents) to make healthier choices (Cawley, 2006). Similarly as the market is not providing sufficient, clear, information to allow consumers to make rational, healthy choices, as demonstrated by food labelling confusions, then this also fuels the debate for more heavy-handed government intervention.

If the fact that obesity is associated with increased morbidity and mortality is insufficient by itself to justify bringing obesity on to the government's agenda and for them to take action to reduce it, then maybe the economic implications tip the balance. The health care costs associated with obesity are increasing and are projected to grow rapidly (as discussed above). If prevention measures are not successful then these costs will need to be borne somehow (i.e. by the taxpayer) (Cawley, 2006). Furthermore the costs of obesity are lop-sided. On the one hand, even assuming the consumer has full information about the benefits of physical activity and a healthy diet as well as the detrimental health consequences of obesity, there will still be some people for whom an obesogenic lifestyle has the lowest "cost" (in terms of time, opportunity costs, and money) and so this is their optimal choice (Finkelstein et al, 2005). However it is not

an optional choice from a national viewpoint because, on the other hand, the taxpayer bears much of the monetary cost of obesity (Finkelstein et al, 2005). Accordingly it follows that there is a role for the government to intervene.

The fact that the food industry sells energy dense, high fat, high sugar foods and energy saving / sedentary behaviour devices, such as cars, television sets and play stations, is not good reason for government intervention in the market. They sell these foods because there is demand for them. If consumers demanded healthy products then these would be provided - the strength of the diet industry reflects this. However, the food industry is not "playing fair". They have used complicated marketing and advertising practices to increase the amount people eat, whilst at the same time (in the US at least) lobbying government bodies responsible for providing dietary advice to consumers to ensure the message to reduce their energy intake does not get across (Elliot, 2003). These aggressive sales tactics are at least partly due to overproduction of food leading to intense competition to win sales (Nestle, 2003). This sales competition takes place through new or improved products, increased portions, health claims, advertising, campaigns at special groups such as children (Nestle, 2003), all the while aiming to increase prices as well as sales volumes. Furthermore government subsidies have facilitated the increased manufacture of cheap, high fat and sugar snacks and drinks (Rigby et al, 2004). Industry (i.e. agriculture, food production and retail, restaurants, diet, pharmaceuticals) do not benefit if society were to eat less (Nestle, 2003). Accordingly they strongly lobby government to ensure little (no) action is taken to discourage overeating (Nestle, 2003; Weiss & Smith, 2004), as there would be serious economic consequences for them if obesity reduced. In view of this lack of fair play, industry cannot be relied upon to comply with any voluntary codes of practice to reduce obesity and obligatory policies need to be established. However the obesity issue is highly political. The (potential?) conflict of interest between governmental funding and influences from food companies and the government's responsibility to protect the public need to be borne in mind. This leads to a "policy paradox" whereby governments support food industry as well as making lifestyle recommendations to maximise population health (Rigby et al, 2004) and has lead to governments taking action contrary to best practice for consumer health and more akin to helping boost food companies' balance sheets (Boseley, 2004).

Litigation also has a role to play in protecting public health, particularly when government policy is non-existent or insufficient (Daynard et al, 2004). A classic example of this is with the tobacco industry. Whilst a move towards the prolific litigation culture of the USA is perhaps not desirable, nevertheless litigation can help to increase public awareness of the issue and to improve self-regulation of industry, eventually restraining those practices that are detrimental to consumers. For example, leading food companies to rework their merchandise and marketing methods because, with the obesity issue, potential lawsuits are likely to include "unfair and

deceptive trade practice" (Daynard et al, 2004). Indeed, it was the wrongdoings of the tobacco manufacturers, rather than the health risks of tobacco, that resulted in successful litigation against tobacco companies (Daynard et al, 2004). The food companies consider litigation a very real threat, as demonstrated by their attempts to try to prevent it from being allowed (Kelly & Smith, 2004; Nelson, 2004).

At the present time turning around public perception of the acceptability of overeating and sedentary behaviour leading to obesity might seem an impossible task. Imagine changing social norms about the (un)acceptability of using a car for a short journey rather than walking. However, so too was changing attitudes to smoking and drink driving, but both have been very successful (albeit not entirely eliminated). It is argued that the only way we will see a radical reduction in obesity rates is to implement radical policy changes, to regulate food production, marketing and consumption (Davey, 2004). This view was corroborated in a recent debate at the International Conference of Obesity held in Sydney where there was an overwhelming majority in favour of a more "heavy hand" of government than that which currently exists across many different countries. Regulation can transform an environment in an instant (Hayne et al, 2004). It could be used to create leptogenic environments, in the same way that we now have smoke-free environments and clean water (Davey, 2004). But is there the political will to do what is necessary to fight obesity? There are many conflicts of interest.

There are many possible government interventions that may help to prevent obesity. It is important that policy does not solely focus on changing individuals' behaviour, but that it also looks at the role of industry and media in order to make changes at these levels as well.

In relation to the impact of television and physical activity on childhood obesity there are several suggestions. Ban (or at least more heavily regulate) advertising of unhealthy foods aimed at children, especially in schools and on television (Butler, 2004; Davey, 2004; Finkelstein et al, 2004; Hayne et al, 2004; Tillotson, 2004; Weiss & Smith, 2004). Additionally any food advertising to children that is permitted could be taxed, with the proceeds being used to fund healthy lifestyle initiatives and education (Hayne et al, 2004). Proactively, public service announcements could be shown during children's programming to promote healthy eating and physical activity (Pratt et al, 2004). Changes to the physical environment may also help to prevent obesity (Hayne et al, 2004). More pavements, less parking, more "park and ride" schemes, more parks, etc, may all promote a more active lifestyle.

The school environment is an important influence on children. Accordingly it may be helpful to ban unhealthy products from school vending machines (or even a total ban) (Butler, 2004; Finkelstein et al, 2004, Hayne et al, 2004, Tillotson, 2004; Weiss & Smith, 2004). Clear,

enforced, nutritional guidelines for healthy school dinners are required (Finkelstein et al, 2004; Hayne et al, 2004; Weiss & Smith, 2004), whilst providing the schools with the tools required to prepare these meals. The implementation of healthy eating schemes may also be beneficial. For example, in the UK, there are several such schemes, including the new Healthy Start Scheme, the continuation of the National School Fruit and Vegetable Scheme, and Food in Schools as part of the Healthy Schools Initiative. Compulsory physical education and nutrition classes in schools may also make a difference to the obesity epidemic (Hayne et al, 2004). The nutrition classes should include how to read food labels, as this will facilitate healthier food choices. This obviously also necessitates clear nutrition labels to be provided by the food industry (Butler, 2004; Hayne et al, 2004), which probably needs legislation to ensure it happens in a coordinated, comprehensible manner.

Whilst the use of tariffs or import bans cannot be used to control consumption due to the implications on global trade (Rigby et al, 2004), this does not prevent the use of taxes to tackle obesity. A "fat tax" could be put on unhealthy (energy dense, high fat or high sugar) foods, which could fund, at least in part, these obesity prevention strategies (Davey, 2004; Finkelstein et al, 2004; Tillotson, 2004; Weiss & Smith, 2004). Whilst this is often dismissed as a "stealth" tax on the poor, if an economic viewpoint is taken, then it is suggested that no amount of increased education or clear nutritional information will change the dietary and activity choices some individuals make (Finkelstein et al, 2005), in which case financial incentives (or disincentives) are required. Also deprivation is strongly correlated with obesity, with an unhealthy diet being an inexpensive diet (Drewnowski, 2004), so there is an argument to implement policies that have a larger effect on low socio-economic groups. A similar tax could be levied on products that promote sedentary activity (Finkelstein et al, 2004; Pratt et al, 2004). The other side of this coin is to change the way agriculture subsidies work to reduce the retail costs of fruit and vegetables and to discourage, rather than support, the marketing of obesogenic foods (Kelley & Smith, 2004). Given the tripartite conflict of interests between consumers, industry and governments, which initiatives will be more successful? A supply side stance (such as restricting food advertising) or demand side (healthy eating education) or a combination of both (Tillotson, 2004). Also can these initiatives work given the existing influential economic and agriculture policies?

2.8 Conclusion

From both the perspective of the increased health risk to the individual and the high economic cost of treatment of obesity and related diseases, it is important that obesity is preferentially prevented from occurring, whilst nevertheless implementing treatment programmes in parallel as current rates of obesity are already high and we can not ignore these patients. However, going forward, prevention will be more effective in children as obese children tend to become obese adults and it may be that behavioural patterns that determine obesity are set in childhood. Further, as the bulk of the population are predisposed to become obese, particularly if living in an obesogenic environment, then obesity is likely to be an important public health issue for some time to come.

The present review has looked at the different levels of behaviours leading to obesity, which helps us to understand why the aetiology is so complex and that potential causal factors should not be considered in isolation as the interaction between these factors is important. Many studies have looked at simple, single or bivariate relationships with obesity, rather than considering the multiple factors that actually comprise the aetiology of childhood obesity and considering their inter-relationship and their relative importance. If we don't understand how these factors interact, or the relative strength of different obesogenic factors, we can't predict the outcome for any one individual.

This review of the aetiology of childhood obesity considered physical activity, diet and various obesogenic environments. Strong predictors of obesity were found to be high amounts of sedentary time, snacking, skipping meals, portion sizes, energy density of foods/meals and potentially a high sugar consumption. Also various obesogenic environments may be impacting a child's risk of obesity. Unfavourable conditions in the foetal environment are a risk factor for subsequent obesity. Infant postnatal weight gain can follow a high-risk pattern - a warning sign for subsequent obesity is when a child is becoming increasingly fat when his/her peers are generally showing a reduction in fatness (i.e. between around 6 months and 5 years old), plus if this fatness is developing when other children are tending to decrease fat it is probably a warning of persistent obesity (Lobstein et al, 2004). Breast feeding may have a protective effect, although this effect may be due to confounding by maternal diabetes or BMI. Similarly longer sleep duration seems to be protective, but may be a marker for other factors. Parental BMI has a strong positive association with childhood obesity and familial similarity in behaviour can predict risk of obesity. The literature supports the view that low SES and/or deprivation in childhood in the home environment is associated with childhood obesity as well as subsequent obesity in adulthood. However many studies take a too simplistic approach to defining deprivation and insufficient consideration of possible confounding factors, such as

parental BMI. Also the school environment may influence prevalence of obesity, although the evidence is weaker, but nevertheless schools can be used as a platform to help prevent obesity. Developments in industry, stemming from economic growth, serve to enhance consumption and are aspired to by developing countries, yet are contributing to our obesity problems. Government led regulation and industry self-regulation can help to level this playing field, albeit many conflicts of interest exist. Further, the extensive, robust impact of the media cannot be ignored.

This review emphasizes the need for multi-level approaches if we truly want to prevent childhood obesity. It also serves to highlight that there is a need to extend the current research base in order to build a well-founded framework to form the basis of a strategy for the prevention of childhood obesity, in particular to be able to address measurable, changeable environments in order that viable, long-term, population level prevention strategies can be successfully implemented.

Chapter 3: The Geography of Obesity

3.1 Introduction
3.2 The geography of health
3.2.1 Compositional or contextual?
3.2.2 Social capital
3.2.3 Scale of analysis
3.3 The geography of obesity
3.4 Discussion

3.1 Introduction

The purpose of this chapter is not to give an exhaustive review of the geography of health, as this would be too broad a remit and would not permit any in depth consideration of pertinent topics. Instead it seeks to review selective areas of the geography of health that can be applied to the geography of obesity. In reviewing the vast literature of geography of health, the focus has been placed on how health geography provides information about spatial variations in health outcomes and risk factors, and how this relates to the environment.

This chapter starts by providing a definition of the broad concept of the geography of health, before moving on to consider selective areas of health geography that can be applied to obesity, in particular examining the composition and contextual effects debate, followed by a consideration of how social capital may add a further dimension to the explanations of differences in health. The issue of spatial scale is also addressed. Next, an overview of the geography of obesity literature follows, focusing on spatial analysis of food access, green space availability, neighbourhood characteristics, and degree of urbanization, as well as examining the effect of multiple environmental determinants on obesity. Obesity has risen rapidly in recent years and there is growing evidence of spatial variations worldwide. These studies increase understanding about aetiology of obesity and can distinguish between individual and environmental impacts. Finally there is a discussion about how this analysis can add value to obesity research and prevention initiatives.

3.2 The geography of health

Health geography is the application of geographical information, perspectives and methods to the study of health, disease and health care. The study is considered a sub-discipline of human geography, however, it draws from many different fields to elucidate the research, such as epidemiology, sociology, ecology and statistics. Health geography research is diverse and wide-ranging (Asthana et al, 2002). It can provide a spatial understanding of a population's health, the distribution of disease in an area, and the environment's effect on health and disease.

It provides the ability to find and comprehend patterns of disease or health behaviours, highlighting any inequalities. It also deals with accessibility to health care and spatial distribution of health care providers and services. This research is undertaken using both qualitative and/or quantitative methods.

3.2.1 Compositional or contextual?

There is a large literature contending that our environment, particularly that of our place of residence or school/work, impacts on health related behaviour and therefore health outcomes (Macintyre et al, 2002; Mohan et al, 2005). This is a relatively new area of research (with no such articles on Medline bibliographic database until 1993 (Pickle et al, 2005)), although this recognition of the impact of geography on disease dates back to the 1800s, with John Snow and his cholera maps. This increased interest in the effect of place on health seems to stem from the publication of the Black Report some 28 years ago (Black et al, 1980). Since then many authors have shown that deprivation is related to mortality (UK examples include: Dorling et al, 2000; Senior et a, 2000; Maheswaran et al, 2002; Woods et al, 2005) as well as to specific health outcomes (Cooper et al, 2000; Lostao et al, 2001; Benigni & Giuliani, 2002; Horne et al, 2004; Shang et al, 2004; Tseng et al, 2006). The health impact of place has wide-ranging effects, including physical, psychological, social, spiritual and aesthetic (Frumkin, 2003).

An important debate within health geography is that of whether the environment has compositional or contextual effects on health. That is, the issue of whether individual or area effects on health predominate. Accordingly, the compositional school of thought is that individuals have risks of ill health, therefore an area's ill health is reflective of that of the individuals who live (or work, as appropriate) there. For example, do obese people congregate in similar locations? Conversely, the contextual theory is that living (or working) in an area imposes ill health on that area's residents. For example, do certain attributes of places cause its inhabitants to become obese? That is, does the environment increase the risk of obesity (or have a protective effect) (contextual), or are risk factors simply inherent to the individual (compositional); are people living in poor communities more obese because of lower socio-economic status, or because there is something intrinsically unhealthy about living in such communities?

The literature is not conclusive on this point. Mitchell et al (2000) argue that mortality is entirely due to compositional effects, and several other authors have also emphasised the significance of standardising the age-sex distribution and the socio-economic factors of an area before comparisons can be made (Hayes, 1991; Shouls et al, 1996). However a further body of literature concludes that area of residence does have an effect over and above effects of

population composition (Reading et al, 1999; Pickett & Pearl, 2001; Subramania et al, 2001; MacIntyre et al, 2002).

3.2.2 Social capital

Extending from the debate regarding the issue of the relative importance of individual- or area-level effects on health, follows the question regarding the impact of social capital. It has been suggested that the social environment may play a role in rationalising the differences in health between places and populations. The sense of worth of the individual and/or the social unity of the community may be detrimentally impacted by social inequalities, particularly in relation to income (Mohan et al, 2005). Accordingly social capital may explain the connection between social inequality and health.

"Trust, connection and reciprocity" are the fundamental concepts of social capital (Ziersch et al, 2005), and the over-riding philosophy of it is that benefits and obligations are bestowed on individuals through membership of a social group (Hawe & Shiell, 2000). High levels of social capital confers benefits on the individuals/communities associated with it (Mohan et al, 2005) for a variety of reasons, including social trust, assisting coordination and communication, and developing a broader sense of self thereby encouraging collective behaviours (Putman, 1995).

It is a hypothesis that has been interpreted at both the area and individual levels (Baum & Ziersch, 2003). As an individual construct, social capital focuses on the resources (potential and actual) that individuals amass due to their belonging to social networks (Bourdieu, 1986). Conversely, social capital as a community construct focuses on the strength and structure of social ties. Putman (1995) defines it as "features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit". Accordingly social capital has been measured at different geographic scales, from the local, neighbourhood to the national level (Kawachi et al, 1997; Ellaway & MacIntyre, 1999; Subramanian et al, 2001).

The differences in definitions make it difficult to assess the relationship between social capital and health. A number of studies have shown a relationship between indicators of social capital and health (Rietschlin, 1998; Gattrell et al, 2000; Joshi et al, 2000; Hyppa & Maki, 2001), but the results are not universal with some studies showing no or inconsistent associations (Ellaway & MacIntyre, 1999; Baum et al, 2000; Veenstra, 2000).

Wallace & Wallace (1997) discuss how the break down of social and community factors can lead to unhealthy behaviours, which in turn loops back further breaking down social &

community factors, thereby exacerbating the vicious circle. Environments (social, political, economic, etc) progressively change (many small changes to many different factors) and one final small change acts as a "trigger point" in unbalancing the network, tipping the balance into "unhealthy". The occurrence of this trigger point is critical in whether a disease becomes endemic or not. Using obesity as an example, this condition could exist in a population for a long time, and as long as each case lead to, on average, just one new case, the prevalence of the condition remains stable. But if circumstances change, say the safe cut-through to school gets built on so children have to be driven to school, drug use increases in an area making it unsafe for children to play outside, a fish and chip shop opens directly outside the secondary schools, then this could cumulatively be sufficient to unhinge the disease and cause a massive increase in obesity. Accordingly we need to understand these environmental factors, or the "upstream" factors (so from the above example, why did drug use increase and why did urban planners not consider the impact on walkability by building near a school?), in order to understand the mechanism of disease (Gattrell, 2005), and in this case the obesity epidemic.

3.2.3 Scale of analyses

Spatial data are often considered at the macro level (e.g. county level), largely because of the absence of routinely available small area health data (e.g. super output area, postal sector), and to a lesser degree due to a lack of spatial analysis skills or awareness amongst health researchers / epidemiologists. This means that monitoring and targeting are being undertaken at relatively crude geographical scales, which are often too large to be considered neighbourhoods (for example, the difference between a ward in the UK with an average population of around 2400 people, with the local "neighbourhood" of probably only a few streets). However, health behaviours are not homogeneous across space and this lack of uniformity is exacerbated the larger the study area unit, and it is debateable at what point the analysis becomes so large that it is "global" rather than "local" (e.g. county vs. postal sector). A key advantage of undertaking spatial analysis of health data at the micro scale is that this facilitates focus on key problem areas, rather than relying on averages for the whole region as per a global analysis. To generalise for a whole city would mean that health professionals / planners could miss small problem areas.

3.3 The Geography of obesity

Health geography concepts are applicable to the study of obesity. In particular, geographic information system (GIS) and spatial analysis techniques can be used to model the obesogenic environment. This is particularly useful as obesity is such a complex interaction between biophysical, social, environmental and psychological factors. There are several different areas

of obesity research where spatial analysis has been undertaken, including food access, green space availability, neighbourhood characteristics, and degree of urbanization, as well as examining the effect of multiple environmental determinants on obesity.

Poor access to affordable, healthy food is considered to be a contributory factor to poor diet, poor health and obesity. It has been shown that these "food deserts" exist (for example, Clarke et al, 2002; Wrigley, 2002; Whelan et al, 2002). Improving access to food can increase fruit and vegetable intake, which also suggests that limited access to healthy, affordable food does affect the diet consumed (Wrigley et al, 2002). Also Sturm & Datar (2005) found that higher fruit and vegetable prices were positively correlated with change in BMI. Yet all the evidence does not agree with other authors not finding a positive relationship. For example, Pearson et al (2005) did not find an association between the amount of fruit and vegetables consumed and food desert features (namely, access to supermarkets, fruit and vegetable price, deprivation) in the UK, neither did Winkler et al (2006) in Australia (considering distance, shop density and opening hours) despite considering areas of differing socioeconomic disadvantage, although the evidence does seem to be stronger in the USA (Cummins & Macintyre, 2006).

Likewise, the relationship between green spaces and obesity has been considered. A start was made at this by Liu et al (2002), reporting in that paper the impact of proximity to "play spaces", however as they used BMI rather than a BMI standard deviation score to define overweight and obesity in children their results are unreliable, but the concept is reasonable. A more thorough examination of park space was undertaken by Coen & Ross (2006), although this study considered the association with health rather than obesity. They concluded that parks in areas of poor health showed material disadvantages compared to those in healthy areas.

It has been shown that neighbourhood characteristics, such as deprivation, impact behaviours that effect health, such as walking and smoking (Ross, 2000). Deprivation is commonly associated with obesity, although the relationship is not straightforward, depending on the timing of the outcome measure of obesity (that is, whether it is in childhood or adulthood) (Parsons et al, 1999; Hardy et al, 2000; Okasha et al, 2003; Monden et al, 2006). Also different authors use different measures of deprivation, ranging from a simplistic single indicator of socio-economic status (SES) as a proxy for deprivation (Romon et al, 2005; Strauss & Knight, 1999; Danielzik et al, 2004; Cecil et al, 2005; Lamerz et al, 2005; Stamatakis et al, 2005) to a more sophisticated indicator of deprivation by ranking several different factors (Kinra et al, 2000; Dummer et al, 2005; Kinra et al, 2005). Few studies have considered specific characteristics of the environment, namely the impact of road safety issues on obesity (Timperio et al, 2005).

Furthermore overcrowding, poverty, migration, pollution, housing, employment can all create environmental changes that may initiate the breakdown of community factors that adversely impact health. This is encapsulated in the term "urbanisation", which relates to the concentration of populations into towns/cities and the corresponding changes associated with this: migration; transformation of the economic and physical organisation of the city; changes in behaviour of populations due to "urban-living" (Asthana et al, 2002). It is likely that these behavioural changes effect health behaviours and so the probability of disease. This potential role of urbanization in the aetiology of obesity also needs to be considered in light of urban regeneration and whether this supposedly positive urban change impacts health (Curtis et al, 2002). In this vein levels of urban sprawl (that is, the amount of developed land for a given constant population), which is a relatively recent phenomena and reduces accessibility of work, school and social activities by foot, have been showed to be positively associated with obesity in the USA (Vandegrift & Yoked, 2004), where walkability is more of a universal issue than in the UK. Similarly the term "globalisation" recapitulates how activities and decisions in one area of the globe are having a considerable impact in locations that are far-away (Asthana et al, 2002). States and societies are now more interconnected and interdependent that they once were. Global pandemics of disease exist, with obesity perhaps being a classic example as prevalence of obesity in Europe rapidly catches up with that in the USA.

Other studies have looked at multiple aspects of an obesogenic environment rather than single attributes, which is more realistic as environmental factors do not operate to shape our health behaviours in isolation. Geographic clustering of environments were determined and compared with different health outcomes, including obesity, in Nashville, Tennessee. Point (individual) data for health outcomes were aggregated to census units, and compared to aggregate census data for that unit. The 136 census data were amalgamated and categorised into 12 definitions of "poverty, crowding, urban core, low SES, aging residential, family residential, born outside Tennessee, immigrants, rental units, rental costs, intelligensia and accumulated wealth" (Schlundt et al, 2006). They found that locations with poor health (albeit self reported) were associated with poverty, crowding, aging infrastructure and low levels of education. Similarly Nelson et al (2006) have modelled the obesogenic environment. Instead of using aggregated census variables to illustrate neighbourhoods, the authors used cluster analysis procedures to identify six different environment patterns related to physical activity levels and obesity using eight different residential variables (income, race, education, SES, crime rates, road type, walkability and recreation facilities). The data showed significant differences in physical activity levels and obesity between the different neighbourhood patterns.

3.4 Discussion

Recent UK White Papers (DH, 2004; DH, 2008) outline a number of actions to tackle key current public health issues; of which obesity is high up the agenda. This shift in government health policy (in the UK at least) away from considering isolated disease groupings towards a population approach that considers the determinants of health will require massive change at all levels as solutions to health problems lie in many places (Hawe & Shiell, 2000).

Furthermore, whilst it may be possible to identify individual level determinants of health, very often it is not possible to modify these factors or behaviours. Accordingly interventions need to occur at a higher level (the "upstream" factors), changing the cultural, social, and physical factors that also affect health, thereby considering the determinants of health at the population level rather than individual level. It should be remembered that area level variables are not the same as individual level variables (Robert, 1999), plus the location of individuals is not random (Propper et al, 2005). In considering the compositional and contextual debate and weighing the evidence, it can be reasoned that the health (or ill health) of an area is composed of a combination of the health profiles and health behaviours of the residents together with environmental factors of the locality e.g. access to green spaces, number of primary health care facilities, pollution levels, etc.

The issue of compositional or contextual effects of the environment on a population's health can be elucidated by establishing the extent of spatial variation of, say obesity, in an area (Yiannakouliasa et al, 2003). Individual level variation in disease is less likely to be explained by contextual effects where there is minimal spatial variation, but if there is significant spatial variation, then it is possible to consider the contextual effects and whether or how such variation is explained by place. Accordingly it is preferable that analyses would contain data on both community and individual attributes, modelling the two simultaneously in order to glean the most information about health determinants. Plus, importantly, it is at the contextual level where public health measures can be most effectively introduced, rather than trying to change the behaviour of individuals directly.

A consideration of social capital can make an important contribution to an understanding about health inequalities. The concept of social capital may help to focus interventions to redress health inequalities at the level of community rather than focusing on trying to change individuals' behaviour (Hawe & Shiell, 2000).

There is a strong case for relationships existing between different aspects of the environment and obesity. The obesity studies do not necessarily indicate causality; nevertheless they do

provide more detail about correlations between environmental factors and obesity and also suggest that overcoming social and economic challenges would facilitate the reduction of health inequalities. Furthermore the consideration of environmental patterns, rather than single socio-economic variables, more closely simulates the real world and real environments in which people live and work. Environmental factors cannot, and do not, operate in isolation. As such, a combined effect needs to be considered. A nutritional analogy is looking at dietary patterns (such as a Mediterranean diet, traditional British diet, etc) rather than individual nutrients. Rather than focusing on the "downstream" outcomes of individual behaviours or lifestyles, better understanding of the "upstream" factors that tip environments into obesogenic environments is required, and whether any trigger points or thresholds exist.

Illuminating how people react and interact with the environment to the benefit or detriment of their health may not provide a complete solution to understanding the obesity problem, but it does provide an enhanced platform for analysis, evaluation and decision-making in health planning. Furthermore as public health embraces the micro level spatial analysis concept and moves to a more local approach, this will facilitate more focused and detailed health planning. It enables governments and health professionals to response to local differences in health behaviours, and to develop and implement more targeted interventions and health policies for prevention.

This is a time when many different literatures are coming together, and through this combination of expertise the field of obesity prevention has much to gain with a little innovative thinking and challenging of traditional analysis. Looking forward, the geography of obesity should have a role in improving our knowledge about the aetiology of obesity (i.e. "spatial epidemiology" works alongside, not instead of, traditional epidemiology). These are additional erudite ways of conceptualising and devising population-level and place-level interventions and health policies to help prevent obesity, which may assist in transforming the propensity of interventions to focus too much on individual-level activities.

Chapter 4: Spatial variations in childhood obesity in Leeds

4.1 Introduction
4.2 Background
4.2.1 Use of routinely collected data to monitor childhood obesity
4.2.2 Spatial analysis of health data
4.3 Methodology
4.3.1 Study area
4.3.2 Sources of data
4.3.3 Analysis of data
4.4 Results
4.4.1 Routine data coverage
4.4.2 Global analysis
4.4.3 Spatial analysis
4.5 Discussion
4.5.1 Data collection
4.5.2 Global analysis
4.5.3 Local analysis
4.5.4 Limitations
4.5.5 Conclusion

4.1 Introduction

Obesity in children is a rapidly growing problem in the UK and worldwide and has been increasing at accelerating rates in more recent years. It is associated with a number of co-morbidities in childhood and with increased risk of adult disease, particularly cardiovascular disease, hypertension and type 2 diabetes (Lobstein et al, 2004). On top of this treating obesity and obesity related diseases is expensive; the total cost of treating obesity in the UK was £3.3-3.7 billion in 2002 and is estimated to increase to £7 billion by 2020 (Select Committee Report, 2004). The Foresight report (2007) has now increased this estimate to £45.5 billion by 2050. Reducing childhood obesity and health inequalities are at the centre of the UK government's health policy, with halting growth in childhood obesity is a prime objective. To be able to assess whether this is achieved and whether interventions to reduce the prevalence of childhood obesity are successful, then accurate and comprehensive data are required at both baseline and after the changes are made. To drill down into potential inequalities in health data, these data need to be considered at the micro-level rather than taking averages for a whole city whereby pockets of problem areas could be missed.

There are two key objectives for this chapter. The first is to consider the potential to use routinely collected data to monitor trends in obesity and the second is to increase our understanding about obesity and associated inequalities.

This chapter firstly outlines the use of routinely collected data to monitor childhood obesity, and then the use of Geographic Information Systems (GIS) and spatial statistics to analyse health data together with an exploration of the problems associated with this type of analysis. Then the methodology is described, starting with a description of the sources of data used, followed by an

explanation of the analysis undertaken. These analyses lead to the identification of "hot" and "cold" spots of problem areas at the residential micro-level and enable consideration of how populations in these areas differ. The data are considered at school level in chapter 5. Finally the discussion section puts the results into context of the existing body of work and an indication of the insights that these results give to the understanding of the whole area is given.

4.2 Background

4.2.1 Use of routinely collected data to monitor childhood obesity

In the UK the National Child Health Computer System (NCHCS) is used to record the results of regular health check ups for children, which should be undertaken on all children when the child is at specific ages (which varies across the country), but generally weight and height (amongst other things) are measured at regular intervals from birth until school entry, with some areas also measuring older children (7 and 11 years). Each Primary Care Trust (PCT) is responsible for collecting its own data. Therefore the quality of the data varies across the country (confirmed by interviews with PCT personnel responsible for managing the NCHCS). Also recently the government has introduced a new monitoring of heights and weights of 11 year olds, the National Child Measurement Programme (NCMP) (http://www.ncmp.ic.nhs.uk/), again intended as 100% coverage.

Three studies have specifically looked at using this population data to monitor trends in obesity across the UK, although with differing results. Banerjee et al (2003) showed poor coverage (c.67%) of weight and height measurements in the South Wales (Rhondda & Taff Ely) area, either because no height and/or weight measurements were taken (for example, if the child did not attend school on the day of measurement) or because the data had not been input into the NCHCS (remaining as inaccessible paper records) due to lack of resources to undertake the data input. Height was considered secondary to weight (as not all those weighed had their height recorded), which obviously means BMI cannot be calculated. Similarly Blundred et al (2001) examined data in the Wirral Health Authority of the North West region from 1989 to 1998 and also found low coverage. They had to exclude 25% of the records due to missing or inaccurate data (15% did not have full data recorded; 9% had incorrect age data; 1% had unfeasible weight or height data). Jones et al (2005) had a completely different experience however, also in the South Wales (Swansea, Neath and Port Talbot) area. This database had a much higher coverage of height and weight data (87%-99%) than that reported by Banerjee (2003) in neighbouring localities. However there were still many measurement errors, only some of which were obvious (e.g. when a later height measurement was less than the earlier one; or when weight and, more frequently, height data had been recorded incorrectly, usually as the result of

incorrect placing of decimal points). It was suggested (Betts, 2003) that the NCHCS should have been developed with the ability to reject obviously wrong entries, as well as to highlight those in the highest and lowest centiles.

Other authors suggest that routinely collected data are inadequate for monitoring trends in childhood obesity, suggesting that the data are too inaccurate, and that school entry age (which is when routine measurements are taken) is too young as levels of obesity are low relative to later in childhood (Rudolf et al, 2006). In view of the fact that the average child visits their general practitioner 15 times up to the age of 14 years, there are many opportunities for measurement (Betts, 2003). However, as the data are showing significant increases in obesity in young children (i.e. by school entry) (Blundred et al, 2001), then monitoring does (additionally?) need to take place in these young age groups, which will also facilitate the design of interventions targeted at this young age group if there is to be hope of having an impact on the rising prevalence of obesity.

Given the investment already made to measuring height and weight at school entry and in infants, to ignore this source of data would be a huge waste of resources. Accordingly this study reviews the quality of the NCHCS records in the Leeds primary care trusts and considers whether it is possible to use these data to monitor trends in obesity and, importantly the obesogenic environment, in young children in the UK.

4.2.2 Spatial analysis of health data

Spatial scale is important when considering health data, as it may be that variables that illuminate variation in disease at one scale may not be relevant at a different scale (Gatrell, 2002). For example, differences in disease prevalence between countries may be a function of levels of government spending on health; variations within a small region may be due to environmental variables. Areas of high and low rates of disease may give clues regarding causation (Gatrell & Loytonen, 1998). The scale of analysis (or level of aggregation) is a trade-off between specificity and precision (Wilkinson et al, 1998). The smaller the area the more specific the results to that population, but the greater the imprecision and potential for bias due to small number problems.

There is much discussion in the health geography literature (Gatrell 2002 is a good starting point) about whether health variations from place to place are due to compositional or contextual effects. The former being that these spatial health variations are due to different types of people living in different places and the latter that they are due to places themselves differing in terms of the quality of the environment or other attributes. In reality it is likely to be

a combination of both effects: individual health behaviour choices will vary from person to person (although one might expect them to be similar in a small, homogenous area such as output area, where the population have broadly similar census-type characteristics) (compositional), and will be impacted by the environment around them (contextual) (see Figure 2.7, the ecology model of health behaviours in chapter 2). The key point is that if a particular area has a high prevalence of disease, then rather than taking a total population approach, it makes more sense to strategically target those high-risk populations (a good example is the dental caries programme where 90% of the disease is concentrated in 10% of the child population and targeting has successfully controlled the problem (Tickle, 2002)).

A map of disease incidence or prevalence (the latter being more useful for slow developing, chronic conditions like obesity) may show spatial patterns or clusters (spatial autocorrelation), which in turn may give clues about the disease aetiology. Positive spatial autocorrelation is where similar values inhabit adjacent positions on a map. Negative spatial autocorrelation is where high and low values are located adjacently. That is, spatial autocorrelation exists when the location of disease cases are dependent on the location of other disease cases, such as with an infectious disease where an infected individual is likely to infect those living close by, or if obesity were caused by a lack of walkability in an area then all residents of areas of low walkability are likely to be equally effected. There is no spatial autocorrelation if the arrangement of values is completely random. The smaller the spatial units (finer geography of analysis) the increased likelihood that close spatial units are dependent. It is this lack of independence inherent in spatial data that affects the statistical methodologies that can be used to analyse the data. Events distributed over space tend to not be independent. This leads to problems with statistical analysis that assumes independence (Fischer et al, 1996). If spatial autocorrelation is ignored, the confidence in the risk relationships will be overestimated with biased p-values that are too small, providing "statistically significant" results when none exist (Kulldorf, 2006).

Using a spatial scan statistic to search for hot spots of high prevalence of disease is effectively looking for local spatial autocorrelation (Rigby & Gatrell, 2000). In looking for hot spots of high prevalence of disease, it is geographical clusters of data that should be considered, not simply clusters of data. If it were the latter then neural networks would be a suitable technique to use. The use of spatial statistical techniques, such as Spatial Scan Statistic (SaTScan) (Kulldorff, 1997), FleXscan (Tango & Takahashi, 2005) or WinBUGs (Cowles, 2004) enable us to identify any significant spatial or temporal clusters of high prevalence of the condition or disease (the software for each is freely available from the Internet, see Kulldorff & IMS (2005), Takahashi, Yokoyama & Tango (2005) and Spiegelhalter, Thomas, Best & Lunn (2002), respectively). Of these three, SaTScan is best suited for the purpose of this thesis: WinBUGS is

not "user friendly" and best suited to use by a statistician; the data do not require a flexible cluster shape (such as if disease followed the course of a river), so FleXscan is not appropriate; furthermore SaTScan enables different models to be used, such as a Bernoulli, Normal, Poisson and Exponential.

However, the use of spatial data is not without its problems: confidentiality of patients is a primary concern, which can lead to use of aggregated data, rather than data for each individual, and the problems associated with it (e.g. loss of detail, or imputation to prevent identification of individuals). The methodologies used to integrate data and present the results are important as they can affect the interpretations. For example, the scale used in creating a map can generate very different visual patterns. Finally, and probably the least easily resolved, migration can be a major issue when dealing with chronic disease patterns, although hopefully in the analysis of child data, as in this thesis, this is less of an issue.

Health data can be identified geographically, for example, through use of residential ward or postcode. However the finer the geographical resolution the concern about patient confidentiality increases. Nevertheless full postcode data can be converted to Cartesian coordinates (for the centroid of the postcode) and mapped to an acceptable approximation whilst retaining acceptable levels of patients' confidentiality. Due to this concern over the protection of patients' confidentiality, individual level health data are not always available for analysis. The data supplied for research are often aggregated to areal units, such as wards. Using appropriate spatial analytical techniques, the data can then be used to answer health questions of interest. However, investigating data grouped into areal units is plagued with several problems, the major ones being the Modifiable Areal Unit Problem (MAUP), ecological fallacy and small number problems.

The MAUP occur when analysing data by area (point data aggregated to areal units) and different results are obtained depending how the areal units are configured; that is, where boundaries are, which is a relatively arbitrary choice i.e. areal units are "modifiable" (Openshaw, 1983). In addition to the usual sources of bias that affect individual level data (e.g. selection bias, observation bias, interviewer bias, recall bias) there is the problem of "ecological fallacy" with aggregated areal data. This is the failure of the expected ecologic effect (i.e. at population or community level) estimates to reflect the biologic effect at the individual level (Rothman & Greenland, 1998) and is due to heterogeneity within areal units. The size of areal units is important: smaller areas are more homogenous, but then run into small number problems. A common problem with working with relatively small areal units (or small number of population at risk) is the misclassification of risk estimates. These may lead to extreme risk estimates for areas with small population or small number of cases (Gatrell, 2002). The small

number problem is essentially a problem of scale, because as the size of the areal units increases, so does the number of cases within the unit (for example, there are more cases of obesity in the UK than at ward level). With analysis at the micro-level, where there are only a handful of cases per area, then estimates of disease risk become unstable (since the addition or subtraction of just one case can greatly affect the estimate). Almost by definition there are going to be a small number of events (e.g. cases of obesity) in a small area with a small population; so analysis at the micro-level is going to have problems with producing sufficient numbers for analysis (Carstairs 1981). The data will consequently have large standard errors.

There are two key ways to minimise the effect of this small numbers problem and to increase precision. One is to aggregate data where possible: to extend data collection over several years or to work at a coarser geographical scale (although the latter point is not of help if looking for micro-level hot spots of disease). The other is to use statistical shrinkage methodologies to shrink unstable disease estimates toward the population mean rate; that is, to use Empirical Bayes estimation. Empirical Bayes estimation method is a "smoothing" approach that gives risk estimates based on mean risks of the area unit and that of neighbouring areas (Leyland & Davies, 2005). In this way it prevents undue attention being focussed on areas with small numbers. An Empirical Bayes variation is estimated from the variation in the data itself: the data for small numbers is pooled across areas to provide a more stable estimate of the rate. To put it another way, Empirical Bayes works by "borrowing strength" from areas where there are large numbers measured to improve the data in areas with low numbers measured. There are four possible methods to smooth data using Empirical Bayes; one global and three local techniques. Global smoothing causes the rates to be smoothed based on all geographical areas (i.e. using the mean for the study area). Local smoothing permits smoothing to be limited to a subset of the areas based on different criteria, either based on population limits, distance or a particular field in the dataset. The type of smoothing chosen will depend upon the needs of the user and the characteristics of the data.

Data integration, that is the process of making different data sets compatible with each other, needs to be undertaken with care – at best it can be a headache, and at worst can make the analysis invalid. This problem is caused by various factors, such as different spatial referencing systems, different degrees of generalisation, locational errors, and different temporal coverage (Fischer et al, 1996). Similarly the presentation of data using maps can present a number of difficulties, for example in the choice of number of classes used, class intervals, and colour scheme, as these affect how the map is read or perceived (Cliff et al, 1998; Monmonier 1996). The most commonly used method for visualising areal-based data is the thematic, or choropleth map. In this map the areal units are classified and coloured (or shaded) according to the value of a particular attribute under investigation. However large areas on maps may be sparsely

populated (rural areas) and yet these areas can visually dominate the map to the detriment of the more densely populated, smaller, urban areas, where the attribute being mapped may be of more relevance. Some human geographers, notably Prof D. Dorling, have introduced the use of novel cartographic techniques where areas are re-sized according to the subject of interest, such as population size (cartograms rather than thematic choropleth maps) (Dorling, 1995); however these techniques do not come with the standard mapping software and their use is beyond the scope of this thesis.

Last but not least, migration can be a major issue when looking at chronic disease data, such as obesity. This time lag, or latency period of disease, between exposure to pathogen and appearance of symptoms can be many years. Current health may reflect where the person used to live, perhaps many years ago, rather than their current place of residence. This issue is acerbated when looking at small areas, as people tend to move small distances. The achievement of reliable results in epidemiological research without the use of accurate personal histories (which would be too time consuming and costly to obtain, not to mention unfeasible), including migration data, is very difficult.

In conclusion, despite the many potential problems of analysing spatial data, with careful choice of statistical methods and choosing relatively homogenous areal units to examine, taking health data down to the micro-level can add invaluable knowledge about the location of hot spots of high prevalence of disease, which can help with the design and implementation of health policies that best target resources to those most at need or the most high-risk populations, and also with increasing knowledge about disease causation.

4.3 Methodology

4.3.1 Study Area

The study area covers the 33 Leeds wards (see Figure 4.1). Leeds is a large city in the North of England with a diverse population in terms of affluence and ethnicity. This area has been selected for largely pragmatic reasons associated with proximity to and connections with the University of Leeds. Also it is a large enough area to highlight differences, covering a variety of different demographic and socio-economic factors, as well as to make an important contribution to national knowledge about childhood obesity. The analysis is undertaken at ward (N=33), lower super output area (N=476) and output area (N=2440) level.

Figure 4.1. Map of the study area, which consists of the 33 wards in Leeds, Yorkshire. The study area is highlighted within the context of the outline of England and Wales. The top left corner illustrates the ward boundaries at a larger scale.

4.3.2 Sources of Data

Childhood obesity will be described using body mass index (BMI) (weight (kg)/height (m^2)), which is a measure for obesity (as described in chapter 2). There are three key sources of child BMI data used in this study:

1 The BMI data for 3 to 6 year olds were obtained from primary care trusts' (PCTs) records of routinely collected data for children born since 1995 (the PCTs' data on older children are sporadic).

The National Child Health Computer System (NCHCS) has been used since 1986 to record the results of regular health check ups for children. PCTs routinely measure (heights and weights) children to monitor their progress, comparing it to "norms" using growth charts for the British dataset (see Figure 2.1 in chapter 2) for gender and age. These charts include nine centile curves based on divisions of two thirds of a standard deviation, thus ranging from the 0.4th to 99.6th centile. Age and gender specific BMI cut offs define overweight as above the 91st centile and obesity as above the 98th centile (Cole et al, 1995). The specific ages at which children are routinely measured varies from PCT to PCT and can vary from year to year. Until September 2004 Leeds health visitors would measure weight and height (amongst other things) at birth, 6 weeks, 7 months, 18 months and 3 years old. Since then these examinations are only done at 6 weeks and 2 years old. Similarly until the end of 2003 Leeds school nurses were encouraged to measure children at reception (5 years old), Year 3 (7 years old) and Year 7 (11 years old), and since then they only measure children at reception and Year 6. These (primary) school-entry medicals have been consistently offered, but are not obligatory, either on the part of the school or the family. It is not clear how aggressively families who fail to turn up for appointments are followed up.

Data on older children are sporadic and unreliable. Unfortunately many of these Leeds records, particularly for 3 year olds, have not (yet) been transposed onto the NCHCS, remaining as inaccessible paper records (we did not examine any paper records). Recently the UK Department of Health took the (controversial) step of introducing monitoring for obesity in primary schools (NCMP), which may prove useful for future analysis, but are not included in this study as the data were not available in time.

BMI data for 3 to 6 year olds were obtained from the NCHCS records of routinely collected data for children in the Leeds PCTs born since 1995, in order that serial cross-sectional analyses could be undertaken. The only exclusion criteria was if the child had subsequently died, primarily because the cause of death would be unknown yet it might have affected their growth or weight and so bias the results. As well as height and weight data for each child, we also collected data on gender, age (calculated from the date of birth and date of the examination), and full postcode. Data were provided by year of birth. Initially children were excluded who were too young or old for our study (under 2.5 years; over 6.5 years). We also removed duplicate entries for the same measurement date for a child (this was due to a different entry being made on the system for each health practitioner who was present at the examination). This left us with the number of children who were measured by year of birth, which was required to calculate the proportion of total children who had been measured, determined using annual births and by census data. Then the data were sorted by year of measurement using Microsoft Excel 2000. Children's records with no height and/or weight data were excluded. If children had been measured more than once in the same year, the duplicate entries (randomly chosen) were excluded. These repeated measurements were scrutinised for cases in which the second height measurement was less than the first as an indication of the frequency of at least one obvious measurement error (in which case both records were excluded). The results for BMI were scrutinised for outliers. It was clear that weight and, more frequently, height data had at times been logged inaccurately, giving rise to unlikely BMI results. These mistakes seemed to arise due to the incorrect placing of decimal points. Accordingly a correction formula (multiplication of the weight entry by 10, 100, or 1000, as appropriate, and division of the height entry by 10) was applied to adjust for this. Following on from these adjustments, in order to remove the effect of remaining outliers that were also due to incorrect data entry, the data were reexamined and entries with an unfeasible BMI (BMI standard deviation score > +/- 4.00) were excluded. Finally entries where the child lived outside of the study area or where no, or incomplete, postcode information was provided were also excluded.

2 The "Trends" study measures BMI in a sample of 5, 9 and 13 year olds in Leeds (Rudolf et al, 1999; Rudolf et al, 2006). Data are available for use in this study that was collected in 2004 and 2005.

The Trends project in Leeds has collected growth data on school children in thirty-five primary schools and five secondary schools during 2004 and 2005 (the 2006 data were not available in time for this study). The aim of the project was to construct a simple, reproducible method to monitor trends in childhood obesity in light of the governmental target to stop the rise in childhood obesity by 2010 (Rudolf et al, 2006). Ten of the primary schools (for the 2004 study) and the secondary schools were selected as a purposive sample; a further 25 schools were randomly selected by computer from the remaining 230 state primary schools across the city for the 2005 sample. Agreement had been obtained from the head teachers and governors of each school for a specially trained health care assistant and scribe to measure children in reception class (age 5 years) and in Year 4 (age 9 years) and Year 8 (13 years old). "Opt out" consent was obtained from parents, and measurements were made in the summer term (April – June). Children, wearing light clothing only and no shoes, were measured. Height was taken to 0.1cm accuracy using a freestanding stadiometer (Raven Dunmow) and weight to 0.1kg, as previously described (Rudolf et al, 2003). As well as height and weight data for each child, data on gender, age (calculated from the date of birth and date of the examination), and full postcode were also collected. Children's ethnicity was determined from school records. All children with no or incorrect postcode data were excluded from the dataset, as were children living outside of the study area.

3 The "RADs" (Rugby League and Athletics Development Scheme) study measures BMI (amongst other things) in a sample of 11 year olds in Leeds. Data are available for use in this study that were collected in 2004-06.

The RADs project in Leeds has collected growth data on school children in thirty-nine different secondary schools during 2004, 2005 and 2006. Originally the RADs programme was set up (in 2004) to identify athletically talented children starting secondary school in Leeds who were then offered a place on a sports development programme. Given that each year approximately 5000 children commence their secondary school studies, it became clear that these data also provided a means to evaluate risk factors associated with prevalence of overweight and obesity, so the purpose of the programme evolved. Agreement was obtained from each school to participate in the programme and "opt out" consent was obtained from parents. Children, wearing light clothing only and no shoes, were measured. Height was taken to 0.1cm accuracy using a freestanding stadiometer and

weight to 0.1kg. The same equipment was consistently used by especially trained staff. Also measured was percentage body fat using bio-electrical impedance, waist circumference (at the top of the hip), plus there were a range of sport performance tests to assess speed, power and coordination. The following additional information was also collected from East Leeds PCT for each child: gender, age (calculated from the date of birth and date of the examination), and full residential postcode were also collected. All children with no or incorrect postcode data were excluded from the dataset, as were children living outside of the study area. Similarly if no date of birth, height or weight data were collected, then the child was excluded from this analysis.

No children were contacted, and no new measurements were taken. All data were anonymised and a unique identifying number given to each child. Residential Census 2001 low super output area (SOA) and output area (OA) zones were allocated to each child using the X/Y coordinates of their full postcode. Ethical approval was obtained from Leeds (East) Research Ethics Committee.

This analysis also uses census data to facilitate an understanding of the trends and patterns of obesity that emerge. In the absence of individual socio-economic status (SES) data, micro-level area based SES was used to proxy the individual socio-economic characteristics. There were two types of these data: the Index of Deprivation and the Open Area Classification "Super Groups". The Index of Deprivation 2004 (Communities and Local Government, 2004) was used to assign a measure of SES based on the child's full postcode of residence. A higher Index of Deprivation score indicates a more deprived area. The Index of Deprivation contains 7 domains built from 37 social, economic and environmental indicators, derived from the 2001 census data. The domains and their weights (the degree of importance attached to each domain in constructing the index) are as follows: Income deprivation 22.5%; Employment deprivation 22.5%; Health deprivation 13.5%; Education skills and training deprivation 13.5%; Barriers to housing and services 9.3%; Living environment deprivation 9.3%; Crime 9.3%. This index is provided at SOA level (of which there are 476 in the study area), which is a new census boundary designed to facilitate comparisons across the country. They have a minimum population of 1000 individuals, with a mean of 1500. They are built from groups of output areas (typically 4-6) and contain homogenous populations with similar socio-economic characteristics.

Demographic data is available from the 2001 census. However there is a vast amount of census data and the inter-relationships between multiple variables is important, as none exist in an area in isolation. Accordingly this analysis also used another index in the analysis that, like the Index of Deprivation, also seeks to summarise multiple characteristics of a micro-locality: the

National Statistics 2001 Open Area Classification of Output Areas (Vickers & Rees, 2007). This is a powerful and effective way of summarising the complexity of census data, providing a straightforward indicator of the features of the area and of the likeness between areas. Forty-one census variables are used in the index, namely: demographic attributes (including, age, ethnicity, country of birth, and population density); household composition (including, living arrangements, family type, and family size); housing characteristics (including, tenure, type & size, and quality/overcrowding); socio-economic traits (including, education, socio-economic class, car ownership & commuting, and health & care); and employment attributes (including, level of economic activity, and employment class type) (Vickers, 2007). The index is created by firstly finding, for each output area, the other output areas with which it is most similar, then by secondly grouping these output areas into a three-tier hierarchy: seven clusters of Super Groups (OAC super groups); twenty one clusters of Groups; fifty two clusters of Sub-Groups. This analysis uses the seven Super Groups, which are classified as follows: 1: blue collar communities; 2: city living; 3: countryside; 4: prospering suburbs; 5: constrained by circumstances; 6: typical traits; 7: multicultural. These titles broadly summarise the characteristics (41 census variables) of the Super Groups (for full details see slides 15-21, http://www.reallifemethods.ac.uk/ training/workshops/geodemographics/documents/vickers-geodemographics.pdf).

4.3.3 Analysis of Data

There are three components of this analysis: Firstly an analysis of the coverage and usability of the routinely collected data is undertaken. Secondly the data are considered on a global basis; that is, for the whole of Leeds without consideration of spatial scale. This includes an analysis of the change in childhood obesity in Leeds over time and for different age groups of children. Thirdly the data are investigated at a local level; the data are mapped to visualise any spatial patterns using different definitions of childhood obesity (proportion of obese children; proportion of overweight and obese children; mean BMI standard deviation score) at both SOA and ward levels. The use of different scales enables the pros and cons of micro-level analysis to be considered. Also the data are tested for significant micro-level hot spots of childhood obesity using a spatial scan statistic, again using all three definitions of childhood obesity. Plus a multi-level model is constructed to allow for covariates (deprivation and OAC super groups) and the resulting adjusted hot and cold spots are described.

1. Routine Data Coverage:

Coverage (the proportion of children on the system who had a record of sex, height, and weight) was calculated for each measurement year for the routinely collected data. This calculation was not pertinent to the other two BMI datasets as they were samples. Three different measures of

coverage were calculated. The "normal" method is to calculate the proportion of children with height and weight data against all of the children for that age / year on the NCHCS system. However in this study the "birth" and "census" coverages were also calculated. The birth coverage uses the total number of children born in Leeds in the corresponding year of birth as the denominator. Similarly the census coverage uses the total number of children of corresponding age living in Leeds in the measurement year.

2. Global Analysis:

Body mass index (BMI) was calculated and converted to standard deviation scores (BMI SDS) using the UK 1990 growth references (Freeman et al, 1995; Cole et al, 1995) to allow for statistical computation of children of different ages and sex. Mean age, height, weight, BMI and BMI SDS was calculated for each measurement year of the routinely collected data, and for each age group for the three childhood "obesity" datasets. The proportions of children falling into each BMI category were calculated. These categories were defined as follows: "obese" is above the 98th centile (equivalent to BMI SDS > 2.00); "overweight" is above the 91st centile (equivalent to BMI SDS ≥ 1.33); "acceptable" was taken as below the 91st centile (equivalent to BMI SDS < 1.33) (see Figure 2.1).

As the routinely collected data were provided over six years, this enabled serial cross-sectional analyses of the proportion of children in the various BMI categories over time to be considered using a logistic regression model; time trends were investigated using "year of measurement" as a covariate. Using data from all three childhood obesity datasets enabled analysis of the proportion of children in the various BMI categories in five discrete age groups (namely, 3 years, 5 years, 9 years, 11 years and 13 years), which were analysed with logistic regression models. In both regression models differences between boys and girls were also allowed for.

Descriptive analysis of the two census derived indices was also undertaken. Mean deprivation score was calculated. The global relationship between deprivation and mean BMI SDS was analysed using a linear regression model. Similarly the odds ratios between deprivation (dichotomised as affluent / deprived i.e. a score greater or less than 30) and numbers of obese, and overweight and obese, were calculated. The OAC super groups were tabulated and mean BMI SDS and deprivation calculated for location and for study population.

3. Spatial Analysis:

The spatial analysis complements and extends the global analysis by considering the data at the residential micro-level. Small number problems prevent individual age group or measurement year analysis (even after allowing for an Empirical Bayes smoothing technique), so the data will be considered cumulatively using all three aggregated datasets.

A GIS (ArcGIS version 9.0) was used to summarise the childhood obesity datasets for the residential analysis at various spatial scales (ward, SOA, and output area levels, as appropriate). The residential location of each measured child was determined by linking each child's residential full postcode centroid to the relevant OA using GIS software. Ward and SOA data ties in with the OA data. Mean BMI SDS was calculated for each residential SOA and ward, as was the proportion of (measured) children in each BMI category (obese, overweight, acceptable). In view of the small aggregate numbers involved (which remains an issue despite using cumulative data) the data were smoothed using an Empirical Bayes technique. Given the reason for the smoothing is concern over the small numbers measured, a local smoothing based on population limits (using the number of children measured as the population variable) was used. In this way the smoothing searches the neighbouring areas (for each area, e.g. SOA, in turn) until it is at least equal to the population limit set. Then each population area and its neighbouring areas that meet the population limits are referred to as a group and smoothing takes place within this group. The process is iterative.

Firstly the data were mapped (using ArcGIS version 9) at residential level showing the proportion obese, proportion overweight and obese, and the mean BMI SDS. This identifies any spatial patterns across the study area. These residential analyses were undertaken at SOA and ward level, to highlight the extra information provided by undertaking analysis at the smaller scale. Deprivation and the OAC super groups were also mapped at SOA and OA levels respectively. The boundary data for the study area were downloaded from UKBORDERS (an online digital UK boundaries provider) in a form compatible for use with ArcGIS V.9.0 software. Secondly, to examine whether any residential spatial clusters of high prevalence of childhood obesity, "hot spots", are significant, a Spatial Scan Statistic software (SaTScan) was used to search for spatial clusters only (no temporal analysis were attempted due to small number issues) and to test whether they were statistically significant. The SaTScan User Guide (Kulldorf, 2006) provides a comprehensive explanation of how to use the software.

SaTScan will run analysis to search for hot spots at the individual as well as the aggregate level. Arguably, analysis at the individual level is more precise as there are no MAUPs (that is, aggregate boundaries are, at the end of the day, arbitrary and use of different boundaries could create different clusters). With the SaTScan software the data may be either aggregated (e.g. at the census tract) or there may be unique coordinates for each observation. The "individual" data in this study used the postcode to identify each individuals' residential location (which are not unique). With 33,594 children in the dataset, there are 9,651 different postcodes; accordingly hot spot analysis at the "individual" level would really be aggregated at postcode level. By using a micro-scale aggregation minimises the effect of arbitrary geographical aggregation of

the data; the finer the geographical resolution of the input data, the better able SaTScan is to evaluate different, smaller or larger, aggregations of data through its continually moving circular window. Note, the stability of the results of SaTScan depend upon the population sizes of the circles the model creates and not on the geographic resolution of the input data. However the resolution of the postcode unit was too fine for SaTScan (it crashed). In addition, these boundaries do not align with the OA/SOA/ward boundaries used by other data in the analysis. As such, it is better to aggregate to OA level (which still have 2440 unique locations), which is a micro-level geographical boundary that can tie in with other data that we are considering (rather than confusing the issue by using non-aligning geographic boundaries).

SaTScan can run many different types of probability model. Applicable to this study's data set were the Bernoulli model and the Normal model. The Bernoulli model can be used when the data consists of individuals with or without the disease in question (i.e. obesity, or obesity and overweight). For the purpose of the SaTScan model, those without the disease are "controls". The Normal model is suitable for continuous data that can be positive or negative values (i.e. mean BMI SDS). As clustering is obviously going to occur in highly populated areas, such as towns, the methodology to identify hot spots needs to be able to determine if the clustering is above that expected based on the population distribution, which both the Bernoulli and Normal SaTScan models do take into account (by considering the location of controls as well as cases, and by using a standardised statistic, respectively). Three SaTScan models were run. Two Bernoulli models were run: cases were number of obese children, and obese and overweight children (using the Empirical Bayes smoothed data); controls were the rest of the population in each OA (so all non-obese children and all acceptably categorised children respectively); the "geo" file was the Cartesian coordinates for the centre of each OA that had children measured. The third model run was a Normal model: cases were the number of children measured in each OA, with an attribute for mean BMI SDS (smoothed using Empirical Bayes) for each OA; no control file was required; same "geo" file as the previous models. All models used a maximum spatial cluster size of "50% population at risk", which is the default setting and the least arbitrary choice, as in this way SaTScan will evaluate very small and very large clusters plus everything in between. No geographic overlap of clusters was permitted. Maps showing the location of all significant clusters were produced using ArcGIS v.9.

Last, but not least, in order to investigate whether there is any relationship between residential obesity and/or overweight and deprivation or OAC super groups, multi-level modelling techniques (Rasbash et al, 2004) were then employed. This allows the dependency inherent in child observations nested within the same SOA to be taken into account. That is, it partitions the variation in obesity across each of the hierarchies, thus respecting the natural aggregation of the data (children at level 1, nested within SOAs at level 2) leading to less erroneous inferences

than if the natural ordering had not been accounted for. MLwiN version 2.0 was used based on a 2-level (child within SOA) hierarchical linear model. The independent variable was the BMI SDS of each child. The explanatory variables used in the model were deprivation score and OAC super group (7 categories). The analysis was undertaken at SOA level rather than OA level partly due to the increased small number problems at OA and also because MLwiN software could not handle that number of children and area levels. The residuals were examined to highlight any resulting hot and cold spots of childhood obesity (i.e. areas where BMI SDS was higher/lower than expected given the deprivation and OAC super groups). These results were mapped and the cluster characteristics summarised to facilitate comparisons.

4.4 Results

The routinely collected data were cleaned. Table 4.1 shows by year of measurement the number of children present on the NCHCS (for example, in 2000, n=5069) (total n=42396), the number with a record of weight and height (n=30811), and the number that were excluded for various reasons: due to duplicate entries in the same measurement year (n=317); because their BMI SDS was outside of a plausible range (> +/- 4.00) (n=322); postcode issues (i.e. postcode outside the study area, and no or incomplete postcode provided) (n=796). Whilst in eliminating BMI SDS outliers there is a slight risk that exceptionally sized children are omitted from the analysis, thus underestimating the childhood obesity problem, it is more important that the numerous incorrect data entries are excluded. The total number of child weight and height records deemed usable was 29376. A large proportion of the "clean" data were adjusted for errors in data entry (the most common being height entered in mm rather than cm).

Similarly the Trends and RADs datasets were also cleaned; but as these data were from recent, well-organised and run, studies there was much less cleaning required. On the other hand, there were fewer records to use. Table 4.2 shows the number of children measured as part of the Trends and RADs studies by age group (n=17,938), the number that were excluded for various reasons: due to no or incomplete postcode (n=5,959); living outside the study area (n=583); those with missing height, weight or date of birth data (n=7,493). The total number of child weight and height records deemed usable for this analysis was 10,458 (79% of useable records). All children measured by the RADs project in 2004 were excluded (not included in Table 4.2) because no date of birth data were collected. It should be remembered that these studies were not designed with this analysis in mind, which is why key data were missing in relatively large numbers. The response rates for both studies was high: the Trends project approached a sample of 35 primary and secondary schools and the RADs programme had participation from 22 (52%), 33 (79%) and 32 (76%) secondary schools in Leeds in 2004, 2005 and 2006

respectively. For both studies, the majority (over 95%) of pupils within the relevant classes in those schools participated.

3-6 year olds	Year of measurement						
	1998	1999	2000	2001	2002	2003	Total
No. of children recorded on the system (no duplicates)	3636	6268	5069	7832	8335	11256	42396
No. with record of weight & height	2324	3107	3779	6208	6333	9060	30811
% with weight & height data	63.9%	49.6%	74.6%	79.3%	76.0%	80.5%	72.7%
No. of records excluded:							
Duplicate entries	5	41	33	47	72	119	317
BMI SDS too low / high	42	67	35	37	63	78	322
Postcode outside study area	31	28	61	103	127	152	502
No or incomplete postcode	21	32	34	56	64	87	294
Total no. of records excluded	99	168	163	243	326	436	1435
% excluded	2.7%	2.7%	3.2%	3.1%	3.9%	3.9%	3.4%
Final no. of records	2225	2939	3616	5965	6007	8624	29376
% on system that are clean	95.7%	94.6%	95.7%	96.1%	94.9%	95.2%	95.3%

Table 4.1. Details of the 3-6 year old children with height and weight records on the NCHCS (routinely recorded data) by year of measurement, highlighting entries that were excluded.

	5 years	9 years	11 years	13 years	Total
No. of children measured	1120	1247	10830	904	14101
No with height, weight & age data	1120	1247	9892	904	13163
% with full data	100%	100%	91.3%	100%	93.3%
No. of records excluded:					
Postcode outside study area	28	39	512	4	583
No or incomplete postcode	28	19	1935	140	2122
Total no. of records excluded	56	58	2447	144	2705
% excluded	5.0%	4.7%	24.7%	15.9%	20.6%
No of usable records	1064	1189	7445	760	10458
% of useable records	95.0%	95.3%	75.3%	84.1%	79.4%

Table 4.2. Details of the children with height and weight records measured as part of the Trends project (5, 9, and 13 years old) in 2004 and 2005 and the RADS project (11 years old) in 2005 and 2006, highlighting entries that were excluded. 667 and 271 children measured in the RADs study did not have height/weight data or date of birth (respectively).

4.4.1 Routine Data Coverage

The routinely collected data's coverage numbers for 3 and 5 year olds (as these are the age groups targeted by the primary care trusts for measurement) were calculated by three different methods: normal, birth and census. The birth and census coverage figures are much lower than the "normal" calculation coverage values (see Table 4.3). For 3 year olds, normal coverage varied from 51% to 79% (mean 72%). However birth coverage was only in the range of 12% to 40% (mean 22%). For 5 year olds, normal coverage varied from 57% to 95% (mean 92%). Birth coverage had a wide range at 5% to 54% (mean 35%), although if the latter two years of data collection, where coverage suddenly fell to 5% (most likely due to the data for the current year not being input into the NCHCS yet), are excluded, the range for 5 year olds is a more robust 44% to 54% (mean 50%). The census and birth coverage figures are similar.

3 year olds	Year of birth							
	1995	1996	1997	1998	1999	2000	2001	Total
No. of children recorded on the system (no duplicate entries)	3530	2567	1175	1118	956	2541	995	12882
Total number of children born in the Study Area in this year	8860	8793	8431	8478	8270	7764	7949	58545
Total number of 3 year olds from the 2001 census	8405	8405	8405	8405	8405	8405	8405	58835
% of children born recorded on system	39.8	29.2	13.9	13.2	11.6	32.7	12.5	22.0
% of children on 2001 census recorded on system	42.0	30.5	14.0	13.3	11.4	30.2	11.8	21.9
No. of children on system with record of weight and height	2484	2039	703	568	583	2070	758	9205
% of children on system with record of weight and height	70.4	79.4	59.8	50.8	61.0	81.5	76.2	71.5
% of children born with records	28.0	23.2	8.3	6.7	7.0	26.7	9.5	15.7
% of children on census with records	29.6	24.3	8.4	6.8	6.9	24.6	9.0	15.6
5 year olds	**Year of birth**							
	1995	1996	1997	1998	1999	2000	2001	Total
No. of children recorded on the system (no duplicate entries)	3914	4211	4569	4447	386	362	na	17889
Total number of children born in the Study Area in this year	8860	8793	8431	8478	8270	7764		50596
Total number of 5 year olds from the 2001 census	8857	8857	8857	8857	8857	8857		53142
% of children born recorded on system	44.2	47.9	54.2	52.5	4.7	4.7		35.4
% of children on census recorded on system	44.2	47.5	51.6	50.2	4.4	4.1		33.7
No. with record of weight and height	3471	4012	4264	4224	221	328		16520
% of children on system with record of weight and height	88.7	95.3	93.3	95.0	57.3	90.6		92.3
% of children born with records	39.2	45.6	50.6	49.8	2.7	4.2		32.7
% of children on census with records	39.2	45.3	48.1	47.7	2.5	3.7		31.1

Table 4.3. Coverage of routinely measured data. The shaded row indicates the "normal" coverage calculation for routine data (i.e. that reported by other authors); this underestimates the number of missing children. Note, this table is presented by year of birth

4.4.2 Global Analysis

For the serial cross-sectional routinely measured data mean age, height, weight, BMI and BMI SDS at the beginning of the period were 3.1 years, 95.4cm, 15.0 kg, 16.4 kg m^{-2} and 0.231 respectively, compared with 4.7 years, 106.4 cm, 18.8 kg, 16.6 kg m^{-2} and 0.464 respectively at the end of the study period (1998-2003) (see Table 4.4). Similarly using the data from all three obesity datasets, the mean for age, height, weight, BMI and BMI SDS was calculated by age group; these statistics rise as the children get older (see Table 4.5).

3-6 year olds	Year of measurement					
	1998	1999	2000	2001	2002	2003
Final no. of records	2225	2939	3616	5965	6007	8624
% on system that are clean	95.7	94.6	95.7	96.1	94.9	95.2
Mean age (years)	3.1	3.5	4.7	5.1	5.1	4.7
Age range	2.6; 3.9	2.6; 5.0	2.7; 6.0	2.8; 6.4	2.7; 6.5	2.5; 6.5
Mean height (cm)	95.4	98.7	105.9	109.0	108.9	106.4
Height range	79; 114	74; 124	67; 130	. 78; 132	71; 150	81; 141
Mean weight (kg)	15.0	16.0	18.4	19.5	19.6	18.8
Weight range	10; 26	10; 34	8; 36	10; 42	9; 37	10; 47
Mean BMI (Kg/m2)	16.4	16.4	16.4	16.4	16.5	16.6
BMI range	12; 23	12; 25	12; 25	12; 27	11; 25	12; 27
Mean BMI SDS	0.231	0.273	0.383	0.429	0.525	0.464
BMI SDS range	-3.8; 4.0	-3.9; 4.0	3.8; 4.0	-3.9; 4.0	-3.7; 3.9	-3.9; 4.0

Table 4.4. Summary of cleaned routinely recorded data with gender, age, height and weight data by year of measurement for 3-6 year olds in Leeds.

	3 years (RMD)	5 years (RMD)	5 years (Trends)	9 years (Trends)	11 years (RADs)	13 years (Trends)
N	7881	15257	1064	1189	7445	760
Mean age (years)	3.1	5.1	5.3	9.3	11.6	13.3
Age range	2.5; 3.5	4.6; 5.5	4.5; 6.4	7.4; 10.4	10.1; 12.8	12.5; 14.4
Mean height (cm)	95.5	109.1	111.9	135.2	148.3	158.0
Height range	67; 149	83; 130	91; 140	108; 158	120; 178	132; 186
Mean weight (kg)	15.0	19.6	20.2	32.4	42.9	52.5
Weight range	8; 37	11; 40	12; 39	18; 90	22.2; 97.7	28; 118
Mean BMI (Kg/m2)	16.4	16.4	16.1	17.6	19.3	20.8
BMI range	12; 24	11; 25	11; 26	10; 36	11; 45	14; 45
Mean BMI SDS	0.236	0.501	0.258	0.394	0.473	0.539
BMI SDS range	-3.95; 3.95	-3.93; 3.99	-4.77; 4.48	-5.68; 3.95	-6.16; 3.96	-3.09; 4.12

Table 4.5. Summary of cleaned obesity datasets with gender, age, height and weight data by age groups for children in Leeds. Note, the routinely measured data for 3 and 5 year olds relates to measurements in 2000-2003; the Trends data for 5, 9 and 13 year olds was measured in 2004 and 2005; the RADs data for 11 year olds was measured in 2005 and 2006.

The childhood obesity data were analysed over time (serial cross-sectional analyses, not longitudinal). The proportions of children who were in the Acceptable, Overweight (but not obese) and Obese ranges of BMI according to the British reference dataset age and gender specific cut-offs at the start of the period (1998) were 84.8%, 10.7% and 4.5%, which by the end of the period (2003) had changed to 82.0%, 11.4% and 6.6% respectively (see Figure 4.2). Logistic regression analysis showed that the proportion of obese 3-6 year old children rose significantly between 1998 and 2003 (see Table 4.6); as did the proportion overweight (R^2 = 0.001 (Nagekerke); λ^2 (5) = 18.8; p = 0.002) and the proportion overweight and obese (R^2 = 0.002 (Nagekerke); λ^2 (5) = 34.4; p < 0.001) (full results not shown). Children were 1.5x more likely to be obese in 2003 than 1998; similarly overweight was 1.1x more likely and overweight and obese was 1.2x more likely.

When the data are considered by gender, logistical regression analysis shows significant rising trends for all categories except overweight boys (see Table 4.7 for full results for obese data; see Table 4.8 for summary results for all three models). Whilst the proportion of obese girls was lower than that for boys across all years, this difference was not significant (p = 0.056) (2-tailed 2 sample equal variance t-test); although this same proportion difference was significantly lower for overweight (p<0.001) and overweight and obese (p<0.001). Girls were 1.8x more likely to be obese in 2003 than 1998; similarly overweight in girls was 1.1x more likely, and overweight and obese was 1.3x more likely. The corresponding figures for boys were 1.3x, 1.0x and 1.1x respectively.

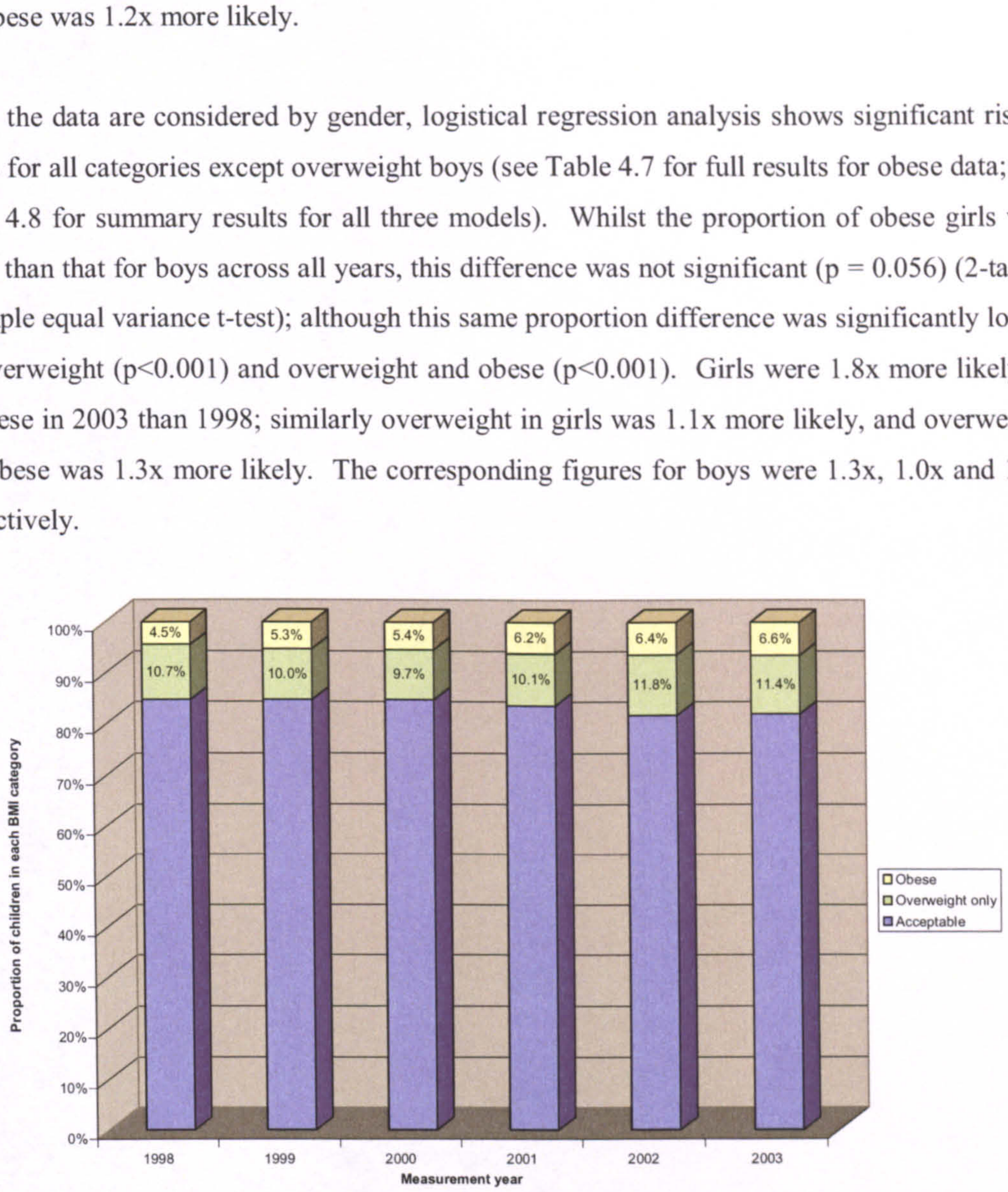


Figure 4.2. Proportion of children with Acceptable, Overweight (but not obese) or Obese BMI SDS' for 3-6 year olds children routinely measured between 1998 and 2003, showing an increase in childhood obesity over time.

Year Measured	N	B	S.E.	Wald	df	Sig.	Exp(B)	95%CI lower	95%CI upper
1998	2225			20.839	5	.001			
1999	2939	-.406	.111	13.357	1	.000	.666	.536	.828
2000	3616	-.238	.093	6.519	1	.011	.788	.657	.946
2001	5965	-.209	.085	6.010	1	.014	.811	.686	.959
2002	6007	-.072	.069	1.077	1	.299	.931	.813	1.066
2003	8624	-.037	.068	.287	1	.592	.964	.843	1.102
Constant	na	-2.650	.043	3732.638	1	.000	.071	na	na

Table 4.6. Summary of logistic regression results considering whether children are obese (or not) over time (serial cross-sectional analysis from 1998 to 2003). $R^2 = 0.002$ (Nagekerke); λ^2 (5) = 21.9; p = 0.001

Boys: year Measured	N	B	S.E.	Wald	df	Sig.	Exp(B)	95%CI lower	95%CI upper
1998	1116			17.810	5	.003			
1999	1500	-.269	.145	3.422	1	.064	.764	.575	1.016
2000	1814	-.317	.131	5.872	1	.015	.728	.563	.941
2001	3136	-.389	.125	9.733	1	.002	.678	.531	.865
2002	3125	-.071	.093	.581	1	.446	.931	.775	1.118
2003	4481	.019	.091	.042	1	.837	1.019	.852	1.219
Constant	na	-2.599	.059	1949.632	1	.000	.074	na	na
Girls: year Measured	**N**	**B**	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95%CI lower**	**95%CI upper**
1998	1109			11.789	5	.038			
1999	1439	-.578	.173	11.109	1	.001	.561	.399	.788
2000	1802	-.151	.133	1.298	1	.255	.859	.662	1.115
2001	2829	-.035	.118	.089	1	.766	.965	.766	1.216
2002	2882	-.074	.103	.517	1	.472	.929	.759	1.136
2003	4143	-.106	.103	1.067	1	.302	.899	.735	1.100
Constant	na	-2.708	.064	1780.307	1	.000	.067	na	na

Table 4.7. Summary of logistic regression results considering whether boys or girls are obese (or not) over time (serial cross-sectional analysis from 1998 to 2003)

Boys	R^2 (Nagekerke)	df	λ^2	p
Obese	0.003	5	18.7	0.002
Overweight	0.001	5	8.1	0.152
Obese & overweight	0.002	5	21.7	0.001
Girls	**R^2 (Nagekerke)**	**df**	**λ^2**	**p**
Obese	0.003	5	13.3	0.021
Overweight	0.002	5	13.4	0.020
Obese & overweight	0.002	5	16.4	0.006

Table 4.8. Summary of key statistics from logistic regression for whether boys or girls are obese and/or overweight (or not) over time.

The childhood obesity data were also analysed by age group of the children measured. The proportions of children who were in the Acceptable, Overweight (but not obese) and Obese ranges of BMI at age 3 were 85.2%, 9.9% and 4.8%, changing for children aged 13 to 72.4%, 14.3% and 13.3% respectively (see Figure 4.3). Logistical regression analysis showed that the proportion of obese children was significantly higher as the children got older (see Table 4.9); similarly for the proportion overweight ($R^2 = 0.003$ (Nagekerke); λ^2 (4) = 50.9; $p < 0.001$) and the proportion overweight and obese ($R^2 = 0.017$ (Nagekerke); λ^2 (4) = 349.0; $p < 0.001$) (full results not shown). Older children were 3.0x more likely to be obese; similarly overweight was 1.5x more likely in older children and overweight and obese was 2.2x more likely.

When the data are considered by gender, logistical regression analysis shows significant rising trends with age for all categories (see Table 4.10 for full results for obese data; see Table 4.11 for summary results for all three models). The proportion of obese girls was significantly lower than that for boys across all age groups ($p = 0.002$) (2-tailed 2 sample equal variance t-test); similarly for overweight girls ($p = 0.012$) and overweight and obese girls ($p = 0.001$). Girls were 2.7x more likely to be obese at age 13 than at age 3; similarly overweight in girls was 1.8x more likely, and overweight and obese 2.3x more likely. The corresponding figures for boys were 3.3x, 1.3x and 2.1x respectively.

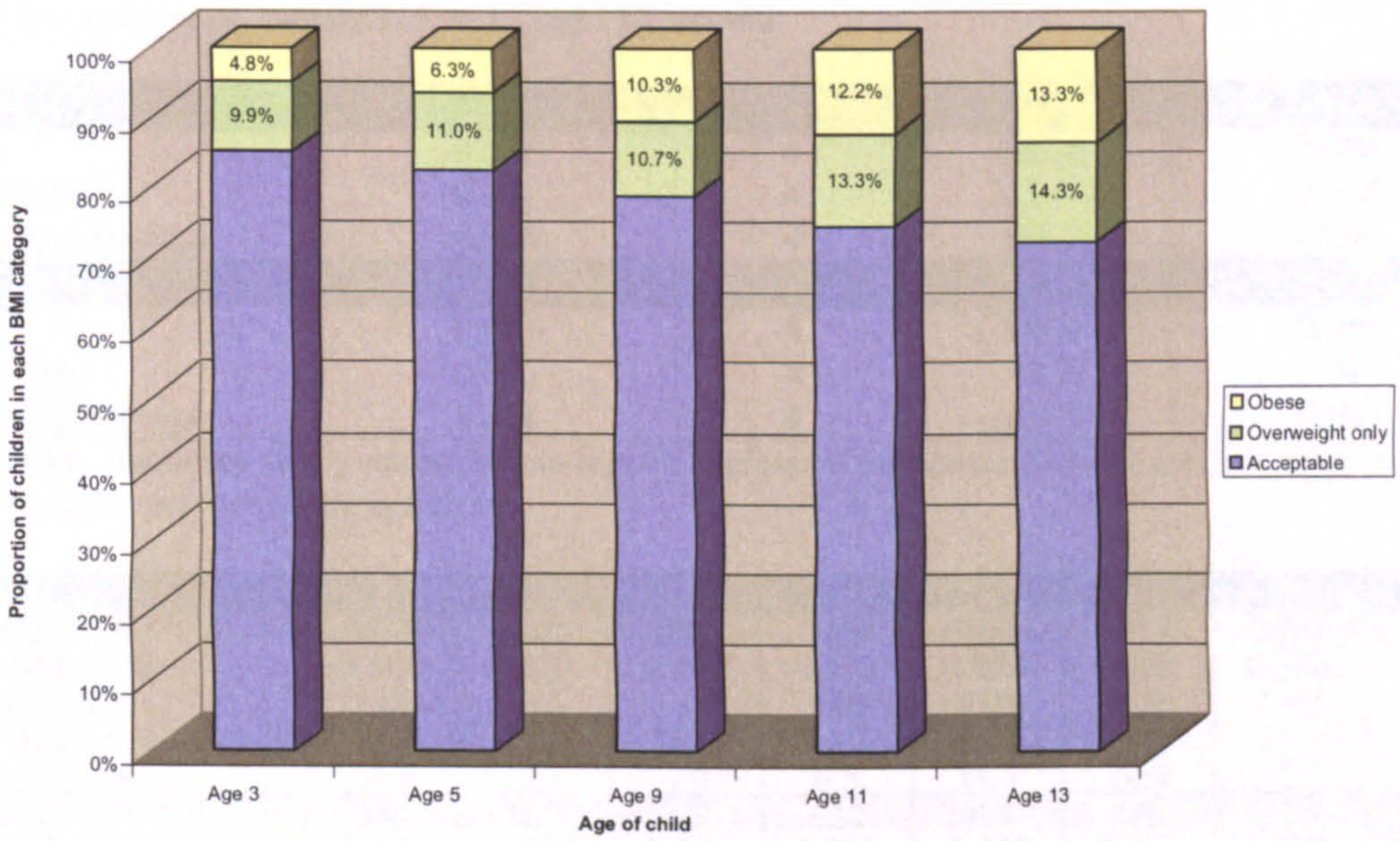


Figure 4.3. Proportion of children with Acceptable, Overweight (but not obese) or Obese BMI SDSs by age groups for children from all three obesity datasets. The routinely measured data for 3 and 5 year olds was measured in 2000-03; the Trends project data for 5, 9 and 13 year olds was measured in 2004-05; the RADs project data for 11 year olds was measured in 2005-06. This shows an increase in childhood obesity with age.

Age Group	N	B	S.E.	Wald	df	Sig.	Exp(B)	95%CI lower	95%CI upper
3 years	7881			376.413	4	.000			
5 years	16321	-1.104	.119	86.015	1	.000	.331	.262	.419
9 years	1189	-.822	.112	54.265	1	.000	.440	.353	.547
11years	7443	-.284	.143	3.934	1	.047	.753	.569	.997
13 years	760	-.102	.113	.818	1	.366	.903	.724	1.126
Constant	na	-1.876	.107	308.088	1	.000	.153	na	na

Table 4.9. Summary of logistic regression analysis considering whether children are obese (or not) by the five age groups, namely 3, 5, 9, 11 and 13 years old. $R^2 = 0.026$ (Nagekerke); λ^2 (4) = 368.0; $p < 0.001$

Boys Age Group	N	B	S.E.	Wald	df	Sig.	Exp(B)	95%CI lower	95%CI upper
3 years	4093			228.922	4	.000			
5 years	8329	-1.207	.162	55.570	1	.000	.299	.218	.411
9 years	656	-.830	.151	30.291	1	.000	.436	.324	.586
11years	3637	-.191	.187	1.043	1	.307	.826	.572	1.192
13 years	389	-.101	.153	.437	1	.509	.904	.671	1.219
Constant	na	-1.783	.144	152.364	1	.000	.168	na	na
Girls Age Group	**N**	**B**	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95%CI lower**	**95%CI upper**
3 years	3788			154.921	4	.000			
5 years	7992	-.989	.176	31.576	1	.000	.372	.263	.525
9 years	533	-.813	.166	23.964	1	.000	.443	.320	.614
11years	3806	-.453	.225	4.056	1	.044	.636	.409	.988
13 years	371	-.094	.167	.314	1	.575	.911	.656	1.264
Constant	na	-1.980	.159	155.056	1	.000	.138	na	na

Table 4.10. Summary of logistic regression results considering whether boys or girls are obese (or not) by the five age groups, namely 3, 5, 9, 11 and 13 years old.

Boys	R^2 (Nagekerke)	df	λ^2	p
Obese	0.031	4	227.2	0.000
Overweight	0.001	4	12.8	0.012
Obese & overweight	0.016	4	169.2	0.000
Girls	**R^2 (Nagekerke)**	**df**	**λ^2**	**p**
Obese	0.023	4	149.0	0.000
Overweight	0.006	4	45.0	0.000
Obese & overweight	0.018	4	185.9	0.000

Table 4.11. Summary of key statistics from logistic regression for whether boys or girls are obese and/or overweight (or not) across the age groups

	1	2	3	4	5	6	7	Total
No.children	6637	1002	607	7989	5655	7172	4532	33594
Mean BMI SDS	0.488	0.379	0.492	0.392	0.478	0.423	0.366	0.422
SD BMI SDS	1.08	1.15	1.00	1.06	1.09	1.06	1.19	1.09
Mean deprivation	38.8	21.9	12.7	12.3	46.1	19.5	51.4	30.3
SD deprivation	16.9	12.2	4.9	6.8	14.7	10.4	17.7	20.0
No. OAs	349	167	62	528	436	570	270	2382
Mean OA BMI SDS	0.473	0.389	0.490	0.398	0.480	0.419	0.381	0.429
SD OA BMI SDS	0.36	0.80	0.48	0.34	0.43	0.39	0.45	0.44
Mean OA deprivation	35.3	23.5	12.2	12.3	43.5	19.6	49.0	28.1
Range OA deprivation	6; 77	5; 73	4; 26	2; 53	7; 77	2; 67	9; 79	2; 79
SD OA deprivation	16.1	12.2	5.0	6.9	15.1	10.4	17.2	18.5

Table 4.12. Summary of the 7 OAC super groups by mean BMI SDS for each child. Standard deviation and the number of children in each sub-dataset are included. Also provided is a summary of the data by OA, in order that locational mean BMI SDS and deprivation can be compared to that of the study population.

A global analysis of the covariates of interest (namely, deprivation and OAC super groups) was also undertaken. The average deprivation score for the study area based on the deprivation score for the children measured was 30.3, and based on geography (i.e. taking each OA only once) was 28.1. A simple linear regression of deprivation score for each SOA against each SOA's mean BMI SDS shows a positive relationship; with higher deprivation associated with a higher mean BMI SDS (see Figure 4.4). Considering proportions of children who are overweight and/or obese against deprivation shows that children in the most deprived areas are 2.1x more likely to be obese (1.8x for overweight and obese) than those living in the most affluent areas (deprivation score < 30). In relation to the OAC super groups and obesity, if the BMI SDS is summarised by super group (see Table 4.12) it is clear to see that OAC's 2 and 7 (city living and multicultural) have the most variability (in terms of BMI SDS). Also OAC's 1 (blue collar communities), 3 (countryside), and 5 (constrained by circumstances) have above average BMI SDS, with OACs 2, 4 (prospering suburbs), and 7 having below average. These results are not simply tied to deprivation. For example, OAC super group 7, with low obesity levels, are largely found in the city centre, which generally has higher deprivation (average locational deprivation score 49.0; mean child deprivation score 51.4). Conversely OAC super group 3, with high obesity levels, are largely found in rural areas, which are generally affluent areas (average locational deprivation score 12.2; mean child deprivation score 12.7).

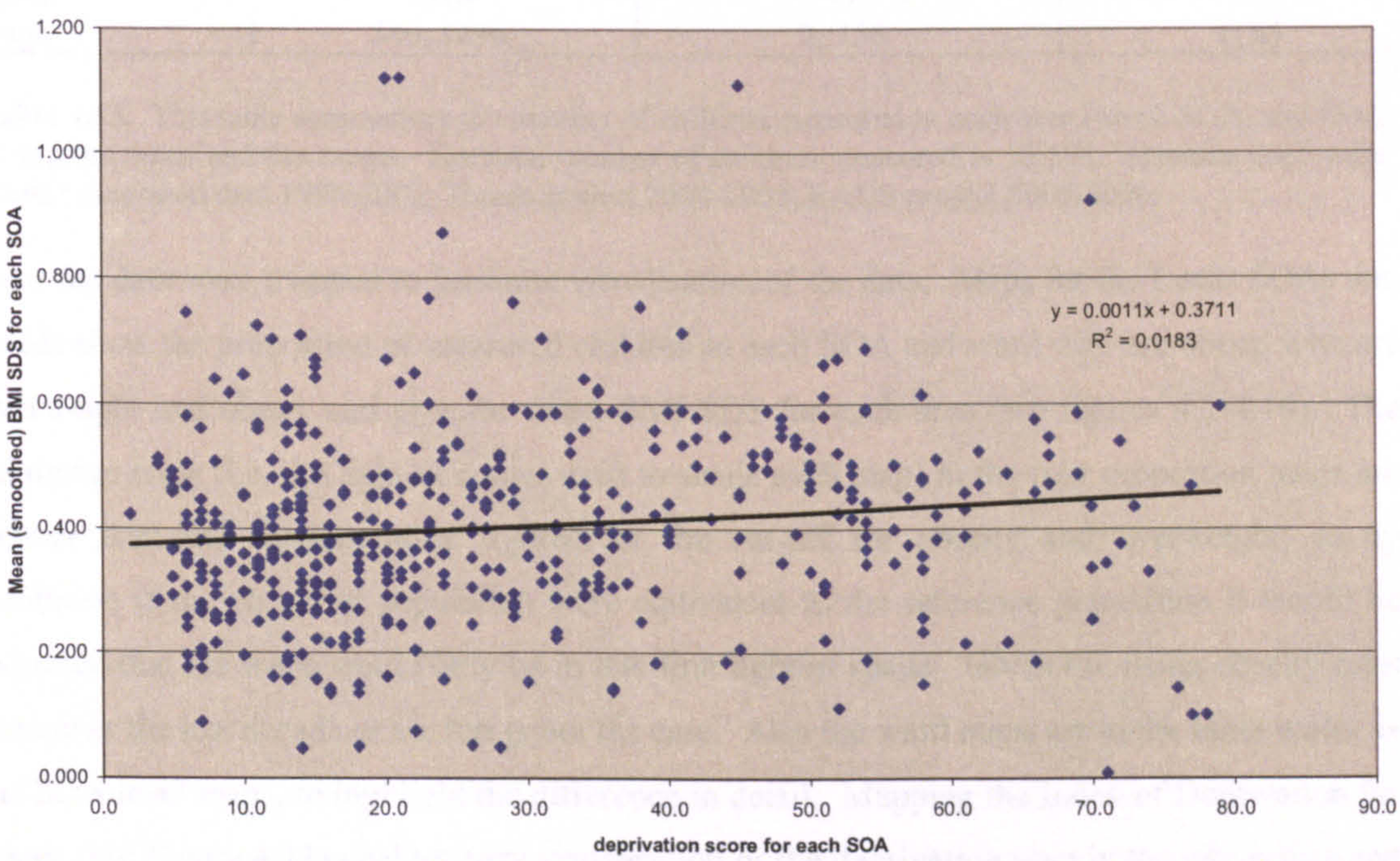


Figure 4.4. Scatter plot of deprivation score for each SOA against mean (smoothed) BMI SDS for each SOA. Linear regression line shows positive relationship.

4.4.3 Spatial Analysis

All three BMI datasets (i.e. PCT 1998-03; Trends 2004-05; RADs 2005-06) were combined and summarised by home (residential SOA, OA & ward) locations. Calculations were undertaken for mean BMI SDS as well as number of children categorised as Obese, Overweight (and not obese) and Acceptable. These data were smoothed using Empirical Bayes (EB) with a population limit set at equal to the mean for wards and twice the mean for OAs and SOAs. These limits were selected based on the range of children measured in each area (see Table 4.13). The ward level data are fine; EB was used to smooth the data and unsurprisingly had little impact. The SOA level data had children measured in all 476 areas. Few areas had less than 25 children measured and the average number measured per area is above 50; accordingly EB was sufficient to resolve these small number problems. The OA level data had 57 OAs within the study area with no obesity data (mostly in the north east part of the study area) and many areas with very few numbers of children measured. This suggests that this scale is too small (without a much larger dataset) to do analysis. Nevertheless the data excluding these "zero" areas were smoothed using EB, as it is necessary to use OAs to consider the OAC super group variables.

	Ward	SOA	OA
N	33	476	2382
Mean	1018	71	14
range	340; 1496	2; 156	1; 65

Table 4.13. This table summarises the number of children measured in each area (ward, SOA, and OA), giving the mean and the range. The total number of children measured is 33,594. Datasets used were routine measured data 1998-2003; Trends project 2004-2005; RADS project 2005-2006.

Next the data were mapped to facilitate visualisation of the data. Maps for the Leeds SOAs and wards show the proportion of measured children in each SOA and ward who are obese, who are overweight and obese, and give the mean BMI SDS for each area (see figures 4.5-4.10). The minimum scale (i.e. the lightest colour used to shade each map) in the four proportion maps are greater than the corresponding figures for the cut-off for obesity and overweight; so by definition if the observed population were equivalent to the reference population it would be expected that the maps would only be in this first lightest shade. Given the rising obesity rates seen over the last decade or so, this is not the case. Also the ward maps are at the same scales as the SOA level maps, to highlight the difference in detail. Mapping the Index of Deprivation for Leeds (see Figure 4.11) enables easy visualisation of the deprivation seen in the city centre and the relatively wealthy rural surroundings. Similarly a map of the Area Classification Super Groups (Figure 4.12) gives an indication of the homogeneity and heterogeneity within and between neighbourhoods (respectively) in Leeds. These two maps quickly give an overview of the population composition across Leeds.

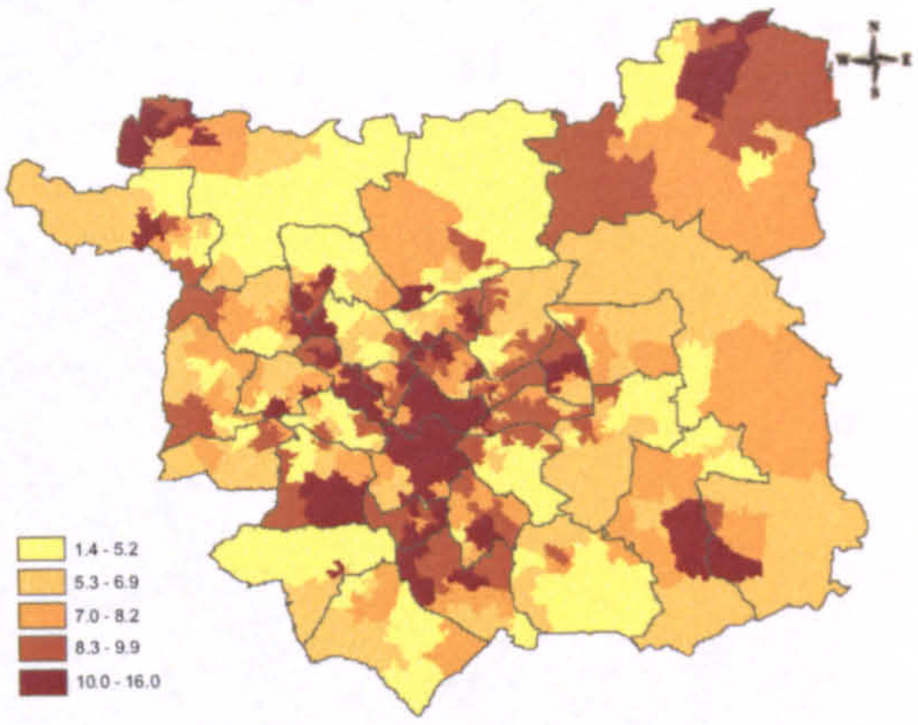


Figure 4.5. Map of proportion of measured children in each SOA who are obese (EB smoothed; quintile scale). The darker the shading the higher the proportion. The lines represent the ward boundaries.

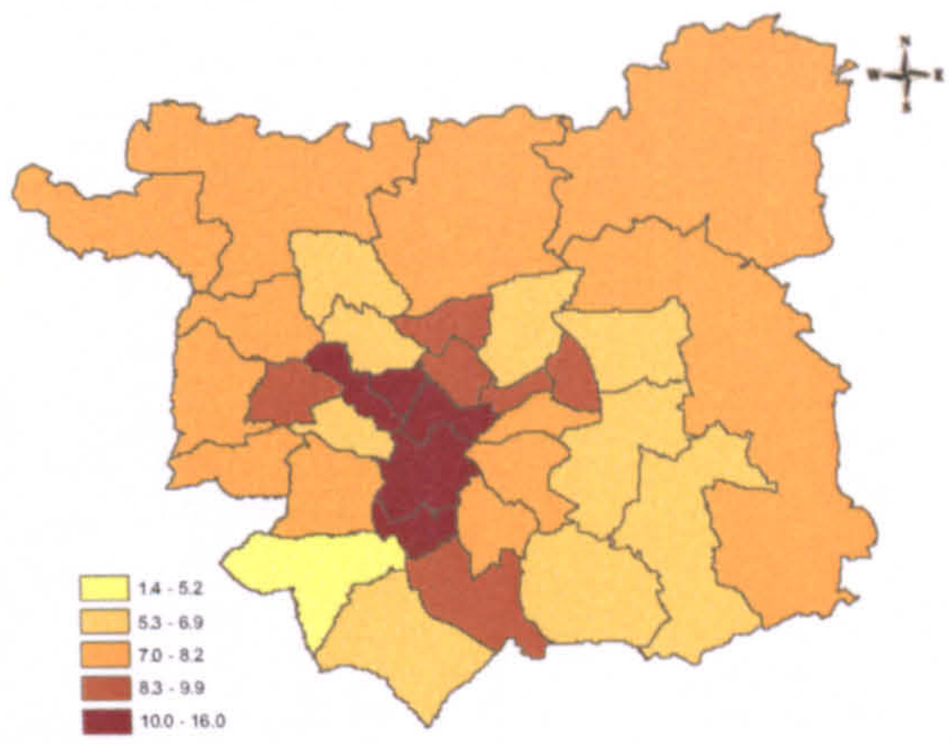


Figure 4.6. Map of proportion of measured children in each ward who are obese (EB smoothed; same scale as the SOA level map). The darker the shading the higher the proportion.

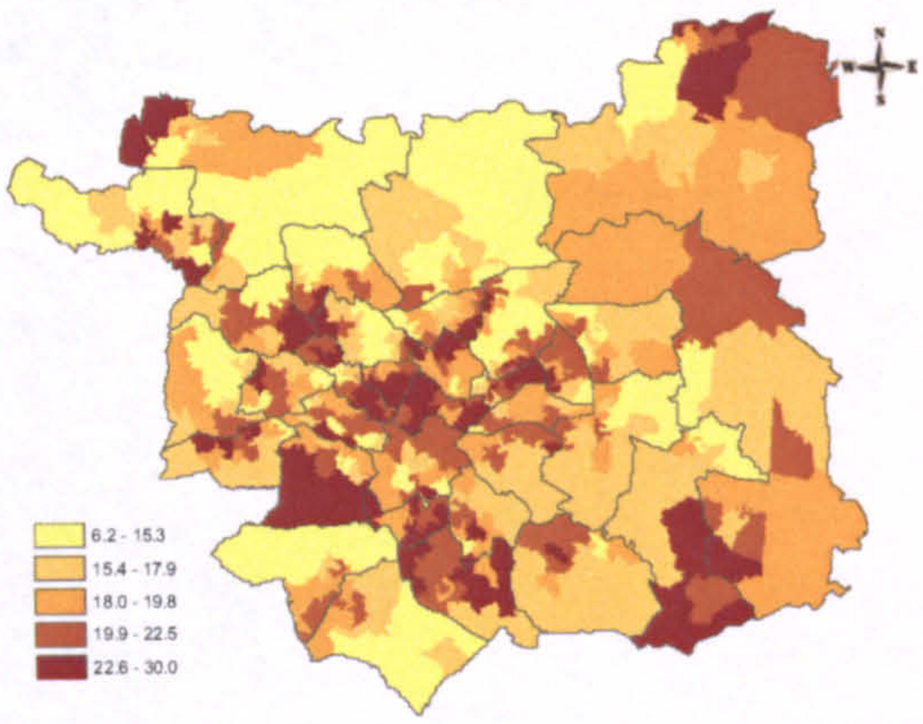


Figure 4.7. Map of proportion of measured children in each SOA who are overweight and obese (EB smoothed; quintile scale). The darker the shading the higher the proportion. The lines represent the ward boundaries.

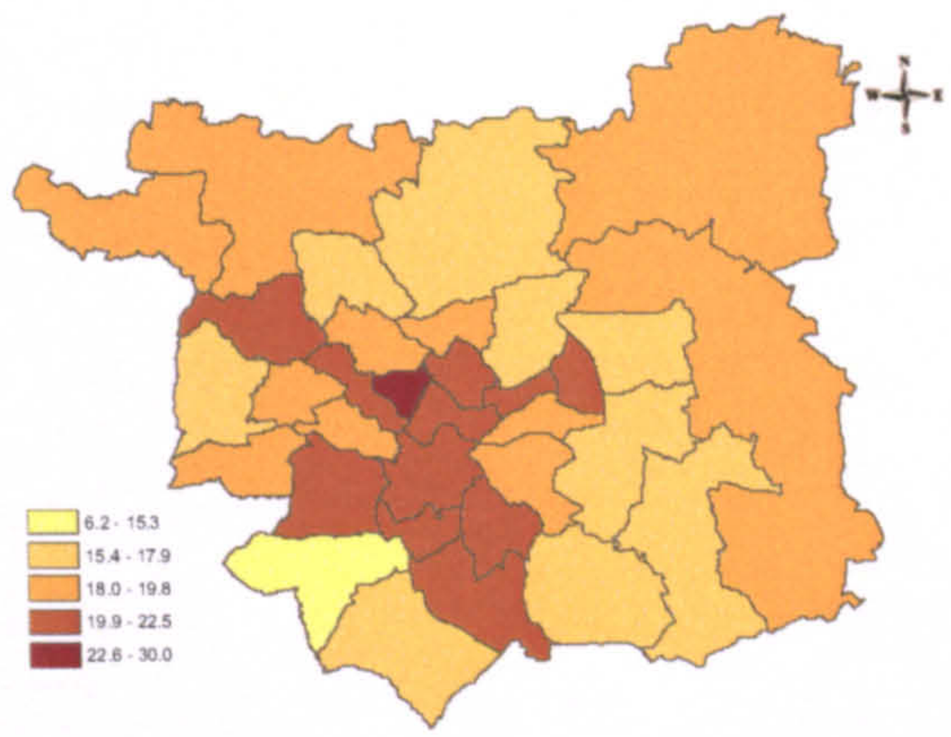


Figure 4.8. Map of proportion of measured children in each ward who are overweight and obese (EB smoothed; same scale as the SOA level map). The darker the shading the higher the proportion.

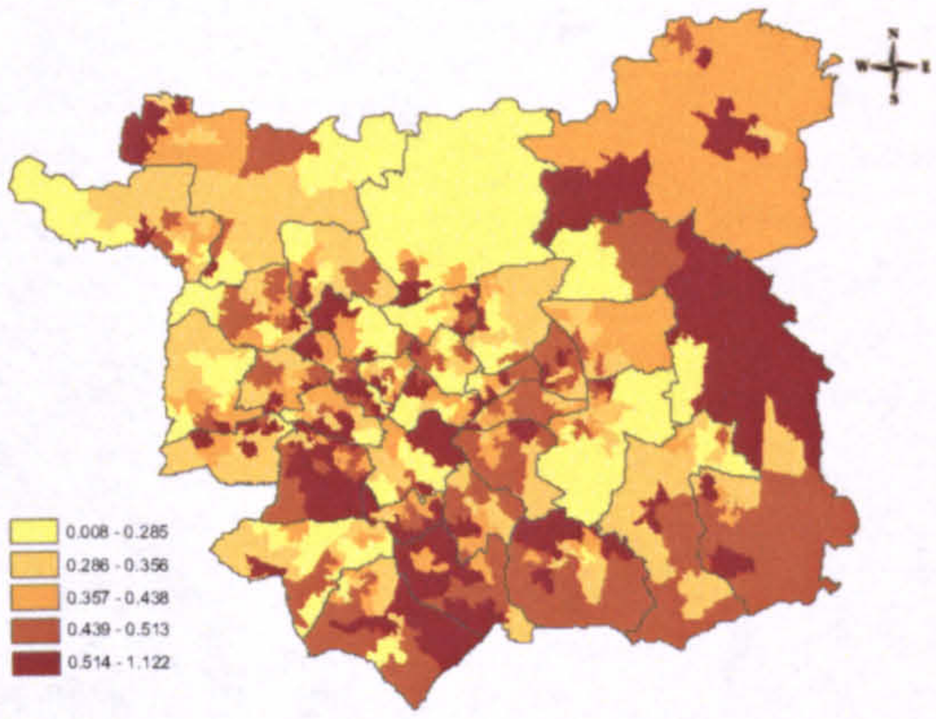


Figure 4.9. Map of mean BMI SDS in each SOA (EB smoothed; quintile scale). The darker the shading the higher the mean. The lines represent the ward boundaries.

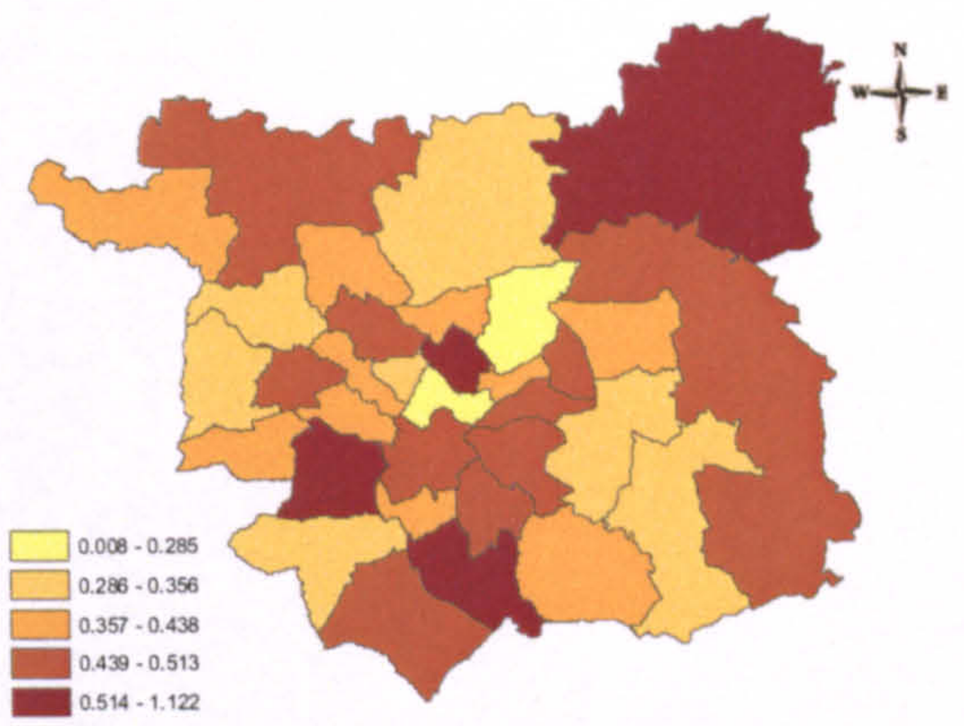


Figure 4.10. Map of mean BMI SDS in each ward (EB smoothed; same scale as the SOA level map). The darker the shading the higher the mean.

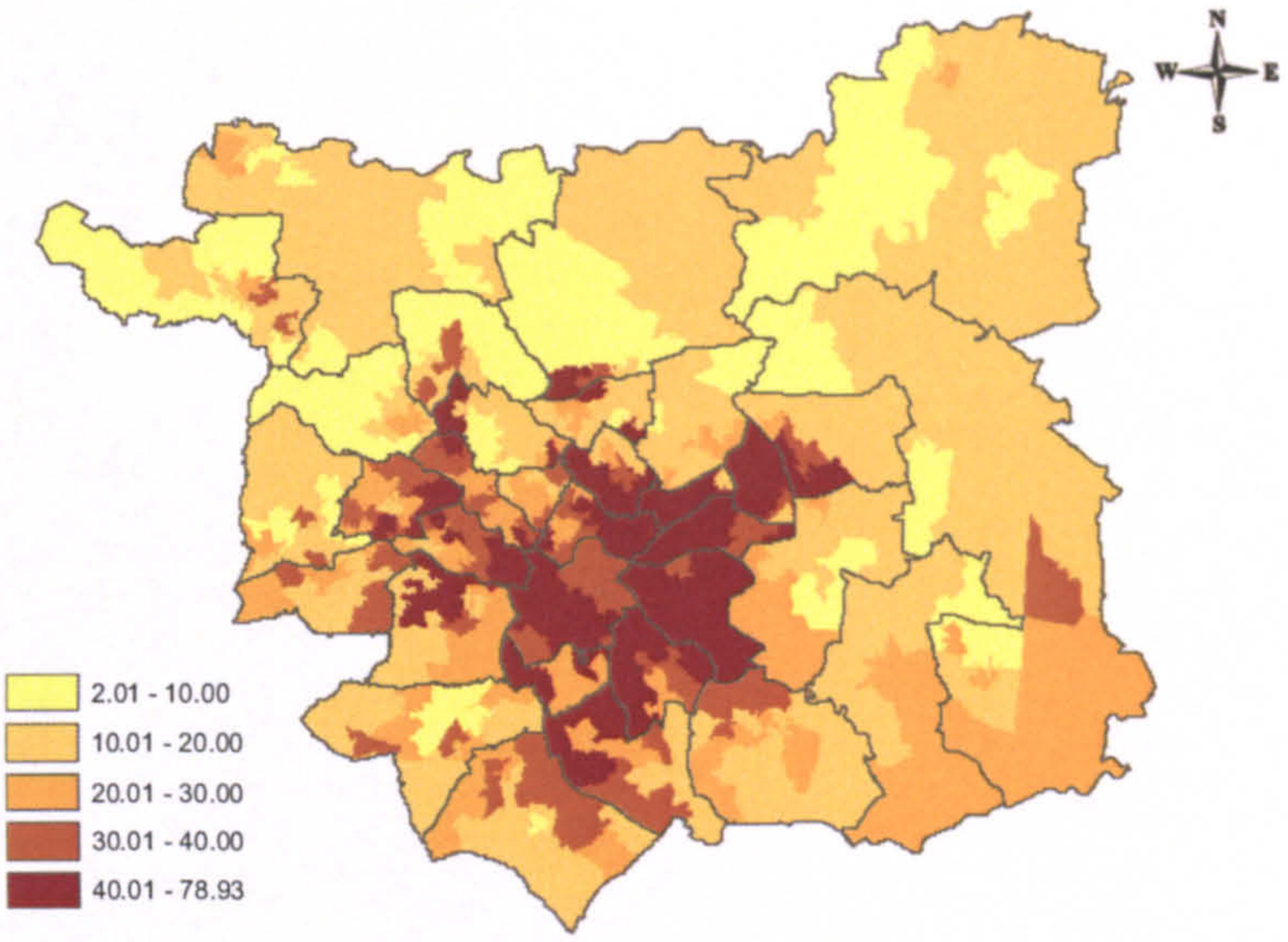


Figure 4.11. Map of the Index of Multiple Deprivation (2004) across Leeds by SOA (manual scale). The darker the shading the higher the deprivation. The lines represent the ward boundaries.

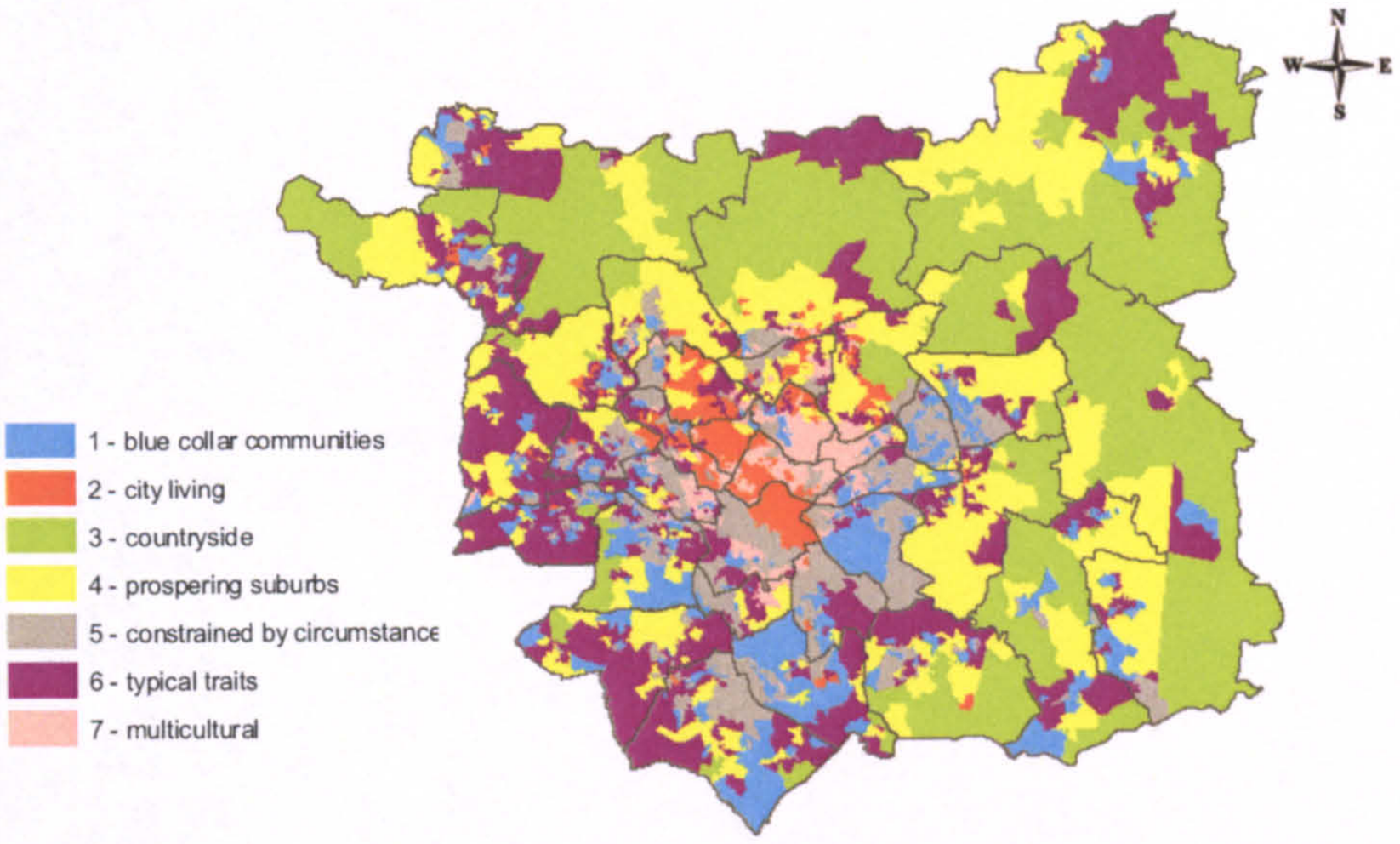


Figure 4.12. Map of the Area Classifications for Leeds (which are currently only available at output area level). Each colour represents a different OAC super group. The lines represent the ward boundaries.

The maps of the childhood obesity data show that some areas have higher prevalence of obesity and/or overweight (or mean BMI SDS) than others. In order to determine whether these differences are significant a spatial scan statistic, SaTScan, was used. This model permits the identification of any statistically significant hot spots of childhood obesity. Three separate SaTScan models were run: 1. Bernoulli obese (or not); 2. Bernoulli overweight and obese (or not); 3. Normal mean BMI SDS. These models identified several significant clusters, which are shown in Figures 4.13-4.15. Children living in Cluster 1 were 1.5x more likely to be obese than children outside of the obese clusters (p =0.001), and those in Cluster 2 were 6.1x more likely (p = 0.40). Children in Cluster 3 were 1.2x more likely to be obese or overweight than children

elsewhere (p = 0.002), and in Cluster 4 were 2.2x more likely (p = 0.018). Children in Clusters 5-15 had a significantly higher mean BMI SDS than children outside of these clusters (p=0.001, p=0.001, p=0.001, p=0.017, p=0.017, p=0.017, p=0.038, p=0.038, p=0.040, p=0.040, p=0.041 respectively). Clusters 1, 2 and 3 are large clusters in approximately the same location (i.e. central Leeds) as the smaller clusters 6, 7, 8, 12, 14 and 15. Cluster 9 is in a similar location, albeit shifted slightly south-east. This leaves cluster 4, which is for overweight and obesity; cluster 13, which is for mean BMI SDS, in the north west corner of the study area; and finally, clusters 10, 5 and 11 in the west of the study area. This is summarised in Figure 4.16.

After identifying significant hot spots of childhood obesity in Leeds, the next step was to take the census derived indices, namely deprivation and OAC super groups, into account, to see if these covariates were impacting the data. In order to be able to separate out the effects at individual and area level, a multi-level model was constructed (see Figure 4.17). Deprivation did not have a significant effect on the fit of the model but OAC super groups did, at least for OAC 7 (multicultural). Whilst only about 1% of the total variance in BMISDS may be attributed to differences between SOAs, there was nevertheless significant variation between SOAs (σ^2_{u0}), even after adjusting for super groups and deprivation ($p<0.001$). Accordingly most of the variation in the data was at child level (σ^2_e). Analysis of the residuals (see Figure 4.18) identified four hot spots of childhood obesity (where mean BMI SDS in the area was higher than expected after taking deprivation and OAC super group into account) and five cold spots (where mean BMI SDS was lower than expected).

The location of these hot and cold spots is shown in Figure 4.19 and their characteristics are summarised in Table 4.14. The hot spots were all in affluent areas (maximum deprivation 28.9). Only OAC super groups 2 and 3 (city living and countryside) were not represented in hot spots. The cold spots spanned both affluent and deprived areas; the most deprived being Cold Spots 4 and 5 with deprivation scores of 71 and 64 respectively. OAC group 7 was represented in both affluent and deprived cold spots. However OAC groups 3, 4 and 6 (countryside, prospering suburbs and typical traits, respectively) were in the affluent cold spots and OAC group 1 and 5 (blue collar communities and constrained by circumstances, respectively) were in the deprived cold spots. OAC group 2 was not represented in the cold spots. The average age was similar across all groups, with the same age range of children measured, from all possible measurement years, with the proportion of boys ranging from 41-57%.

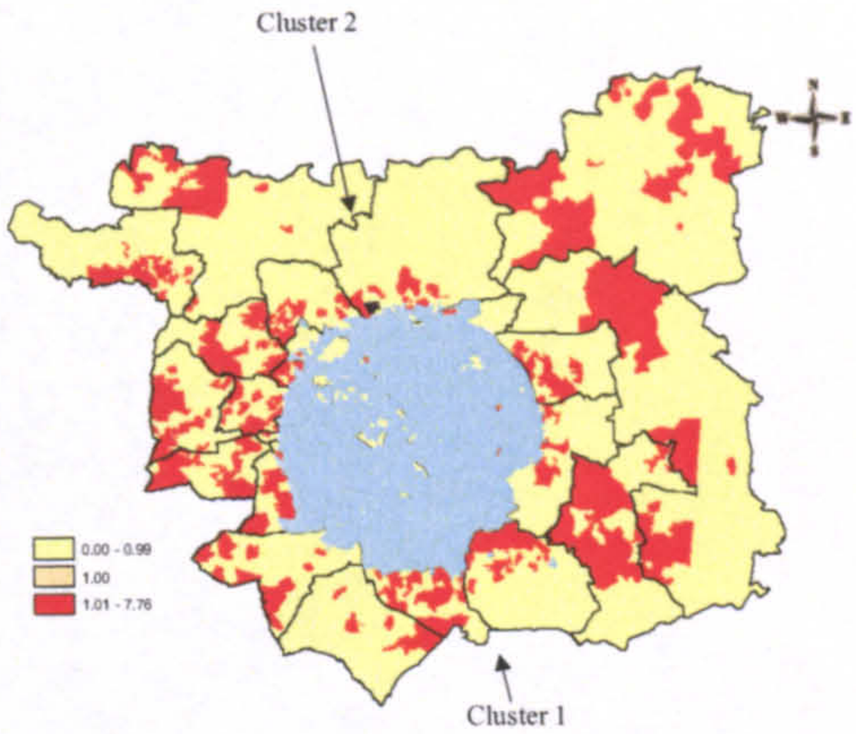


Figure 4.13. Map of the relative risk of obesity across OAs in Leeds. A manual scale was used to highlight those areas with a relative risk of greater than or less than 1 (red shading indicates the relative risk is greater than one). The two significant "hot spots" are highlighted Bernoulli SaTScan analysis): Cluster 1 is in blue (p=0.001) and Cluster 2 (just north of Cluster 1 – it's very small) is in purple (p=0.040). Children in these hot spots were 1.5x and 6.1x (Cluster 1 and 2, respectively) more likely to be obese than children living in other parts of the study area.

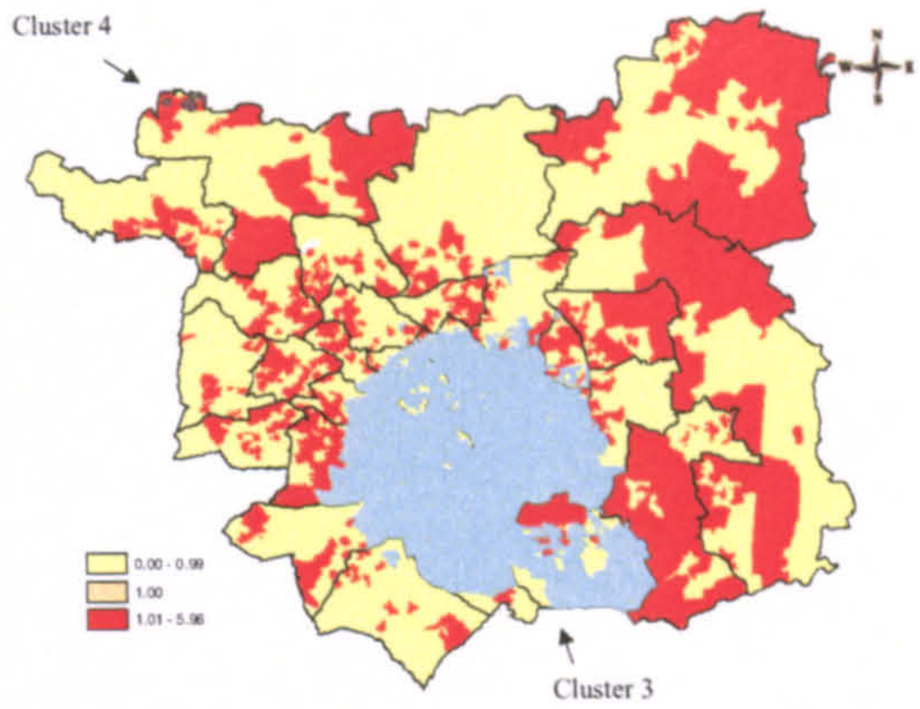


Figure 4.14. Map of the relative risk of overweight and obesity across OAs in Leeds. A manual scale was used to highlight those areas with a relative risk of greater than or less than 1 (red shading indicates the relative risk is greater than one). The two significant "hot spots" are highlighted (Bernoulli SaTScan analysis): Cluster 3 is in blue (p=0.002) and Cluster 4 (small cluster in top north west corner) is in purple (p=0.018). Children in these hot spots were 1.2x and 2.2x (Cluster 3 and 4, respectively) more likely to be overweight and obese than children living in other parts of the study area.

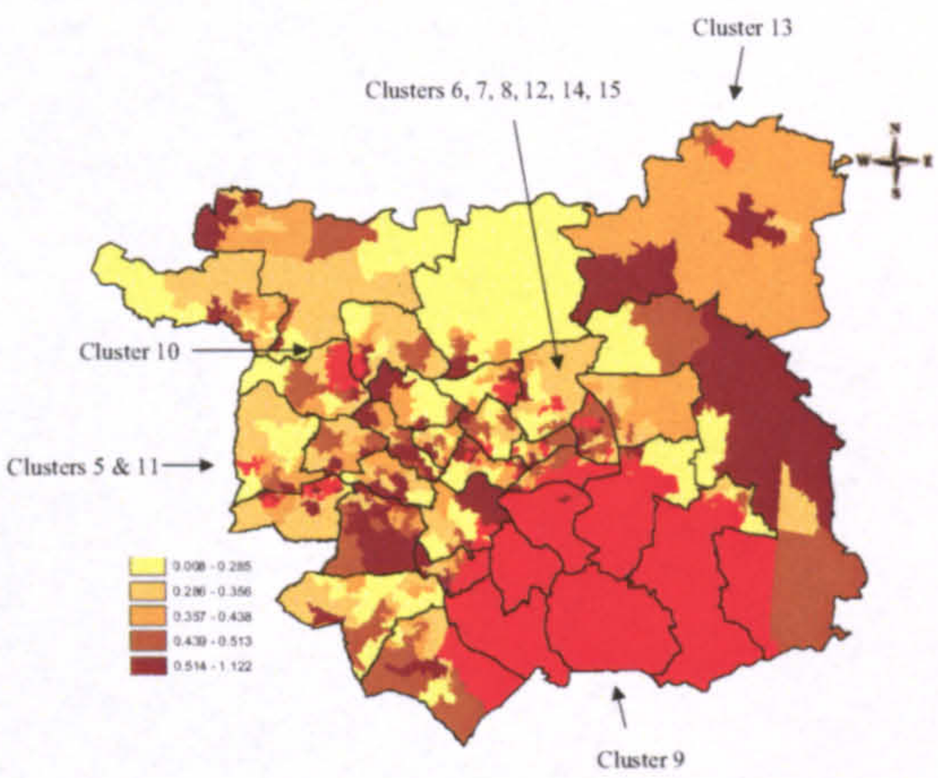


Figure 4.15. Map of mean BMI SDS (smoothed) across SOAs in Leeds (i.e. same background as Figure 5). 15 significant hot spots are highlighted (Normal SaTScan analysis): Cluster 9 is the obvious large cluster in the south of the study area (p=0.017); clusters 10 and 13 are the next largest (p=0.017, p=0.040 respectively), with many of the remaining clusters only representing a single OA (all p values < 0.05).

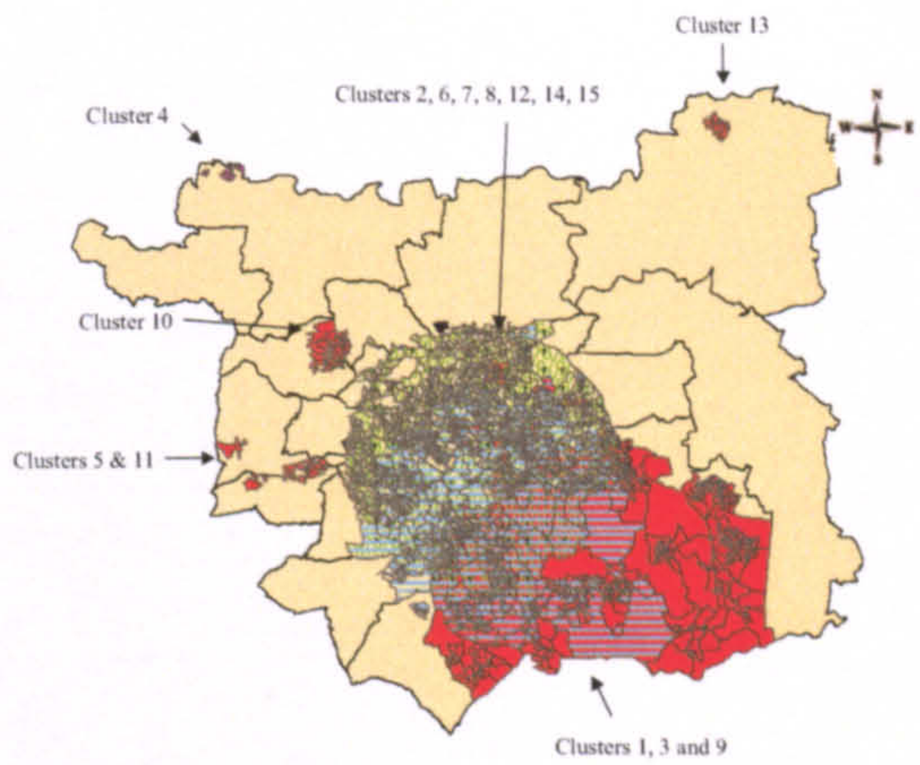


Figure 4.16. Map showing all the hot spots against a blank background (SOA boundaries). Note cluster 1 is now in green pattern and cluster 3 is stripy blue (both were solid blue colour). Highlights the overlap between clusters 1, 2, 3, 6, 7, 8, 9, 12, 14 and 15. Clusters 4, 5, 10, 11 and 13 stand out separately. The lines represent the ward boundaries.

$$y_{ij} = \beta_{0j} + \beta_1 OAC1_{ij} + \beta_2 OAC2_{ij} + \beta_3 OAC3_{ij} + \beta_4 OAC4_{ij} + \beta_5 OAC5_{ij} + \beta_6 OAC7_{ij} + \beta_7 Dep_{ij} + e_{ij}$$

$$y_{ij} = \beta_{0j} + 0.008\ (0.022)\ OAC1_{ij} + -0.049\ (0.039)\ OAC2_{ij} + 0.055\ (0.050)\ OAC3_{ij} + -0.023\ (0.020)\ OAC4_{ij} + 0.044\ (0.025)\ OAC5_{ij} + -0.058\ (0.028)\ OAC7_{ij} + 0.000\ (0.001)\ Dep_{ij} + e_{ij}$$

$$\beta_{0j} = 0.418\ (0.019) + U_{0j}$$

$$U_{0j} \sim N(0, \sigma^2_{u0}) \quad \sigma^2_{u0} = 0.012\ (0.002)$$

$$e_{ij} \sim N(0, \sigma^2_e) \quad \sigma^2_e = 1.175\ (0.009)$$

-2*loglikelihood = 101001.600 (33594 of 33594 cases in use)

Figure 4.17. 2 level multi-level model, with children at level 1 and SOA at level 2, adjusted for the OAC super group (categorised against OAC super group 6, which was closest to the mean) and deprivation score. The value for the coefficient for each variable is given and the number in brackets is the standard error.

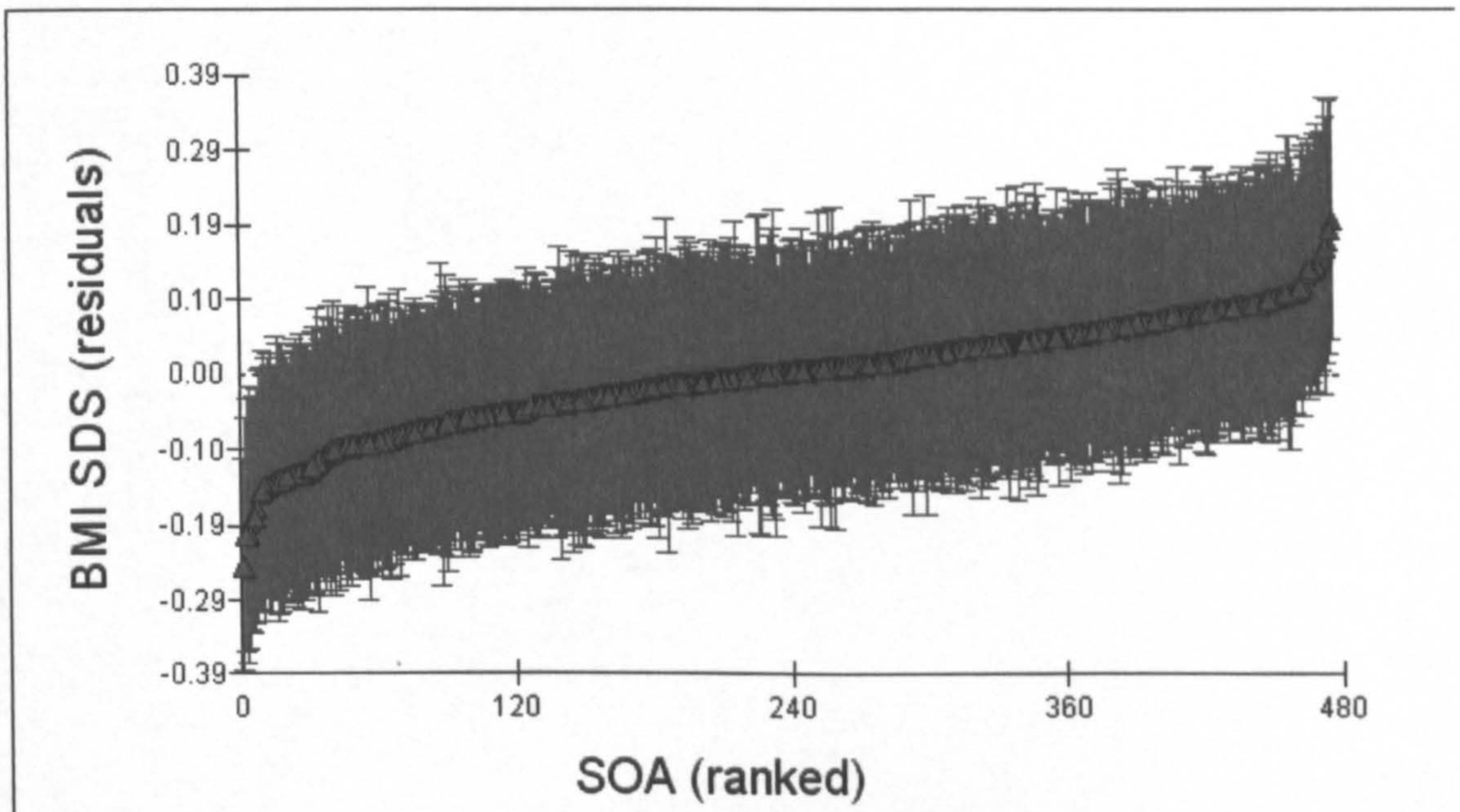


Figure 4.18. Plot of the residuals at SOA level from the multi level model using 1.96x standard deviation to calculate 95% confidence intervals. Significant SOAs are those that do not cross the 0.00 (BMI SDS residual) line. On the right hand side, SOAs are more obese that the model would predict (given the deprivation and super group number of the area). Vice versa on the left hand side. These are the "hot" and "cold" spots.

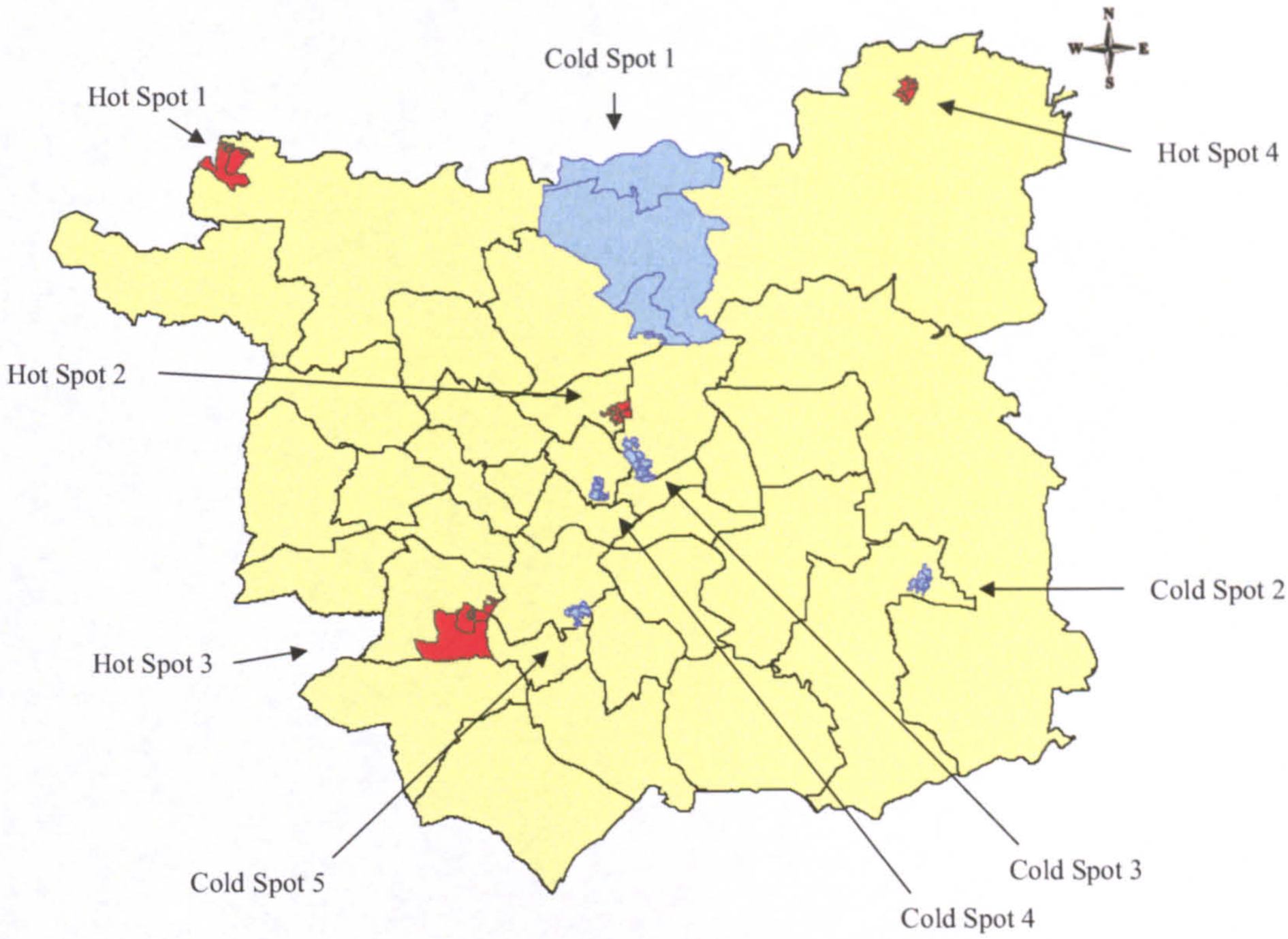


Figure 4.19. Map of Leeds highlighting those OAs that the MLM identified as having a significantly higher (red) or lower (blue) BMISDS than expected, after taking deprivation and OAC super group into consideration. The lines represent the ward boundaries.

	HS1	HS2	HS3	HS4	CS1	CS2	CS3	CS4	CS5	Total
No children	91	81	70	87	56	106	209	156	62	33594
% obese & overweight	30.0	25.8	24.8	26.9	6.2	13.1	14.9	13.5	13.8	11.3%
% obese	14.3	14.4	11.7	14.4	1.7	3.0	5.8	7.1	8.2	7.6%
Mean BMISDS	0.760	0.450	0.600	1.122	0.159	0.345	0.103	0.008	0.216	0.422
Mean Deprivation	28.9	19.5	23.2	20.6	12.7	6.1	21.5	70.9	64.2	30.3
OAC super group	1,5	6,7	1,4,6	1,5	3,4,6	4	4,6,7	7	1,5,7	all
% male	44.0	54.3	57.1	48.3	41.1	47.2	49.8	55.1	43.5	50.9
Mean age	6.3	6.9	6.5	6.4	5.2	5.3	7.1	6.5	6.8	6.4
Age range	3; 12	3; 13	3; 12	3; 13	3; 12	3; 12	3; 12	3; 12	3; 14	3; 14
Years of measurement	1998; 2006	1998; 2006	1998; 2006	1998; 2006	1998; 2006	1998; 2006	1998; 2006	1998; 2006	1998; 2006	1998; 2006

Table 4.14. Summary of characteristics of those hot and cold spots identified by MLM after adjusting for deprivation and OAC super group.

4.5 Discussion

Childhood obesity is a serious condition with serious consequences, with little or no adequate treatment available. Treatment is not a viable (no good treatment available) or affordable (expensive) option. Obese children will most likely become obese adults, with all the corresponding increased health risks. Accordingly childhood obesity is a major public health problem that needs to be monitored and addressed. Monitoring child BMI trends is important because of the rapid and substantial rise in prevalence of obesity. This process would be assisted by a policy of universal measurement when children start school (Hall, 2000). Recently the government has commenced blanket measuring of 11 year olds in order to monitor the obesity epidemic. Nevertheless school entry still offers an excellent opportunity to monitor trends in early cases of obesity in children (and these data are being collected anyway, whether it is used to monitor obesity or not). As with the 11 year olds data almost 100% of children can be examined (in theory) and these large datasets would enable micro-level analysis to be undertaken to identify the high-risk populations. These data can contribute to a core dataset to benefit child public health.

4.5.1 Data collection

This study examined childhood obesity in Leeds using BMI as a measure of obesity, which isn't perfect (for reasons discussed in chapter 2) but is a suitable measure for epidemiological purposes. The number of children excluded from this analysis was low and largely related to postcode issues; 58% of the 4140 children excluded from the study had no or an incorrect postcode and 26% lived outside of the study area. This was because the BMI data was collected for different original purposes than that of this study.

The routinely collected data showed that the "normal" coverage calculation of height and weight data was on a par with that reported elsewhere (Banerjee et al, 2003; Jones et al, 2005). However in terms of coverage of the whole population, the "birth" and "census" coverage figures more truly represent the proportion of the population that was measured (corroborated by the fact that both measures produced similar figures), which were low. Birth and census coverage for 5 year olds was more robust (particularly if the last two years of measurements, where many data had not been recorded on to the NCHCS yet, were excluded) at an average of 46.2% and 45.1% (birth and census coverage respectively). This is probably because this age group are a more captive target as they are measured at school rather than relying on attendance at a clinic. It should be noted that the coverage is lumpy; i.e. some schools have particularly high or low measurement rates. At least this means the data are more like a random sample, although it is not known whether schools do not participate because they think obesity is a

problem for them or because there isn't the funding or time. Qualitative interviews with the system administrator and school nurses suggests the latter rationale.

As there was no standardised procedure for selecting the children that were measured (such as cluster analysis) for the routinely collected data, by definition, then the children that were measured may not be representative of the population. Also the issue of non-response bias must be addressed for this dataset. However as coverage increases, these concerns diminish. Given the low non-measurement percentage (i.e. where height and/or weight data not provided) (8% for 5 year olds; albeit higher for 3 year olds at 28%), the magnitude of any non-response bias is also probably low; but may still be present, depending upon the reasons for children not being measured. Informal qualitative interviews discovered that the principle reason for missing data was that a large proportion of the records are still paper based because the primary care trusts have insufficient resources (both time and money) to transpose the data onto the computer system. This was particularly true for 3 year olds and the more recent 5 year olds' measurements. To a lesser degree another possible reason for missing data was families not turning up for assessment, with no subsequent follow up (as health staff were not aware of the importance of obtaining as close to 100% coverage as possible). For the 5 year olds this was simply due to absence from school on the day of measurement. However for the younger children it is not known why parents did not attend their clinic appointment and can not distinguish between "whether they felt their child is "normal" and doesn't need checking", or conversely, "too thin or fat and not wanting to get into trouble". It was also reported that if health visitors or school nurses were busy and the child looked "normal", then height and weight measurements were not carried out.

Aside from the problem of missing data, measurement bias may also be present as many different people carried out the measurements. Whilst it is understood that the primary care trusts aimed to standardise the measuring methodologies, the author was unable to ascertain the specific nature of the standardised methodology. Accordingly, it is necessary to conclude that differences in techniques between examiners are possible. A small percentage (<2%) of data were identified, retrospectively, as having a measurement error: this was undertaken by searching the dataset for children who had been measured twice in the same year and whose second height measurement was less than the first. In these circumstances at least one of the height measurements must be wrong, but it is impossible to determine which; these children were excluded from the dataset. Whilst this isn't an exact method of identifying measurement errors it gives an indication of the size of the problem (i.e. small). This only applied to the routinely collected data. Finally there is also the possibility of information bias, which in the original routinely measured dataset was very likely as many entries had used the wrong unit of measurement (e.g. mm rather than cm). Many of the routinely recorded data records were

adjusted for input errors and this is indicative of the low quality of the data input, perhaps because of a lack of priority (in funding and time terms) given to this task by the primary care trusts. The two studies' data did not suffer from high levels of data entry errors. However it was possible to use a formula to correct these errors, although some may have slipped through. Also data entry errors were minimised by excluding children with implausible BMI SDS. Whilst this may serve to underestimate the obesity problem, given the high level of data entry errors it was strongly felt that it was more important to remove these likely errors (as it was likely that more children were in this category incorrectly than correctly). In summary, for the routinely recorded information the data for the 5 year olds is likely to have relatively little non-response, measurement and information bias, although the 3 year olds' data are less reliable. This was not a problem for the two study samples.

Some authors have been critical of using the routinely collected data as a means of monitoring trends in childhood obesity (Rudolf et al, 2006) although other authors do advocate its use for monitoring (Blundred et al, 2001; Jones et al, 2005). This argument is still important, despite the population-level measuring of 11 year olds that has recently commenced, as it facilitates monitoring and analysis of childhood obesity in different age groups. A review of the evidence for the benefit and/or harm of growth monitoring (i.e. not limited to monitoring obesity trends) came to the conclusion that there was no reliable evidence to either support or refute the claim that investment in the activity has worthwhile health benefits (Garner, 2000). The theoretical benefits of using this routinely collected data are low marginal cost (when taken together with other school entry screening procedures), potentially high coverage, and the additional benefits of identifying children with other growth related disorders (Hall, 2000). Routinely collected data, if amassed accurately, with comprehensive (as close to 100% as possible) coverage, and with high quality measurement and data entry, can be used to examine trends in childhood obesity (equally so could an adequately sized and randomized sample with a high response rate (Levine et al, 2007)). However the measurement must be done to a high standard, using a standardised methodology with trained examiners, so reliable equipment must be provided and used, and staff training is vital (Hall, 2000). Yet for the NCHCS data to be useable to monitor trends in childhood obesity in the UK, there is a need to improve the quality of the database, because any method of monitoring trends in the population needs to be simple and accurate and to have comprehensive coverage (Jones et al, 2005). Accordingly the quality of the routinely collected data needs to be improved, particularly focusing on the data entry – improving the accuracy of the typed entries as well as transposing the backlog of paper records and keeping up the conversion onto the NCHCS up to date - if it is to be reliable in monitoring trends in obesity. Measurement error must also be measured and monitored. This dataset can also be used to examine the relationship between obesity and obesogenic factors. To do temporal spatial analysis takes very large datasets to avoid small number problems, so high coverage is essential

if this examination of trends and health inequalities can be undertaken on a routine basis as an essential part of public health investigations. A thorough review of the NCHCS across the country is needed to identify those areas that are under-performing and need urgent attention (and probably additional resources) to facilitate this process. As significant resources are already committed to measuring height and weight at school entry (and also now at age 11) to ignore this resource, to continue to leave the data as paper records unavailable to researchers and public health professionals, is a massive waste of resources and potentially valuable information.

4.5.2 Global analysis

The global analysis showed obesity (and overweight) rates increasing over time and as children age (in boys and girls alike). Over time there was a small increase in average BMI (0.2 kg/m^2), although the range remained fairly static. This modest average gain converts into a 50% increase in the proportion classified as obese; that is, children were 1.5x more likely to be obese in 2003 than 1998. The finding that obese (and overweight) children have increased in frequency in Leeds over this time scale is not unexpected given the worldwide (Kumanyika et al, 2002; Ebbeling et al, 2002; Lobstein et al, 2004) and national rising trends in childhood obesity: in the UK between 1995 and 2002 the prevalence of obesity in children aged 2 to 15 years increased from 10.4% to 16.6% (60% rise) in boys and 11.7% to 16.7% (43% rise) in girls (Sproston & Primatesta, 2002). The mean age of children did increase over time, but as the variable being examined (BMI SDS) was already standardised for age and gender, the rise in obesity (and overweight) with time cannot be attributed to the children being slightly older towards the end of the time period than at the beginning. When considering age, 13 year old children were 3.0x more likely to be obese than 3 year olds; with children gaining, on average, 4.4 (0.303) kg/m^2 in BMI (BMI SDS). Similarly, other authors have shown rising rates of obesity across all age groups in the UK (Blundred et al, 2001; Sproston & Primatesa, 2002). Higher rates of obesity in older children indicates that it is a cumulative problem and once obese the children are remaining so, corroborating the assertion that obese children are likely to become obese adults (Guo et al, 1994; Freedman et al, 2002), with all the corresponding health and social disadvantages. The largest datasets were for 3, 5 and 11 year olds. Age, height and weight obviously increased in the older children. When considering gender there were significant rising trends for boys and girls individually as well as cumulatively by time and age (except overweight boys over time). The proportion of obese (and overweight) girls was lower than that for boys, by age group and over time, which differs from results from other studies (Jones et al, 2005). Whilst the difference between genders was significant it was only a small effect size, so probably not clinically important (e.g. from a potential public health intervention perspective).

The average deprivation score for the study area based on the deprivation score for the children measured was 30.3, and based on geography (i.e. taking each OA only once) was 28.1. This suggests our study population lived in slightly more deprived areas than the average for Leeds. A linear regression showed deprivation was positively associated with childhood obesity. Those children living in more deprived areas were twice as likely to be obese (or overweight) than children living in more affluent areas. This concurs with work by other authors. The following single indicators of SES, as a proxy for deprivation, all showed that children with lower SES / more deprived backgrounds have an increased risk of childhood obesity: household income (Strauss & Knight, 1999; Stamatakis et al, 2005); entitlement to free school meals (Cecil et al, 2005); families with lower education levels (Danielzik et al, 2004; Lamerz et al, 2005; Romon et al, 2005). Additionally other studies using the Townsend Deprivation Score (an index score based on a combination of adult unemployment, household size, and car and home ownership) have shown that children from more deprived areas have higher risk of obesity (Kinra et al, 2000; Kinra et al, 2005). The increased prevalence of obesity in children from more deprived backgrounds could be due to a multitude of factors, for example: dietary differences are often apparent; no safe play area for the child; lack of opportunity / funds for activities, so TV viewing is the primary leisure activity by default; presence of food deserts (lack of accessible, affordable, healthy (low energy dense) food); constraints on calories per pound, which focuses purchases on energy dense foods.

4.5.3 Local analysis

Significant spatial differences in childhood obesity were found across the study region, identified as hot spots (before adjusting for covariates). Children living in Leeds city centre (a highly deprived area) were between 1.5x-6.1x more likely to be obese than children outside of the obese clusters; similarly for overweight. Nevertheless hot spots were found in both deprived and affluent areas, suggesting either a spread of obesity across socio-economic groups and/or something special about those areas affects the aetiology of obesity, whether that is an obesogenic environment factor (e.g. a lack of parks) or something about the individuals (e.g. these children may be very inactive).

Analysis of the OAC super groups is novel. It is a relatively new classification system that is likely to prove extremely useful, as it gives more and different information about an area than simply deprivation alone, plus, importantly, is open about the methodology used to determine the index (unlike other geo-demographic classifications, such as Mosaic) (Vickers and Rees, 2007). Some geo-demographic groups were more likely to be obese than others. Children living in areas classified as OAC 1 (blue collar communities), 3 (countryside), and 5

(constrained by circumstances) were generally more obese than children living in OAC 2 (city living), 4 (prospering suburbs), and 7 (multicultural). Some OAC super groups displayed obesity prevalence contrary to the "deprivation theory"; that is, multicultural neighbourhoods were less likely to be obese but were located in deprived parts of Leeds, and conversely countryside neighbourhoods were more likely to be obese but located in affluent parts of Leeds.

There is evidence of a social class gradient in obesity in the UK (Kinra et al, 2000) and in France (Romon et al, 2005), with children from lower socio-economic groups, showing higher prevalence or change in prevalence than children from higher socio-economic groups. Yet this study is showing a spread of obesity across socio-economic groups, which has been noted by other authors in more recent work (Buchan I, personal communication with author, September 2006). It may be that other variables in the OAC index construction may be having this effect with obesity, such as the demographic or household attributes that are not included in the index of deprivation.

Hot (& cold) spots remained even after adjusting for deprivation and OAC super groups. In the final model, the hot spots were all in affluent areas (maximum deprivation 28.9) (removing the formerly large city-centre hot spot), suggesting that all the areas of high obesity prevalence in deprived areas had been adjusted for and excluded. Only OAC super groups 2 and 3 (city living and countryside) were not represented in hot spots, suggesting these are good areas to live if obesity is to be avoided. The cold spots spanned both affluent and deprived areas. OAC group 7 (multicultural) was represented in both affluent and deprived cold spots. However OAC groups 3, 4 and 6 (countryside, prospering suburbs and typical traits, respectively) were in the affluent cold spots and OAC group 1 and 5 (blue collar communities and constrained by circumstances, respectively) were in the deprived cold spots. OAC group 2 was not represented in the cold spots; so whilst "city living" is not obesogenic, neither is it leptogenic. This analysis is novel.

As crude maps of counts of cases by area, reflect the population distribution or age structure and do not allow for any natural variation in the underlying population, all mapped data should be standardised in some way. Accordingly focusing on mean BMI SDS achieves this. However the proportion obese (and/or overweight) maps do not adjust for a different age or gender structure of the population, so these hot spots could simply be a reflection of older children and/or more boys being measured. The final adjusted hot spot model used mean BMI SDS as the childhood obesity dependent variable so these hot (& cold) spots are not affected. The average age for each (final, adjusted model) hot and cold spot was similar to the total dataset average, with the same age range of children measured, from all possible measurement years. The proportion of boys measured ranged from 41-57% (against the dataset average of 50.9%)

and there was no clear pattern (some hot and cold spots had lower proportions and some had higher).

This study also enabled analysis of childhood obesity at different spatial scales to be considered. The ward level proportion obese (and overweight) maps show only the central (highly deprived) areas of high prevalence of childhood obesity. Averaging obliterates the potential hot spots in the north-east, north-west and south-west corners of Leeds. Quite a different picture is created for the mean BMI SDS maps, showing a lot more variation over the region at both SOA and ward level. This may be an example of the difference scale can make to a map. In presenting the results the maps used either a quintile scale (to evenly represent the distribution of the data) or a manual scale to facilitate comparisons across maps. In the proportion maps (Figures 4.5-4.8) the minimum categories (i.e. the lightest shade) are greater than the proportion of children who would be obese (or obese and overweight) by definition (i.e. 2% and 9%, respectively). Accordingly we would expect that if childhood obesity were not an issue then these maps would be entirely coloured in the lightest shade, which is clearly not the case. However in the BMI SDS maps even the darkest colour does not represent either obese or overweight (standard deviation of over 2.00 or 1.33 respectively). This is because this is an average for the area (so it would be very shocking if the average was even classified as overweight as this would suggest that a very high percentage of children were at least overweight). Nevertheless this does indicate that using a micro-level unit of analysis does enable small pockets of problem areas (i.e. high prevalence of obesity) to be identified, which otherwise get "averaged out", and potentially ignored increasing health inequalities, if the city as a whole is considered. Analysing data at micro-level gives more information about patterns of disease, enabling targeted interventions to be designed and implemented.

4.5.4 Limitations

There are many potential problems with the analysis of spatial data as on top of the usual sources of bias, a whole new stream of problems emerge, such as MAUP and small number problems. As the spatial scale of the data reduces, by using a finer geography for analysis, researchers can run into confidentiality problems. Accordingly data must be anonymised and sensitivity is required over maps published to ensure children can not be identified. Data were aggregated to the census boundaries of OA and SOA, and these aggregated data were used as a proxy for the individual in the absence of true individual data. As these Census boundaries are created to select homogenous populations rather than otherwise arbitrary boundaries for other purposes (such as the postcode, which is created for postal convenience, or health care boundaries) this effectively minimises the MAUP. Similarly undertaking analysis at the micro level, approximating to individual level data, should serve to limit the potential ecological

fallacy errors. A larger problem with this analysis was that of small numbers. Even if a large dataset such as this one is taken down to the micro-level, the populations in each area get very small. This problem was resolved in two classic ways. Firstly the data were aggregated, both over time and across age groups, to increase the populations and cases in each area. Also Empirical Bayes smoothing techniques were applied to the data to reduce any instability in the results.

Often, as in this case, the spatial element of health data is the postcode. This is obviously a different aggregate geography than the census OAs and SOAs; that is, the boundaries are not aligned. This issue was minimised by using point data for the full postcode (not postal sector) to determine exactly which OA (and accordingly SOA) each child lived in. As this data integration uses the centre of the postcode as the child's address (yet the child may live at the far end of the postcode area), accordingly it is possible that some children were allocated to the wrong, adjacent, OA. This use of "uncertain" locations (i.e. using the postcode centroid to determine residence) can reduce statistical power if the locations are inexact (Jacquez et al 1996; Jacquez & Waller, 1996). At this small scale this effect is likely to be minimal. Further there is always the potential for a child to give the wrong postcode. However if this happens it is likely to be an incorrect (i.e. non-existent) postcode, which would be identified by the GIS; although it is possible (but not likely) that a small number of children gave wrong but actual postcodes, so their location in the spatial analysis is incorrect. Or if children spend time with separated parents, at different locations, there may be an effect from two different postcode areas. Another potential data integration problem is that of the dates of measurement of the data. The obesity data related to measurement periods varying from 1998 to 2003, yet the census indices related solely to 2001. This is not problematic because the census areas, OA and SOAs, are very static, and tend to stay homogenous over time (if people change they move home, rather than the characteristics of the area changing).

Where the disease has long latency period from onset of the process to clinical diagnosis (such as obesity), the residence of the patient at the time of diagnosis may not be relevant in terms of disease causation. This is the problem of migration. However this study is looking at children, so maximum length of latency period is 13 years and for many of the subjects much less than this (not decades as can be the case with adults). This minimises the migration problem, but it remains a possibility that some of the children did live in the study residence their whole life and some did not (percentages unknown).

4.5.5 Conclusion

Both aims of this study have been fully met:

Firstly in relation to the routinely collected data, this chapter has described an analysis of the local (the five Leeds PCTs) National Child Health Computer System (NCHCS) records to examine the difficulties associated with using these data. The routine data underlying this analysis for 3 year olds is not robust and it should probably not be relied upon as a sole source of information about trends in childhood obesity in young children, but the routinely collected data for 5 year olds (and the other BMI datasets) are reliable. With a little effort on the part of the primary care trusts in relation to data entry, the routinely collected data could become a valuable source of information for monitoring obesity and, with the use of spatial analysis techniques, the impact of the obesogenic environment. This latter point is important as large datasets, such as one with nearly 100% coverage such as the routinely recorded dataset, are required to undertake analysis at the micro-level.

This chapter has also looked to increase understanding about childhood obesity and associated health inequalities. It has shown that obesity in 3 to 13 year olds in Leeds has risen since 1998 and that prevalence of obesity increases with age. These results serve as a base from which future trends in Leeds can be monitored and also for comparison with trends elsewhere in the UK.

Also childhood obesity was significantly associated with deprivation, with children living in highly deprived areas being twice as likely to be obese (or overweight) than children living in more affluent areas. Nevertheless some hot (and cold) spots were found in affluent (and deprived, respectively) areas, suggesting either a spread of obesity across socio-economic groups and/or something special about those areas affecting the aetiology of obesity. Some homogenous geo-demographic groups were more likely to be obese than others. Furthermore some such groups displayed obesity prevalence contrary to the "deprivation theory"; that is, some groups that were less likely to be obese were located in deprived parts of Leeds, and vice versa. These are important demographic differences between areas of high and low prevalence of obesity.

Spatial analysis techniques used at the micro-level can help us to understand the variations in obesity within an area more thoroughly, allowing us to gain a better understanding of the driving factors and to identify key problem areas, emphasising pockets of high (or low) prevalence and identifying high-risk populations defined spatially or by any of the covariates analysed. It can show how the prevalence of obesity is linked to various spatially defined

covariates, such as those contained within census data. Also mapping and analysing data down to the micro-level (e.g. SOA level) facilitates the focus on key problem areas, rather than relying on averages for the whole of an area, e.g. a whole city / town or by ward, as per a global analysis. To generalise data for, say, the whole of Leeds would mean that health professionals would miss small problem areas. Together, this increases our understanding about patterns of childhood obesity, thereby facilitating the development of focused and targeted interventions to prevent childhood obesity. This analysis provides an enhanced environment for analysis, evaluation and decision-making in health planning.

This analysis methodology is also generalisable to other chronic diseases and to other localities.

Next steps:

Spatial microsimulation modelling can further enhance analysis and evaluation of environments by allowing us to synthese data (whether spatially linked or not) down to individual (or household) level, permitting us to map and analyse synthesised data at individual (or household) level. It enables us to build on existing demographic data contained within the model, as additional covariates entered into the microsimulation model provide further details about residents' environments and develops the relationship between the risk factors of interest and health data - see chapters 6 and 7. Also it is possible to drill down into the differences between the final model hot and cold spots of obesity, to gain an understanding of the potential causal factors. What is it about some areas that make children more likely to be obese if they live there, and some areas children are less likely? Are there any place-specific reasons for the high (or low) prevalence? The challenge will then be to develop targeted preventive measures aimed at childhood obesity using sources such as the routinely collected BMI data to monitor future success (or otherwise). These issues will be addressed in chapter 8. It is also important to consider the school environment, as both home and school are important potentially obesogenic environments for children - see chapter 5.

Chapter 5: Measuring the school impact on child obesity

5.1 Introduction
5.2 Background
5.3 Methodology
5.3.1 Measurement of pupils' growth
5.3.2 Analysis of data
5.3.3 Development of model
5.4 Results
5.5 Discussion

5.1 Introduction

This chapter explores the impact that schools have on their pupils' obesity, identifying those where targeted input is most needed. Section 5.2 provides some background information regarding obesity and the obesogenic environment, briefly overviewing interventions to prevent childhood obesity (more detail is given in Chapter 8) and the role that schools may be able to play to facilitate the prevention of obesity in children. Section 5.3 describes the modelling process that was developed using data on a socio-economically and ethnically representative sample of 2319 school pupils aged 5 and 9 years old attending 35 Leeds primary schools, collected over two years. The three steps in the multi-level model involve calculating the "Observed" level of obesity for each school using mean BMI SDS; adjusting this using ethnicity and census-derived deprivation data to calculate the "Expected" level; and calculating the "Value Added" by each school from differences in obesity at school entry and transfer. The results are set out in section 5.4. There was significant variance between the schools in terms of mean BMI SDS (range -0.07 to +0.78). Residential deprivation score and ethnicity accounted for only a small proportion of the variation. Expected levels of obesity therefore differed little from the Observed, suggesting robustness in the data (i.e. schools are ranked in line with our expectations), and the Value Added step produced very different rankings, which we explore. In conclusion (section 5.5), there is variation between schools in terms of their levels of obesity. Our modelling process allowed us to identify schools whose levels differed from that expected given the socio-demographic make up of the pupils attending. The Value Added step suggests that there may be a significant school effect. If this is validated in extended studies, the methodology could allow for exploration of mechanisms contributing to the school effect, and identify schools with the highest unexpected prevalence. Resources could then be targeted towards those schools in greatest need.

5.2 Background

Obesity is a significant medical and social problem. Its prevalence has escalated over the last two decades, reaching pandemic levels in the developed world and is also increasing across the developing world (Wang & Lobstein, 2006).

The UK Department of Health has recently taken the controversial step of introducing monitoring for obesity in Primary Schools in the UK. With the introduction of widespread monitoring, it makes sense to explore whether the process might be utilised to bring benefits to the evidence base in other ways. It is, for example, possible that the data might prove useful in improving our understanding of the determinants of obesity and exploring the impact that schools may have on the problem.

The environment we live in has been described as "obesogenic" – an environment that hinders sufficient physical activity and promotes excess intake of food; thereby making obesity more likely to occur. The term obesogenic environment has been defined more precisely as "the sum of influences that the surroundings, opportunities or conditions of life have on promoting obesity in individuals or populations" (Swinburn et al 1999). These interrelate and determine health behaviours which are influenced by many different factors: social, economic, regulatory, cultural and physical.

Primary schools provide an obvious setting where children can be targeted to reduce obesity. They spend a considerable part of their waking hours at school, eat meals there and are required to have physical exercise as part of the curriculum. The influences and behaviours of peers build social "norm" expectations and may also impact on the eating and physical activity behaviours of the individual thereby affecting obesity (Monge-Rojas et al, 2002; McCabe et al, 2002; Gilmer et al, 2003; Zabinski et al, 2006; Salvy et al, 2007). Furthermore, primary schools often play an important role in a community. They may be much more than an educational facility, and frequently offer opportunities and activities for those who do not have children at the school.

There are a number of examples whereby policy and initiatives may impact on pupils' risks of obesity. It has been shown, for example, that a school's food policy to promote healthy eating can impact the nutrient intake of children, reducing the intake of high fat and high sugar foods (Neumark-Sztainer et al, 2005), which in turn may reduce obesity levels (see Chapter 2). Conversely, other food policies can lead to unhealthy eating behaviours and increases in overweight / obesity, such as the use of breakfast clubs (Belderson et al, 2003), the consumption of packed lunches rather than school meals (Whincup et al, 2005), and the presence of vending

machines in schools (New & Livingstone, 2003). A population-based approach in primary schools in Leeds was successfully implemented showing positive changes in playground activities, tuck shops and school meals (Sahota et al, 2001a; Sahota et al, 2001b). Singapore has launched a national healthy lifestyles programme, called "Trim and Fit", using similar methods to the Sahota study, and has seen a fall in obesity levels (Toh et al, 2002). One school-level intervention has influenced children's actions outside of school, with a decrease in the amount of television watched by children and their adiposity (Robinson, 1999).

Schools have recognised their role in the fight against childhood obesity and a plethora of initiatives are springing up within schools and on their premises. However, at this stage the initiatives are based on a small amount of evidence. Recent systematic reviews of the evidence base has shown that research into the effectiveness of school based interventions is very limited and generally of inadequate quality (Summerbell et al, 2005; Connelly et al, 2007). These reviews show that there have been a small number of randomised controlled trials and of these very few have had an effect in preventing childhood obesity. Clearly, there is a need for good quality randomised controlled trials, but they take time and are expensive. While we await their development and results, we contend that we might be able to learn some valuable lessons from the natural experiments that are being carried out in schools across the country and elsewhere.

In considering the impact that schools may have on their pupils' risks it is important to take into account those social factors that are known to be associated with obesity. Of these social disadvantage and ethnicity are key. Studies show that children with lower socio-economic status (SES) and/or from more deprived home environments have an increased risk of childhood obesity (Strauss & Knight, 1999; Kinra et al, 2000; Danielzik et al, 2004; Cecil et al, 2005; Kinra et al, 2005; Lamerz et al, 2005; Romon et al, 2005; Stamatakis et al, 2005 – see Chapter 4). In the West, non-white children are more likely to be obese than white children, although this is largely thought to relate to socio-economic factors, such as parental education and family income (Strauss & Knight, 1999; Strauss & Pollack, 2001; Lobstein et al, 2004). This has been shown in children of south Asian origin in the UK (Whincup et al, 2002).

This paper proposes a methodology for identifying the differences between schools based on a detailed analysis of growth measurements interpreted against readily available geographical data. Our hypothesis is that by exploring differences between schools we may be able to determine school factors that are, for better or worse, having an impact on children's risks of obesity. At the same time we may be able to highlight "hot" and "cold" spots of obesity so allowing better targeting of resources to those communities in greatest need.

5.3 Methodology

A modelling process was undertaken using growth data collected on school children during 2004 and 2005 in Leeds. The growth data had been collected on primary school children in thirty-five schools as part of the Trends project - a project aimed to construct a simple, reproducible method to monitor trends in childhood obesity in light of the governmental target to stop the rise in childhood obesity by 2010 (Rudolf et al, 2006). Ten of the schools were selected as a purposive sample in 2004, and a further 25 were randomly selected by computer from the remaining 230 state schools across the city.

5.3.1 Measurement of pupils' growth

Agreement had been obtained from the head teachers and governors of each school for a specially trained health care assistant and scribe to measure children in Reception class (age 5 years) and in Year 4 (age 9 years). "Opt out" consent was obtained from parents, and measurements were made in the summer term (April – June 2004 and 2005) with children wearing light clothing only and no shoes. Height was taken to 0.1cm accuracy using a freestanding stadiometer (Raven Dunmow) and weight to 0.1kg, as previously described (Rudolf et al, 2003).

5.3.2 Analysis of data

Body mass index (BMI) was calculated and converted to standard deviation scores (SDS) using the UK 1990 growth references (Cole et al, 1995) to allow for statistical computation of children of different ages and sex (Rudolf et al 1999). In the absence of individual SES, micro-level area based SES was used to proxy the individual socio-economic characteristics. A measure of SES was assigned using the Index of Deprivation 2004 (Communities and Local Government, 2004), which is derived from the 2001 census data. It is provided at lower Super Output Area (SOA) level (of which there are approximately 500 in the study area) of the residential location of each observed pupil, and was determined by linking each child's residential postcode centroid to the relevant SOA using GIS software (ArcGIS V9.0). We chose this method, rather than calculating the deprivation score of each schools' location, as home deprivation has been associated with childhood obesity (as discussed above). When deprivation is based on the location of the school, no such relationship has been shown to date (Dummer et al, 2005). Ethnicity was divided into two categories (south Asian and non south Asian) for pragmatic reasons, as south Asian (covering Indian, Pakistani and Bangladeshi) was the largest ethnic minority group, which is reflective of Leeds as a whole (Unsworth & Stillwell, 2004). Children's ethnicity was determined from school records using Nam Pehchan software to

identify south Asian pupils from their names (Cummins et al, 1999), which shows very high sensitivity and reasonable specificity.

5.3.3 Development of model

The steps involved in analysis of the growth data and development of the methodological model are shown in Figure 5.1. BMI SDS was used, rather than percent obese children, as this is the preferred measure of obesity (Rudolf et al, 2006).

- ***Step 1 – Ascertaining that the sample was representative of primary school aged children in the city***

  A subgroup analysis by ethnicity (south Asian or non south Asian) and socio-economic status was undertaken and compared with the 2001 census data, in order to ascertain that the sample of children were representative of the city as a whole.

- ***Step 2 – Ranking of schools according to observed measures of obesity ("Observed ranking")***

  The mean BMI SDS for children in both Reception and Year 4 was determined for each school and this was taken to be their *observed measure* of obesity. Schools were ranked according to their mean BMI SDS, and were then mapped against the Index of Deprivation of their location. This provided a preliminary identification of hot and cold spots of obesity across the city.

- ***Step 3 – Ranking of schools according to how much their measure of obesity deviated from the expected ("Expected ranking")***

  We then adjusted the data (for children in both Reception and Year 4 combined) to take the measured deprivation and ethnic mix of the pupils into account to determine the school's *expected measure* of obesity. As discussed earlier, both of these variables have been highlighted as potential confounders in the literature.

  Multi level modelling (MLM) techniques (Rasbash et al, 2004) were then employed, as this allows the dependency inherent in pupil observations nested within the same school (or neighbourhood) to be taken into account. MlwiN V2 was used based on 2 level hierarchical linear model (pupil within school and neighbourhood (Super Output Area), which were cross classified at level 2). The independent variable was the BMI SDS of each child. The explanatory variables used in the model were SES (deprivation score) and ethnicity (binary variable: south Asian / non south Asian). A cross-classified approach was used because children living in the same neighbourhood do not necessarily go to the same school, and

likewise children attending the same school may or may not live in the same SOA, which creates different levels of dependency in the data. Accordingly both school and SOA are at level 2 in the model, with the pupil at level 1. That is, it partitions the variation in obesity across each of the hierarchies, thus respecting the natural aggregation of the data (pupils at level 1, nested within schools at level 2, cross-classified with SOAs also at level 2) leading to less erroneous inferences (i.e. the modelling framework improves the accuracy of the estimates and leads to fewer erroneous inferences) than if the natural ordering had not been accounted for.

Schools were then ranked according to the extent that the Observed mean BMI SDS levels differed from the Expected levels (residuals). This Expected ranking is of more interest than the first stage Observed ranking as it allows identification of schools that have levels of obesity that are different from expected given deprivation and ethnicity of pupils, and so permits identification of real hot and cold spots.

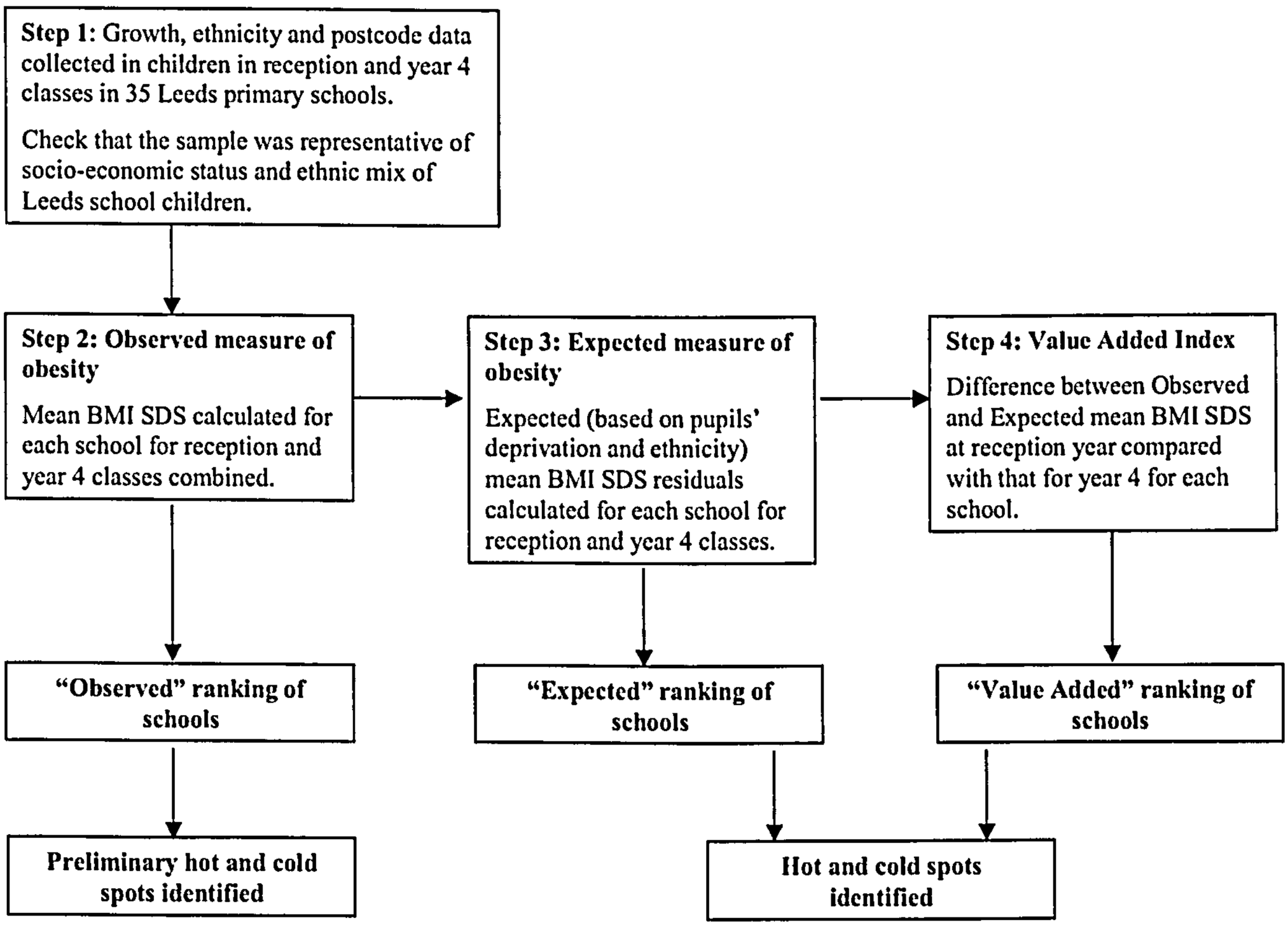


Figure 5.1: The steps involved in data analysis in the development of the methodology for identification of hot and cold spots of obesity in schools.

- ***Step 4 – Calculation of the Value Added index ("Value Added ranking")***

  The final stage examined whether it was possible to determine the "value added" by a school. In theory, measures of obesity at entry to school might be taken to be a reflection of the community from which the children are drawn. As children at the age of five years have

been exposed to school for a very short period the school effect can be considered to be minimal. Accordingly we calculated a Value Added index (to use an academic analogy derived from SATS testing), as the difference between the Observed and Expected mean BMI SDS at Reception year compared with that for Year 4 for each school, given deprivation and ethnicity of each schools' pupils. This was achieved by firstly calculating the Expected residuals for Reception pupils and then separately for Year 4 pupils, using the same cross-classified MLM as in Step 3. Then determining the difference between these two sets of Expected residuals (Year 4 residuals – Reception residuals). This Value Added figure was used to rank the schools. This ranking was compared with the previously described rankings, and also used to identify hot and cold spots.

5.4 Results

There were a total of 2911 children in Reception and Year 4 classes in the 35 schools, 2367 of whom (81.3%) were measured; no systematic bias was identified across the schools regarding the unmeasured children. There was no valid postcode or ethnicity data available on 48 children, so data on 2319 children were available for this analysis. These children attended 35 schools with a mean number of 66 children (range 36 to 117 children) measured in each school. 13.5% were of south Asian origin (Pakistan, Bangladesh and India). Comparison of the population with census data confirmed that the sample was representative of the socio-economic and ethnic make up of the city. There was little difference in the gender breakdown of children measured between the two year groups: Reception classes were 47% female, 53% male; Year 4 classes were 45% female, 55% male. Higher deprivation was associated with higher mean BMI SDS (mean 0.003, 95%CI <0.001, 0.006; p=0.046, Expected model), although the size of the effect was of doubtful clinical importance. South Asian children were found to have a significantly lower BMI SDS than non south Asian children (mean –0.199, 95%CI -0.352, -0.046; p=0.011, Expected model).

There was a wide range of obesity across different schools, with Observed mean BMI SDS for schools ranging from -0.07 to +0.78 (see Table 5.1). As expected, given the rising rates of obesity in the UK, the sample had a higher average BMI than the reference population. A number of schools had higher or lower measures of obesity than we expected given the Index of Deprivation of their location (this is highlighted in Figure 5.2, which uses Observed BMI SDS data).

The revised Expected rankings of the schools given the ethnicity and residential deprivation score of the pupils are also shown in Table 5.1. It can be seen that there is little difference between the Observed rankings and Expected rankings for the schools with the highest and

lowest obesity (mean BMI SDS). That is schools with the highest and lowest observed rankings are very similar to the schools with the highest and lowest expected rankings.

The last step in the process was to calculate the Value Added index. This was calculated by deducting the extent that the Observed mean BMI SDS differed from the Expected mean BMI SDS for children in Reception classes (i.e. the Reception residuals) from the same residual difference for children in Year 4 classes (i.e. the Year 4 residuals). Schools with a large negative Value Added figure have higher measures of obesity in Year 4 than might be expected, suggesting that the "value added" by the school for obesity risk was detrimental. Schools with a large positive Value Added figure have a lower level of obesity than expected in Year 4, suggesting that the schools may have a beneficial impact on their pupils' risk of obesity. The difference between these two sets of residuals forms each schools' Value Added figure, which we ranked. The last column in Table 5.1 shows the ranking of the schools according to this new index.

This Value Added ranking was markedly different from that derived from the previous analyses (a comparison of the Value Added index with the Expected ranking system shows little relationship: Spearman correlation coefficient -0.39, $p = 0.02$). The last column in Table 5.1 shows the schools ranking according to this measure. Comparing the figures in the right-most three columns for each school in Table 5.1 demonstrates how similar the Observed and Expected rankings are, and how different the Value Added ranking is to both the other rankings. Figure 5.3 shows the schools mapped against the Index of Deprivation for Leeds again, however this time using the Value Added ranking for the schools (as opposed to the Observed data used in Figure 5.2). The location of the schools in the top quartile using the Value Added ranking (schools ranked 1-9) are indicated as blue spots (cold spots: predicted obesity is lower in Year 4 than Reception classes), the middle two quartiles (schools ranked 10-26) are indicated as green spots, and the bottom quartile (schools ranked 27-35) are shown as red spots (hot spots: predicted obesity is higher in Year 4 than Reception classes). This mapping process clearly illustrates schools that are hot spots (suggesting factors at those schools may be adding to the obesity problem of their pupils), such as the school identified as "X" which is a hot spot in an affluent area, or cold spots (suggesting a positive school effect), such as the school identified as "Y" which is a cold spot in a disadvantaged area. Those schools that may be having a beneficial or detrimental effect on their pupils' obesity levels are therefore highlighted (in red or blue) and are likely to be of particular interest for further study.

Table 5.1: Data on 35 schools contributing data on children in Reception class (age 5 years) and Year 4 (age 9 years), with ranking of the schools by three methods

Schools[2]	Deprivation (mean of children)	Deprivation (location of school)	% south Asian	Observed Mean BMI SDS (SD) (Reception)	Observed Mean BMI SDS (SD) (Year 4)	Observed Mean BMISDS (SD) (both years)	Expected Residual[3]	Value Added index[4]	Observed Ranking[5]	Expected Ranking[6]	Value Added Ranking[7]
A	14.3	13.7	3.1	-0.09	-0.05	-0.07 (1.04)	-0.14	0.05	1	1	29
B	52.8	51.0	88.0	-0.11	0.17	0.03 (1.44)	-0.10	0.05	2	3	30
C	11.0	14.1	0.0	0.18	-0.11	0.04 (1.18)	-0.08	-0.03	3	4	13
D	17.1	16.9	5.1	-0.40	0.44	0.05 (1.05)	-0.11	0.30	4	2	35
E	21.1	14.9	37.9	0.10	0.09	0.09 (1.08)	-0.07	-0.01	5	5	19
F	35.5	22.3	0.0	0.34	-0.03	0.14 (0.92)	-0.07	-0.07	6	6	4
G	22.1	15.3	3.6	0.06	0.22	0.14 (1.21)	-0.06	0.04	7	7	28
H	13.7	12.5	4.0	0.13	0.26	0.19 (1.01)	-0.03	0.02	8	12	25
I	56.6	53.3	5.6	0.04	0.33	0.21 (0.95)	-0.05	0.04	9	8	27
J	8.5	6.6	0.0	0.20	0.28	0.24 (1.06)	-0.02	0.01	10	16	22
K	21.2	8.4	10.6	0.11	0.34	0.25 (1.00)	-0.02	0.02	11	15	24
L	13.3	13.9	26.7	0.25	0.26	0.25 (1.15)	-0.00	-0.04	12	19	10
M	31.4	16.0	65.6	-0.02	0.43	0.26 (1.28)	0.01	0.07	13	21	33
N	35.0	41.4	2.2	0.07	0.43	0.28 (1.04)	-0.04	0.09	14	11	34
O	45.7	25.5	0.0	0.27	0.30	0.28 (0.96)	-0.04	-0.01	15	9	16

Table continued on next page...

[2] The list of schools are anonymised and ranked according to the Observed mean BMISDS
[3] After adjustment for deprivation and ethnicity of children
[4] Value Added index = $[O_i - E_i]\ Y_4 - [O_i - E_i]\ Y_0$
[5] Ranking according to *Observed* mean BMI SDS
[6] Revised ranking according to *Expected* mean BMI SDS (following adjustment for deprivation and ethnicity of children)
[7] Revised ranking according to Value Added index

Schools[8]	Deprivation (mean of children)	Deprivation (location of school)	% south Asian	Observed Mean BMI SDS (SD) (Reception)	Observed Mean BMI SDS (SD) (Year 4)	Observed Mean BMISDS (SD) (both years)	Expected Residual	Value Added index	Observed Ranking	Expected Ranking	Value Added Ranking
P	49.2	30.3	1.7	0.39	0.19	0.30 (0.87)	-0.04	-0.06	16	10	6
Q	59.5	66.3	18.0	0.26	0.34	0.31 (1.27)	-0.02	-0.01	17	14	18
R	58.9	78.1	7.4	0.43	0.23	0.33 (1.30)	-0.03	-0.07	18	13	3
S	32.0	39.0	1.6	0.13	0.51	0.34 (0.95)	-0.00	0.06	19	17	31
T	14.1	14.2	21.3	0.17	0.59	0.36 (1.21)	0.03	0.07	20	24	32
U	40.0	34.4	0.0	0.49	0.30	0.37 (1.12)	-0.00	-0.06	21	18	7
V	23.8	12.1	0.0	0.24	0.51	0.39 (1.01)	0.01	0.02	22	22	26
W	27.3	12.8	11.7	0.31	0.48	0.40 (1.04)	0.03	-0.00	23	23	20
X	54.8	70.5	17.3	0.30	0.51	0.43 (1.11)	0.01	0.00	24	20	21
Y	13.3	18.0	2.1	0.38	0.57	0.47 (1.08)	0.05	-0.01	25	29	17
Z	8.1	2.0	0.0	0.43	0.51	0.48 (0.98)	0.08	-0.03	26	30	14
AA	34.0	48.7	11.6	0.69	0.37	0.50 (1.11)	0.04	-0.08	27	26	2
AB	25.3	28.6	0.0	0.56	0.54	0.55 (1.08)	0.05	-0.04	28	27	9
AC	62.2	66.2	5.8	0.59	0.53	0.55 (1.18)	0.04	-0.06	29	25	5
AD	45.9	48.4	3.4	0.45	0.65	0.56 (1.20)	0.05	-0.02	30	28	15
AE	22.5	19.6	3.6	0.52	0.72	0.62 (1.03)	0.12	-0.04	31	34	8
AF	11.1	11.0	2.5	0.55	0.75	0.64 (0.99)	0.09	-0.04	32	31	11
AG	52.3	53.1	2.6	0.53	0.80	0.66 (1.18)	0.10	-0.03	33	32	12
AH	38.2	73.2	0.0	0.75	0.61	0.68 (1.36)	0.12	-0.14	34	35	1
AI	38.3	30.7	0.0	0.54	0.91	0.78 (1.12)	0.11	0.02	35	33	23

[8] The list of schools are anonymised and ranked according to the Observed mean BMISDS

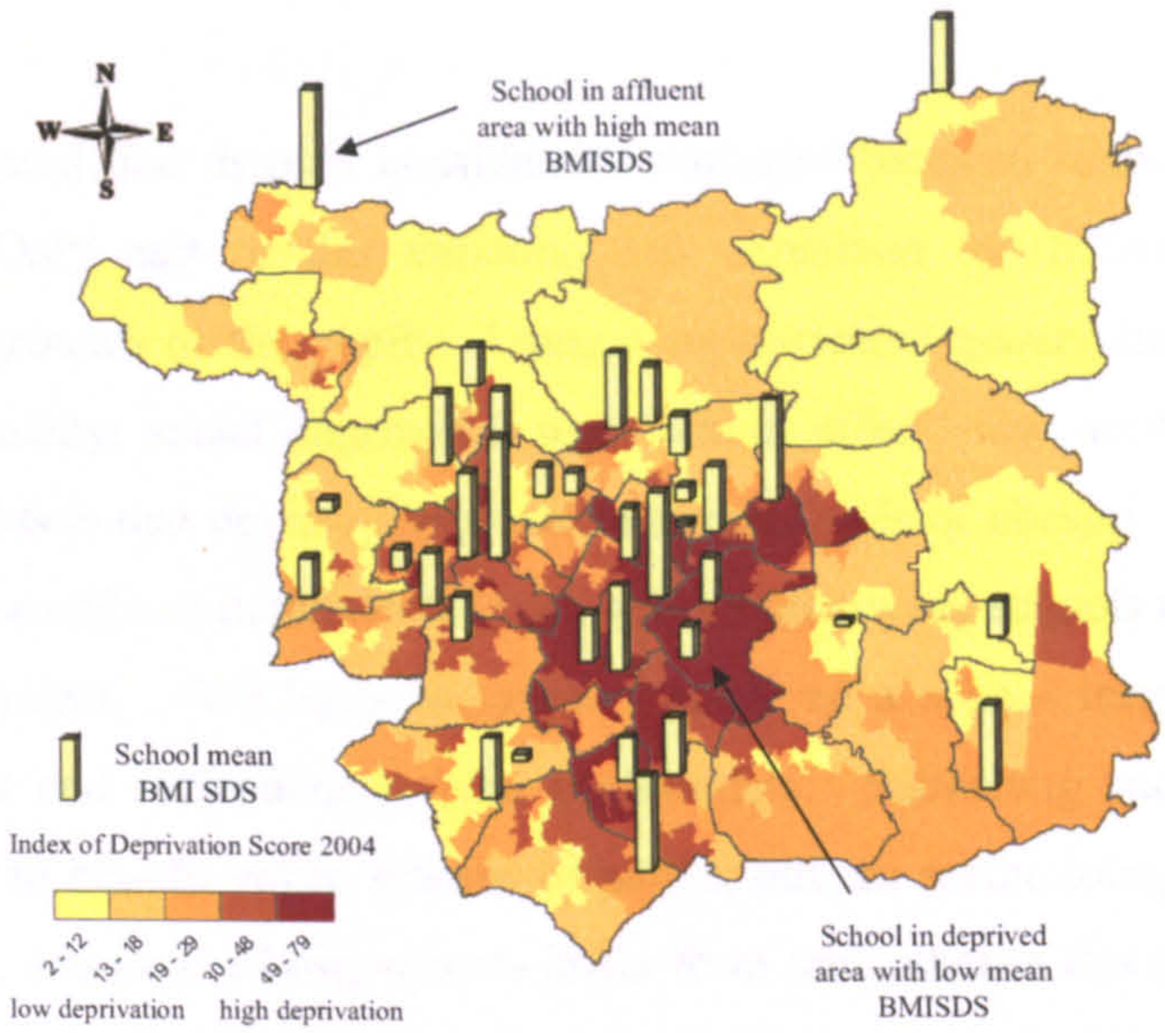


Figure 5.2: The geographic location of the primary schools related to indices of deprivation in Leeds. The Observed mean BMI SDS of each school is shown by the height of the bar. Each boundary area on the map represents a ward and the shading is undertaken at SOA in Leeds. The darker the background shading, the higher the level of deprivation in that area. This mapping exercise illustrates how some schools have discrepant measures of obesity than might be expected from the part of Leeds where they are located.

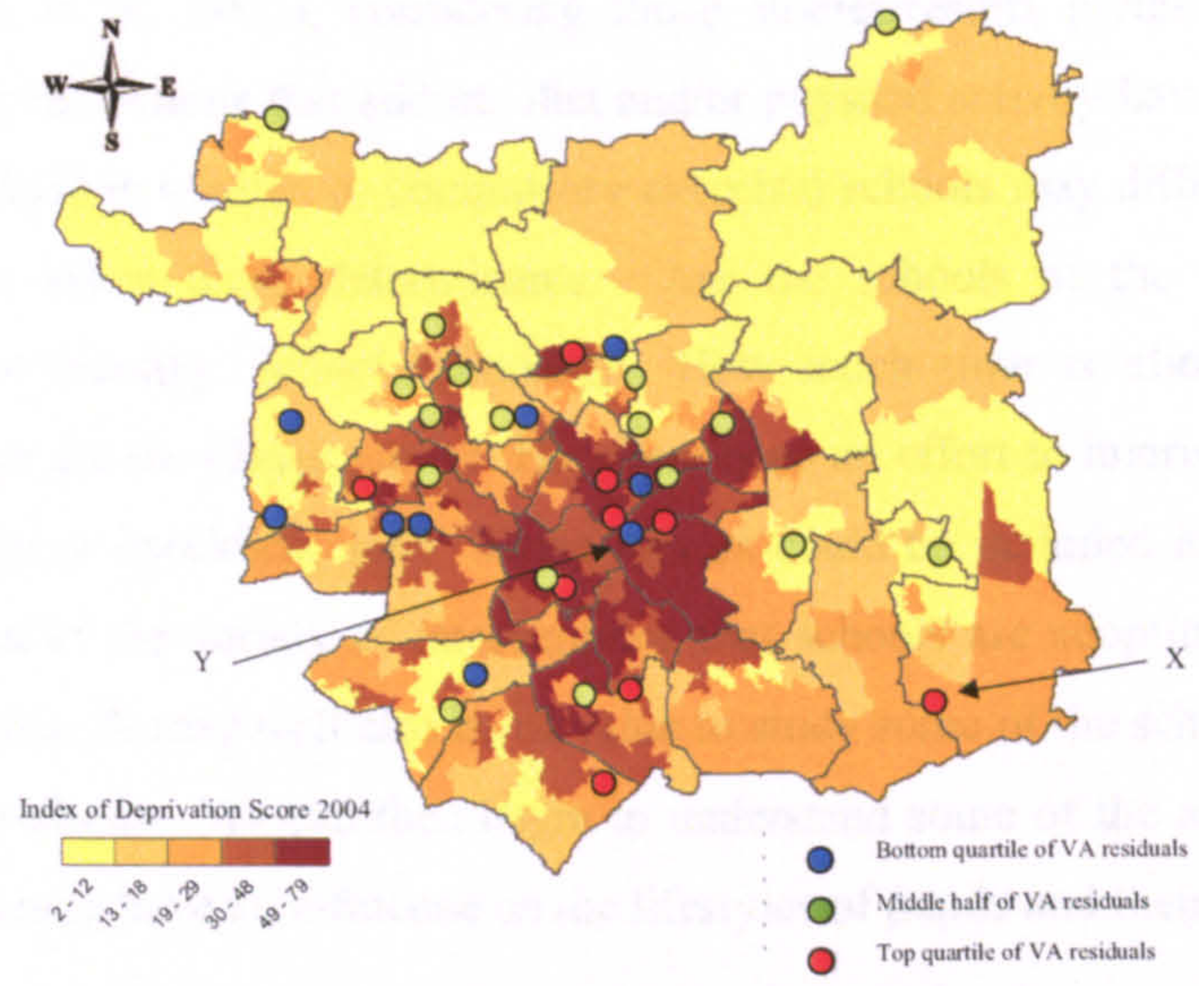


Figure 5.3: Hot and cold spots of obesity in primary schools in Leeds (as per the Value Added model). The green spots locate the primary schools whose BMI SDS were in the expected range for the deprivation indices and ethnicity of their pupils; the red spots where BMI SDS was higher (bottom quartile of ranked Value Added residuals) and the blue spots where the BMI SDS was lower (top quartile of ranked Value Added residuals). This allows identification of schools with unexpected measures of obesity. Two schools are highlighted. X is a hot spot of obesity in an affluent area. Y is a cold spot of obesity in disadvantaged areas. Please note, the boundary and deprivation data in this figure are the same as for Figure 5.2.

5.5 Discussion

We have demonstrated that there is considerable variation between schools in terms of their pupils' obesity. Only part of this variation was explained by differences in the socio-demographic backgrounds of the pupils. Using a sophisticated geographic modelling process we have taken ethnicity, social deprivation and "school effect" into account and so allowed identification of schools that deviate from their expected levels of obesity. The strength of our results is that they provide an indication of how well, or otherwise, schools are functioning in an obesogenic environment. We have, in essence, produced a school league table for obesity adjusted for school and socio-demographic confounders. Publishing such a table would of course be unlikely to benefit either schools or pupils, but the methodology does allow for an exploration of why, and maybe how, schools differ from each other in this regard. The next step must be to study the exceptional schools in depth, and to try to understand why some schools manage to combat the obesogenic environment to the benefit of its pupils while others do not.

We can begin to hypothesise why schools may differ in their impact, facilitated by reviews of interventions to prevent obesity in children (Summerbell et al, 2005; Flynn et al, 2006; Brown et al, 2007; Connelly et al, 2007), considering those studies/results pertaining to the school environment. As interventions that address diet and/or physical activity have been shown to be successful (particularly in relation to compulsory exercise) schools may differ in their approach to parameters that affect these determinants. Are the schools at the top of the league participating in the Healthy School Standard? How much time is allocated for physical activity? How large are the playgrounds? Has there been an effort to improve school lunches? Is there a tuck shop or breakfast club? These factors could be included in the model and so allow an evaluation of the variety of interventions that schools are adopting, as well as other environmental factors. It may well also be of value to study some of the schools in more depth, using qualitative methods. We can then begin to understand some of the attitudes and culture within schools that may have an influence on the lifestyles of pupils and their families too.

This chapter only offers a preliminary exploration of a methodology. Further study is required to validate the findings, and should involve all primary schools in the city, rather than a representative sample. The biggest concern must relate to the small size of the samples used to produce the "league tables". Measurements were only available for children in Reception classes and Year 4, and so the power may well be inadequate to identify hot and cold spots reliably and with confidence. Missing data accounted for on average just over 10% of pupils but did not seem to be systematically biased across schools. One way to confirm the league table's validity would be to repeat the process in subsequent years to see if there is some consistency in the results. The analysis would also be improved if measures were available on

children in every year rather than just at school entry and exit. If "school effect" is a valid hypothesis we would expect to see a dose response with increasing number of years at school. Increasing the number of children measured (and similarly the number of schools assessed) would also reduce the sampling error.

In this chapter we have explored the concept of a school Value Added index. This analytical process has its attractions as in theory it allows simple correction for deprivation and ethnicity as well as a number of other unknown confounding factors. It is based on the assumption that children at school entry are a reflection of socio-demographic and biological factors that influence the development of obesity. The argument in this chapter is that the Recpetion class children have not been at school long and so any effect of the school is minimal on these children. Conversely the Year 4 children have been attending school for some time and are more likely to be affected by school policies.

When we compared the Value Added index with the Expected ranking system we found little relationship. The Expected model predicts each school's average obesity for all children (i.e. Reception and Year 4 classes combined) given the deprivation and ethnic mix at each school. The Value Added model goes one step further and is based on how each school's current population of pupils in Reception and Year 4 vary (allowing for the relevant deprivation and ethnic mix). It is suggested that the difference in predicted obesity between year groups is due to the "school effect", however it may be due to other confounding factors. The models do not take into account changes that may be occurring at neighbourhood level that are impacting on obesity on top of any schools' influence, although adding the cross classification system to our model resulted in very little effect on the model estimates so any impact in relation to area deprivation is small. Nevertheless we recognize that there is a growing literature about the potential impact of the built and social environment on health and it may be that the Index of Deprivation is not an adequate description of neighbourhood. Accordingly in any future, larger, study it may be useful to record additional potential confounding variables in order to add them into the model on top of SES and ethnicity to improve its fit. For example, it may be possible to ask residents to complete a simple questionnaire regarding their perceptions of the neighbourhood, such as safety, availability of facilities and public transport. We suggest using data on a person's perception of a particular feature of their neighbourhood (e.g. how safe it is to use the local park), as this is arguable more important than the actual data (e.g. local crime statistics) in affecting people's behaviours. Further, geo-demographic data could be added (such as the Output Area Classification system (Vickers & Rees, 2007)). The data showed one potential outlier school due to a low BMI SDS for Reception year pupils: however the same trained measurer and equipment was used suggesting that there may be a genuine difference at this school in that year. Qualitative in-depth analysis of schools that do not follow the typical

pattern will help to elucidate on potential reasons for such discrepancies and whether (or not) it is a true difference.

This study uses cross-sectional, not longitudinal, data. Accordingly it is possible that there is a period effect where the children in the Reception year classes are fundamentally different to the children in Year 4 classes, for example, due to living in different neighbourhoods or to changes in school policies over time. However the spatial distribution of subjects in Reception and Year 4 classes are similar. That is, children in both class years at the same school live in the same neighbourhoods. Plus the effect of neighbourhood on the children is taken into account in the cross-classified multi-level model structure. Data regarding school policies was not collected, but in light of the recent spotlight nationally on childhood obesity it is likely that school policies have changed during the time the pupils in Year 4 have been at school. Accordingly it would be useful for a future qualitative study in the "outlying" schools to consider changes in schools' policies to see whether differing changes in policy before the measurement year appear to have affected BMI outcomes at the point of measurement.

The ranking of schools according to the impact they have on their pupils' risks of obesity may be worthwhile for another reason. Political focus often tends to be directed towards a total population approach. From a public health stance, that is only cost effective where there is a fairly uniform distribution of a condition. Otherwise it makes more sense to strategically target populations with the highest prevalence. (A good example is the dental caries programme where 90% of the disease is concentrated in 10% of the child population and targeting has successfully controlled the problem (Tickle, 2002)). Our identification of primary schools where the problem is greatest could allow a concentration of resources to the best advantage.

This chapter proposes a methodology for analysing growth data on children that has potential for producing public health benefits at relatively little cost. The cost of the measuring exercise itself has been estimated at £2.00 per child (Levine et al, 2007). The identification of hot and cold spots of obesity can allow allocation of resources to those communities who would most benefit. In addition, in depth study of schools which are particularly successful in curbing the development of obesity in its pupils may well lead to an understanding of those qualitative and quantitative factors that have an impact on children's lives. It could also provide interim direction regarding the benefits of interventions until more solid evidence emerges from randomised controlled trials. In this way, we may be able to learn a great deal more about how schools can promote resilience to the pervasive obesogenic environment to which children are exposed today. (This qualitative work was outside of the scope of this thesis, but hopefully funding for further work in this area will be awarded in due course).

Chapter 6: A spatial microsimulation model of obesogenic environments and behaviours in Leeds: SimObesity

6.1 Introduction
6.2 Small area population estimation methods
6.3 SimObesity Model Specification
6.3.1 Choice of obesogenic variables
6.3.2 Choice of constraint variables
6.3.3 Structure of SimObesity
6.3.4 Input data
6.3.5 Algorithm Methodology
6.3.6 Output data
6.4 Validation of synthetic micro-data
6.5 Discussion
6.5.1 Validation / Calibration
6.5.2 Methodology limitations
6.5.3 Uses of SimObesity
6.5.4 Conclusion

6.1 Introduction

Childhood obesity is an increasing problem in the UK and worldwide. It has been shown that there are hot spots of high prevalence of obesity in children in certain parts of Leeds (chapter 4). Accordingly there is a need to investigate micro-area data in order to be able to respond to local differences in the prevalence of health related behaviours and obesity in order to be able to develop targeted interventions and health policies that are based on locally relevant evidence. Some risk factors are already geo-referenced (e.g. deprivation), however many others are not (e.g. dietary variables, physical activity, social capital) but are available in national level surveys (such as the Health Survey for England). SimObesity is a spatial microsimulation model that is designed to synthesise obesogenic variables, including obesogenic environment and behaviour variables (the distinction is explained below), at the small-area level in Leeds in order that the relationship between actual obesity data and synthetic obesogenic covariate data can be analysed at the micro level. This is a novel application of spatial microsimulation modelling.

This chapter explains how the SimObesity model was built and how it facilitates the identification of obesogenic covariates (including obesogenic environments and behaviours) for children. The resulting synthetic data can be mapped and analysed. This spatial microsimulation model could also be used to address other research questions, depending on the data available in the various large, longitudinal, national level surveys, such as National Diet and Nutrition Survey or the British Household Panel Survey, that are available in the UK. The key benefits of using spatial microsimulation are: to add more attributes to the population under analysis by adding census data to the survey data thereby creating a richer dataset; to get data to a smaller geographical scale in order to identify "hot spots" of problem areas; and it is cheaper

and quicker than commissioning a survey of the local area. Furthermore it may be possible to use spatial microsimulation modelling to undertake "what if" scenario analysis to theoretically evaluate the potential impact of potential interventions on the prevalence of childhood obesity in say 5, 10 or 20 years time, which is cheaper and much quicker than running a pilot study.

SimObesity combines individual micro-data from the Health Survey for England 2002 (HSE), and separately from the Expenditure and Food Survey 2005 (EFS), which both only have location data at the scale of large areas, with census statistics for lower Super Output Areas (SOAs) to create synthetic micro-data estimates for SOAs in Leeds using a reweighting deterministic algorithm. The new, synthesised, micro dataset includes all the attributes from both the survey and the census datasets. This allows non geo-referenced variables, such as diet and social capital from the HSE and the amount spent on food from the EFS, to be directly estimated for SOAs, enabling detailed micro-level analysis of environments to be undertaken. It is not possible to determine childhood obesity from the HSE survey despite the inclusion of BMI data because exact age data are not included.

This chapter firstly explores the alternatives for estimating small area populations (section 6.2). Then it describes the SimObesity model specification, including the choice of variables to use in the model, and how the model is structured and works. The result is the creation of 715,169 / 715,167 individuals (from the HSE and EFS simulations respectively) in Leeds whose characteristics match as closely as possible the characteristics of the 715,402 actual individuals living in Leeds as shown in the 2001 Census (section 6.3). Finally, the validation of the synthetic micro-data is discussed in section 6.4 with a discussion of the results and methodology in section 6.5.

6.2 Small area population estimation methods

Microsimulation modelling is a methodology to synthesise large-scale population micro-datasets. A key feature in microsimulation models is whether they are aspatial or spatial. Aspatial microsimulation models have been successfully developed since their original conception in the late 1950s by Orcutt (1957), and are used extensively by economists e.g. budget analysis, measurement of poverty, impact of tax changes, policy impact assessment, etc. However there are also many medical applications, largely to analyse the effectiveness of medicines. These models focus on "who" is impacted (for example, by a proposed policy change) rather than "where" the impact occurs. Accordingly they are normally constructed at large spatial scales, such as national or regional level. Spatial microsimulation modelling is a more recent development, the history of which has already been thoroughly covered by other

authors (e.g. Ballas et al, 2005). They are concerned with where people reside and as such tend to be built down to the local level.

Spatial microsimulation modelling involves building spatially disaggregated large-scale micro-datasets on the attributes of individuals or households, often using a combination of information sources (such as, census data, hospital records, surveys). That is, spatial microsimulation models enable the production of synthetic micro data, typically combining data from the census with sample surveys to create synthetic households with lists of attributes. It models real life events by simulating the characteristics and actions of households in the area under consideration. Accordingly, when data of a characteristic of interest are not available at the required spatial scale, then these data can be synthesised using microsimulation: spatial microsimulation enables the generation of synthetic household estimates for non-spatially defined obesogenic factors or to bring spatially defined data down from higher spatial levels to household level.

However there are other methods to undertake small area population estimation, most notably the use of multi level modelling and agent based modelling (ABM). On the face of it use of multi level modelling to produce small area estimates for obesity is appealing as both individual and ecological factors affect the health behaviours that cause obesity. However to do this it is necessary to use data from a survey that includes both the disease variables and the risk factors data. For example, Twigg et al (2000) used the Health Survey for England to predict smoking and excessive alcohol consumption by age, gender and marital status. No survey exists that has childhood obesity data within it. The closest is the Health Survey for England, which includes body mass index data on many of the participants, but as there is no exact age given for the children (only age rounded to the nearest year), then it is not possible to determine whether children are obese, overweight or acceptable weight (by definition it is age and gender specific). It may be possible to use multilevel modelling to determine various obesogenic factors, but only for the few factors that are influenced at both the individual and ecological level, so this would be restrictive. Furthermore the multilevel models are constrained by the requirement for particular cross-tabulated breakdown of census data, which is not always available as desired (Twigg et al, 2006). To date work has been undertaken at the ward level, but there is no reason that this could not be undertaken at a finer geography, such as lower super output area, as census data are available at this level. ABM is a technique for modelling social behaviour. The software agent is constructed as a "self-directed object" using a set of predefined rules to decide what action to take (Brown & Xie, 2006). That is, a person will behave in a particular way given certain circumstances. However in this study it is not only health behaviour per se that is being modelling but also the broader obesogenic environment. This does not involve modelling

people's decisions or how they react to a particular situation, which is more akin to agent based modelling. Therefore spatial microsimulation modelling is more appropriate.

6.3 SimObesity Model Specification

Many different factors may affect the populations' risk of becoming overweight or obese, including the "obesogenic" nature of an environment as well as individuals' behaviours (see Figure 6.1). This study is looking to model these obesogenic covariates in Leeds by synthesising data for all individuals/households in Leeds using SimObesity, which is a spatial microsimulation model.

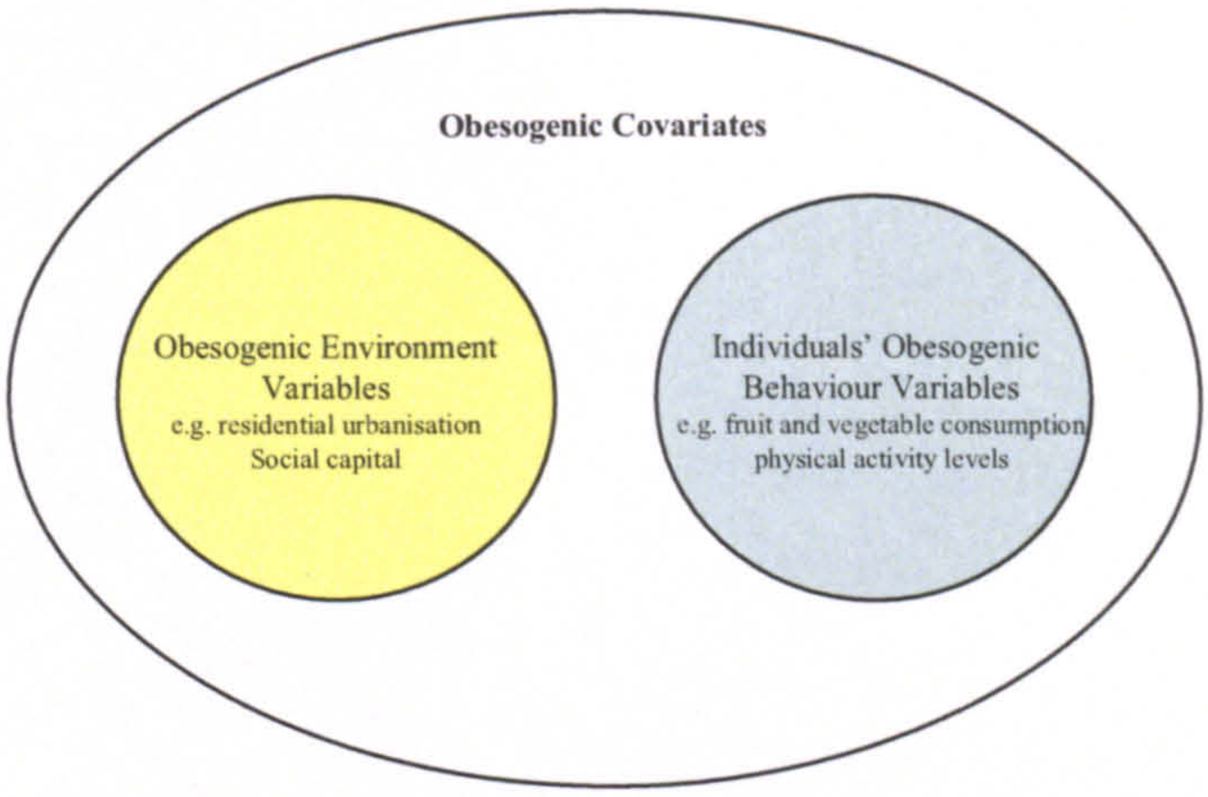


Figure 6.1. Venn diagram of obesogenic covariates - illustrating how both obesogenic environment variables and obesogenic behaviour variables are incorporated in the definition.

SimObesity was implemented in Java, an object-oriented programming language, which has been accepted as the most suitable type of programming language for spatial microsimulation modelling (Ballas, 2001). It can be operated on any computer system and platform without amending any code (i.e. it is platform independent). Note, a (large) part of this PhD involved writing the code for this programme[9] (see Appendix B for the structure of the java code).

The program implements a deterministic combinatorial optimisation reweighting approach to generate spatially disaggregated population micro datasets at the SOA level, synthesising the 715,402 people in the 2001 Census living in Leeds. Specifically the implementation of the spatial microsimulation approach to model obesogenic covariates involves selecting the combination of individuals from the population datasets that best fit the 2001 Census aggregate statistics for each of the small areas (i.e. SOAs in Leeds). This results in an individual-level dataset constrained by some census statistics with a list of attributes from both datasets. The key to the success of this modelling is the choice of input variables ("constraints") in the model, which must be predictors for the output variables.

---

[9] SimObesity was developed within the School of Geography, University of Leeds - particular thanks to Mr Kirk Harland who provided the bulk of the java code for the programme

In modelling obesogenic covariates for Leeds at the micro-level it was firstly necessary to establish a "wish list" of childhood obesity obesogenic variables to be simulated (e.g. TV viewing) and then to review and shorten this list based on being able to find a sufficiently disaggregated micro dataset to use as a population base dataset. This is described in section 6.3.1. Then the obesogenic variable list is further refined as it is only possible to synthesise those obesogenic variables that have a correlation with census-type variables (e.g. tenure) (within the same dataset). This analysis correspondingly affects the choice of constraint (input) variables and is detailed in section 6.3.2. Section 6.3.3 outlines the overall structure of SimObesity, highlighting the input and output variables, with section 6.3.4 detailing all of the input files required to run SimObesity. Section 6.3.5 explains the algorithms (and java code) underlying SimObesity. Finally section 6.3.6 outlines the output files from SimObesity and illustrates the resulting output from the two simulations.

The data for SimObesity comes from a combination of sources: data from the 2001 Census (Office for National Statistics, 2001a; 2001b) are combined in two separate simulations with the Health Survey England 2002 (HSE) (National Centre for Social Research, 2004) and the Expenditure and Food Survey 2005 (EFS) (Office for National Statistics and DEFRA, 2006). The fact that more than one national survey is being used is novel and it enables more obesogenic variables to be simulated. Census data has 100% coverage and is at a fine geographical scale, although availability of small area data is limited in order that confidentiality is preserved. Conversely the HSE and EFS are samples, albeit nationally representative, at a coarse geography. All are cross sectional.

The HSE and EFS surveys were selected to use in this study because both contain many obesogenic variables. Further these variables of interest were correlated with the census variables that will be used as constraints in the simulation, which is obviously key. These studies were sourced from the University of Essex UK Data Archives (http://www.data-archive.ac.uk/). Many other surveys were considered, including UK Time Use Survey (Ipsos-RSL & ONS, 2003), British Household Panel Survey (University of Essex, 2006), National Diet and Nutrition Survey (Office for National Statistics, 2005), Television: the public's view (Independent Television Commission, 2005), and SARS (individual or household level census data) (SARS, 2001), but they did not include relevant variables and/or the variables were not correlated with any census-type variables contained within the survey (analysis not shown). As Huang and Williamson (2001) point out, the quality of the synthetic micro-data is likely to be affected by the size of the sample used as a parent population. The larger the sample size, the more possible combinations of individuals exist and the better the fit is likely to be. Both the HSE and the EFS are large datasets containing many thousands of records.

The purpose of SimObesity is not to predict childhood obesity in Leeds, as there is actual data available for this; rather SimObesity is to provide data on the obesogenic environment and behaviour (such as diet and physical activity) covariates in order that relationships with obesity in Leeds can be determined. By simulating both types of covariates, it may be possible to control the analyses for the behaviour covariates. This "big picture" needs to be born in mind when looking for datasets to use to simulate obesogenic covariates in Leeds.

6.3.1 Choice of obesogenic variables

There are three stages to the process of deciding which variables to simulate at the micro-level for spatial analysis. Firstly the aetiology of childhood obesity is complex and multi factorial, so it is impossible to study all obesogenic factors. Accordingly it is necessary to focus on several key obesogenic factors identified as having a significant relationship with BMI from a review of the literature. Secondly even if a variable is identified as having an important relationship with childhood obesity, the ability to simulate it depends upon being able to find a reliable dataset that includes that variable. In choosing a dataset to use for spatial microsimulation, it is important that the dataset is as disaggregated as possible. That is, as well as data on obesogenic variables it also includes demographic and census information about the interviewees. This is because it is these census variables that are used as the "input variables" (or constraint variables) for the simulation. The third step is to ensure that within the chosen dataset the obesogenic variable has a strong correlation with the constraint (input) variables for the spatial microsimulation model. (This last step is intertwined with the choice of constraint (input) variables and is discussed separately in section 6.3.2).

The literature review on the aetiology of childhood obesity (Chapter 2) determined many potential obesogenic variables that were of interest to simulate, which were born in mind as national datasets (from the University of Essex data archives) were trawled looking for current disaggregated datasets that included these variables for children. Accordingly datasets such as the British Household Panel Survey were discarded, as there was little, if any, data on children. Three datasets that were determined as containing suitably disaggregated childhood obesogenic variables were the Health Survey for England, the Expenditure and Food Survey and the UK Time Survey. The variables included dietary and physical activity variables, social capital (such as access to supermarkets, transport links and neighbourhood safety), parenting style and expenditure on food (see Table 6.1). Additionally, three obesogenic variables that will not be included in this analysis but that deserve a special mention are parental BMI, fast food consumption, and access to green spaces.

Health Survey for England 2002 (HSE) (National Centre for Social Research, 2004):
1. Energy density is the key issue with consumption of fast food. Accordingly a good compromise may be to use a figure for average daily energy intake per day for each child instead, although this figure would need to be adjusted for age and gender (as calorific intake is dependent upon these factors).
2. Diet is obviously associated with obesity. The HSE dataset has many different dietary variables: number of portions of fruit and/or vegetables per day; amount of fat consumed; amount of cheese/red meat/fried food/fish/chocolate, crisps, nuts, biscuits/ cakes eaten per day
3. Social capital ▪ Transport links may be associated with childhood obesity. The HSE asks participants for their opinion on how good local public transport is. ▪ Similarly safety in a neighbourhood may affect childhood obesity, perhaps by reducing the propensity to spend time outdoors engaging in physical activities, such as walking or playing, or possibly it may affect shopping habits in some way. The HSE asks participants whether there are problems with teenagers hanging around in their neighbourhood, or whether there are vandalism or graffiti problems. These could be used as markers of perception of safety. ▪ It will also be worth considering the HSE data on the participants' perception of the quality of local leisure facilities. ▪ Similarly re the HSE data on participants' perception of the ease of getting to the supermarket.
4. This dataset also includes data on physical activity levels. For children this is described as the number of hours of activity per week, and for adults as the number of days per week when they are active for at least 30 minutes.
5. Degree of urbanisation may affect the obesogenic nature of the environment and can also be considered.
Expenditure and Food Survey 2005 (EFS) (Office for National Statistics and DEFRA, 2006):
6. Food purchasing patterns and expenditure on food may be linked with changing rates of body mass and the development of obesity in children. The EFS dataset has the amount spent on many different foods (such as bread, cakes, beef, milk, vegetables, fruit, etc) as well as food eaten outside the home and all data is given per household.
7. There are many other variables in the EFS dataset that may allude to an obesogenic environment, including various indices or markers of wealth or deprivation, and ownership of various goods. Considering these factors will also enable the consideration of whether the correlation between obesogenic factors and the prevalence of childhood obesity varies with household income levels.
UK Time Survey 2000 (Ipsos-RSL & ONS, 2003):
8. A number of different surveys have information on hours of TV viewing and computer time (collectively "media time"). TV: The Publics View 2003 (Independent Television Commission, 2005), even though the data are very disaggregated (e.g. SEG, ethnicity, economic activity, etc), actually only has media time data for adults (and we can't assume that adult viewing is the same as child viewing). Similarly the British Household Panel Survey (the current latest version is 2005) (University of Essex, 2006) only has data for children older than 11 years. The UK Time Survey 2000 could be used, as it has media time data for children from 8 years old (albeit only whole years of age, no date of birth information to calculate exact age).
9. Other possible markers of physical inactivity include reading and studying, which are also included in the UK Time Survey 2000. This could be added to total media time to calculate the total sedentary time.
10. Levels of children's physical activity could be reviewed. The UK Time Survey 2000 has information on several physical activity variables, as minutes per day. This includes total physical activity plus the two sub-categories of sport and walking/cycling as travel.
11. Similarly sleep may be associated with obesity. The individual minutes per day of sleep are included in the UK Time Survey 2000.
12. Parenting style may have an impact on childhood obesity rates. A marker of family environment factors is the amount of contact time between parent and child. The TV dataset has data on the time (minutes per day) spent with parents (proxy for parenting style), with family and with friends.

Table 6.1. Details of obesogenic variables available in UK national level surveys.

Parental BMI is strongly associated with childhood obesity. Adult (and child) BMI are disclosed in both the Health survey for England 2003 and the National Diet and Nutrition Survey (16-64) 2000. The HSE has the advantage of having children of all ages in the dataset, which enables a more accurate population (i.e. adult and child) to be simulated. However the simulated population whilst generating an accurate adult:child, and male:female population, doesn't specifically tell us which adult lives with which child (assuming that defines a parental relationship). Accordingly it will not be possible to analyse parental BMI as an obesogenic covariate.

Fast food consumption is associated with obesity. Accordingly ideally this study would have included a variable for the number of fast food meals a child eats per week (fast food frequency), however it was not possible to find a national survey dataset (searched University of Essex data archives) on frequency of fast food consumption of children. For example, the HSE survey also does not include fast food data. Similarly neither does the National Diet and nutrition Survey (4-18 years) 1997. Data from a study in Leeds (Cadet) gives daily diet nutrient content (e.g. total saturated fat), frequency of consumption of daily foods consumed (e.g. crisps, sugar sweetened drinks, etc) as well as total daily energy intake (Cade et al, 2006). It does not include any data on fast food intake. The UK Time Survey (Ipsos-RSL & ONS, 2003) has information on time (minutes per day) spent in restaurants, cafes or pubs. It also has details of total eating time (individual minutes per day of eating). However it is not valid to use external eating time as percentage of total eating time as a proxy for frequency of eating out because, for example, a long meal in a restaurant is very different to a quick meal in a fast food venue. The amount spent on external food vs. home food (using the Expenditure and Food Survey) would also be an inaccurate proxy of fast food consumption (as restaurant meals are often/normally more expensive than fast food meals). Also direct approaches to fast food outlets asking for data were rebutted (unsurprisingly perhaps, given the focus of this study on childhood obesity). An alternative way of analysing fast food information (rather than simulating the data) would be to use a geographic information system (GIS) to calculate the fast food outlet density. In this vein a list giving the geographical location of all fast food outlets in Leeds has been obtained, which could be used to determine fast food outlet density (as simple location is not necessarily associated with consumption). But this dataset is inaccurate (a quick ground-truth analysis showed many fast food outlets were missing from the list). Also even if the list were accurate it would be a massive task to accurately calculate the fast food outlet density, which would need to be in relation to a buffer zone from each individual child's home (as census boundaries are not relevant when deciding where to shop, in that the adjacent SOA may be closer or more convenient than that within a person's "own" SOA), which is not really the onus of this study.

Finally there is varying data on the association of availability of green spaces and childhood obesity, although the consensus seems to be moving towards no relationship (Pearce J, personal communication, paper in process of publication). Again details of the geographic information for the location of all council run parks and green spaces in the Leeds area has been obtained, although this does exclude any privately run public green spaces in Leeds. There is also a small survey regarding the participants' favourite parks and safety in those parks. However the main problem is how to determine "access" as opposed to "proximity", which would require spatial interaction modelling to include variables such as car ownership and distance, as well as extensive GIS work. So this variable is also not included in this study.

6.3.2 Choice of constraint variables

A key difficulty with spatial microsimulation modelling is determining which constraints will be used. This spatial microsimulation model, SimObesity, works by selecting individuals (or households) from a study population (i.e. HSE and EFS) to match certain "constraint variables" of the actual population in that area (i.e. census data) (the algorithms used are explained in section 6.3.5 below). Accordingly the possible constraint variables are limited to what census-type variables are included in the study population datasets.

This section explains the third step in the choice of variables, which involves two intertwined processes. Firstly to identify the obesogenic variables that will be simulated in this study - only those obesogenic variables in the "shortlist" described in Table 6.1 that have a correlation with the input (constraint) variables can be simulated. The reason for this is explained below. As the census provides the known data necessary for the constraint variables, the same census-type value must be in the constraint files (available in Census) as well as the population files (available in HSE and EFS). Secondly it is necessary to identify potential input variables (i.e. census-type variables) that are contained within the population datasets, and to choose which ones will be used to simulate the obesogenic variables – which should be the census-type (input) variables with the strongest correlations with the obesogenic (output) variables. Both processes occur concurrently. In summary, the HSE and EFS population datasets did show correlations between some of the obesogenic variables and census variables whereas the UK Time Survey did not (analysis undertaken on the UK Time Survey is not described here). Accordingly it is possible to simulate obesogenic variables from the HSE and EFS datasets, but not the variables from the UK Time Survey.

The choice of which variables to use as constraints is key. Obviously different choices result in different final synthetic populations (Huang and Williamson, 2001), although the aim is always to replicate the real population as closely as possible. Generally, the more constraint variables

used the better the synthetic micro dataset produced (i.e. more similar to the real population). However, it has to be born in mind that the more constraint variables added in, the more comparisons with the real data would be required, which means more time will have to be spent running the model. In practice, the number of constraints would be restricted to no more than about six, with a limited number of categories, otherwise it would be easy to end up in a situation where each synthesised household may fall into one of millions of different possible states. Obviously, disaggregated data that are not used as constraints ("non-constraints") would remain attached on the final list of attributes and could be used for validation purposes. Accordingly SimObesity has been set up to accommodate a maximum of six constraint variables.

Initially in this study, the determination of which variables to use as constraints in the microsimulation model was based on using the variables that produced the lowest errors; synthetic micro-populations were simulated using many different combinations of constraint variables, and the levels of errors were compared to see which permutation produced the lowest results (not reported here). However this methodology is flawed; overlooking the crucial fact that a spatial microsimulation model only accurately and reliably simulates the constraint variables and any variables correlated with the constraint variables. That is, the constraint variables and obesogenic variables being simulated must be strongly correlated for an accurate simulation to occur. Note, no one particular paper or author emphasizes this point, although the concept is alluded to: for example, "difficulty lies in identification of which characteristics are dependent upon others" (Clarke, 1996).

Accordingly, non-constraint variables (i.e. census-type variables included in the population dataset that are not used as constraint variables) may or may not simulate well – one just doesn't know. Other authors (e.g. Ballas et al, 2005; unpublished discussions in CSAP, School of Geography, University of Leeds) discuss using these non-constraint variables to validate the model. This methodology is problematic. It might be that a non-constraint variable simulates well due to correlations with the input variables or due to chance; more likely it won't simulate well, resulting in high errors. Furthermore, whether or not the non-constraint variables have simulated well has no bearing on how well the constraint variables and all variables correlated with these input variables have simulated. Calibration (as opposed to validation) of the model by comparing simulated and actual data for the constraint variables will confirm whether the simulation of the constraint variables is good and the errors low. Furthermore, if the obesogenic variables have a strong correlation with the constraint variables, then if the constraints simulate well it is a safe statistical assumption that the obesogenic variables simulate well. This is important as it is not possible to compare the simulated obesogenic variables to actual data (by definition; else why would one be simulating the values?).

This concept is illustrated graphically in Figures 6.2 and 6.3. Both graphs are plots of simulated data against actual (census) data for the same census-type variable, i.e. tenure. Obviously with a perfect simulation the data would lie on a 45° regression line (i.e. y = x). Figure 6.2 shows the results of the simulation when tenure is included as a constraint (input) variable; this shows an almost perfect fit of the simulated against actual. When census-type variables were not used as constraint variables, the data simulated very badly. For example, Figure 6.3 shows the results of the simulation when tenure is not included as an input variable, with all other things equal; the fit around the 45° regression line is poor and the R square value (describing the fit of the data around its actual regression line) is also low.

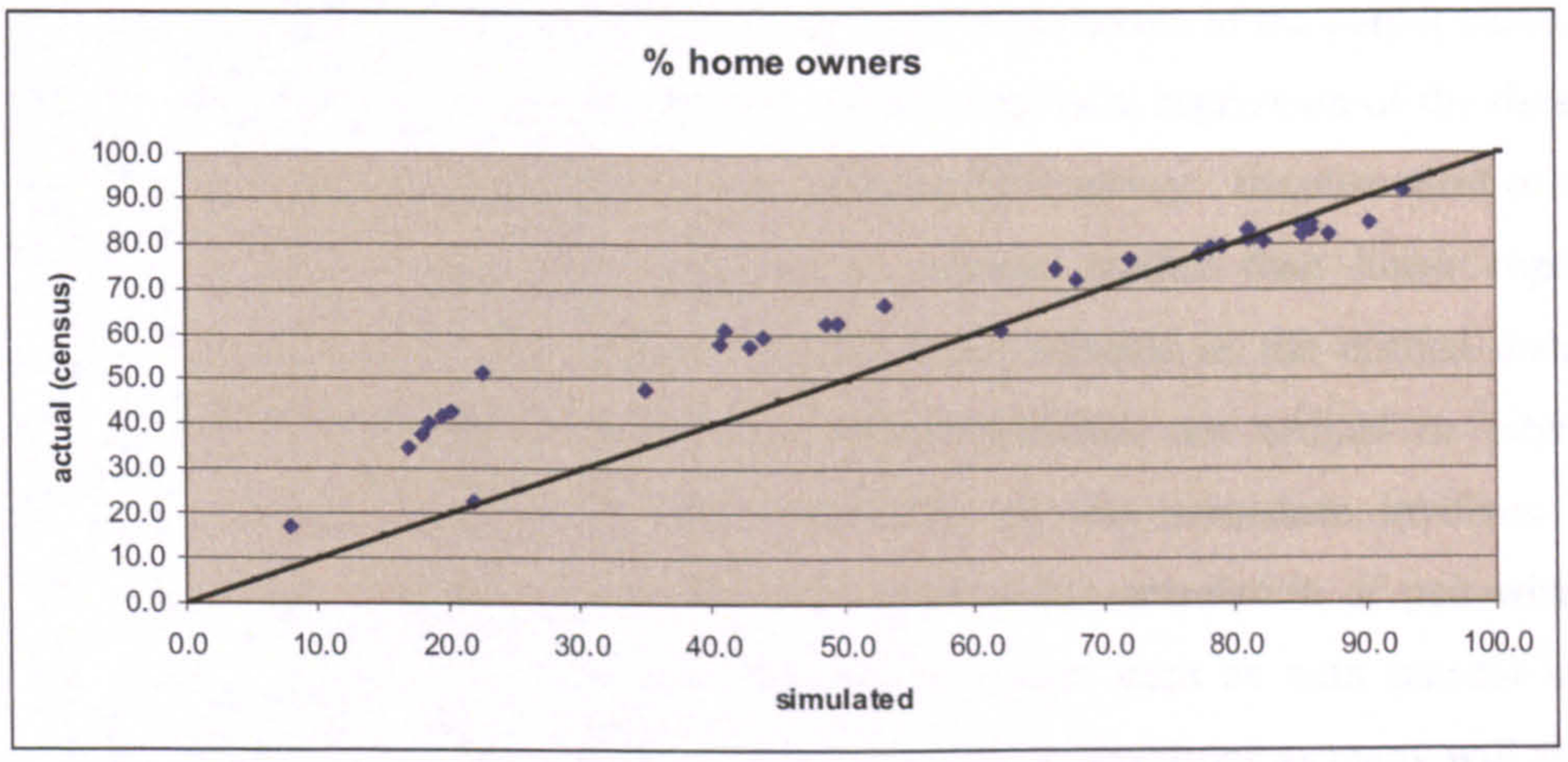


Figure 6.2. Scatterplot of simulated (x axis) and actual (census) (y axis) values for the proportion of households that own their own home (with or without a mortgage) for the Leeds wards. It is based on a simulation of the HSE dataset, using gender, age, household type and tenure as constraints. The line is at 45° (i.e. y = x), as if the simulation was perfect the actual value and simulated value would be the same. $R^2 = 0.93$. This clearly shows that the simulated proportion of home-owning households is very similar to the actual proportion of such households.

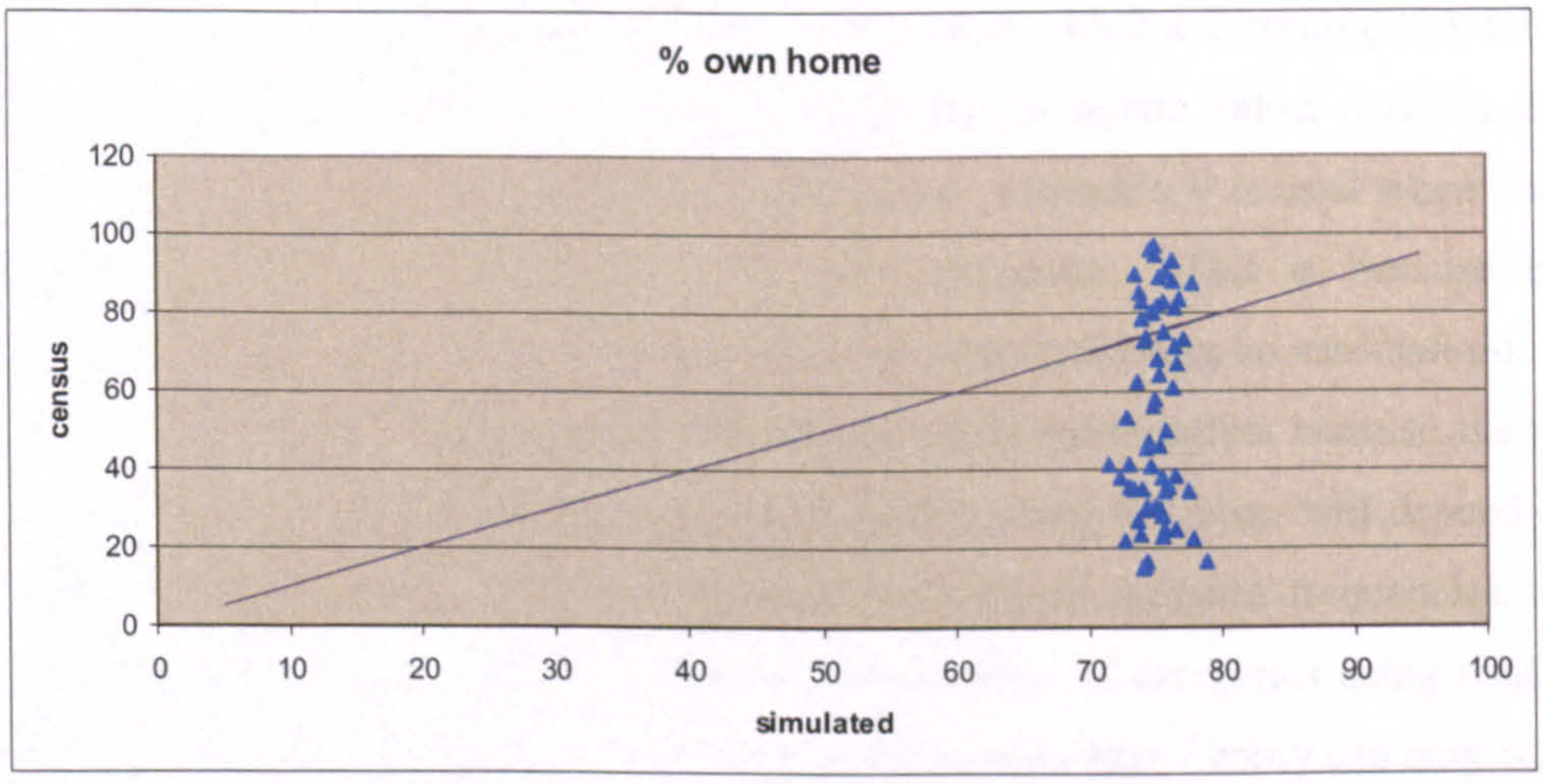


Figure 6.3. Scatterplot of simulated (x axis) and actual (census) (y axis) values for the proportion of households that own their own home (with or without a mortgage) the Leeds wards. It is based on a simulation of the HSE dataset, using gender, age and household type (i.e. not tenure) as constraints. The line is at 45° (i.e. y = x), as if the simulation was perfect the actual value and simulated value would be the same. $R^2 = 0.02$. The simulation values are limited to the range of 70-80% of the population being home-owners, rather than the actual range of 10-90%.

As such, the choice of constraint variables should be determined by the existence of a strong correlation between the factor of interest to be simulated (i.e. obesogenic variables, such as food expenditure) and census-type data that are available in that survey (i.e. the constraint variables such as tenure). These constraint variables are termed "optimisation constraint variables". However it may also be that some variables need to be accurate in the simulated dataset (e.g. child's age and gender in order to define obesity), so these data may be included as constraint variables irrespective of whether a correlation exists with the output obesogenic variables. These constraint variables are termed "control constraint variables".

To establish which census-type variables available in each population dataset would make the best (optimisation) constraint variables (that is, be the best predictors of the output variables) we undertook a number of correlation analyses as well as a logistical regression of the data, which are described below, firstly for the HSE survey then the EFS survey. Because most of the data are nominal, the use of logistical regression is required (rather than linear regression). Similarly, use of Pearson's correlation coefficient is not suitable as the normal distribution assumption fails. Furthermore the data are mainly nominal, not ordinal or interval, so Spearman's correlation coefficient is also unsuitable, as this procedure involves ranking observations for each variable in turn, and then involves the correlation of pair-wise ranks. Thus where one variable is dichotomous (binary) in nature, such as with marital status or household type, it may not be able to discriminate between observations as many will share tied ranks on these variables. Accordingly for this dataset we could use a Pearson's chi square test to detect whether there is any significant association between two categorical variables. However it does not say anything about the strength of the association. Therefore it is also necessary to use a Phi or Cramer's V test to measure the strength of the association between two categorical variables (whether nominal or not). Phi is used with 2 x 2 contingency tables (i.e. both variables are binary); it is calculated by taking the chi square value, dividing it by the sample size and then taking the square root of this value. Cramer's V is used when one of the two categorical variables contains more than two categories. This is because in these circumstances phi fails to reach its minimum value of zero (indicating no association). Trying to calculate the mean of a categorical variable is obviously meaningless because the numeric values we attached to the different categories are arbitrary and the mean will depend on how many members each category has. Accordingly we need to measure frequencies. So we analyse the number of things that fall into each combination of categories using contingency tables. The key assumptions for this analysis are: each person / item / entity can only contribute to one cell of the contingency table (e.g. can not use on a repeated measures design); expected frequencies should be greater than five.

The strength of the correlation also needs to be taken into account. Effect sizes are useful because they provide an objective measure of the importance of an effect. Correlation coefficient of 0 means there is no effect; a value of 1 means there is a perfect effect. The parameters of what constitutes a large or small effect is widely accepted: r = 0.10 (small effect) – the effect explains 1% of total variance; r = 0.30 (medium effect) - the effect explains 9% of total variance; r = 0.50 (large effect) - the effect explains 25% of total variance (Cohan, 1992).

Constraint choice for Health Survey for England

The first step in identifying the constraint variables to use in SimObesity is to understand the dataset. This enables potential output variables (i.e. obesogenic variables) and constraint (input) variables (i.e. census-type variables) to be identified. The Health Survey for England (HSE) is a multi-stage stratified random sample designed to monitor trends in the nation's health. The survey focuses on different health issues in different years, although a number of core questions are included every year; including a focus on cardiovascular disease, minority ethnic groups, the health of older people, and, in 2002, a focus on children and young people (with a boost sample of young people). The population surveyed is adults (aged 16 and over) and children (aged 0-15 years) living in private households in England. Data is collected by interview with the household members as well as self-completion questionnaires, followed by a nurse visit, which was extended to include additional procedures at one sixth of the households. The survey was conducted throughout the year to take into consideration seasonal differences.

The HSE 2002 has a child boost survey, so has data on more children (9461 children from a total sample of 18398). Also it contains extra information that is not included in other years of this survey; in particular it has data on children's physical activity levels. However it does not have the detailed dietary survey that the HSE 2003 has (although it transpires that this data is missing for most of the 4358 children in that sample of 18553 people; so is not of use to this study anyway). The HSE 2002 does still have details of fruit and vegetable consumption. Whilst the HSE dataset does include anthropometric measurements, such as BMI, it does not include the exact age, only a rounded age on the date of the nurses visit. Accordingly it is not possible to calculate an accurate BMI SDS (a few months can make a big difference - see Table 6.2) and thus to determine which children are obese or not. The HSE 2002 also has information about degree of urbanisation and social capital. These latter data include questions on "does the area have good local transport", "... good leisure things for people", "...problem of teenagers hanging around", "...problem of vandalism, graffiti or deliberate damage", and "is it easy to get to the supermarket". There is also data on household income and socio-economic group (but in a way that does not tie up with census data), but there are many people with missing data, particularly for total income.

Any variables that are also covered by the census can be included as a constraint variable; this is because it is essential that all constraint variables are both included in the population dataset (i.e. HSE dataset) and in the census. There are ten census-type variables available in the HSE dataset (i.e. variables that are compatible with the classifications given in the census): household type (simplified to with or without children); car (or van) availability (car/van available or not); household tenure (home owner or not); gender; age (0-2 years, 3-6 years, 7-9 years, 10-12 years, 13-15 years, 16-18 years, adults: these categories were chosen as they give more detail about children's ages and tie in with the ages of the BMI data that is available for this study i.e. 3-6 years, 9 years, 11 years and 13 years); marital status (married/cohabiting or not); ethnic group (white or not); deprivation score (five categories from least to most deprived); top qualification achieved by the individual (degree or equivalent; A or O levels or equivalent; none or other qualifications; full time student or child); economic activity (economically active or not). All of these variables can be compared to known census data. So these are the variables that could be used as constraint variables.

Age at last birthday	Age at nurses visit (rounded)	sex	BMI	Potential exact age	BMISDS	Centile	Obese?
0 years	0 years	female	19.256	0.1	2.932	99.8	Yes
				0.2	2.229	98.7	Yes
				0.3	1.777	96.2	No
				0.4	1.493	93.2	No
				0.5	1.326	90.8	No
5 years	6 years	male	18.816	5.5	2.071	98.1	Yes
				5.6	2.058	98.0	Yes
				5.7	2.044	98.0	Yes
				5.8	2.029	97.9	No
				5.9	2.012	97.8	No

Table 6.2. This table takes data from the HSE 2002 and demonstrates the difference a few months in age makes to whether a child is defined as obese (> 98th centile) using the British reference dataset classifications.

In "cleaning" the dataset, people with missing data were excluded (see Table 6.3), leaving 9386 children and 8131 adults in the final dataset for analysis. For the urbanisation variable only two people had missing data so they were excluded. For social capital, the other key output variable, these questions were only answered by people aged 16 upwards, and adults without a response for any one of the five social capital questions were excluded (n=728). That means the social capital for an area can be determined from an adult's perspective and applied to children living in that area. There is information about the number of portions of fruit and vegetables consumed, but these data are sporadic for young children (some of whom may not be eating solids yet; some of whom may just have poor diets). So this variable can be analysed for people from age 5 upwards. The physical activity data are good for children aged 0-15 years; the few with missing data were excluded (n=29). All 16-18 year olds (n=1376) and most adults have no physical activity data. Rather than exclude these interviewees, the physical activity analysis can be undertaken for children (aged 0-15 years), and the simulated physical activity data for adults

will need to be examined to ensure those simulated individuals with missing data are excluded from the calculations. The household income and socio-economic group variables are not "key" obesogenic variables that this study wishes to simulate, but they may provide some useful information. However total household income has many people with missing data (n=3335), and less so for socio-economic group where less people have missing data (n=702); accordingly rather than exclude the people without these data, it will be analysed if the simulated population do not have too many people with missing data. Also any constraint variables can not have missing data, so although at this stage in the analysis the final constraint variables had not yet been chosen, as the numbers of missing data were small (n=122), all people with missing census-type variable data were excluded. The resulting HSE 2002 dataset for simulation and analysis has a much higher proportion of children compared to the 2001 Census, because of the child boost sample. This affects any variable affected by child status: i.e. age, household type, qualifications achieved, marital status, and economic activity. See Table 6.4.

It is preferable to do one simulation to synthesise one population for all of the output variables of interest (from both a statistical and computational perspective), thus it is necessary to determine the optimal combination of constraint variables that suit all the output variables. As there can only be six constraint variables it is necessary to decide which six are optimal across most of the output variables. It is clear that in modelling the child environment, the gender and age of the child needs to be accurate in the simulation. Accordingly gender and age will be "control" constraint variables. The other type of constraint variable is an "optimisation" constraint. These are variables that are correlated with the output variables in order to ensure the output variables simulate accurately (as described above). This leaves capacity for a maximum of four "optimisation" constraints (as the maximum number of constraints is six) from the remaining census-type variables.

The second step in determining the variables for SimObesity was to undertake a logistical regression analysis in order to determine how well the census-type variables (e.g. tenure) predict the outcome variables. A stepwise logistic regression analysis was used as there were ten potential census-type variables yet only a maximum of four could be used as optimisation constraint variables, plus age and gender as control constraint variables. Accordingly a binary categorisation of each of the output variables was input as the dependent variables and all of the census-type variables were input as covariates. Whilst a stepwise approach can be affected by random variation in the data, which may result in non-replicable results, it is nevertheless appropriate to use in this situation as effectively rather than testing a theory (whereby a forced entry method, with all of the covariates placed into the regression model in one block and parameter estimates calculated for each block, would be more appropriate), this analysis will determine which census variables predict obesogenic variables the most strongly. The results are summarised in Table 6.5.

	Children	Adults	Total
Original number of people	9461	8937	18398
Exclude:			
Child physical activity data	29	0	29
Urbanisation	2	0	2
Social capital	0	728	728
Household type	1	3	4
Car availability	5	3	8
Tenure	22	23	45
Ethnic group	16	5	21
Top qualification	0	35	35
Economic activity	0	9	9
Final dataset	9386	8131	17517

Table 6.3. Details of the number of interviewees excluded from the HSE 2002 dataset, highlighting the variable with the missing data.

Proportion of population	HSE 2002	Census 2001 (Leeds)
N	17517	715402
Who are male	47.2	48.3
Who are female	52.8	51.7
Aged 0–2 years	8.1	3.4
Aged 3-6 years	11.1	4.8
Aged 7-9 years	8.8	3.8
Aged 10-12 years	8.9	4.1
Aged 13-15 years	8.7	3.9
Aged 16-18 years	7.9	3.8
Aged 19+ years	46.4	76.2
Live in household without children	42.5	53.7
Live in household with children	57.5	46.3
Least deprived	16.1	13.3
Deprived 2	15.1	16.5
Deprived 3	18.0	18.4
Deprived 4	21.8	19.1
Most deprived	29.0	32.7
Highest educational qualification is	7.3	14.0
HEQ is A or O levels, or equivalent	25.8	31.3
No or other qualifications	11.0	34.7
Full time student / child	55.9	20.0
Who are white	89.0	91.9
Who are other ethnic group	11.0	8.1
Who are home owners	69.6	65.3
Who rent or "other"	30.4	34.7
With at least 1 car or van available	82.8	66.4
Without car or van available	17.2	33.6
Who are married or cohabiting	27.9	37.7
Who are not married	72.1	62.3
Economically active	29.2	42.9
Not economically active	70.8	57.1

Table 6.4. Comparison of HSE 2002 (cleaned) to the 2001 Census data for the census-type variables contained within the HSE dataset. The HSE 2002 contains a child boost, so the child variables are higher than for the census proportions.

Predictors	-2LogLikelihood	Chi sq	Nagelkerke's adjusted R^2
Fruit and vegetables			
qualifications	14142	318	.034
age	14018	442	.047
Ethnicity	13968	492	.052
Tenure	13909	551	.058
Deprivation	13871	589	.062
Economic activity	13855	605	.064
Household type	13842	618	.065
Marital status	13827	633	.067
Car availability	13810	650	.068
sex	13805	655	.069
Child Physical activity			
Age	8591	1417	.227
Ethnicity	8540	1468	.235
sex	8503	1506	.240
deprivation	8492	1516	.242
Urbanisation			
Deprivation	16969	1122	.096
Ethnicity	16594	1498	.127
Car availability	16505	1587	.134
Qualifications	16482	1606	.136
Economic activity	16472	1619	.137
Age	16449	1642	.139
Martial status	16438	1653	.140
Socio Economic Group			
deprivation	21970	1099	.085
Qualifications	21109	1961	.148
Tenure	20761	2308	.172
Ethnicity	20719	2350	.175
Age	20699	2371	.176
Economic activity	20690	2379	.177
Marital status	20683	2386	.177
Public transport			
Deprivation	10909	166	.025
Sex	10872	203	.031
Ethnicity	10831	244	.037
Qualifications	10815	260	.039
Household type	10804	270	.041
Car availability	10794	281	.042
Marital status	10789	285	.043
age	10785	289	.044
Access to leisure facilities			
Age	12413	105	.015
Deprivation	12353	164	.024
Ethnicity	12309	209	.030
Tenure	12283	235	.033
Qualifications	12259	259	.037
Marital status	12247	271	.038
Access to supermarket			
Car availability	4111	112	.032
Economic activity	4064	158	.046
Marital status	4052	170	.049
Qualifications	4031	191	.056
tenure	4026	196	.057
Problem Teenagers			
Deprivation	11984	448	.063
Marital status	11919	513	.072
Tenure	11887	545	.077
Qualifications	11847	585	.082
Economic activity	11834	598	.084
Age	11824	608	.085
Ethnicity	11819	613	.086
Household type	11814	618	.087
Vandalism			
Deprivation	11485	612	.087
Tenure	11431	665	.094
qualifications	11389	708	.100
Age	11379	718	.101
Economic activity	11371	726	.102
Ethnicity	11363	734	.104
Car availability	11355	742	.105

Table 6.5. Results of the stepwise logistic regression analysis for the HSE dataset using all possible constraint variables as predictors for the output variables. It shows the ranking of the predictors for each output variable, as well as some key statistics for each covariate.

The difficultly with this dataset is that different output variables have different optimal combinations of constraints, so choosing a combination that works for all outputs is very difficult; the solution is to find the best compromise for all output variables. That is, to find the variables that are important predictors for many of the output variables in order that that combination of constraint variables that will permit all output variables to be synthesised in the same simulation, minimising the errors for all variables. On a "popularity" basis, that is those constraints that most often appear highly ranked, then the four optimisation constraints would be deprivation, ethnicity, qualifications and tenure.

Accordingly this "optimal compromise" of constraint variables (sex, age, deprivation, qualifications, ethnicity, and tenure) was run through a logistical regression using an entry method to assess the fit; i.e. how well these constraints predict each of the output variables. Table 6.6 summarise this analyses. Ideally for a good fit of the regression model the following is sought. The –2 log-likelihood statistic should be as low as possible as this shows that the model is improving at predicting the output variable more accurately, but this isn't comparable across models for different output variables. Instead look for a high chi square, which measures the difference between the model as it currently stands and the model when only the constant was included. The degrees of freedom allows the determination of whether the chi square is significant, and all the chi squares in Table 6.6 are clearly strongly significant. The column giving Nagelkerke's adjusted value of R^2 shows to what degree this model accounts for the variability in the data; so the higher the better. Similarly the last column gives the Hosmer-Lemeshow goodness-of-fit statistic, which assesses how well the chosen model fits the data, testing the hypothesis that the observed data are significantly different from the predicted values from the model; so a high number (a non-significant value – indicating that the model does not differ significantly from the observed data) is desirable. Table 6.6 shows how this constraint combination is a good fit for all of the output variables. Also the model using this combination is explaining much of the variability in the data for many of the output variables; in particular, physical activity 24%, urbanisation 13%, SEG 18%, vandalism 10%.

Output variable	-2Log Likelihood	Chi sq	Degrees of freedom	Nagelkerke's adjusted R^2	Hosmer-Lemeshow statistic
Fruit and vegetables	13855	605	15	.064	.133
Physical activity	8490	1519	11	.242	.094
Urbanisation	16547	1545	16	.131	.066
Socio Economic Group	20691	2379	16	.177	.173
Transport	10813	262	11	.040	.621
Leisure facilities	12258	260	11	.037	.222
Supermarket access	4117	105	11	.031	.100
Problems with teenagers	11855	577	11	.081	.128
Problems with vandals	11367	730	11	.103	.325

Table 6.6. Results of logistical regression (entry method) analysis using sex, age-full, ethnicity, deprivation, qualifications and tenure as constraints to predict each of the output variables.

A further variable that should be considered in the final model is the Wald statistic; this has a chi square distribution and tells us whether the b-coefficient for that predictor is significantly different from zero. If the coefficient is significantly different from zero then we can assume that the predictor is making a significant contribution to the prediction of the outcome variable. As we might expect, given that the constraint combination in these models is a compromise across all the output variables, the Wald statistic is not significant for all constraints for most of the output variables (analysis not shown here). However, positively, all output variables had at least three of the six constraint variables with significant Wald values (most had four or five), suggesting that for these output variables these significant constraints were predicting the output variable well. Three constraints are sufficient for a good simulation.

In conclusion, gender, age, deprivation, qualifications, ethnicity and tenure all play important roles and are important predictors for the output variables seen in this dataset (namely fruit and vegetable consumption, children's physical activity levels (both behavioural covariates), urbanisation, SEG, public transport, leisure facilities, shop access, teenagers and vandals (obesogenic environment covariates)). Therefore these variables will be included as constraint variables in the spatial microsimulation model for obesogenic covariates, SimObesity.

Constraint choice for Expenditure and Food Survey

Similarly the first step in identifying the constraint variables to use for the EFS simulation with SimObesity is to understand the dataset. This enables potential output variables to be identified (i.e. obesogenic variables) as well as the constraint (input) variables (i.e. census-type variables). The Expenditure and Food Survey (EFS) brings together the former Family Expenditure Survey (FES) and the National Food Survey (NFS) since 2001. As with the FES and NFS, the EFS continues to be primarily used to provide information for the Retail Prices Index, National Accounts estimates of household expenditure, analysis of the effect of taxes and benefits, and trends in nutrition. There are two questionnaires, one about income and one about expenditure, as well as covering demographic information. In addition every household member (aged 16 years and over) completed an expenditure diary for two weeks. The latest version of the EFS was utilised, which was from 2005. It has data on a total of 6798 people. No one was excluded from the analysis as there was no missing data.

The principal output variable of interest was how much each household spends on food per two-week period (which can also be adjusted for the number of people living in the household). This is the variable that the choice of constraint variables was based on. However, there are other output-type variables that may be of interest to simulate (assuming the constraint variables chosen show a strong enough correlation) and this is discussed shortly.

Any census-type variable included in the EFS could be used as a constraint variable; this is because it is essential that all constraint variables are both included in the population dataset (i.e. EFS dataset) and in the census. Note as the survey is undertaken at household level, individual details, such as age and gender, are not available. There are five census-type variables available in the EFS dataset that are compatible with census definitions: household type (simplified to with or without children); car (or van) availability (car/van available or not); household tenure (home owner or not); type of property (detached, semi-detached, terraced, flat, and other); number of people living in the house (1 person, 2 people, 3 or 4 people, and 5 or more people). All of these variables can be compared to known census data. Note the EFS also has data on economic activity, but it is not structured in a manner that makes it comparable to the census economic activity data. So these five variables could be used as constraint variables. Table 6.7 compares the EFS 2005 dataset with the 2001 Census. Note, the EFS dataset also has data on the number of males/females and children/adults living in each household. The way the algorithm is set up (discussed in section 6.3.5 below) means these data cannot be used as input variables.

Proportion of population	EFS 2005	Census 2001 (Leeds)
N (households)	6798	301614
N (individuals)	16257	715402
Who are male	48.5	48.3
Who are female	51.5	51.7
Who are children	24.6	23.8
Who are adults	75.4	76.2
1 person lives on own	27.5	30.6
2 people live in the household	36.4	33.3
3 or 4 people live in household	29.0	28.8
5+ people live in household	7.1	7.3
Live in household without children	68.2	53.7
Live in household with children	31.8	46.3
Who are home owners	70.2	65.3
Who rent or "other"	29.8	34.7
Proportion with at least 1 car available	75.2	66.4
Proportion without car available	24.8	33.6
Who live in a detached house	22.6	15.5
Who live in a semi or terraced house	60.3	67.5
Who live in a flat or other accommodation	17.1	17.0

Table 6.7. Comparison of EFS 2005 and 2001 Census data. Note, the EFS dataset has 6798 households, comprising of 16257 individuals.

The second step in determining the variables for SimObesity was to undertake a logistical regression analysis in order to determine how well the census-type variables (e.g. tenure) predict the outcome variables. Accordingly a binary categorisation of food expenditure (i.e. greater than or less than £60 per two weeks) was input as the dependent variable and all five of the census-type variables were input as covariates. This analysis is only being undertaken for food expenditure as this is the primary output variable; the other variables are all "extras" that could add further detail to the simulated population. As there were only five census type variables

which is less than the maximum number SimObesity is constrained to, the theory was tested that these five covariates would predict food expenditure significantly well. Accordingly a forced entry logistical regression method, with all of the covariates placed into the regression model in one block and parameter estimates calculated for each block, was utilised.

All of the predictor variables are added to the model as each one significantly improves the predictor power of the model. As described above for a good fit of the regression model the–2 log-likelihood statistic should be as low as possible (-2LL=6664), which is better demonstrated by a significant chi square value (chi square (df 8) = 2745; $p < 0.001$). Accordingly this new model is significantly better at predicting food expenditure than it was with only the constant included (the model chi-square is an analogue of the F-test for the linear regression sum of squares). R square (Nagelkerke's adjusted value) showed that this model, with all five constraint variables, accounted for 44% of the variability in the data. The number of people in each household accounted for the most variability, with home tenure following. The remaining three covariates accounted for a small proportion of variability. The Hosmer-Lemeshow goodness-of-fit statistic assesses how well the chosen model fits the data, testing the hypothesis that the observed data are significantly different from the predicted values from the model; so a high number (a non-significant value – indicating that the model does not differ significantly from the observed data) is desirable, which is the case here (HL=0.500). Further key statistics are summarised in Table 6.8. In particular the Wald statistics are mostly significant: the b-coefficient for all predictors except the type of property (semi-detached/terraced) is significantly different from zero, suggesting that the predictors (with the noted exception) are making significant contributions to the prediction of the outcome (food expenditure).

However, as alluded to earlier there are many other variables that may be of interest to simulate, in order to create a more powerful obesogenic environment/behaviour picture. However these variables can only be simulated if they are correlated with (most of) the five chosen constraint variables. As such a simple correlation analysis (using Cramer's V or Phi, due to the nominal nature of the data) was undertaken to ascertain if this was the case. The results for the variables that do have correlations with these census-type variables are given in Table 6.9. This means that ownership of a PC and more than one television, internet access, income level, source of income and spending on school meals were also correlated with the census-type variables and will be synthesised using SimObesity.

	B	(SE)	Wald	df	Sig.	Exp (B)	95% CI Exp (B) (lower)	95% CI Exp (B) Upper
SimHH(1)	.422	.104	16.365	1	.000	1.525	1.243	1.870
SimCars(1)	-.609	.086	50.118	1	.000	.544	.459	.644
Ten(1)	.764	.078	96.122	1	.000	2.147	1.843	2.502
hse2			24.574	2	.000			
hse2(1)	.164	.112	2.161	1	.142	1.178	.947	1.467
hse2(2)	-.193	.095	4.146	1	.042	.824	.684	.993
Nofolks_cd			872.766	3	.000			
Nofolks_cd(1)	-4.573	.189	582.601	1	.000	.010	.007	.015
Nofolks_cd(2)	-2.369	.167	201.034	1	.000	.094	.067	.130
Nofolks_cd(3)	-.784	.145	29.386	1	.000	.457	.344	.606
Constant	1.488	.161	85.532	1	.000	4.428		

Table 6.8. A summary of the key statistics for the logistic regression analysis of how well the five census-type variables available in the EFS 2005 predict expenditure on food. The b-value represents the change in the logit of the outcome variable associated with a one unit change in the predictor variable. The logit of the outcome is simply the natural logarithm of the odds of Y occurring. So if it is greater than 1 then as the predictor variable increases, the odds of the outcome (in this case spending less than £60 on food) occurring increases, and vice versa. For this interpretation to be reliable the confidence interval of Exp(B) should not cross 1. The crucial statistic is the Wald statistic, which has a chi-square distribution and tells us whether the b-coefficient for that predictor is significantly different from zero.

Outcome	Census	Phi / Cramer's V	R^2	Chi sq	df	p	Strength
PC	HH	.304	0.09	629	1	< 0.001	Medium
	No_cars3	431	0.19	1261	2	< 0.001	Medium
	Ten	.211	0.05	301	1	< 0.001	Weak
	HseType	.203	0.04	279	5	< 0.001	Weak
	SimFolks	.419	0.18	1194	3	< 0.001	Medium
Internet	HH	.250	0.06	425	1	< 0.001	Weak
	No_cars3	.439	0.19	309	2	< 0.001	Medium
	Ten	.254	0.07	439	1	< 0.001	Weak
	HseType	.220	0.05	329	5	< 0.001	Weak
	SimFolks	.371	0.14	937	3	< 0.001	Medium
Income	HH	.189	0.04	243	2	< 0.001	Weak
	No_cars3	.329	0.11	1473	4	< 0.001	Medium
	Ten	.337	0.11	772	2	< 0.001	Medium
	HseType	.203	0.04	560	10	< 0.001	Weak
	SimFolks	.201	0.04	547	6	< 0.001	Weak
Source2	HH	.288	0.08	563	2	< 0.001	Weak
	No_cars3	.369	0.14	1854	4	< 0.001	Medium
	Ten	.337	0.11	772	2	< 0.001	Medium
	HseType	.194	0.04	512	10	< 0.001	Weak
	SimFolks	.289	0.08	1203	6	< 0.001	Weak
SchSum2	HH	.511	0.26	1775	1	< 0.001	Strong
	No_cars3	.168	0.03	191	2	< 0.001	Weak
	Ten	.046	0.00	14	1	< 0.001	None
	HseType	.101	0.01	70	5	< 0.001	Weak
	SimFolks	.468	0.22	1490	3	< 0.001	Medium
TV+	HH	.201	0.04	275	1	< 0.001	Weak
	No_cars3	.301	0.09	617	2	< 0.001	Medium
	Ten	.193	0.04	253	1	< 0.001	Weak
	HseType	.286	0.08	555	5	< 0.001	Weak
	SimFolks	.357	0.13	865	3	< 0.001	Medium

Table 6.9. Details of output variables that are included in the EFS dataset that also have correlations with the five constraint variables chosen to simulate food expenditure, and so can also be simulated as part of the same model. The strength of the correlation was determined using the following cut-offs: $r > 0.10$ (small effect); $r > 0.30$ (medium effect); $r > 0.50$ (strong effect).

Other variables that were analysed (not shown here) but that had no or weak correlations with the census type variables included council tax payments, wealth index, ownership of telephone, television, digital television, video or DVD player, microwave, tumble dryer, washing machine, fridge and gas mains. Any output variable that does not have a good correlation with the input variables (constraints) is likely to be synthesised with high errors (and there would be no way to check this, so we would have to assume a poor fit). The EFS dataset also has details on the amount each household spends on individual foods, such as cake and soft drinks. However these data were not considered; partly due to inconsistencies across the data (impossible to tie up the definitions and avoid double counting) and partly due to time constraints. These individual food variables will not be simulated.

In conclusion, household type, car availability, tenure, property type and number of household members are important predictors for food expenditure (a behavioural covariate), and these variables will be included in the spatial microsimulation model, SimObesity. Using these constraint variables, SimObesity will also simulate spending on school meals (behavioural covariate), ownership of a PC and more than one television, internet access, income level, and source of income. The latter variables are arguably behavioural covariates (e.g. individuals choose whether to buy a computer), but given a person's income (which affects disposable income and purchasing ability) is not necessarily a 'choice', these are categorised as environment covariates (see Figure 6.1).

6.3.3 Structure of SimObesity

This study is seeking to model the obesogenic environment and behaviours for Leeds at the micro level. No one survey contains all the variables of interest. Accordingly two surveys are separately being used as population datasets for two separate simulations in order to simulate more obesogenic variables. SimObesity combines the outputs of the UK 2001 census with the Health Survey for England 2002 data, and separately with the Expenditure and Food Survey 2005 data, using deterministic algorithms (see Figure 6.4). That is, fruit and vegetable consumption, children's physical activity levels, urbanisation, SEG, and social capital (public transport, access to facilities, neighbourhood safety) will be simulated using the HSE 2002 as the population dataset, with gender, age, deprivation, qualifications, ethnicity and tenure as constraint variables. Similarly, household expenditure on food, ownership of a PC and more than one television, internet access, income level, source of income and spending on school meals will be simulated using the EFS 2005 as the population dataset, with household type, car availability, tenure, economic activity, and number of household members as constraint variables. These variables are summarised in Table 6.10. The resulting synthetic micro-populations are aggregated to SOA.

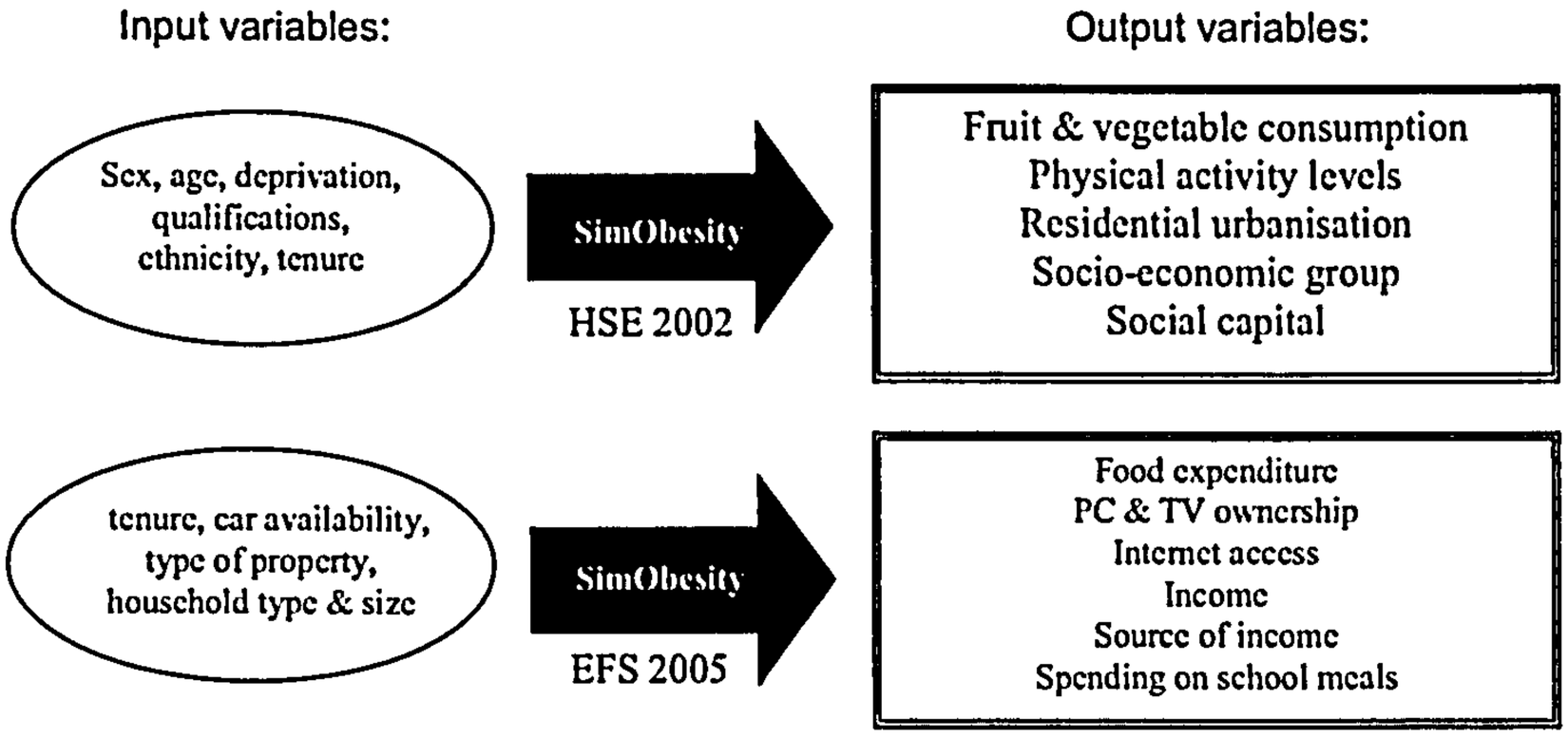


Figure 6.4. Summary of structure of the spatial microsimulation, showing what variables will be simulated and the input variables required to do this.

HSE 2002		Census 2001
Sex	⇨	Sex
Age	⇨	Age
Index of Deprivation	⇨	Index of Deprivation
Qualifications	⇨	Qualifications
Ethnicity	⇨	Ethnicity
Tenure	⇨	Tenure
Economic activity		Economic activity
Household type		Household type
Car availability		Car availability
Marital status		Marital status
Fruit and vegetable consumption		
Physical activity levels		
Urbanisation		
Socio-economic group		
Social capital (public transport, access to facilities, neighbourhood safety)		
EFS 2005		**Census 2001**
Tenure	⇨	Tenure
Household type	⇨	Household type
Car availability	⇨	Car availability
Household size	⇨	Household size
Property type	⇨	Property type
Food expenditure		
PC ownership		
TV+ ownership		
Internet access		
Income		
Source of income		
Spending on school meals		

Table 6.10. Details of the variables used in the two simulations. The variables linked with an arrow were used as constraint variables. Variables that are not also in the census column are the obesogenic variables being simulated. All of these variables are available in the resulting synthetic micro-populations.

An alternative structure to combine two population surveys would be to use a probabilistic algorithm for the second survey. However this method would be very time consuming (as a second java programme, for the probabilistic algorithm, would need to be written and tested). Plus the "deterministic plus deterministic methodology" used in this study is statistically

equivalent to the "deterministic plus probabilistic methodology", because the chain conditional probabilities used in the probabilistic methodology would be determined from the first dataset anyway. The only advantage conferred by a "deterministic plus probabilistic methodology" is that it produces one synthetic micro-population with attributes from both the EFS and HSE at individual level, rather than SOA level as per this study's method. However as all the analysis is undertaken at, at least, SOA level, the extra work is not justified.

6.3.4 Input data

There are four important types of files are needed to run the model (and these files are derived from the above-mentioned data sources, namely the HSE, EFS and census data):

1) Population file
2) Constraint files
3) Summary files
4) Micro-config file

1) Population file

There are two population datasets: one is the Health Survey for England 2002, the other is the Expenditure and Food Survey 2005. The population files are derived from these cleaned datasets. There is only one population file per simulation; accordingly each population file is used in a separate simulation (not together). These files consist of a column for a weighting variable to ensure the data is a representative sample (if none, put 1), columns for each of the constraint variables (in the order that the simulation will run in), and an ID number for each individual (see Table 6.11). The order of the columns in this table is vital, plus there can be a maximum of six constraint variables between the weight and ID columns, so a maximum of eight columns. More details of the population files are given in Appendix C.

Weight	Gender	Age	Ethnic	ID
1.00	1	7	2	10109101
1.00	2	7	2	10109102
1.28	2	4	2	10109103
1.46	2	2	2	10109104
1.00	1	7	1	10110101
...	...	...	...	...

Table 6.11. Example of population dataset, assuming the constraint variables are gender (binary), age (7 categories) and ethnicity (binary). This assumes gender is the first constraint variable, and household type is the last. There is a row for each person, so depending on the dataset used this is likely to be thousands of rows. The larger the sample size, the more possible combinations of individuals exist and the better the fit is likely to be.

2) Constraint files

SimObesity requires a number of constraints to be input, which are the variables by which the study population (i.e. HSE or EFS) and the census population are compared, in order to come up with the best selection from the study population to be the synthesised population. That is, the constraints are the census variables that are being used to simulate the output variables, to ensure that the synthetic populations created accurately represent the "real" individuals in that area. For computational ease the maximum number of constraints is six, so there can only be up to six constraint files; the number of categories used should also be kept to a minimum (due to computational, not technical, limitations - too many will result in the computer crashing). In light of this, both the HSE and the EFS simulation use six constraint variables (the choice of constraint variables is described in section 6.3.2), most of which are binary. A particular exception to this is age, where a more detailed breakdown of the children's ages is required in order to consider obesity in children.

The constraint files were derived from the census tables listed in Table 6.12, which were extracted from CasWeb (http://casweb.mimas.ac.uk/). The exception is deprivation, which was determined from the Index of Multiple Deprivation 2004 available from National Statistics (the relevant department is now called Communities and Local Government: http://www.communities.gov.uk/). The categorisations match that in the population file. All data is at SOA.

Table Code	Table Name	Level	HSE simulation	EFS simulation
UV003	Sex	All people	Constraint	Na
UV004	Age	All people	Constraint	Na
UV046	Household composition	All people	Non constraint	Constraint
N/A	Index of multiple deprivation 2004	All SOAs	Constraint	Na
UV024	Qualifications	All people	Constraint	Na
UV043	Tenure	All people in households	Constraint	Constraint
UV062	Cars or vans	All households	Non constraint	Constraint
UV028	Economic activity	All people	Non constraint	Constraint
UV051	Number of people living in households	All occupied household spaces	Na	Constraint
UV042	Accommodation type	All people	Na	Non constraint
UV009	Ethnic group	All people	Constraint	Na
UV007	Marital status	All people	Non constraint	Na

Table 6.12. Details of the CAS univariate census tables used for the constraint data, highlighting which variables were available in the two population datasets and whether the census variable was used as a constraint or not.

These constraints are input into the model using constraint files: see Table 6.13 for an example (the full listing of the constraint files is given in Appendix D). The more categories a constraint has the more columns these tables will have. The order of the columns is important: the first must be the area code, which has to be named in the manner shown in Table 6.13 (so SimObesity can accept either SOAs or OAs, but not postcodes or postal sectors); subsequent columns relate to the categories and must be in the order of the number assigned to the population dataset, that is for example, "white" would be coded 1 and "not white" would be coded 2. The rows represent each area: for example in SOA E01011264 there are 1434 white individuals (the numbers are not integers due to the input adjustment described below).

ZoneCode	White	Not_white
00DAFAE01011264	1434.0	4.0
00DAFAE01011265	1458.0	33.0
00DAFAE01011266	1391.9	65.1
...	...	...

Table 6.13. Example of a constraint variable table: ethnicity. Each column is for a category of the variable (in this case binary) and each row is for each area. This uses the low SOA level so there would be 476 rows (at OA there would be 2440 rows). There is one of these tables for each constraint variable.

The constraint tables "tell" the programme how many people live in each area as defined by any one category, using the census counts data. For example, the number of males, the number of children, etc. This is the ideal aggregate number that the model is trying to simulate by selecting individuals from the HSE (or EFS) datasets.

However it should be noted that there are inconsistencies between the census tables produced by official disclosure control measures in order to protect confidentiality. This means that there are often slightly different numbers of people in the different tables for a given output area. For example, if SOA E01011264, which is one of the 17 SOAs in the Aireborough ward is considered, the sex table (UV003) states that 1438 and 25617 people live in E01011264 and Aireborough respectively, whereas the tenure table (UV043) says 1439 and 25215 people live in the two areas respectively. The reweighting algorithm would not work optimally under these conditions; the totals need to be the same. It should be noted that there is no way of deriving a true estimate of the number of residents or households prior to the imposition of disclosure control.

Accordingly the aggregate values in the constraint tables for each SOA are adjusted to ensure the totals always sum to that of the sex tables. That is, the number of people in the small areas we want to populate are adjusted by using the sex table (UV003) to give the number of individuals for each SOA. The number of people in each cell in the constraint tables is adjusted using the formula in Figure 6.5 (and is demonstrated in Table 6.14):

$$\text{Number of people from the constraint table} \times \frac{\text{Total sum for each area of the sex table}}{\text{Total Sum for each area of the constraint table}}$$

Figure 6.5. Formula used to adjust the constraint data to ensure the total number of people for each variable is the same.

SOA	UV046: own	UV046: rent/other	UV046: total	UV003: total	Adjusted: own	Adjusted: rent/other	Adjusted: total
E01011264	1067	372	1439	1438	1066.26	371.74	1438
E01011265	1090	333	1423	1491	1142.09	348.91	1491
E01011266	1288	70	1358	1457	1381.90	75.10	1457
…	…	…	…	…	…	…	…

Table 6.14. This is an example of the input adjustment for the tenure constraint table. For each SOA the number of people in each of the categories (e.g. people who own their home) is multiplied by the sum for that SOA in the sex table and divided by the original sum for that SOA in that constraint table. So for home owners in E10101264: 1067 / 1439 * 1438 = 1066.26.

This adjustment seeks to minimise discrepancies between the totals of the constraint tables (note, it is not important which table is used to adjust the totals; sex was an arbitrary choice: what is important is that all the totals are the same). Whilst the adjusted tables may not be more accurate than the original census tables, the adjustment method ensures the constraint tables are more consistent or at least can be guaranteed to produce the smallest discrepancy.

3) Summary files

The third type of file needed to run the model is a summary of the totals of the population dataset for each constraint variable (see Table 6.15). For each constraint variable a summary of the aggregate number by each category is given in a summary table, as shown in Table 6.15. So if there are 6 constraint files for a simulation, there would be 6 summary files for that population file (the full listing of the summary files are given in Appendix E).

Area	White	Not_white
1	15595	1922

Table 6.15: Example of a summary table for ethnicity. There is always only one row in these tables and there is one table for each constraint variable. This table shows that there are 15595 individuals in the population table who are white and 1922 who are not white (note this sums to 17518, the total population in the HSE dataset).

4) Micro_Config file

The java code to run SimObesity needs to be able to read these files in and to read out the output files. This is facilitated by the use of an Access database called Sim_data; this contains

the population file and all of the constraint and summary files, as well as a Micro_Config file. See Table 6.16 for an example (the full listing of the two Micro_Config files used are given in Appendix F).

Array	Identifier	ItemName
CONST1	ID	ZoneCode
…	…	…
CONST6	ID	ZoneCode
SUMMARY1	ID	area
…	…	…
SUMMARY6	ID	area
CONST1	TABLE	Name of constraint 1 "constraint file"
…	…	…
CONST6	TABLE	Name of constraint 6 "constraint file"
POPULATION	TABLE	Name of population file
SUMMARY1	TABLE	Name of constraint 1 "summary file"
…	…	…
SUMMARY6	TABLE	Name of constraint 6 "summary file"

Table 6.16. Example of a Micro_Config file (contained within an Access database with the population file and all the constraint and summary files). The order and number of rows of a Micro_Config table is very important; if there are six constraints then there will be 25 rows (plus the header row); if two constraints, then 9 rows.

Sim_data must be stored directly on the c drive of the computer running the simulation, under the java source code directory (src). Here there are four folders: database (where Sim_data sits); Dist; Docs; Src; as well as the four batch files required to run SimObesity (Framework, JarFramework, MicroSimulation and SharedObjects). The java code are stored in four folders of the same names as the batch files also under the source code directory, although it is really only the "microsimulation" java files that need to be amended (for example if the number of constraints changes). It is not particularly user-friendly.

6.3.5 Algorithm Methodology

This section describes the algorithms that are used in SimObesity. The principal task in spatial microsimulation is to select individuals from a micro dataset to fill small census areas (e.g. SOAs). The information about individuals to create the micro-populations for SimObesity comes from the HSE and EFS datasets (which are only available at coarse levels of geography). A simulation was run on the HSE dataset and another on the EFS dataset separately, using different constraint variables (the choice of these constraints was described in the previous section) in order to synthesise a micro-population with a list of attributes. These attributes include census input variables as well as several different obesogenic output variables. That is, SimObesity synthesises a micro-dataset of individuals living in the Leeds area (by SOA) such that a tabulation of the individuals in each SOA sum to approximately the same as the census tabulations. In effect this process inputs "real people" into census areas as a synthetic sample.

Each micro area has a list of individuals/households with a corresponding list of attributes. For example:

- Person 1, male, white, 42 years old, married without children, lives with his wife in their rented house, they have a car, he has O levels, low SEG, employed with a low income, living in a deprived suburban area with good public transport and easy access to the supermarket, in a neighbourhood he perceives as safe, access to leisure facilities is poor but he still manages to exercise 3-4 days per week for at least 30 minutes but eats less than 5 portions of fruit and vegetables per day, he is not overweight or obese.
- Person 2, female, white, 8 years old, lives in a household with 3-4 people, her parents have a car and own their own home, high SEG, household income is high, live in an affluent suburban area, she ate no fruit and vegetables during the survey, but does spend at least 7 hours a week being physically active.

Building a static spatial microsimulation model involves constructing a micro-dataset; that is, a population of households in conjunction with their related characteristics. This entails using a combination of available known data and mathematical algorithms to estimate the micro-level population. The algorithms for a static spatial microsimulation model can be either probabilistic (i.e. random sampling) or deterministic (i.e. rule based – if A then B). With this in mind, a static spatial microsimulation model can be split into three categories (Ballas et al, 2005): 1. synthetic probabilistic reconstruction models (random sampling); 2. reweighting probabilistic methodology; 3. reweighting deterministic methodology. A synthetic probabilistic reconstruction model (option 1) involves the use of random sampling. They are only really useful if there isn't any small area known data available. This is not the case in the UK (as we have Census data), so SimObesity does not need to use this methodology. Options 2 and 3 both typically involve reweighting an existing national micro dataset to fit a geographical area description on the basis of random sampling and/or optimisation techniques. Simulated annealing (Kongmuang, 2007), linear programming models (Zhang & Chambers, 2004) and complex combinatorial optimisation (reweighting) methods (Ballas et al, 2006) are examples of techniques that fall into the Option 2 category for the construction of a micro dataset. Iterative proportional fitting (IPF) techniques (Norman, 1999; Rees et al, 2004; Simpson & Tranmer, 2005), and reweighting of a parent sample of micro data such as with SimLeeds (Ballas, 2004) are examples of Option 3.

Previous work (Ballas et al, 2005) has shown that the reweighting technique works effectively in terms of finding the combination of records which best fits known small area statistical constraints. It is theoretically preferable to use a deterministic approach because, as no random sampling procedures are involved, this method will produce "one" solution no matter how many times the model is run. With a stochastic (probabilistic) mechanism each time the

microsimulation model is run, the string of random numbers employed to create choices will be changed and consequently different results achieved. As no two runs will (most likely) produce the same results, the customary method to remedy this is to run the model several times and to use an average of the results (Clarke, 1996). However a probabilistic method is preferable if the base population dataset is very large (e.g. the individual SARS dataset has millions of records). This is because the deterministic methodology does not use any random sampling (by definition); it reweights the whole of the dataset, which for the HSE/EFS surveys is manageable. This is not the case for the SARS data, so it would be necessary to randomly sample from this huge dataset. When using a deterministic reweighting technique the synthetic micro population is selected based on all constraints used, and not on any one constraint used: effectively fitting all constraints simulataneously to each micro area in turn. This means that the order that the input variables are entered into the algorithm does not affect the resulting population (so a simulation using sex, tenure and deprivation returns the same population as one using deprivation, sex and tenure). This is a further advantage of deterministic over probabilistic methodologies. Therefore, in this study, combinatorial optimisation is achieved by using a deterministic reweighting algorithm.

Two deterministic algorithms are used (based on Ballas et al, 2005) to create the micro-population. The first, the reweighting algorithm, compares the entire HSE (or EFS) base population with the census population for each SOA for each constraint variable and reweights each individual based on this (iterating just once through each constraint in turn). That is, the survey records are re-weighed to fit the census data tabulations. Then the second algorithm, an integerisation procedure, converts these reweight values into integers to determine how many of each of the individuals from the base population synthetically "live" in each SOA (as it is not possible to have fractions of people without creating a mess!). The task is to ultimately select those individuals in the base population that best match the descriptions of the population who actually live in the micro area. The descriptions are determined by the census variables that are used as the input variables.

These two algorithms create the micro dataset, stating how many of each individual synthetically "live" in each SOA. They give appropriate weights for the individuals to synthetically exist in each SOA based on the census tables used as "constraints" in this methodology. However these same weights may translate into relatively high over or under estimates of some variables that were not used as constraints in the simulation. It is possible to add a third algorithm to swap out poor (badly fitting) individuals for others to improve the match with the census statistics to adjust areas with high over and under estimations (Ballas et al, 2005). It is based on swapping suitable matching simulated individuals (that is individuals that have all attributes but one in common). This swapping algorithm has not been

implemented in SimObesity. The reasons for this are three-fold (in order of priority): Firstly all the output variables of interest (i.e. the obesogenic variables) show a correlation with the constraint variables, so if the constraint variables have simulated accurately (which is reviewed in the validation section of this chapter) then it is highly likely that the output variables have also simulated with low errors. Secondly it is only possible to run this third, swapping, algorithm for non-constraint census variables; i.e. it does not improve the simulation for the output (obesogenic) variables (which is why the correlation with the constraint variables is so important). Finally it could be argued that it is over adjusting the data. Furthermore, if SimObesity is used for "what if" scenario analysis then it is not possible to do this algorithm to test the efficacy of proposed public health interventions as the choice of which areas to run the algorithm on is somewhat subjective.

A java model was built to run these two algorithms, as the datasets are too large to run this in excel. As this was the first programme written by the author in java, to test whether the algorithm was functioning as expected a mini spatial microsimulation model (due to excel memory constraints) was built to ensure that the java code and excel formulas produced the same populations – which the final versions (results not shown) did. A summary of the key parts of the code is given below and a full version is available from the author upon request.

(1) Reweighting algorithm

The first algorithm in SimObesity reweights the existing micro datasets (HSE and EFS) so that they fit small area population statistic tables (2001 census area statistics univariate tables). A deterministic reweighting methodology was used to select the HSE (and separately the EFS) individuals that best matched the SOA census individual variables (six variables for the HSE simulation: sex, age, deprivation, qualifications, ethnicity, tenure; and five variables for the EFS simulation: tenure, car availability, type of property, household type and size). Note, SimObesity is set up to run on either OA or SOA. This analysis is undertaken at the SOA because small number problems are greater at OA level.

When using the single census variable tables (e.g. just sex or tenure) as in this analysis, rather than cross-tabulated census tables (e.g. tenure by sex) as used by Ballas and colleagues (2005), there are two stages to calculating the reweight value (i.e. reweighting the data). The first step of the reweighting algorithm uses the equation in Figure 6.6. The java code for the reweighting algorithm all falls within the reweighting class. The corresponding part of the java code is highlighted in Figure 6.7. The second step of the reweighting algorithm uses the equation in Figure 6.8 and the relevant exert from the reweighting class of the java code is given in Figure 6.9.

For HSE ij

X ij = Weight ij x constraint ij / sum ij

Figure 6.6. Equation for step 1 of the reweighting algorithm (where i = person ID and j = micro area). Note HSE ij represents each person in the population dataset, thus is HSE survey for the HSE simulation and EFS survey for the EFS simulation. X is the resulting reweight from this step of the algorithm for person i in area j. Weight is the person's original weight in the population table for the first constraint in the algorithm and is the resulting weight (Yij) from the previous constraint for all subsequent constraint variables. Constraint ij is element ij of the corresponding constraint table (Table 6.13). Sum ij is element ij of the corresponding sum table (Table 6.15).

```
newArray[oaCounter][personCounter]=oldArray[oaCounter][personCounter]*constraint[oaCounter][cat-
1]/summary[0][cat-1];  }
```

Figure 6.7. Exert from the reweighter class java code that represents step 1 of the reweighter algorithm. I.e. For HSE ij, X ij (i.e. newArray) = Weight ij (i.e. oldArray) x constraint ij / sum ij. Note, "old array" is the weight column in the population table for the first iteration (i.e. the first constraint). For subsequent iterations (constraints) it becomes the output from the previous iteration (the Y value).

For HSE ij

$$Y_{ij} = X_{ij} \times \Sigma \text{ constraint } j \;/\; \Sigma X_j$$

Figure 6.8. Equation for step 2 of the reweighting algorithm (where I = person ID and j = micro area). Xij is the resulting value from step 1 of the reweighting algorithm. Σ constraint j is the sum of the relevant area column for the constraint variable. Similarly Σ X j is the sum of the relevant area column for the reweight value calculated in the previous step.

```
//put in  a temp for/next loop to cycle thru the first element
setArray(1);
for(int i=0; i< newArray.length; i++) {
        double totalPop = 0;
        double acPop = 0;
        String s="";

        for(int t=0;t<constraint[i].length;t++){
        acPop=acPop+constraint[i][t];
        }
                for (int j=0;j<newArray[i].length;j++){
                totalPop = totalPop+newArray[i][j];
                }
                        for (int j=0;j<newArray[i].length;j++){
                        newArray[i][j] = newArray[i][j]*acPop/totalPop;
                        }
                        if (i==0){
                        for (int j=0;j<newArray[i].length;j++){

                System.out.println(Double.toString(newArray[i][j]));
                //converts dbl to string for printing, this is the output
                }
        }
}
//Basically this translates as:
//newArray[i][j] = newArray[i][j]*acPop/totalPop;
//      acPop=acPop+constraint[i][t];
//      totalPop = totalPop+newArray[i][j];
```

Figure 6.9. Exert from the reweighter class java code that represents step 2 of the reweighter algorithm.

To build the micro dataset all the individuals in the population dataset are given an initial weight that compensates for error, bias, refusals, etc. Then these weights are readjusted in order to fit the census micro-data using the two equations detailed in Figures 6.6 and 6.8. The reweight algorithm is carried out for the first constraint variable (for all SOAs individually), then the algorithm moves to the next constraint variable, carrying out the same process for the next constraint, and so forth until all constraints have been looped through and reweighted once. It is only necessary to do a proportional fit once, as the HSE and EFS data are a good fit to the general population, so there are good outputs from one run only.

In order to demonstrate how the reweighting algorithm works there follows a worked example. The input data files are given in Tables 6.17-6.19. There are five micro areas in the example, super output area 1-5. In practice this data would be required for all micro areas in the study area (476 SOAs across Leeds). The example runs with two constraint variables, namely sex and tenure, which are detailed in Table 6.17. Both the HSE and EFS simulations use more than two CAS univariate tables (6 constraints for HSE simulation and 5 constraints for EFS simulation). Table 6.18 is an example of a population dataset, such as the HSE or EFS datasets, but with only 10 individuals and two attributes rather than the thousands of individuals in the HSE or EFS datasets and would include many attributes, not just sex and tenure. Table 6.19 gives an example of the summary tables that are used in this example. Again there would be a summary table for each constraint, and it summarises the totals from the population datasets.

As there are two constraint variables there are four stages to the calculation. The order of the constraints in the simulation (given by the order of columns in the population table (Table 6.18) is sex followed by tenure. Thus step 1 of the algorithm is run using sex as the constraint variable (resulting weights are shown in Table 6.20), then step two using sex constraint (resulting weights in Table 6.21). The algorithm then moves to the next constraint, so, next, step 1 of the algorithm is run using tenure as the constraint variable (resulting weights in Table 6.22) and finally step 2 using tenure constraint (resulting weights in Table 6.23). The reweight results from each stage are used in each subsequent stage. After the second step of the algorithm has been run on the last constraint (which in this case is tenure), this is the output from Algorithm 1 that would run into algorithm 2 (i.e. Table 6.23). These reweight values represent the probability that a HSE individual "lives" in that micro area. Note, as there are over 18,000 records in the HSE and perhaps a few hundred people living in each micro area, the actual reweight values will often be less than 1 (although they are all greater than one in the above example because of the small numbers used).

Sex

Area	Male	Female	Total
SOA1	156.0	148.0	304.0
SOA2	145.0	141.0	286.0
SOA3	209.0	227.0	436.0
SOA4	100.0	137.0	237.0
SOA5	146.0	148.0	294.0

Tenure

Area	Own	Rent/Other	Total
SOA1	216.4	87.6	304.0
SOA2	269.0	17.0	286.0
SOA3	379.5	56.5	436.0
SOA4	70.6	166.4	237.0
SOA5	261.4	32.6	294.0

Table 6.17. A hypothetical example of three CAS univariate tables from the 2001 census (by super output area) – the constraint tables. These Census data are aggregated into pre-determined tables, with a count or sum of the relevant constraint factor(s).

Weight	sex	Tenure	ID
1.262	1	1	ID1
1.000	1	1	ID2
1.000	1	1	ID3
0.899	1	1	ID4
1.000	2	1	ID5
1.000	1	2	ID6
1.000	1	2	ID7
1.000	2	1	ID8
1.000	2	2	ID9
1.000	1	2	ID10

Table 6.18. Population dataset – a hypothetical example of the HSE or EFS dataset. Sex: 1=male, 2=female; tenure: 1=home owner, 2=rent/other.

Sex

Area	Male	Female
1	7	3

Tenure

Area	Own	Rent / Other
1	6	4

Table 6.19. A hypothetical example of the summary tables of the population dataset by each of the constraint variables. The first constraint, sex, has been adjusted to take account of the fact that some initial ID weights are not = 1, so sum of weights = 49.3 rather than 50.

Idcode	OA1	OA2	OA3	OA4	OA5
ID1	28.1	26.1	37.7	18.0	26.3
ID2	22.3	20.7	29.9	14.3	20.9
ID3	22.3	20.7	29.9	14.3	20.9
ID4	20.0	18.6	26.8	12.8	18.7
ID5	49.3	47.0	75.7	45.7	49.3
ID6	22.3	20.7	29.9	14.3	20.9
ID7	22.3	20.7	29.9	14.3	20.9
ID8	49.3	47.0	75.7	45.7	49.3
ID9	49.3	47.0	75.7	45.7	49.3
ID10	22.3	20.7	29.9	14.3	20.9
total	307.6	289.3	440.8	239.3	297.4

Table 6.20. The results from implementing the first step of the reweighting algorithm for the first constraint, sex. For example, person 1 (ID1) is male (sex = 1). So for SOA1 his new reweight value (Xij) = his initial weight (Table 6.18) x number of males in SOA1 (Table 6.17) / total number of males in the population table (Table 6.19). I.e. $\mathbf{X}_{(ID1, OA1)} = 1.262 \times 156.0 / 7 = 28.1$. This calculation is repeated for all people in SOA1, then for all people in all other areas.

Idcode	OA1	OA2	OA3	OA4	OA5
ID1	27.8	25.8	37.3	17.9	26.0
ID2	22.0	20.5	29.5	14.1	20.6
ID3	22.0	20.5	29.5	14.1	20.6
ID4	19.8	18.4	26.5	12.7	18.5
ID5	48.8	46.5	74.8	45.2	48.8
ID6	22.0	20.5	29.5	14.1	20.6
ID7	22.0	20.5	29.5	14.1	20.6
ID8	48.8	46.5	74.8	45.2	48.8
ID9	48.8	46.5	74.8	45.2	48.8
ID10	22.0	20.5	29.5	14.1	20.6
total	304.0	286.0	436.0	237.0	294.0

Table 6.21. The results from implementing the second step of the reweight algorithm for the first constraint, sex. For example, person 1 in SOA1, his new reweight value (Yij) = his reweight value from the first step calculation x the sum of the sex constraint in SOA1 (i.e. total number of males and females in that area) / sum of the reweight values from the first step calculation for SOA1. i.e. $Y_{(ID1, OA1)} = 28.1 \times 304.0 / 307.6 = 27.8$. This calculation is repeated for all people in SOA1, then for all people in all other areas.

Idcode	OA1	OA2	OA3	OA4	OA5
ID1	1002.8	1158.7	2357.7	210.1	1134.2
ID2	482.2	87.0	417.0	588.6	167.8
ID3	794.5	918.0	1868.0	166.5	898.6
ID4	714.0	825.0	1678.7	149.6	807.5
ID5	1067.5	197.4	1056.8	1881.5	397.0
ID6	482.2	87.0	417.0	588.6	167.8
ID7	482.2	87.0	417.0	588.6	167.8
ID8	1758.7	2082.9	4734.0	532.2	2125.4
ID9	1758.7	2082.9	4734.0	532.2	2125.4
ID10	794.5	918.0	1868.0	166.5	898.6
total	9337.5	8444.0	19548.2	5404.3	8890.2

Table 6.22. The results from implementing the first step of the reweighting algorithm for the second constraint, tenure. For example, person 1 (ID1) is a home owner (tenure = 1). So for SOA1 his new reweight value (Xij) = the resulting weight from step 2 with the previous constraint variable (Table 6.21) x number of home owners in SOA1 (Table 6.17) / total number of home owners in the population table (Table 6.19). I.e. $X_{(ID1, OA1)} = 27.8 \times 216.4 / 6 = 1002.8$. This calculation is repeated for all people in SOA1, then for all people in all other areas.

Idcode	OA1	OA2	OA3	OA4	OA5
ID1	32.6	39.2	52.6	9.2	37.5
ID2	15.7	2.9	9.3	25.8	5.6
ID3	25.9	31.1	41.7	7.3	29.7
ID4	23.2	27.9	37.4	6.6	26.7
ID5	34.8	6.7	23.6	82.5	13.1
ID6	15.7	2.9	9.3	25.8	5.6
ID7	15.7	2.9	9.3	25.8	5.6
ID8	57.3	70.5	105.6	23.3	70.3
ID9	57.3	70.5	105.6	23.3	70.3
ID10	25.9	31.1	41.7	7.3	29.7
total	304.0	286.0	436.0	237.0	294.0

Table 6.23. The results from implementing the second step of the reweight algorithm for the second constraint, tenure. For example, person 1 in SOA1, his new reweight value (Yij) = his reweight value from the first step calculation x the sum of the tenure constraint in SOA1 (i.e. total number of people in that area) / sum of the reweight values from the first step calculation for SOA1. i.e. $Y_{(ID1, OA1)} = 1002.8 \times 304.0 / 9337.5 = 32.6$. This calculation is repeated for all people in SOA1, then for all people in all other areas.

(2) Intergisation Algorithm

Algorithm 2 is a procedure to select the individuals that synthetically "live" in each micro area and importantly to ensure that no fractions of persons are allocated to an area. This requires an iterative optimisation technique, whereby the decimal weights (e.g. the reweight values in Table 6.23 above) are converted into integer weights. Different integerisation methodologies exist and have been tested and the following methodology has been determined the best in these circumstances (Ballas et al, 2005). This second algorithm selects individuals for each micro area (rather than random sampling as per a probabilistic methodology), assuring that no fraction of persons are allocated to a SOA; people can only be whole people. This algorithm works as follows (and is also explained numerically in Table 6.24 and the relevant java code is outlined in Figures 6.10 and 6.11):

1. For each SOA in turn, rank the data by reweight value from smallest to largest – this is sorting the individuals into ascending order of probability of living in the SOA being populated.
2. Calculate each individual's "cumulative reweight" number by summing the reweight values, one at a time, from the top of the list (i.e. for the lowest reweight values first). Note, for the first individual on the list (i.e. the person with the lowest likelihood of living in that SOA) their cumulative reweight value is just their reweight figure.
3. Each time the cumulative reweight value moves down a row (i.e. to a new, higher re-weighed, individual), if the new cumulative reweight value is > 1, take the whole number (of the cumulative value) as the number of that individual in that area and carry forward the balance. That is, if cumulative value = 2.4, then 2 of that person "lives" in the area and 0.4 would be carried forward to the next person on the list and added to their reweight value to calculate the next cumulative reweight value. It uses a "floor" function, so the integisation always rounds downwards, rather than either up or down.
4. Continue until each person has been allocated an integisation number for each SOA.

SOA	Individual ID	Reweight value		Individuals ranked in ascending order of reweight value	Reweight value	Cumulative reweight value	No. of that individual living in SOA
001	1	0.8		2	0.2	0.2	0
001	2	0.2		4	0.3	0.5	0
001	3	4.5	⇨	1	0.8	1.3 (c/f 0.3)	1
001	4	0.3		5	1.2	1.5 (c/f 0.5)	1
001	5	1.2		3	4.5	5.0 (c/f 0.0)	5

Table 6.24. Simplified illustration of the integisation algorithm. The three left most columns represent the results from the first algorithm, giving each individual a reweight value, which represents how likely they are to live in that SOA. The four left most columns represent the algorithm after the individuals have been ranked in ascending order of reweight value. Then the cumulative reweight value is calculated in order to determine how many of each individual live in the SOA.

```
//start bubble sort here (i.e. sorts array so smallest at top and largest at bottom)
boolean notSorted=true;
while(notSorted==true){   //starts the sorting

        for(int j=0;j<d[0].length;j++){       // For each column (individual)

                if (j==0) {
                        notSorted = false;
                }else{

                        if (d[0][j-1]>d[0][j]){
                        //if the second number in a pair is less than the first
                        // ie if first # > second #

                        double tempWt = d[0][j];
                        //actual weight - remember the second number (create a temp = j #)

                        double tempOrder = d[1][j];
                        //id, to keep track of changes - remember the 2nd number's place

                        d[0][j] = d[0][j-1];
                        //set the second number to equal the first
                        // ie change j so = j-1 (ie change #)

                        d[1][j] = d[1][j-1];
                        //set the second id to the place of the first id
                        // ie change j ID so = ID j-1 (ie change places of the numbers)

                        d[0][j-1] = tempWt;
                        //set the first number to equal the second number
                        // change j-1 so = old j #

                        d[1][j-1] = tempOrder;
                        //set the id of first number to the second number's place
                        // change j-1 ID so = old j ID

        notSorted=true;             //this stops the while loop when all are sorted

        // e.g. say j-1 = 4, j =2
        // then tempwt = 2, tempOrder = ID for 2
        // j(2) = j(4), ditto ID and j-1(4) = tempwt(2), ditto ID
        // So end up with j-1 = 2, j = 4
        // and when you do this for a column of numbers, the sorting gradually
        // moves through the column (bit like a Mexican wave)

                        }               //end of if comparison loop
                }
        }               //End of j loop (using d)
}       //end of while loop

//end bubble sort here
```

Figure 6.10. Extract from reweighter class java code that corresponds to the code to rank the data (step 1).

```
//**/rest of integerisation starts here

// Create a double called cumulativeWt, and set it equal to zero
// Loop through each column j (i.e. each individual)
// Create a 2 dimensional array called sorted, and set it equal to the sorted weights
// from the above code (bubble sort)
// Create a double called r, and set it equal to zero

// Then change CumulativeWt so it equals previous cummulativeWt plus the next result
// in the sorted array.  That is, it is the current record's weight plus the remainder left in
// the counter from the previous record(s)

// Then do the following if statement ...
// If the resulting cummulative weight exceeds 1, then take the floor (i.e. whole number,
// always rounding downwards) - this number is the result for this indiv in this area
// Don't use round, as this rounds up or down
// Then set cummulative weight = previous cummulativeWt less the integer just taken
// as the result for that individual (the floor)

// Then take the array "ra" and give it the floor values (this will be the result array)

// Then the code goes back up to the "for" statement, and moves to the next individual,
// increasing the cummWt by their alg1 result and then looping through the "if" statement
// again, getting floor values whenever the cummWt goes above 1
// (if it doesn't exceed 1, then that individual just gets a zero in the result array, ra)

                double cummulativeWt= 0;

                for(int j=0;j<d[0].length;j++){

                        double sorted = d[0][j];  //equals the sorted weights
                        double r=0;

                        cummulativeWt=cummulativeWt+sorted;

                        if(cummulativeWt>1){
                                r = Math.floor(cummulativeWt);
                    //take whole number part of the sorted weights (leave the fraction / decimals)
                                        // CHANGE ROUND TO FLOOR (see notes below)
                                cummulativeWt=cummulativeWt-(double)r;
                        }       //End if


                        ra[i][(int)d[1][j]] = (double)r;
                        // this gives the floor integer to the results array
                }
        }       //end of for looping thru i in newArray

        return ra;  //  ie gives the number of each individuals in each area

        }
        //end of integerise
```

Figure 6.11. Extract from reweight class java code corresponding to the summing of the reweight values, and when the cumulative value is > 1, then take the whole number as the number of that individual in that area (steps 2-4).

6.3.6 Output data

The output from SimObesity comprises of 33 files, one for each ward. Each ward file has a separate column for each SOA in that ward. The java programme output just lists how many of each individual "lives" in each micro-area (either low SOA or OA, depending how SimObesity is set up – in this study it was set on SOA) (see Table 6.25). Once the list of individuals and their attributes has been estimated, they can be aggregated to any spatial scale. This output was then converted into aggregate (SOA) data for all variables (i.e., constraint, non-constraint and obesogenic variables) for each micro-area using a bespoke aggregation programme (written by the author) in excel. Using an excel spreadsheet (which requires some level of manual input due to memory limitations, but is mostly automated) is more time consuming than if java code calculated it automatically, but in the short term it was quicker to do it this way than to spend a long time writing the appropriate java code. This "conversion" results in a list of SOAs with the sum of each individuals' associated demographic and socio-economic characteristics (see Table 6.26). In addition, the attributes include the obesogenic variables from the HSE 2002 and the EFS 2005. Note, the categorizations of variables used included a "missing data" category so if an individual was simulated with missing data for, say, physical activity, then this could be adjusted for. Subsequent analyses took these absolute numbers and calculated the proportions of the population, so simulated persons with missing data were excluded.

ID	SOA1	SOA2	SOA3	...
10109101	0	0	2	...
10109102	1	0	0	...
10109103	1	0	3	...
10109104	0	0	0	...
10110101	4	1	0	...
...	...	...	...	...

Table 6.25: Example of SimObesity java output. There is a column for each area (this is actually split into separate files by ward, as there are limitations on how many columns a spreadsheet can hold) – so 2440 OA columns or 475 SOA columns in total. There is a row for each "real" individual (17518 in the HSE simulation and 6798 in the EFS simulation). Each number represents how many of that individual synthetically "live" in that area.

Variable	SOA1	SOA2	SOA3	...
Total	1437	1490	1457	...
Male	644	646	678	...
Female	793	844	779	...
Child	267	178	340	...
Adult	1170	1312	1117	...
...	...	...	...	...

Table 6.26: Example of final SimObesity output (after the java output has been converted into aggregate data). There is a column for each area – so 2440 OA columns or 475 SOA columns. There is a row for each category of each variable; this would include constraint variables and the variables being simulated.

The HSE simulation produced 715,169 individuals whose characteristics matched the characteristics of the 715,402 individuals living in Leeds (as per the 2001 Census). Similarly the EFS simulation produced 715,167 matching individuals. Variables that were reliably simulated (this is discussed in more detail in the validation section) are the constraint and

obesogenic variables (i.e. not the non-constraint variables). That is: sex, age, deprivation, qualifications, ethnicity, tenure, fruit and vegetable consumption, physical activity levels, urbanisation, socio-economic group, social capital (public transport, access to facilities, neighbourhood safety) (from the HSE simulation); tenure, economic activity, household type, car availability, household size, food expenditure, PC and TV ownership, internet access, income, source of income, and spending on school meals (from the EFS simulation).

One of the major advantages of spatial microsimulation models is the ability to estimate geographical distributions of socio-economic variables which were previously unknown (Ballas, 2001). It is possible to identify individuals with very specific characteristics: for example, of being a child, living in a highly deprived area, with low perceived neighbourhood safety, low expenditure on food, low fruit and vegetable consumption, low physical activity levels, poor access to shops, low income and a high number of household televisions, i.e. children associated with a higher risk of obesity. Figures 6.12-6.15 show some model outputs by SOA. In particular, Figure 6.12 depicts the estimated spatial distribution of areas with perceived low social capital, namely poor public transport, limited access to leisure facilities and supermarkets, problems with teenagers hanging around and vandalism. This shows that areas where more of the population perceive the neighbourhood as having low social capital is in central Leeds (an area of high deprivation), with separate areas of high prevalence in the north west and south east. Figure 6.13 shows the distribution of populations with low expenditure on food, low income, low SEG, not economically active, no access to a car, rent their home, and have no qualifications. Again there are higher rates in the central, deprived, areas of Leeds, with some isolated cases in the north of Leeds. Figure 6.14 illustrates the distribution of children aged 7-15 years exhibiting obesogenic behaviours; that is, with low daily fruit and vegetable consumption and low physical activity levels. This shows a low percentage of people meeting these criteria in central Leeds, with higher levels of poor diet and low activity amongst children in more affluent, rural parts of Leeds. Finally Figure 6.15 depicts those people living in a highly deprived area, with low perceived neighbourhood safety, and a low income. Anyone outside of the highest deprivation areas is automatically excluded from this classification (thus all the yellow shading outside of the centre of Leeds).

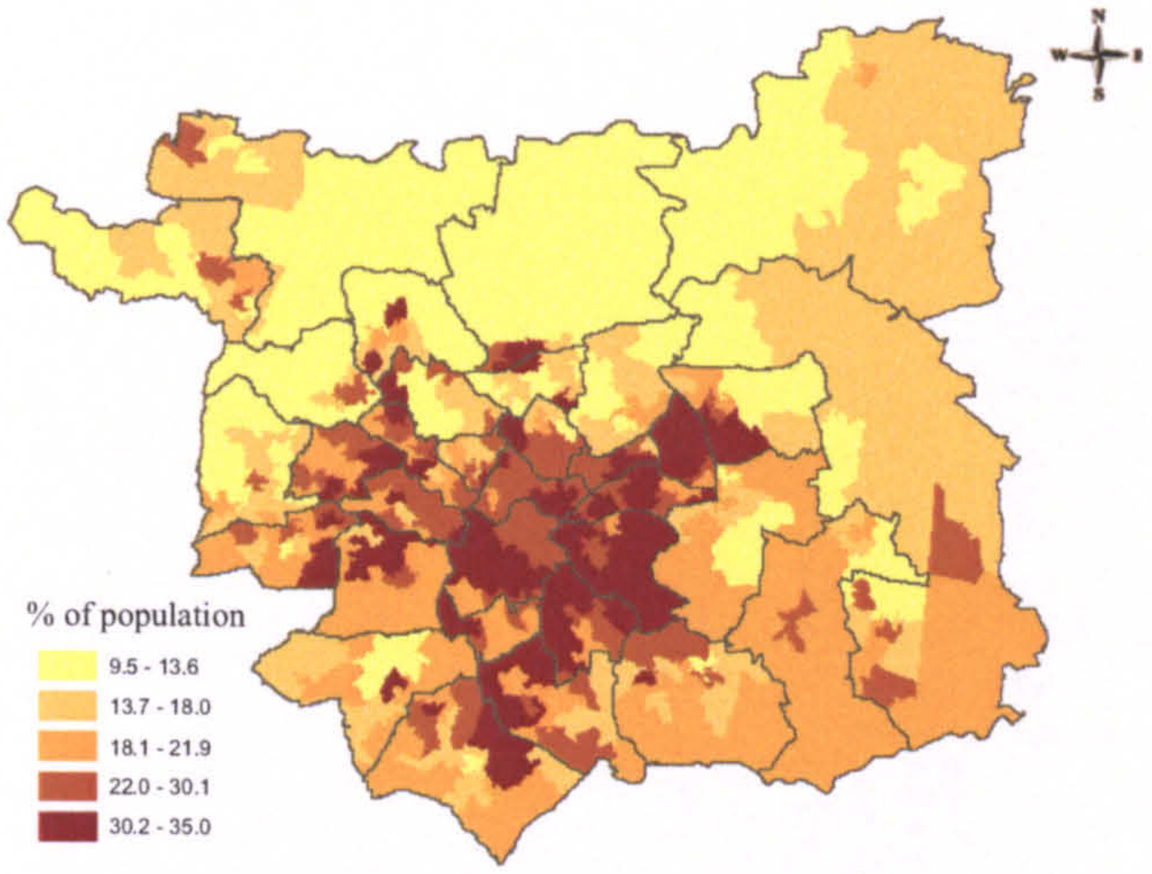


Figure 6.12. Distribution of perceived low social capital by SOA in Leeds: namely poor public transport, limited access to leisure facilities and supermarkets, problems with teenagers hanging around and vandalism.

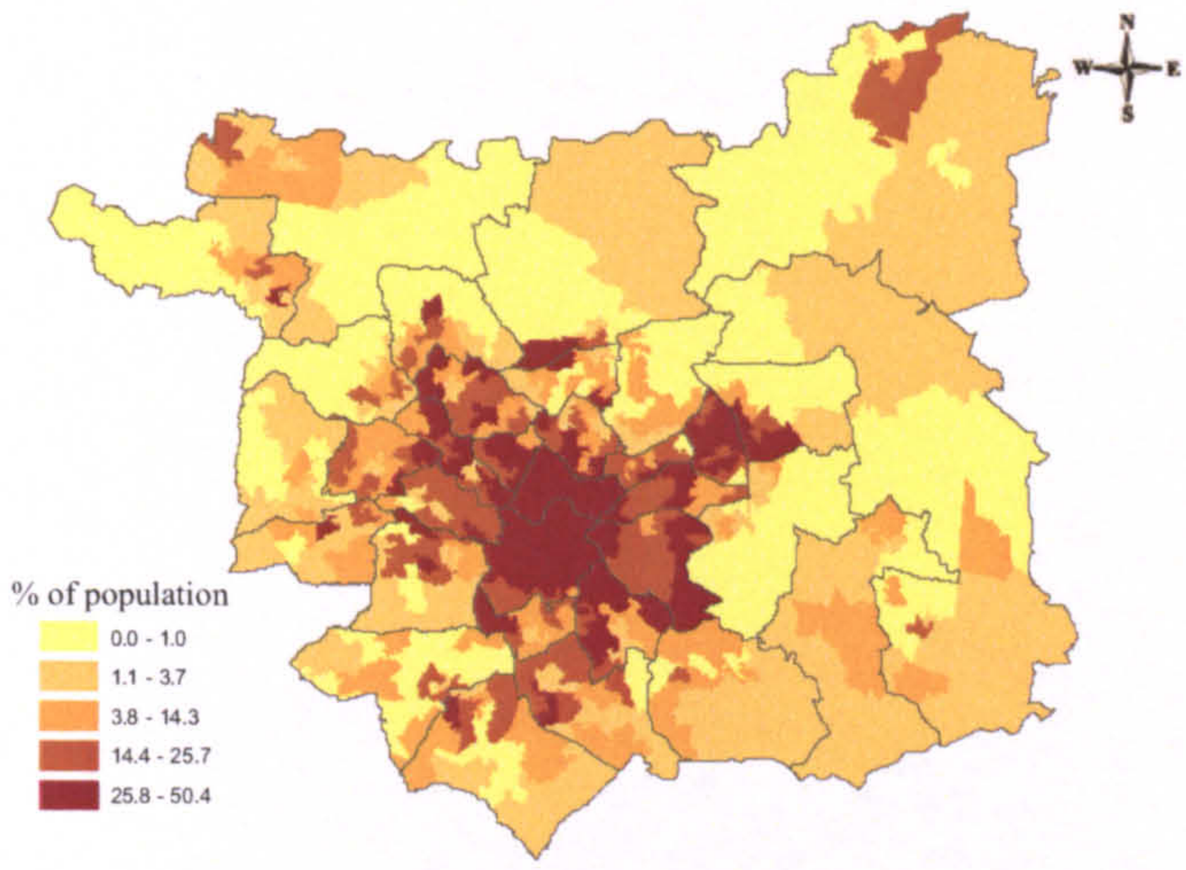


Figure 6.13. Distribution of populations with low expenditure on food, who are on benefits, and no one in the household is economically active, have a low income, no access to a car, and who do not own their home, by SOA in Leeds.

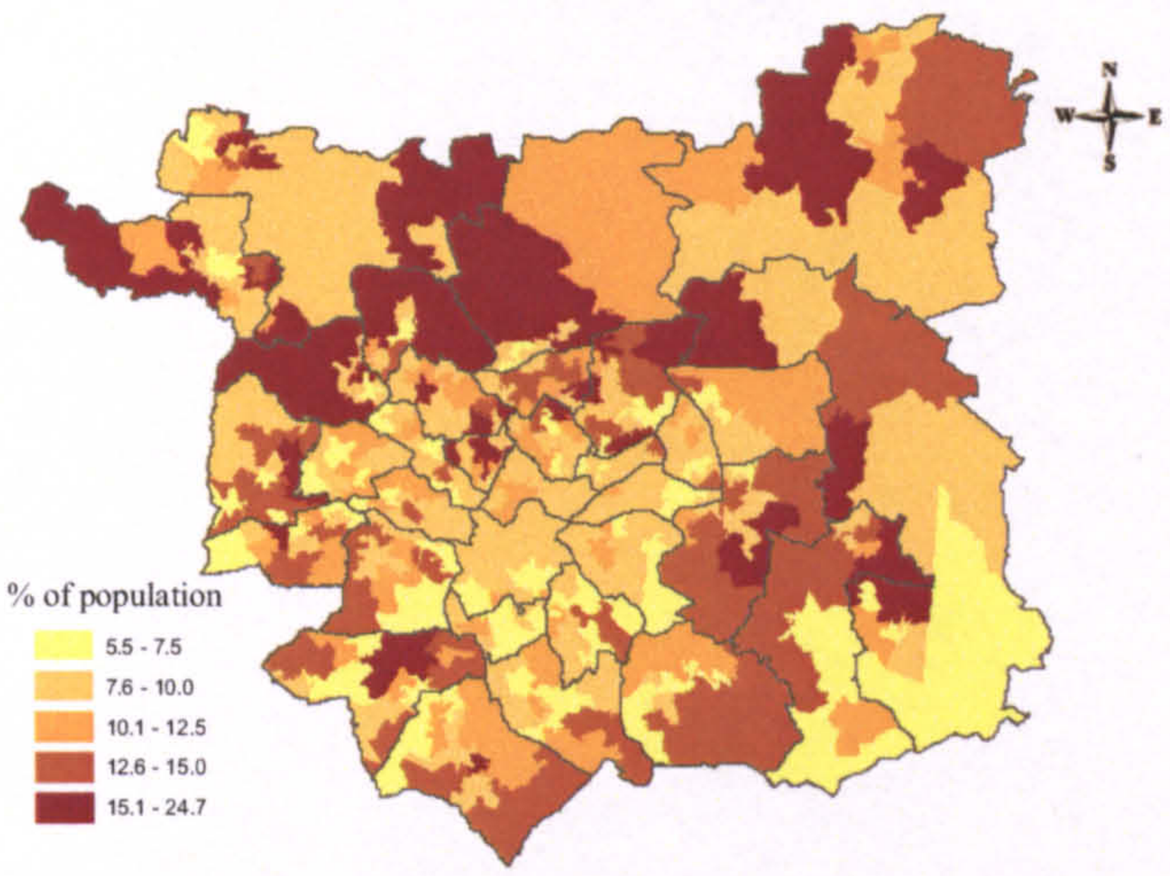


Figure 6.14. Distribution of obesogenic behaviours of children aged 7-15 years by SOA in Leeds: that is, with low daily fruit and vegetable consumption and low physical activity levels.

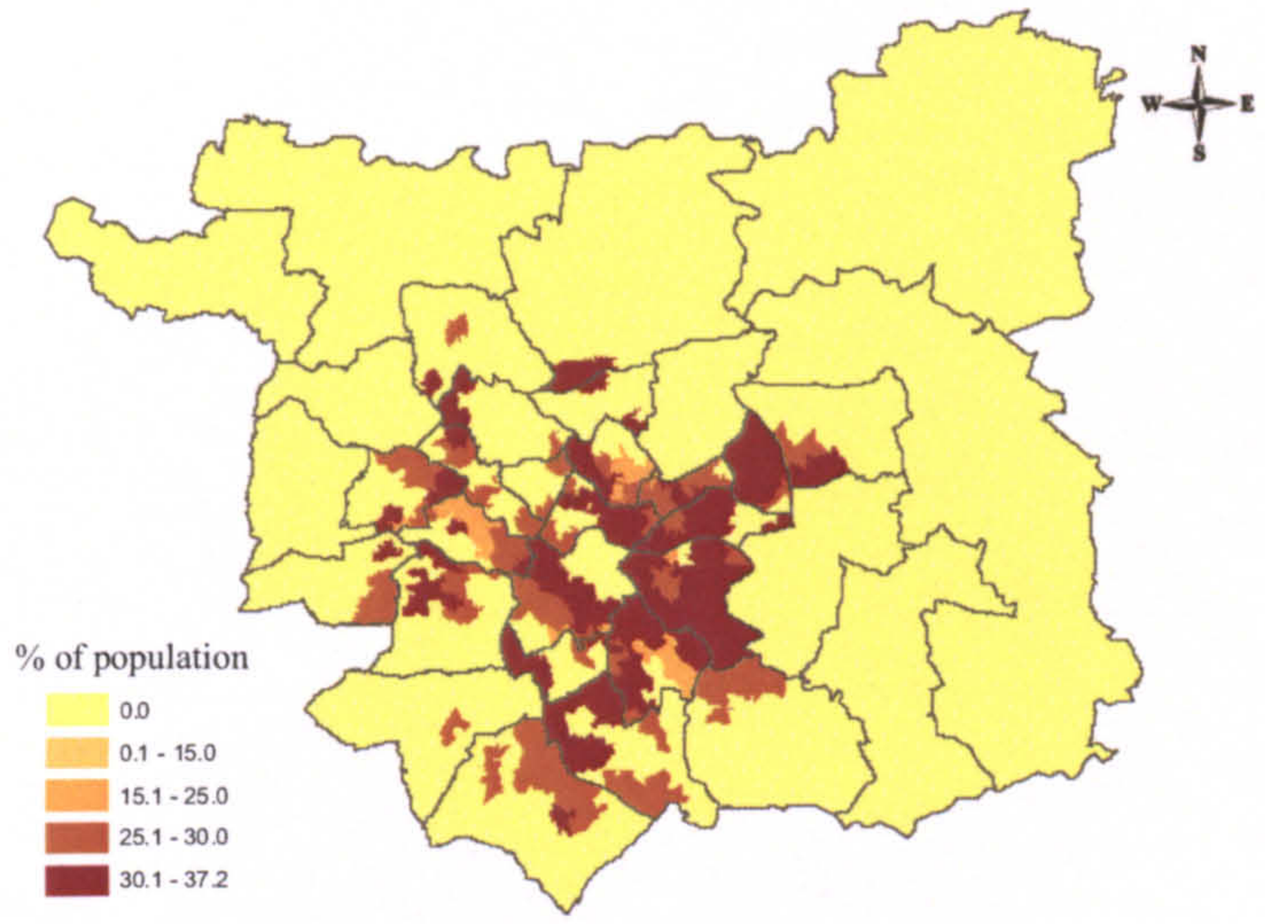


Figure 6.15. Distribution of people who perceive the safety of their neighbourhood as low and have a low income, who live in a highly deprived area, by SOA in Leeds.

6.4 Validation of Synthetic Micro-data

This section discusses ways of validating the simulation outputs, and of showing where SimObesity works well and where it functions less well. As the whole purpose of producing synthetic micro-data is to generate data that do not currently exist for small areas, validation is difficult; which is a drawback of spatial microsimulation. However it is possible to calibrate spatial microsimulation model outputs by aggregating the individual level simulated data to levels at which observed datasets exist and compare the synthetic data with the actual data (Ballas et al, 2005). To do this actual data for Leeds (from Census 2001) are compared to the simulation outputs. This is undertaken at the ward level (due to small number problems at SOA).

A key point of our constraint choice analysis was that we should use variables as constraints that are most strongly associated with the output variables that we are wanting to simulate. We would expect these constraint variables to simulate well. Accordingly simulated data will be "validated" by comparing the actual data to the simulated data for the constraint variables that went into each respective model. Given the correlations between the output (obesogenic) and the constraint variables, we can assume that if the constraint variable simulates well that the output variable is also accurate. Conversely we would expect that non-constraint variables that had no correlation to the input variables to simulate poorly. If they do simulate well, that could be because they are correlated with the input variables or down to luck. A variable that is not correlated to the input variables may or may not simulate well; it is not possible to predict which. There is no point in validating against non-constraint variables as they may (for example if they happen to be correlated to the constraint variables) or may not simulate well. Plus how well these non-constraint variables simulate has no bearing on how well the output

variables of interest (the obesogenic variables) have simulated and so how well the data are simulated.

There are four key analyses undertaken to validate the data. Firstly the simulated population is compared to the corresponding census population for all of the census-type variables to give an approximate indication of the quality of the simulation. Then the error values are calculated. Thirdly a regression of simulated data verses actual data is undertaken to better understand the fit of the simulation, and finally a t-test statistical comparison of the two datasets is done to establish whether there are any statistically significant differences between the synthetic and real populations.

Firstly the population simulated for both the HSE and EFS datasets are (separately) compared with their census counterparts for all of the census-type variables: that is, for both the constraint variables used in the simulation and the non-constraint variables that were not used as input variables in the simulation. This is summarised in Tables 6.27 and 6.28. This very quickly shows that for the HSE simulation out of the constraint variables, gender and tenure appear to have simulated least well, with an underestimation of proportion of males and home owners, but that age and deprivation are almost perfect simulations. The non-constraint variables for the HSE simulation (there are not any non constraint variables in the EFS dataset) have performed poorly. Conversely for the EFS simulation, gender has simulated well, as has age and, again, tenure has simulated less well. However this is simply "eye-balling" the data for Leeds as a whole; it is necessary to determine whether there are any statistical differences between the simulated and actual data at the micro-level.

The next step in the validation process is to calculate the error values for each of the census-type variables in the HSE and EFS simulations. The error report provides information on the aggregate difference between distributions of each actual constraint variable plus other census variables that are synthesised (non-constraint variables), and the synthetic micro-data at the ward level.

The fit of a combination of individuals to known small area constraints is evaluated by the Absolute Error (AE), the absolute difference between observed (simulation) and expected (census) counts for each variable in each area. For example for the number of people who own their home in each of the 33 wards. Ideally, an optimal solution would have an AE of 0, which means there is no difference between the observed and estimated counts, in other words a "perfect fit". In order to compare across tables, the Standardised Absolute Error (SAE) is used. This is the AE divided by the total expected count for each area. An acceptable limit for error is often accepted to be around 10%. Note, it would be reasonable to expect that the performance

Proportion of population:	HSE 2002 simulated	Census 2001 (Leeds)
N	715169	715402
CONSTRAINT VARIABLES		
Who are male	45.2	48.3
Who are female	54.8	51.7
Aged 0–2 years	3.5	3.4
Aged 3-6 years	4.9	4.8
Aged 7-9 years	3.8	3.8
Aged 10-12 years	3.9	4.1
Aged 13-15 years	3.8	3.9
Aged 16-18 years	3.9	3.8
Aged 19+ years	76.2	76.2
Who are children	23.8	23.8
Who are adults	76.2	76.2
Least deprived	13.3	13.3
Deprived 2	16.5	16.5
Deprived 3	18.4	18.4
Deprived 4	19.1	19.1
Most deprived	32.7	32.7
Highest educational qualification is degree, or equivalent	14.7	14.0
HEQ is A or O levels, or equivalent	29.0	31.3
No or other qualifications	33.6	34.7
Full time student / child	22.6	20.0
Who are white	92.1	91.9
Who are other ethnic group	7.9	8.1
Who are home owners	62.5	65.3
Who rent or "other"	37.5	34.7
NON-CONSTRIANT VARIABLES		
No children in household	60.4	53.7
Household has children	39.6	46.3
With at least 1 car or van available	74.8	66.4
Without car or van available	25.2	33.6
Who are married or cohabiting	46.5	37.7
Who are not married	33.6	62.3
Economically active	42.7	42.9
Not economically active	57.3	57.1
Live in household on own	13.2	30.6
Two people in household	30.2	33.3
3-4 people in household	43.1	28.8
5 or more people in household	13.5	7.3

Table 6.27. Comparison of simulated population (from HSE 2002) to the 2001 Census data for the census-type variables contained within the HSE dataset.

Proportion of population:	EFS 2005 simulated	Census 2001 (Leeds)
N (individual)	715172	715402
CONSTRAINT VARIABLES		
Who are male	48.5	48.3
Who are female	51.5	51.7
Who are children	24.6	23.8
Who are adults	75.4	76.2
1 person lives on own	29.2	30.6
2-4 people live in the household	61.6	62.1
5+ people live in household	9.2	7.3
Live in household without children	55.9	53.7
Live in household with children	44.1	46.3
Who are home owners	59.2	65.3
Who rent or "other"	40.8	34.7
Proportion with at least 1 car available	62.8	66.4
Proportion without car available	37.2	33.6
Who live in a detached house	16.4	15.5
Who live in a semi, terraced house, flat or other accommodation	83.6	84.5

Table 6.28. Comparison of simulated population (from EFS 2005) and 2001 Census data for the census-type variables contained within the EFS dataset.

of SimObesity would vary from variable to variable, especially at small areas such as wards and SOAs. When the same combination of constraints are run in different orders, the same errors result. Accordingly it is clear that the order of the constraints does not have effect on the size of the errors; rather the choice (and number) of constraints affects the error size.

For the HSE simulation, very few wards had an SAE > 10% for a constraint variable (see Figure 6.16). In particular, Harehills (DAFM) was the only ward with multiple variables with an SAE > 10%; namely child/adult proportions and the tenure proportions. Other than this, tenure was the only constraint variable with SAE > 10% and it affected eight wards (Armley, Burmantofts, Chapel Allerton, City and Holbeck, Harehills, Hunslet, Richmond Hill and Seacroft). In all cases the proportion of people who own their home was underestimated. For example, for Armley (DAFB), the census values for home owners and for people who rent or otherwise are 13381.7 and 8652.3 (61% and 39% of total ward population) respectively. The simulated values were 11067 and 10957 (50% and 50% of total ward population) respectively. The absolute error was 2314.7 and –2304.7 respectively, giving a SAE of +/-10.5%. Conversely, all of the non-constraint variables (except economic activity) mostly had an SAE > 10%.

For the EFS simulation, the SAEs are generally higher and so are more frequently greater than 10% (see Figure 6.17). Again tenure is the variable that simulates least well, with 13 wards with a SAE > 10% (Armley, Beeston, Burmantofts, Chapel Allerton, City and Holbeck, Harehills, Headingly, Hunslet, Kirkstall, Richmond Hill, Seacroft, University and Weetwood), although car availability also has nine wards with a SAE > 10% (Burmantofts, City and Holbeck, Headingly, Hunslet, Kirkstall, Richmond Hill, Seacroft, University and Weetwood). In all cases the proportion of household who do not own their home was overestimated (correspondingly those who own underestimated) and those without a car available overestimated (correspondingly with a car underestimated).

Then a regression analysis of the percentage of population in each category (simulated vs. actual) is undertaken. The r square statistic (the coefficient of determination) is an indicator that ranges in value from 0 to 1, and reveals how closely the simulated values for the regression trend line fit the actual (census) data. A trend line is most reliable when its R^2 is at or close to 1. It should be noted that because census variables are largely count data (and certainly all the census variables used in this analysis i.e. that are available from the HSE and EFS surveys, are count data), the comparisons relate to proportions.

Figures 6.18-6.20 show the scatter plot for each pair of variables at the ward level, with the simulated proportion on the x axis and the actual (census) data on the y axis. If the data had simulated perfectly, then y = x and all the points would lie on a straight line of gradient 1. For

the HSE simulation, this shows a high coefficient of determination for all constraint variables except gender, and a low coefficient of determination for all non-constraint variables except car availability. As does the EFS simulation, except for households with 5 or more people living there.

Accordingly gender and household size data have been broken down further. Figure 6.21 is a comparison of the simulated (HSE simulation) and actual proportions of males for each ward, and shows that the proportion of males is consistently underestimated, highlighting those few wards with the greatest differences (albeit each still under 5%). Similarly Figure 6.22 compares the simulated (EFS simulation) and actual proportions of large households (5 or more people living there). This shows that for household size the regression analysis suggested just four outlier wards were affecting the coefficient of determination.

However this regression analysis does not give any information about the fit of the simulated data to the "ideal" (i.e. where y=x and the simulated data are the same as the actual data), rather it expresses the fit of the data to the "best fit" line through those data. That is, the coefficient of determination is providing information about precision not accuracy. Accordingly an equal variance 2-tailed t-test is used to determine if there is any significant difference between the two datasets (i.e. simulated and actual). The results of this analysis are summarised in Tables 6.29 and 6.30 for the HSE and EFS datasets respectively. These tables clearly show that as predicted most non-constraint variables simulate very poorly, showing a significant difference between the simulated and expected values. The exception to this is Economic Activity in the HSE simulation, where the simulated values are not significantly different to the census values. Of particular note is how the simulated data for car availability had a high coefficient of determination but is statistically significantly different from the expected data, albeit to a much lesser degree than the other non-constraint variables. Similarly most of the constraint variables simulate very well, showing no significant difference between the simulated and expected proportions. The only exception is gender in the HSE simulation and households with the most people living there, which are significantly different. The proportion of males is underestimated and the number of households with five or more people was overestimated.

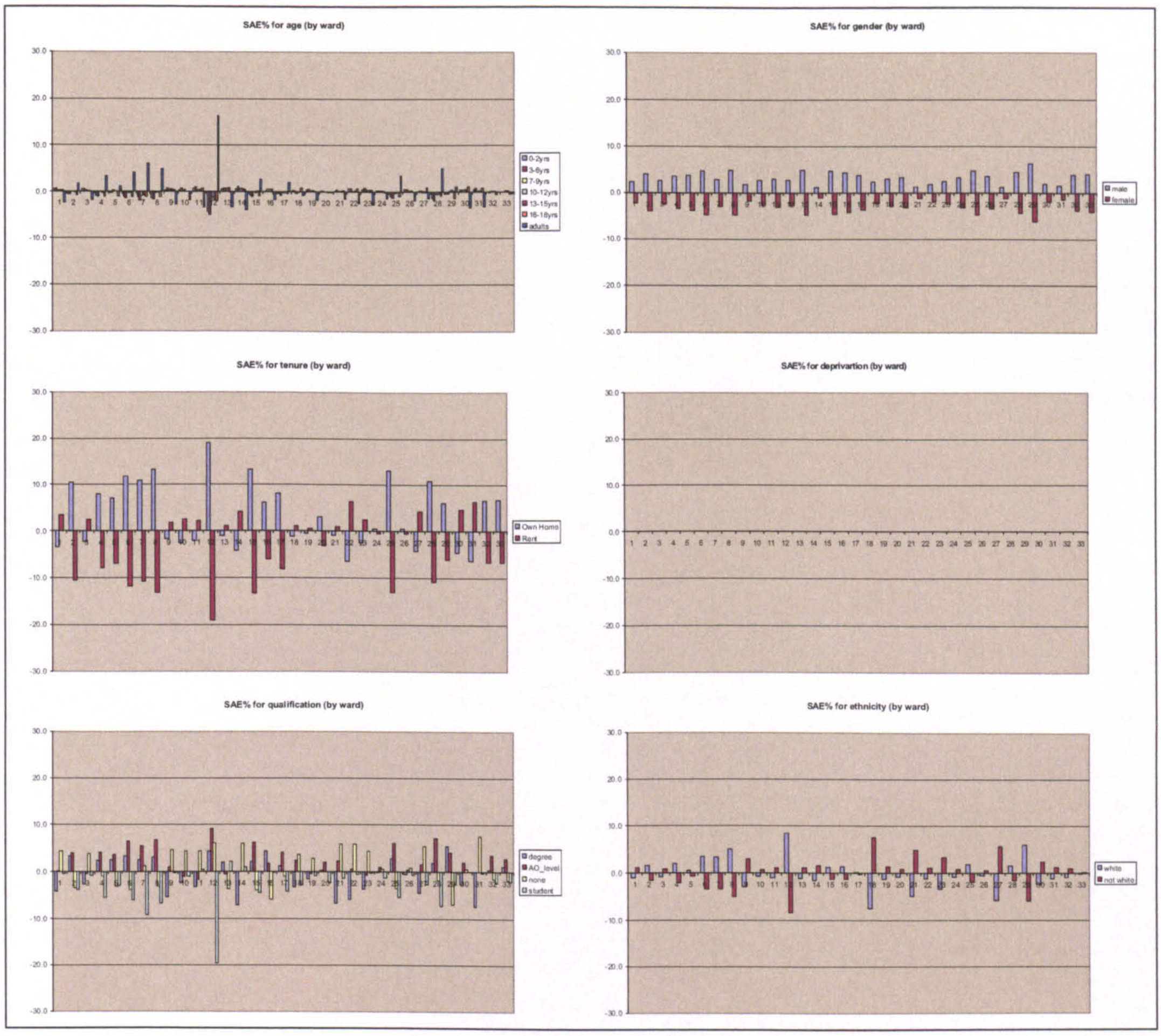


Figure 6.16. Bar charts showing the standardised absolute error percentage for each ward for each of the six constraint variables (sex, age, deprivation, tenure, qualification and ethnicity) in the HSE simulation. Gender, deprivation and ethnicity have no SAE% > 10%. Qualifications (student) and age (adult) have just one ward (DAFM) with a SAE% > 10%. Tenure however has eight wards with a SAE% > 10%.

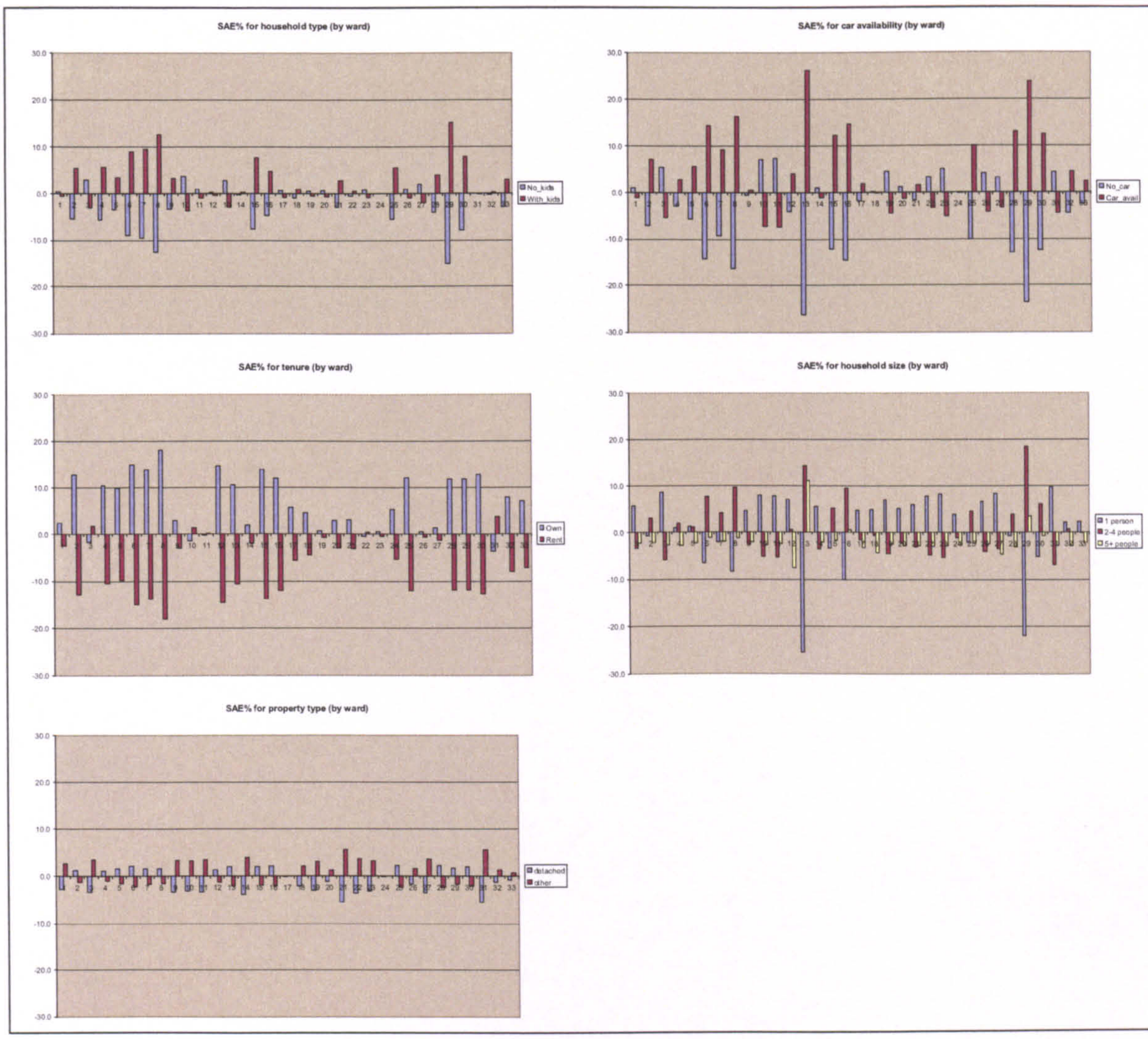


Figure 6.17. Bar charts showing the standardised absolute error percentage for each ward for each of the five constraint variables (household type, car availability, tenure, household size, and property type) in the EFS simulation. Only property type has no SAE% > 10%. Household type and size both have just two wards with a SAE% > 10%. However tenure and car availability have 13 and nine wards (respectively) with a SAE% > 10%.

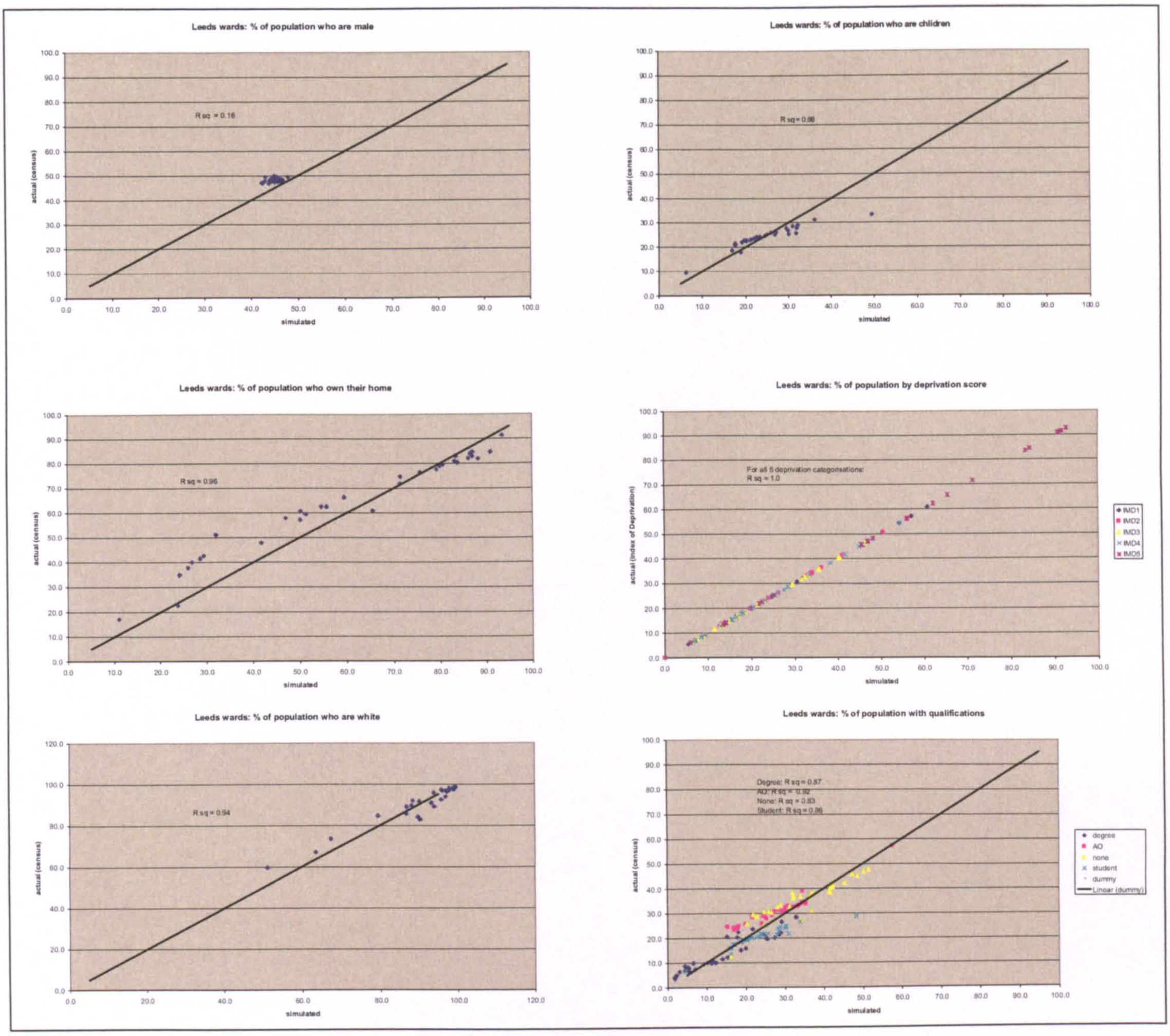


Figure 6.18. Scatter plots for each pair of constraint variables at the ward level from the HSE simulation. The simulated value is on the x axis, and the expected value on the y axis. The dummy regression line indicates the "perfect fit" of y = x. The coefficient of determination for each variable is also given; the closer the number is to 1, the better the fit.

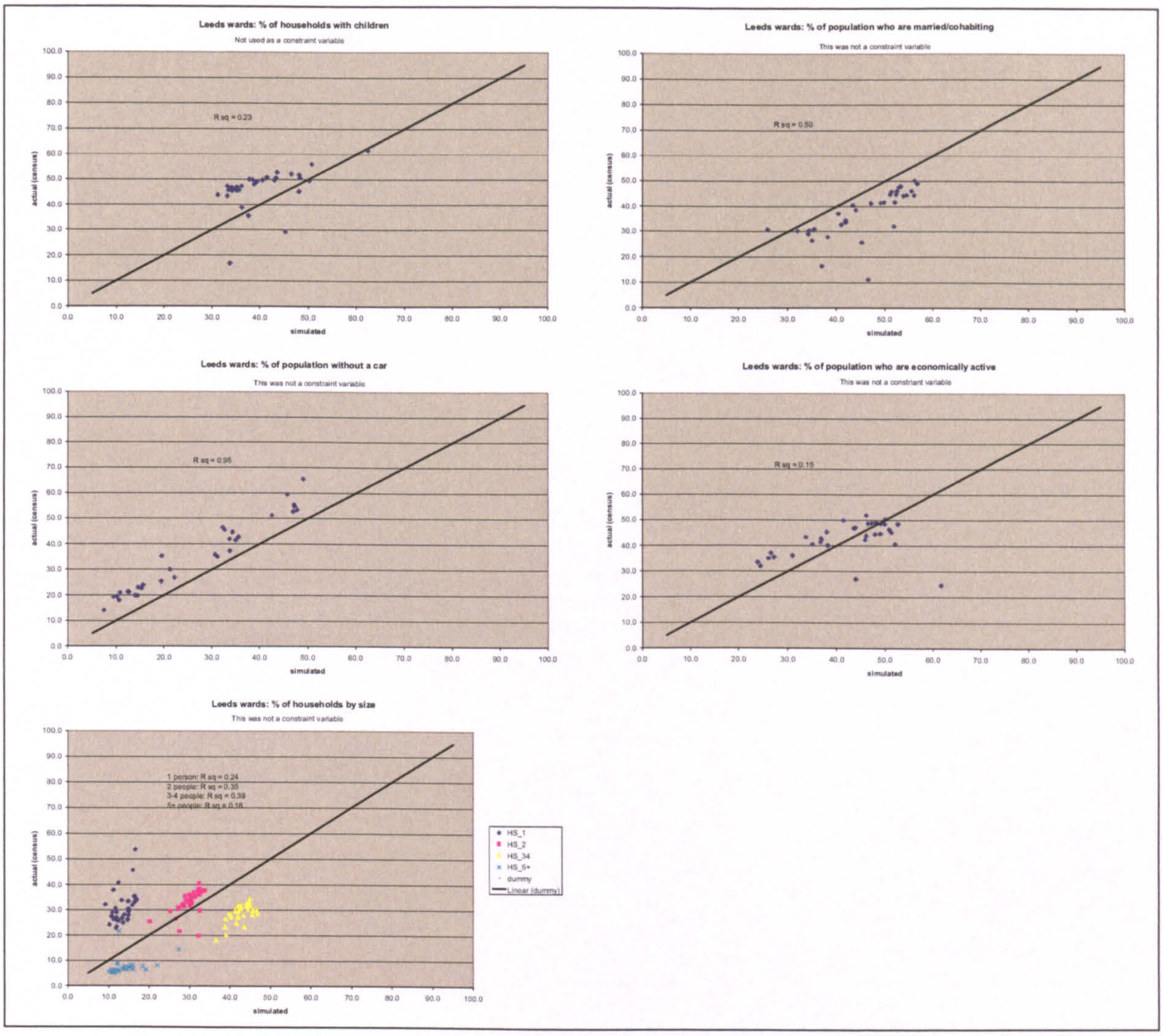


Figure 6.19. Scatter plots for each pair of non-constraint variables at the ward level from the HSE simulation. The simulated value is on the x axis, and the expected value on the y axis. The dummy regression line indicates the "perfect fit" of y = x. The coefficient of determination for each variable is also given; the closer the number is to 1, the better the fit.

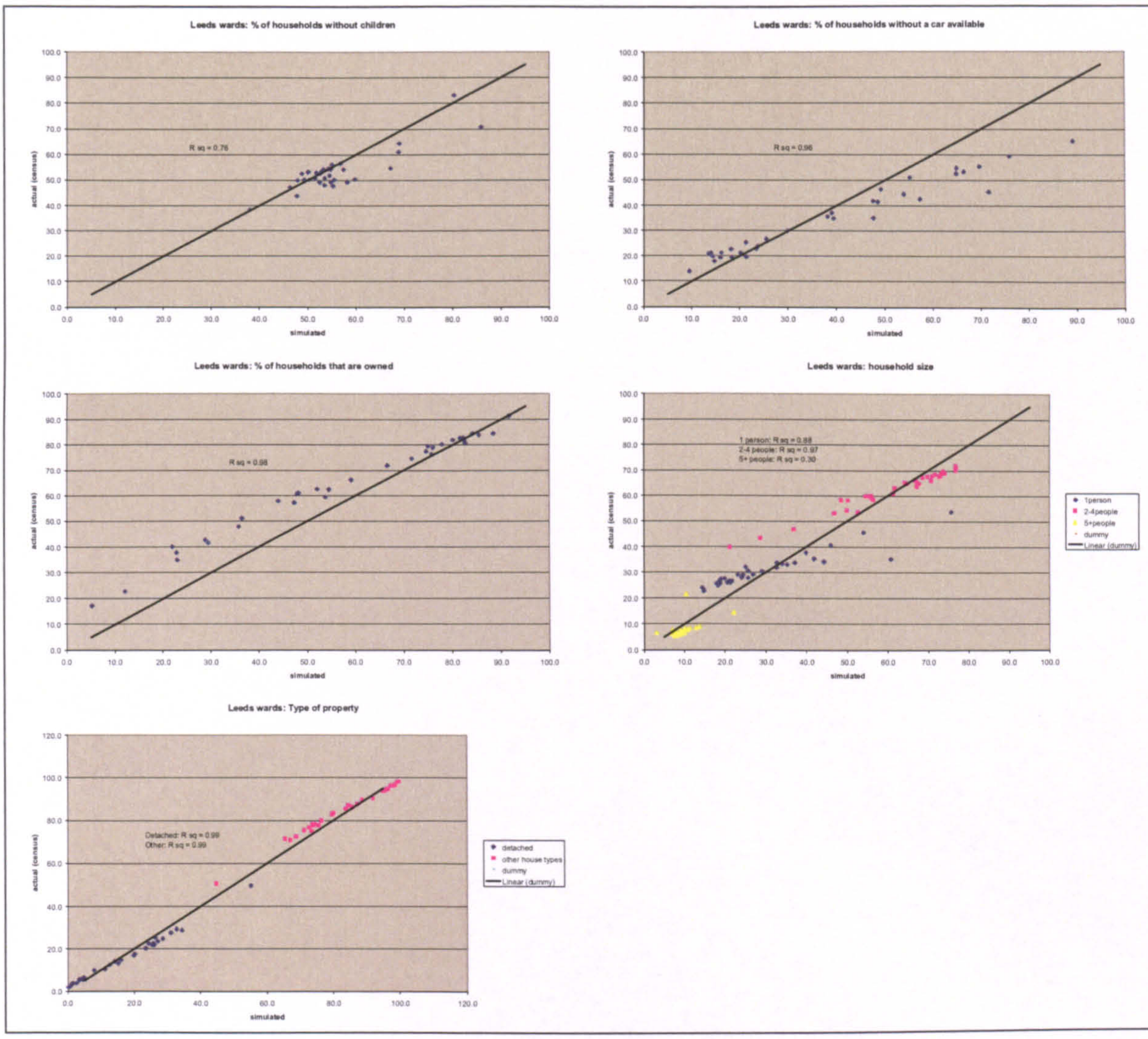


Figure 6.20. Scatter plots for each pair of constraint variables (there are no non-constraint variables in this simulation) at the ward level from the EFS simulation. The simulated value is on the x axis, and the expected value on the y axis. The dummy regression line indicates the "perfect fit" of y = x. The coefficient of determination for each variable is also given; the closer the number is to 1, the better the fit.

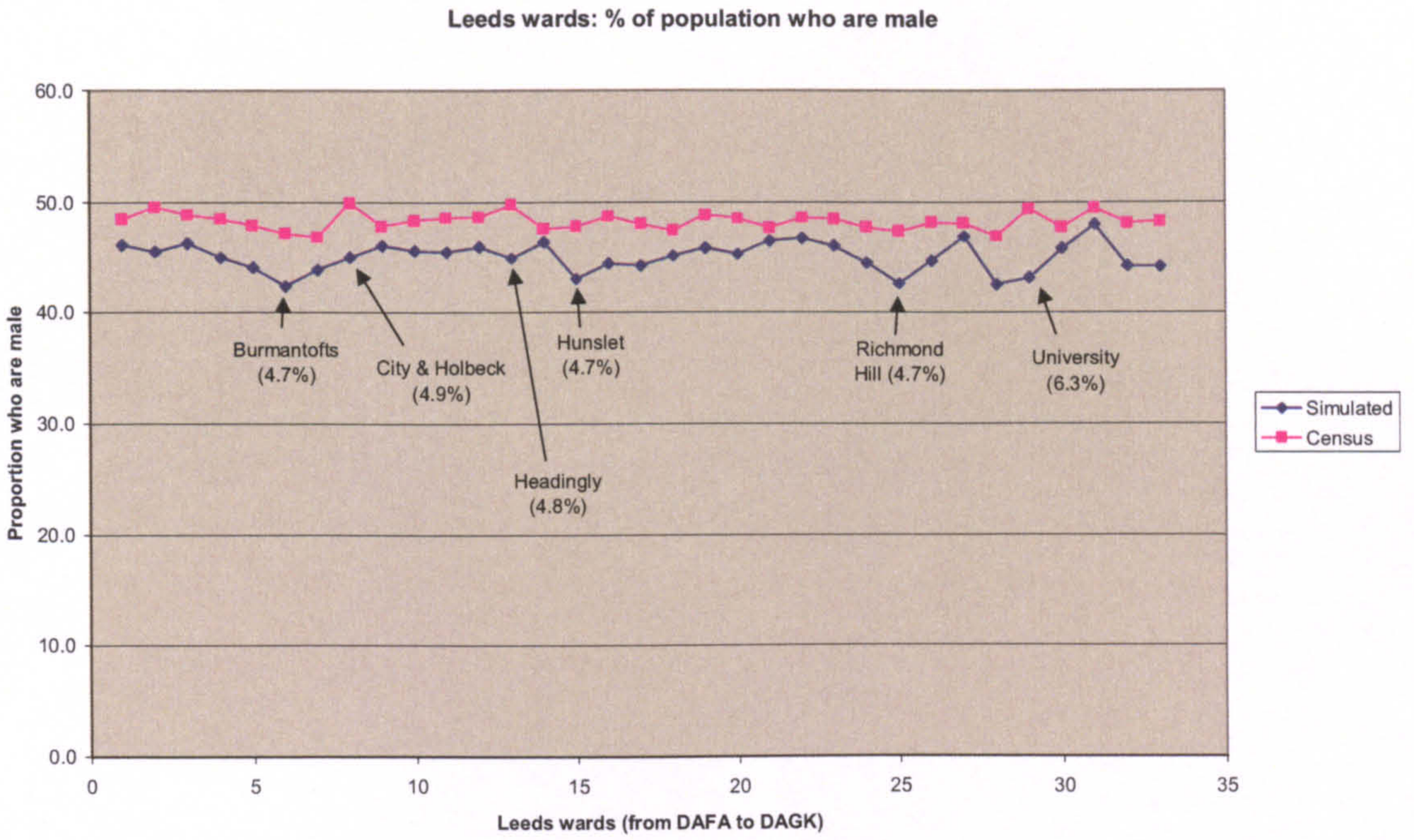


Figure 6.21. Simulated against actual proportion of population who are male (from the HSE simulation). The ward is along the x axis, and both the simulated value and the expected values are on the y axis, to highlight wards that differ the most (with the percentage difference given in brackets).

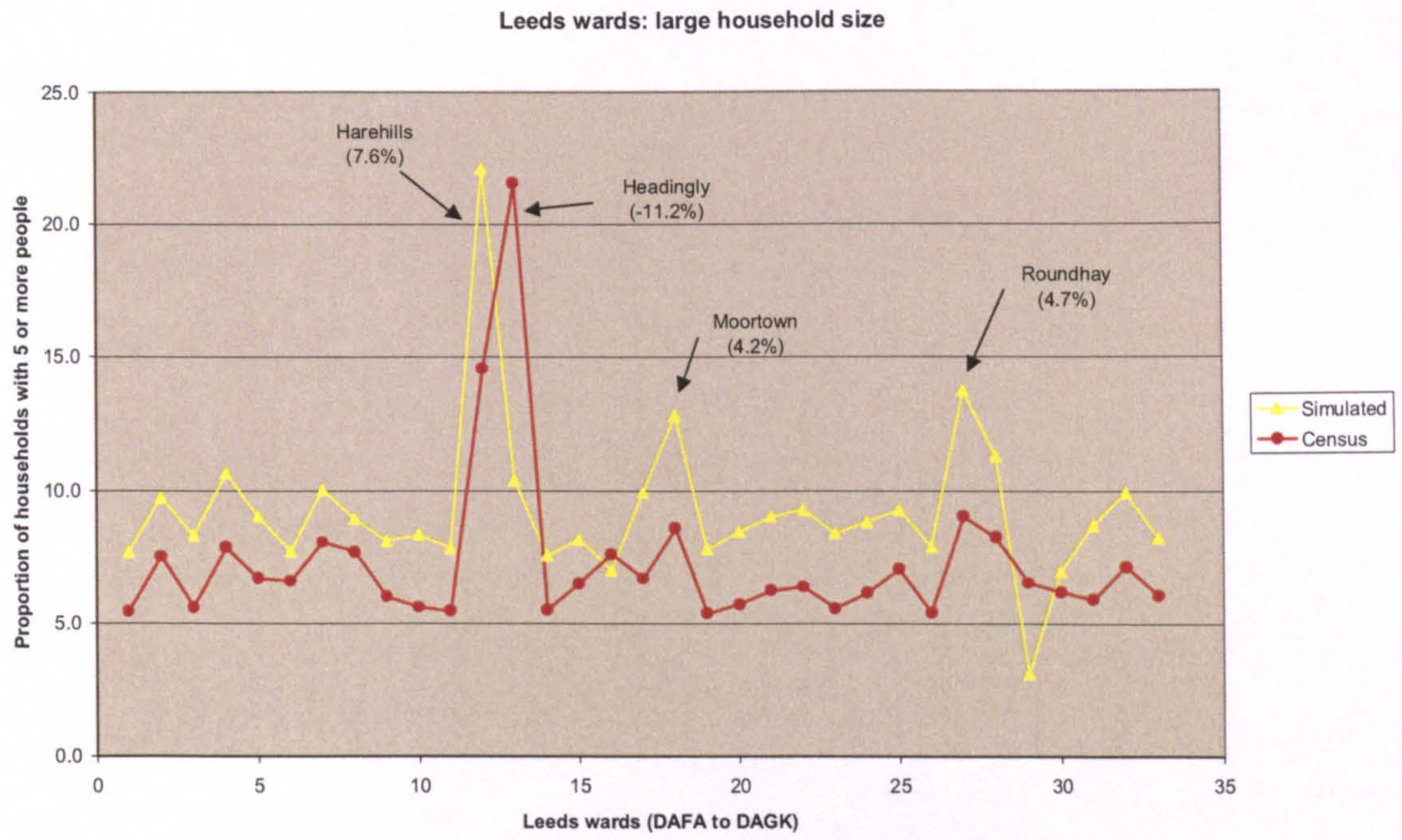


Figure 6.22. Simulated against actual proportion of population who live in households with 5 or more people (from the EFS simulation). Ward is along the x axis. Both the simulated value and the expected values are on the y axis, to highlight wards that differ the most (the percentage difference in brackets).

Proportion of population:	T-Test	
CONSTRAINT VARIABLES		
Who are male / female	0.000	***
Aged 0–2 years	0.616	
Aged 3-6 years	0.639	
Aged 7-9 years	0.821	
Aged 10-12 years	0.572	
Aged 13-15 years	0.644	
Aged 16-18 years	0.254	
Aged 19+ years	0.852	
Who are children / adults	0.852	
Least deprived	1.000	
Deprived 2	1.000	
Deprived 3	1.000	
Deprived 4	1.000	
Most deprived	1.000	
Highest educational qualification is degree, or equivalent	0.798	
HEQ is A or O levels, or equivalent	0.143	
No or other qualifications	0.685	
Full time student / child	0.053	
Who are white / other ethnic group	0.965	
Who are home owners / rent or "other"	0.546	
NON-CONSTRIANT VARIABLES		
No children in household / Household has children	0.001	***
With at least 1 car or van available / without car or van	0.019	*
Who are married or cohabiting / not married	0.000	***
Economically active / not economically active	0.775	
Live in household on own	0.000	***
Two people in household	0.002	**
3-4 people in household	0.000	***
5 or more people in household	0.000	***

Table 6.29. Results of equal variance 2-tailed t-test, comparing simulated population (from HSE 2002) to the 2001 Census data for the census-type variables contained within the HSE dataset. The binary variables have the same result for both categorisations. Stars indicate significance level (*** significant at $p < 0.001$; ** significant at $p < 0.01$; * significant at $p < 0.05$).

Proportion of population:	T-Test	
CONSTRAINT VARIABLES		
Who are male / female	0.067	
Who are children / adults	0.430	
1 person lives on own	0.665	
2-4 people live in the household	0.772	
5+ people live in household	0.010	**
No children in household / Household has children	0.229	
Who are home owners / rent or "other"	0.236	
With at least 1 car or van available / without car or van	0.384	
Who live in a detached house / semi, terraced house, flat or other accommodation	0.793	

Table 6.30. Results of equal variance 2-tailed t-test, comparing simulated population (from EFS 2005) to the 2001 Census data for the census-type variables contained within the EFS dataset. The binary variables have the same result for both categorisations. Stars indicate significance level (*** significant at $p < 0.001$; ** significant at $p < 0.01$; * significant at $p < 0.05$).

6.5 Discussion

This chapter has presented the SimObesity spatial microsimulation model, covering in detail its construction and calibration/validation. It was built using Java: an object-oriented language like Java is the most appropriate for building spatial microsimulation models (Ballas et al, 2006) and enables SimObesity to have the advantage of presenting the data in list form (rather than matrices), with each individual having a list of attributes. A key class in the Java programming for SimObesity is a class called "reweighter", which includes the code for both the reweighting and integisation algorithms, resulting in the estimation of new information (including obesogenic environment and behaviour data) at the micro level. It is a novel application of spatial microsimulation modelling and enables the spatial analysis of survey and census data combined.

The key part of the process of building SimObesity is in the selection of the input (constraint) variables. This chapter has shown that these variables must be correlated with the output variables that are being synthesised if these output variables are to be simulated accurately (i.e. to closely match the characteristics of the actual population). By definition (as we are deliberately synthesising data that we do not otherwise have at the micro-level) there is no way to directly validate these data. Instead we can calibrate the input data (which by definition have actual data, in this case from the 2001 Census); plus we could validate non-constraint variables, although if they are not correlated with the input variables then it is not possible to guarantee that they simulate reliably. Validation of the output variables stems purely from the correlation analysis done in choosing the input variables, and then in determining whether these input variables simulate well; if yes, then it can be assumed that the correlated output variables also simulated well.

6.5.1 Validation / Calibration

Four methods were used to calibrate the data. The first two highlight different variables as less well simulated (gender, tenure and car availability) than the latter two methods (gender and large households). The first method, "eyeballing" the data, is an interesting exercise to get a global view of the results, but is not a valid method to investigate the effectiveness of the micro-level (e.g. ward / SOA) simulation as it summarises the data for the whole of Leeds. Secondly the use of SAE figures uses an arbitrary cut off figure of 10% to highlight outlying wards for particular variables. This analysis is rather simplistic and is not statistically considering the significance of any differences between wards. For the EFS simulation the SAE proportions were generally higher than for the HSE simulation. This may be due to the smaller population dataset, which only had around 7000 people, or due to variations between the census and EFS survey definitions, as the EFS survey was undertaken at household, not individual, level.

Combining two or even three years of EFS data, to create an initial population dataset of around 20,000 people, may produce lower errors. Thirdly the regression analysis visually provides information regarding how well the data fit the ideal simulation (i.e. y = x regression line) but the coefficient of determination describes how well the data fit the best fit regression line, rather than the ideal one. Thus this considers precision rather than accuracy. Finally the equal variance t-test compares each ward's simulated and actual results in turn, in order to determine whether any differences are statistically significant. On balance, this is probably the better method.

The calibration methodologies draw similar conclusions about the calibration of the simulated data. Most of the constraint variables simulate very well, showing no significant difference between the simulated and expected proportions (the key exceptions being gender and large households). However, spatial microsimulation may not work well for synthesising variables that are affected to a large degree by "external and localised factors", such as by public transport, the existence of a large university or prison or employer in an area (Ballas et al, 2005). This will be of particular relevance to the handful of outlying wards where two of the constraint variables (from a total of 11) simulated less well (statistically significant differences between the simulated and actual data).

.

For gender the underestimate of males is true across all wards (and vice versa for women) and there are six wards in particular with a large difference between simulated and actual values (albeit still only approximately 5% difference). For example, in Harehills 9675 males are simulated (total ward simulated population 21045) against 10244 males from Census data (total ward census population 21052). Of these six wards, three (City & Holbeck, Headingly, University wards) have particularly high actual proportions of males, all of which are in the residential catchment for the Leeds Universities, suggesting that possibility student populations are skewing these results. The other three wards (Burnmantofts, Hunslet, Richmond wards) have noticeably low simulated proportions of males, the rationale and impact of which is less clear. It may be that there is something geographically unusual about these wards; a localised feature (such as a safe house for drug users) that increases the actual male population above that expected by the simulation.

Household size is a further variable that warrants some discussion. In the EFS simulation (where it was used as a constraint variable), SimObesity under- and over-estimates some wards and the apparent differences are largely due to just four wards (differences of –11%, 8%, 5% and 4%) (Headingly, Harehills, Roundhay, Moortown, respectively). The ward with the largest difference, underestimating the proportion of large households in the simulation, is in the catchment area for the Leeds Universities. Accordingly houses in this area are likely to have

students living in households with many other students – to increase rents for landlords and reduce per capita rent for students. This would lead to exceptionally high actual proportions and may account for the underestimated simulated number of large households in this area. The other three outlying wards are all fairly central (and adjacent to each other) and all overestimate the simulated proportion of large households. Accordingly perhaps the house types in this ward are unusually concentrated on smaller occupancies (e.g. more ex-council houses or large houses converted to flats), thus leading to a low actual proportion of large households. However if the data for all wards are compared the actual proportion of large households in these three wards is higher than in other wards, just that the simulated proportions are even higher.

Tenure was a constraint variable for both simulations. In the HSE simulation the proportion of people who own their own home was underestimated in eight wards. Also the EFS simulation underestimates the proportion of households who own their own home in 12 wards. That is, for example, the census shows that 13,000 people own their homes and 8,000 people rent, yet the simulation results in 11,000 home owners and 10,000 renters. These wards largely overlapped; so it is the same wards with problems with the simulation of tenure data. The underestimate of home ownership (analogous to overestimate of households that rent their home) may be due to peculiarities about the populations in these wards. A lack of council estates or business buildings in these areas may account for the differences, as the census data shows that less people are renting than the simulation expects. Nevertheless there were no significant differences found between the simulated and actual values for tenure (equal variance t-test, $P = 0.546$).

The availability or otherwise of a car (or van) was a non-constraint variable for the HSE simulation and a constraint variable in the EFS simulation. Given this, it might be expected to see lower errors for this variable in the EFS simulation, which is the case. A little confusingly the HSE simulation underestimated the proportion of individuals without a car, whereas the EFS simulation overestimated. Most wards had very different results from the two simulations; further the HSE simulation showed significant differences between the simulated and actual proportions of people with and without a car available (equal variance t-test, $P = 0.019$), whereas the EFS simulation showed no significant differences (equal variance t-test, $P = 0.384$). These different simulation results highlight the need for correlated input and output variables (as only constraint variables simulate well reliably). The nine wards where the proportion of people without a car was overestimated may be due to localised features, such as the quality of the public transport or deprivation levels. Again, there are no significant differences between the simulated and actual values for car availability where used as a constraint variable (equal variance t-test, $P = 0.384$).

As might be expected most non-constraint variables simulated very poorly, showing a significant difference between the simulated and expected values. The exception to this was Economic Activity in the HSE simulation, where the simulated values were not significantly different to the census values. This is probably because economic activity is highly correlated to two of the input variables, namely age (as no children are economically active) (Cramer V = 0.619; $p < 0.001$) and qualifications (again most children, and all young children, are classified as students as they have not finished their education) (Cramer V = 0.713; $p < 0.001$).

Whilst it is not possible to explicitly quantify the impact of the under/over-estimate of proportion of males/large households on the obesogenic variables simulated, these differences were limited to a few wards. Further, more importantly perhaps, the use of five/six constraint variables, which did simulate well the bulk of which did simulate well, would serve to minimise any errors with the simulated obesogenic variables. The overall results from SimObesity for the two simulations show that the wards match the socio-economic characteristics from the 2001 Census very well. This means that the characteristics of the synthetic population closely match that of the census 'real' population.

6.5.2 Methodology limitations

Spatial microsimulation is both an art and science; which can be frustrating. The quality of the synthetic population is likely to be affected by the size of the sample used as a micro-data database, the number of constraint variables, the consistency of constraint tables and, vitally, the correlation of the output variables to the input variables.

It may be that the sizes of the sample populations were too small. The parent populations, the HSE 2002 and EFS 2005, were both very large datasets, incorporating 17517 and 6798 participants respectively. However these sample sizes may still be too small to generate a synthetic population that truly matches the actual population; error sizes may have been reduced with larger parent population files. It may have been worth increasing the size of the files by amalgamating two year's worth of data. However it was only the HSE 2002 (not 2001 or 2003) that included reliable data on child fruit and vegetable consumption and child physical activity levels, so this output variable would not have increased in size. The reweighting deterministic algorithm has been shown to be a useful tool for simulating micro-area data. However it is computationally intensive. Most computing time is spent on evaluating the difference between constraint tables and synthetic micro-data. Due to computational restrictions, SimObesity is limited to a maximum of six constraint variables. It may be that with better java programming this maximum could be increased, however this is beyond the possibilities for this thesis. However with too many constraints and/or categories, the model can run into small number

problems. If any constraint categories were to have zeros or low numbers (in the census population), the less likely the synthetic population is to be a close match. Accordingly a trade off is required. Six constraints, mostly binary, was a good compromise between too few and too many; the EFS survey did not even have six possible constraint variables to use.

It was explained that the constraint tables needed to be adjusted to minimise discrepancies between the total population summations for each micro area. The adjustment method utilised to standardise the consistency of the census tables (i.e. equating all the totals to those in the sex census table, UV003) was simplistic, but it did ensure that the constraint tables were more consistent. It may have been better to use a mean of the total number of individuals living in Leeds from either all the census tables used, or even all possible individual level census UV tables. However, the end result was the same, in that consistency was achieved. Plus at the small areas considered, the small differences involved would probably not make any difference to the output that resulted.

The difficultly with the HSE dataset is that the SimObesity simulation is going to synthesise many different output variables from the same simulation. However the output variables have different optimal combinations of constraints. Accordingly it is necessary to choose a combination of constraint variables that predict all the output variables. The step wise logistic regression analysis determined the optimal combination of constraint variables as being sex, age, deprivation, qualifications, ethnicity and tenure. Using a entry method logistic regression for this optimal combination showed that all nine output variables (fruit and vegetable consumption; physical activity levels; degree of urbanisation; socio-economic group; public transport facilities; leisure facilities; access to supermarket; problems with teenagers hanging around; problems with vandals or graffiti) showed a highly significant improvement in fit when the predictor variables were added to the model. The predicted values do not differ significantly from the observed values for any output variable (as shown by the non significant Hosmer-Lemeshow statistics) and the predictors are all making a significant contribution to the prediction of the outcome variables (from the mostly significant Wald statistics). Also the model using this combination is explaining much of the variability in the data for many of the output variables; in particular, physical activity 24%, urbanisation 13%, SEG 18%, vandalisation 10%. Accordingly one can say that these six constraint variables predict the output variables well.

Whilst more than one output variable is being analysed from the EFS dataset, the principal variable was household expenditure on food, and so this was the variable used to determine the choice of constraint variables. Other output variables that also showed a correlation with this combination of constraint variables would also be simulated. This is different to the HSE

dataset constraint methodology, but has the same end result (i.e. the output variables are correlated to the input variables). Logistic entry method regression analysis showed that the five constraint variables were all important predictors of expenditure on food. Cramer V correlation analysis showed that ownership of a PC and more than one television, internet access, income level, source of income and spending on school meals were also correlated with the census-type variables.

Whilst the constraint variables chosen were all important predictors of the output variables, it would have been preferable to see stronger correlations ($r > 0.50$) between input and output variables as there is not a linear relationship between effect size and predictive power. Lower correlation figures have almost no value in prediction, with only a small improvement for medium correlation figures. It is only the high range of correlation coefficients where the predictive values become useful and reliable.

A key choice to be made early on in the building of SimObesity was whether to use a deterministic or probabilistic algorithm. This is largely determined by the datasets to be used to simulate the data and software availability. With hindsight, why write software for a new model if other software exists? Spatial microsimulation models are not simple to write/build. Whilst conceptually relatively easy to design, not everyone has the time and expertise required to actually build the model. This can be avoided by either contracting the work out or using a generic model, which are now available at the University of Leeds (but not when this thesis was started). That said, writing the code for the model gives the massive advantage of truly understanding the maths behind the numbers; and if the user does not understand exactly how the model is running this makes it difficult to adapt or update the model and encumbers validation and calibration. However the author could not find any papers directly comparing differences in these methodologies with conclusive evidence as to which is "best".

6.5.3 Uses for SimObesity

The SimObesity simulations result in the creation of a small area population micro dataset consisting of a profusion of health-related variables and socio-economic variables, providing the appropriate environment for an examination of the interdependencies between these variables. These data can be used to undertake analyses of the risks of being obese and the impact of living in an obesogenic environment at the small area level in Leeds (as well as looking at the relationship between obesity and obesogenic behaviours), particularly by reading the output data into ArcGIS software for further analysis and mapping (see chapter 7). Microsimulation modelling facilitates this analysis as it brings the data, whether spatially linked or not, down to the required micro-level. In particular, it gives individual/household estimates of obesogenic

variables, thereby developing information about the relationship between obesogenic factors and childhood obesity. A key advantage of this type of analysis is that mapping health data down to micro levels enables us to focus on key problem areas, rather than relying on averages for the whole region as per a global analysis. To generalise for the whole of Leeds would mean that health professionals could miss small problem areas. Highlighting them enables more focused interventions to be designed and implemented, levelling health inequalities.

Another great advantage of spatial microsimulation is the ability to link data from different sources. The resulting synthetic populations had attached details of many different variables, including sex, age, deprivation, qualifications, car availability, tenure, diet, physical activity, income, expenditure on food, ownership of goods, perceived social capital, etc (see Table 6.10). Some of these variables corresponded to those found in the 2001 Census and some did not; the former variables were used to calibrate the model and the latter variables were the ones that we were interested in simulating as we do not have data at the micro-level for these variables. This creates new population micro-level data unavailable from the published sources. Also it is a good substitute for conducting a detailed survey to produce such a rich dataset at the micro level, which would be expensive and time consuming. It should be noted that once the list of individuals and their attributes has been simulated, the individuals can be aggregated to any new geographical scale. SimObesity used data from the HSE 2002 and EFS 2005 to add obesogenic covariate data (all non-census attributes) to the simulated micro-level database. In the context of this thesis these attributes are important for investigating the relationship between obesity and the obesogenic environment/behaviours and this process will be described in the next chapter. This spatial microsimulation model could also be used to address other research questions, depending on the data available in the various national level surveys, such as National Diet and Nutrition Survey or British Household Panel Survey.

In this way spatial microsimulation modelling also facilitates policy formation and evaluation. By using the output to identify clusters of obesogenic attributes and hot spots of problem areas, spatial statistical analysis can show which, if any, hot spots are significant. This increased understanding of obesogenic patterns will facilitate effective policy formation to reduce obesity prevalence by enabling targeted (to the local area) interventions and health policies to be determined.

On the whole, it can be asserted that the spatial microsimulation methodology detailed in this chapter can be used to give constructive information on obesogenic-related variables (as well as other health-related variables and socio-economic trends) that could be extremely useful in health-related applications. The use of spatial microsimulation techniques provides a rich dataset that can facilitate the understanding of variations in the obesogenic

environment/behaviours at the micro level and to enable the consideration of the relationship between these factors and childhood obesity, and identifying pockets of problem areas. This increases our understanding of childhood obesity, providing an enhanced environment for analysis, evaluation and decision-making in health planning.

6.5.4 Conclusion

This paper explains how SimObesity, a spatial microsimulation model, was built and how it facilitates the identification of obesogenic covariates, including both obesogenic environment variables and behavioural variables.

Methodology

- This paper clearly explains the construction and calibration/validation of a spatial microsimulation model.
- The choice of constraint variables is critical to validation (low errors). The need for a strong correlation to exist between the input (constraint) and output variables is emphasised and rationale explained. This key statistical factor has not been stressed in other microsimulation work.
- The simulated values were a good match to the census data.

Application

- This paper demonstrates a methodology to predict micro-level obesogenic variables, highlighting "hot spots". Subsequent analysis of these data will increase understanding of obesogenic patterns and facilitate effective policy formation to reduce obesity prevalence
- SimObesity also has implications for policy evaluation, assessing the likely future impact of policy change at the local level.
- Generalisable to other health conditions / areas

Chapter 7: Micro-level analysis of childhood obesity, diet, physical activity, residential socio-economic and social capital variables: where are the obesogenic environments in Leeds?

7.1 Introduction
7.2 Background:
7.2.1 Obesogenic Environments
7.2.2 Geographically Weighted Regression
7.3 Methodology
7.3.1 Sources of data
7.3.2 Analysis of data
7.4 Results
7.4.1 Descriptive analysis
7.4.2 Global analyses
7.4.3 Geographically weighted regression analyses
7.5 Discussion
7.5.1 Global analyses
7.5.2 Local analyses
7.5.3 Relationship with income
7.5.4 Limitations
7.5.5 Conclusion

7.1 Introduction

This chapter describes global (the whole of Leeds) and local (at Census low super output area (SOA)) analyses of the relationship between childhood obesity and many potential obesogenic variables, including obesogenic environment variables such as deprivation, urbanisation, socio-economic group, income, access to local amenities and perceived neighbourhood safety, as well as individual obesogenic behaviour variables such as dietary and physical activity behaviours. The covariate data are all synthesised at the individual level using spatial microsimulation (SimObesity). This builds on the work undertaken in chapters 4 and 6.

In chapter 4 the actual (sample) obesity data were analysed in order to identify the locations with high prevalence of childhood obesity. Spatial auto-correlation (which exists when the location of disease cases are dependent on the location of other disease cases) was not adjusted for initially in order to enable the detection of clusters that exist due to these correlations. If they are adjusted away, important hot spots may go undetected and interventions not targeted to or implemented in the right areas. The null hypothesis was effectively that obesity cases are geographically randomly distributed (allowing for population density etc) and the alternative hypothesis is that some hot spots exist due to either differences in underlying risk factors or spatial auto-correlation. Chapter 4 then went on to consider the impact of deprivation and OAC super-groups to determine a potential determinant of the clusters. However this chapter will consider whether various neighbourhood features or behaviours, such as perceived supermarket access or fruit and vegetable consumption, contribute to a higher risk for childhood obesity. It is essential to use spatial regression in order to adjust for the inherent spatial auto- correlation in the data. If not, the confidence in the risk relationships will be overestimated, leading to too

small biased p-values (Kulldorff, 2006). That is, relationships would show to be statistically significant when in fact they are not.

Accordingly these analyses will identify the covariates with the strongest relationships with obesity, as well as highlighting the variation in these relationships across the study area. For example, some SOAs will have a positive relationship and others may have a negative relationship, within the bounds of an overall (global) positive or negative relationship. Thus identifying "at-risk" populations, which could be defined spatially and/or by any of the covariates analysed. It seeks to demonstrate the importance of analysis at the micro-level in order to provide health planners with additional information with which to tailor interventions and health policies to prevent childhood obesity. These analyses will also briefly consider whether the correlation between obesogenic factors and the prevalence of childhood obesity varies with household income or deprivation levels.

This analysis will not incorporate a structured assessment of how obesogenic factors can be changed to reduce childhood obesity. This will be addressed in Chapter 8.

7.2 Background

7.2.1 Obesogenic Environments

Obesity is a significant social, medical and economic problem. Its prevalence has escalated over the last two decades, reaching pandemic levels in the developed world (Ebbeling et al, 2002; Sproston & Primatesta, 2002; Lobstein et al, 2004). Ignoring the obesity epidemic would have serious detrimental public health consequences, in terms of the financial burden on health care systems, individual morbidity, and ultimately (early) mortality. Why are some people more obese than others? The "it's my/his/her genes" is often cited by individuals as a primary cause. However the literature does not support this view. A review by Allison et al (2001) suggests that only about 20% of the population have strong genetic predispositions to be obese or slim regardless of the environment; accordingly why do some places have more obese people than others? Do obese people congregate in similar locations (compositional) or do certain attributes of places cause its inhabitants to become obese (contextual)? These are compelling questions for researchers.

The very rapid rise in childhood obesity suggests that environmental factors rather than single gene defects are the primary cause. Population studies also suggest a strong influence of environmental factors on obesity rates: migrants often have higher obesity than those still living in the country of origin (Ravussin et al, 1994; McDermott et al, 1998; Popkin & Udry, 1998);

immigrants' offspring have higher rates of obesity than their parents (Popkin & Udry, 1998); second generation children have higher obesity rates than first generation children (Popkin, 1998); and, when developing countries' populations have switched to more Western diets and reduced physical activity levels, a rise in the prevalence of childhood obesity is shown (Drewnowski & Popkin, 1997; Wang et al, 2002). Obesogenic environments are thus one of the major explanations for the increasing prevalence in obesity. The definition is a broad concept, being "the sum of influences that the surroundings, opportunities or conditions of life have on promoting obesity in individuals or populations" (Swinburn et al, 1999). Obesity is a result of chronic positive energy balance and yet we all have choices over how much we eat and how much exercise we take. An obesogenic environment is one in which the easy choices are unhealthy choices, promoting a high energy-intake and sedentary behaviours, which makes obesity more likely to occur.

If the environment is affecting peoples' behaviour, leading to over-eating and/or under-exercising, then examining diet and physical activity in isolation is not sufficient to prevent obesity. Rather than focusing preventative strategies on education or behaviour change, a population based, long term, line of attack may be to identify (and then change) the obesogenic environments. In considering whether place affects prevalence of childhood obesity it is important to consider the scale of the analysis. Operating at purely a global scale, say for a whole city, will "average out" small areas of high prevalence such that the mean can be deemed acceptable and the pockets of problem areas are ignored, or rather, not noticed. By considering a smaller scale, the impact of obesity at the micro-level can be highlighted. This is important, because it may well suggest different obesogenic factors are important in different locations, thus not only increasing our understanding about geographical patterns of obesity, but also providing additional information to the health planner, enabling more targeted interventions to prevent childhood obesity to be drawn up and implemented. The aim of this chapter is thus to explore whether variations in the obesogenic environment exist across the neighbourhoods of a major UK city, namely, Leeds.

7.2.2 Geographically Weighted Regression

It is important that spatial data are not simply aggregated and analysed globally. This is illustrated in Figure 7.1, which shows how aggregating spatial data can lead to an incorrect conclusion about the underlying relationships, in this case leading to a supposition that obesity and household income are positively related (i.e. higher income is associated with higher obesity rates) whereas if the data are disaggregated into the two areas that the example data stem from, it is clear that both areas have a negative relationship between income and obesity (this example uses made-up data to illustrate a point).

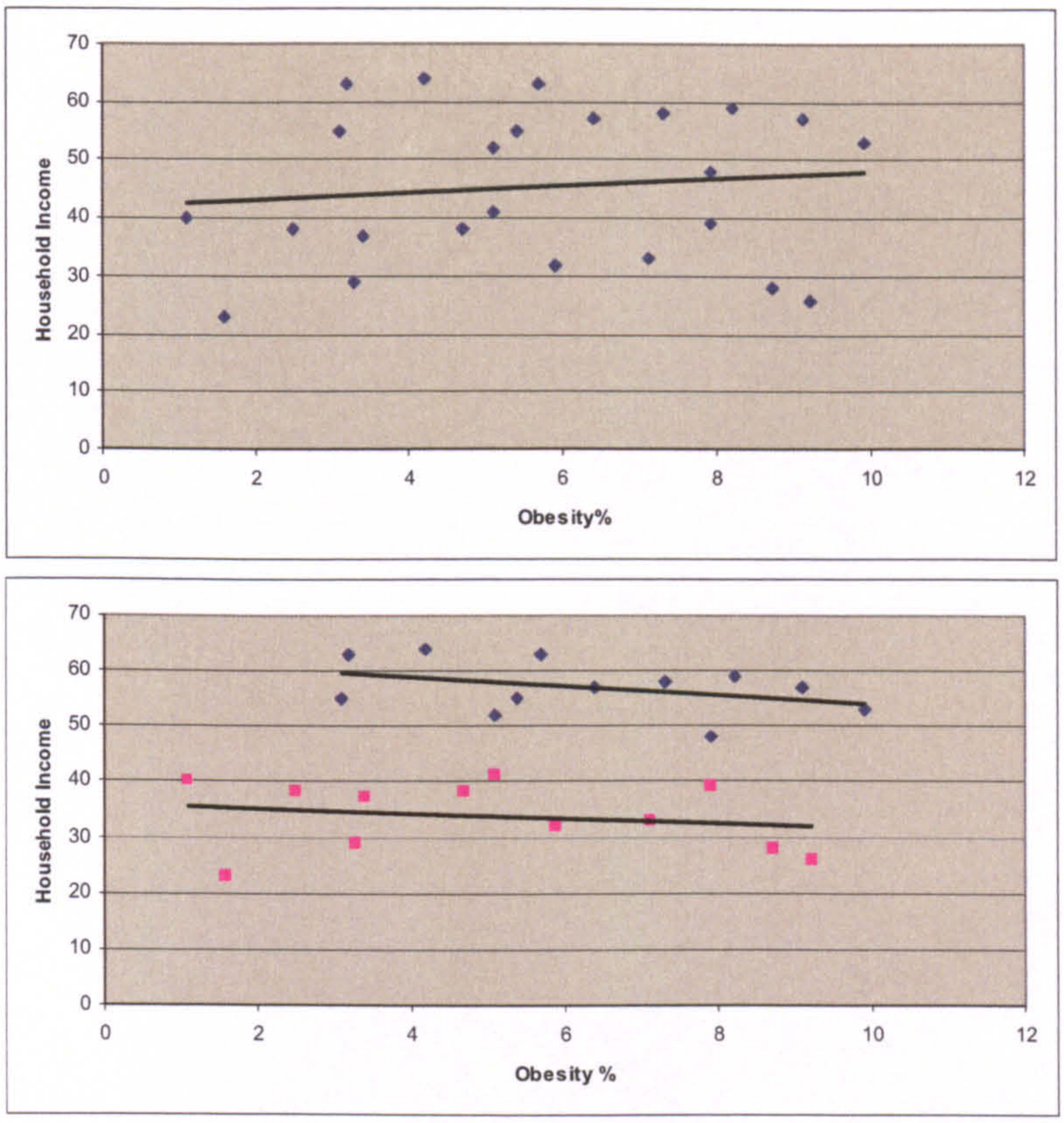


Figure 7.1. Graphic illustration of the importance of analysing spatial data correctly (adapted from Fotheringham et al, 2005). Importantly both scatterplots use the same (theoretical) dataset. The top diagram shows the data spatially aggregated, showing a positive trendline. However if the two datasets are disaggregated, as in the bottom diagram, say to reflect two different areas, the both areas actually show a negative relationship between obesity and household income.

A "normal" linear regression model assumes processes are stationary and the parameter estimates obtained are constant over space (i.e. the results are location independent), which does not highlight potentially important local variations in relationships (Brunsdon et al, 1998; Shi et al, 2006). That is, it assumes the relationships being modelled are the same across the entire study area. Conversely, a local model is a spatial disaggregation of a global model, and the results are, accordingly, location specific; a local model allows the processes under investigation to vary spatially. Spatial non-stationary exists when the same stimulus provokes a different response in different parts of the study region.

This non-stationarity is explained further using an example (from Fotheringham et al, 2005). Suppose a non-stationary process was being modelled using (incorrectly) a global model. Then the model would incorrectly assume the values of β in the regression equation are the same in all localities across the whole study area. This means that any spatial non-stationarity in the relationship can only be seen through the residuals (the difference between the observed and predicted data) of the model. Example residuals from an incorrectly applied global model to spatial data are shown in Table 7.1. As such, whilst it is possible to determine whether there are

any spatial patterns by simply mapping a global model's residuals (or alternatively to calculate a residuals' auto-correlation statistic), it is better to address the issue of spatial non-stationarity directly using geographically weighted regression software that allows the relationships being measured to vary over space.

1.

0.9	0.8	0.8	0.7	0.5
0.8	0.7	0.6	0.5	0.4
0.7	0.6	0.5	0.4	0.4
0.6	0.5	0.4	0.3	0.2
0.5	0.4	0.3	0.2	0.1

2.

0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5

3.

+	+	+	+	0
+	+	+	0	-
+	+	0	-	-
+	0	-	-	-
0	-	-	-	-

Table 7.1. Example of the problems of using a global model to predict a non-stationary process. The top table (1) represents the (example) real values of β_i. Accordingly if a global model is used, the estimated values of β_i will be as per the middle table (2) leading to incorrect residuals ($y_i - y_i'$) as per the bottom table (3) (Fotheringham et al, 2005).

The main output from GWR is a set of local parameter estimates for each relationship. The extent of the spatial variability of any relationship can be ascertained by comparing the range of the local parameter estimates with a confidence interval around the global estimate of the equivalent parameter. By definition (assuming Gaussian distribution) 50% of the local parameter estimates will lie within the inter-quartile range and approximately 68% of the values will lie between +/- one standard deviation of the mean. As the global parameter estimate is the global mean, if the inter-quartile range of the local estimates is greater than two standard deviations of the global mean (i.e. twice the standard error), then this suggests the relationship might be non-stationary (Fotheringham et al, 2005). A Monte Carlo test is used to assess whether the spatial variation in the measured relationships are statistically significant.

The key equations for a geographically weighted regression model are given in Figure 7.2. The main output from GWR is a set of location-specific parameter estimates, which can be mapped across the study region to provide information on spatial non-stationarity in relationships (as well as producing a more accurate regression model). GWR is able to undertake both a global

and a local regression analysis, which is important because what GWR can not do is take a badly fitting global model and suddenly create a well fitting local model. The difference is purely to take account of spatial non-stationarity, so the initial global model must be a good fit before the local model can be considered.

$$y_i = \beta_0(i) + \beta_1(i) X_{1i} + \beta_2(i) X_{2i} + \ \ldots \ + \beta_n(i) X_{ni} + \varepsilon_i$$

Regression equation

$$\beta'(i) = (X^T W(i) X)^{-1} X^T W(i) Y$$

Estimator equation

Figure 7.2. Key equations for GWR regression model, where i refers to a location at which data on y and x are measured and at which local estimates of the parameters are obtained. The estimator where W(i) is a matrix of weights (which can be defined as fixed or adaptive) specific to location i such that observations nearer to i are given greater weight that observations further away. That is, W_{in} is the weight given to data point n for the estimate of the local parameters at location i (Fotheringham et al, 2005).

7.3 Methodology

7.3.1 Source of data

The study area covers the 476 Leeds Census Super Output Areas (SOAs) (see Figure 1 in chapter 4). Childhood obesity was defined using age and gender specific body mass index (weight (kg)/height (m^2)) standard deviation scores (BMI SDS) using the British 1990 growth reference dataset (Freeman et al, 1995; Cole et al, 1995); overweight is defined as above the 91st centile and obesity as above the 98th centile (Cole et al, 1995) (as previously described in Chapter 2).

There are three key sources of child BMI data from Leeds used in this study: for 3 to 6 year olds was obtained from primary care trusts' records of routinely collected data for children born since 1995; a sample of 5, 9 and 13 year olds collected as part of the "Trends" study in 2004 and 2005 (Rudolf et al, 2006); and, a sample of 11 year olds collected as part of the "RADs" study in 2005 and 2006. Further details regarding these studies and also regarding the cleaning of these data is described in chapter 4. No children were contacted directly, and no new measurements were taken. All data were anonymised. A geographic information system (GIS) (ArcGIS v.9.0) was used to summarise the childhood obesity datasets by residential SOA. The SOA of the residential location of each measured child was determined by linking each child's residential full postcode centroid to the relevant SOA using GIS software. To minimise small number problems the aggregate obesity data (percentage of obese children; percentage of obese

and overweight children; mean BMI SDS) for each SOA was smoothed using Empirical Bayes (with a local smoothing based on population limits, i.e. children measured). Ethical approval was obtained from Leeds (East) Research Ethics Committee.

A number of "obesogenic factors" were considered - namely, obesogenic behavioural variables: fruit and vegetable consumption, purchasing school meals, spending on food per household (eaten at home and externally), physical activity levels; obesogenic environment variables: perceived social capital, urbanisation, socio-economic group (SEG), size and source of household income, number of televisions per household, PC ownership, and internet access. These data were obtained from the microsimulation model, SimObesity (as described in chapter 6), using the Health Survey for England 2002 (National Centre for Social Research and UCL, 2002) and the Expenditure and Food Survey 2005 (Office for National Statistics and DEFRA, 2005) as the base populations as well as data from the 2001 Census (Office for National Statistics, 2001a; 2001b). Deprivation score was determined using the Index of Deprivation, 2004 (Communities and Local Government, 2004). The lists of synthetic individuals in each SOA and their lists of attributes, were aggregated to SOA level in a GIS (ArcGIS v.9.0) and the results given as proportions of the population in that SOA (rather than absolute numbers) for analysis. All data are categorised to facilitate this aggregation (see Appendix G). As the simulated data represent the whole population of Leeds (over 715,000 individuals), there are no small number problems.

7.3.2 Analysis of data

The proportion of children who were obese, or overweight and obese, plus the mean and range for BMI SDS, number of children measured, and each of the simulated variables were calculated. The obesity data were mapped, at ward and SOA level, using ArcGIS v.9.0 to visually identify any spatial patterns across the study area. Similarly the obesogenic covariates data were mapped, showing the proportion of the population in each SOA for each covariate.

A series of geographically weighted regression (GWR) analyses, at both global and local level, were undertaken (GWR3 software, version 3.0.1) (Charlton et al, 2003). GWR analyses spatial variations in relationships in the data - permitting the investigation of whether any non-stationary relationships between obesity and the predictor variables exist. It also allows the spatial auto-correlation that is inherent in the data to be accounted for. Percentage of measured children who were obese was the dependent variable (although percentage of overweight and obese children and mean BMI SDS were also separately considered). The simulated variables were the independent variables. Univariate analyses were undertaken (multivariate analysis was

complicated by the extensive collinearity in the data, given the strong correlations ($r > 0.9$) between many of the predictor variables – results not shown).

The global regression model constructed for each covariate against childhood obesity was evaluated first, in order to gain an understanding about the overall relationship between the covariates and childhood obesity in Leeds. This process also enabled the identification of the covariates with the strongest global relationship with obesity. The Pearson correlation coefficient was also calculated (using SPSS v12.0) for each pair of simulated predictor variables and obesity to provide further information regarding any global correlations between the covariates and childhood obesity. A 2-tailed test was used to determine significance.

Next the local regression model for each covariate against childhood obesity was assessed to determine two things: firstly, whether the relationship with obesity was non-stationary, and, secondly, where different (to the global perspective) determinants of obesity were key. Location was input as the x/y coordinates for the centre of each SOA polygon. Kernal type was fixed, kernel shape Gaussian, bandwidth selected by AICc minimisation, with a Monte Carlo test for spatial variation significance. The main output are a set of location-specific parameter estimates, which were mapped (using ArcGIS v.9.0) to provide information on spatial non-stationarity in relationships. Also, the parameter estimates for each covariate were ranked for each SOA in order to prioritise the most important local determinants of childhood obesity in each area, and which may help to understand what drives obesity in specific localities. Maps of the areas where the key local determinants of childhood obesity differed to the key global determinants were created to visualise the geographic differences. The boundary data for the study area were downloaded from UKBORDERS[10] in a form compatible for use with ArcGIS v.9.0 software.

The use of multi-level modelling techniques (Rasbash et al, 2004) to analyse the spatial data was considered as this technique allows the dependency inherent in child observations nested within the same SOA to be taken into account. That is, it would partition the variation in obesity across each of the hierarchies, thus respecting the natural aggregation of the data (children at level 1, nested within SOAs at level 2) leading to less erroneous inferences than if the natural ordering had not been accounted for. However it is not possible to use multi level modelling analysis on this dataset, even though the obesogenic factors data are available at individual level and deprivation and obesity are available at SOA level. This is due to two reasons. First and foremost, MLWin (the MLM software) can not cope with a dataset with over 715,000 individuals (rows), which a model of Leeds has; there are just too many data, even if the number

[10] UKBORDERS is an online provider of digitised boundary datasets of the UK, funded by the ESRC (http://edina.ac.uk/ukborders/index.shtml)

of variables (columns) are reduced (maximum c.1 million cells). Secondly, arguably, there is some mixing of levels. That is, the actual obesity data, which has been aggregated to SOA level is really individual data, but for different individuals than the synthetic ones in this dataset. Similarly one of the (many) obesogenic factors that are attributed to the synthetic individuals, degree of urbanisation, is a neighbourhood level factor rather than individual factor. Finally, maps of census variables for low income, car ownership, low socio-economic group, and deprivation were prepared and compared to the local parameter estimate maps in order to consider any potential relationships.

7.4 Results

7.4.1 Descriptive analysis

The data show that the percentage of obese or overweight children in Leeds is 7.6% and 11.3% respectively, for children aged 2.5 years to 14.4 years. Mean BMI SDS was 0.4 (range 0.0 to 1.1). The mean number of children measured in each SOA was 71 (range 2 to 156). The map of these data at ward and SOA shows considerable variation in obesity across Leeds (see Figure 7.3).

SimObesity simulated data on many different variables made up of census-type variables (such as tenure or age) and obesogenic-type variables (such as fruit and vegetable consumption and supermarket access). The census-type variables have been discussed in the validation section of chapter 6. The global means and ranges for the obesogenic covariates are given in Table 7.2, which shows that there is a wide range of values for each covariate, suggesting there is considerable variation in the different SOAs. However these data are also available at a local level, i.e. for each of the SOAs in Leeds. Accordingly the obesogenic covariates were mapped at this level in order to visually identify any patterns (see Figures 7.4-7.6). Eyeballing the maps indicates that many variables have a strong central pattern, with a few outlying SOAs also with high/low proportions. The notable exception to this is the physical activity maps, which show the most dispersed pattern, with no real central pattern.

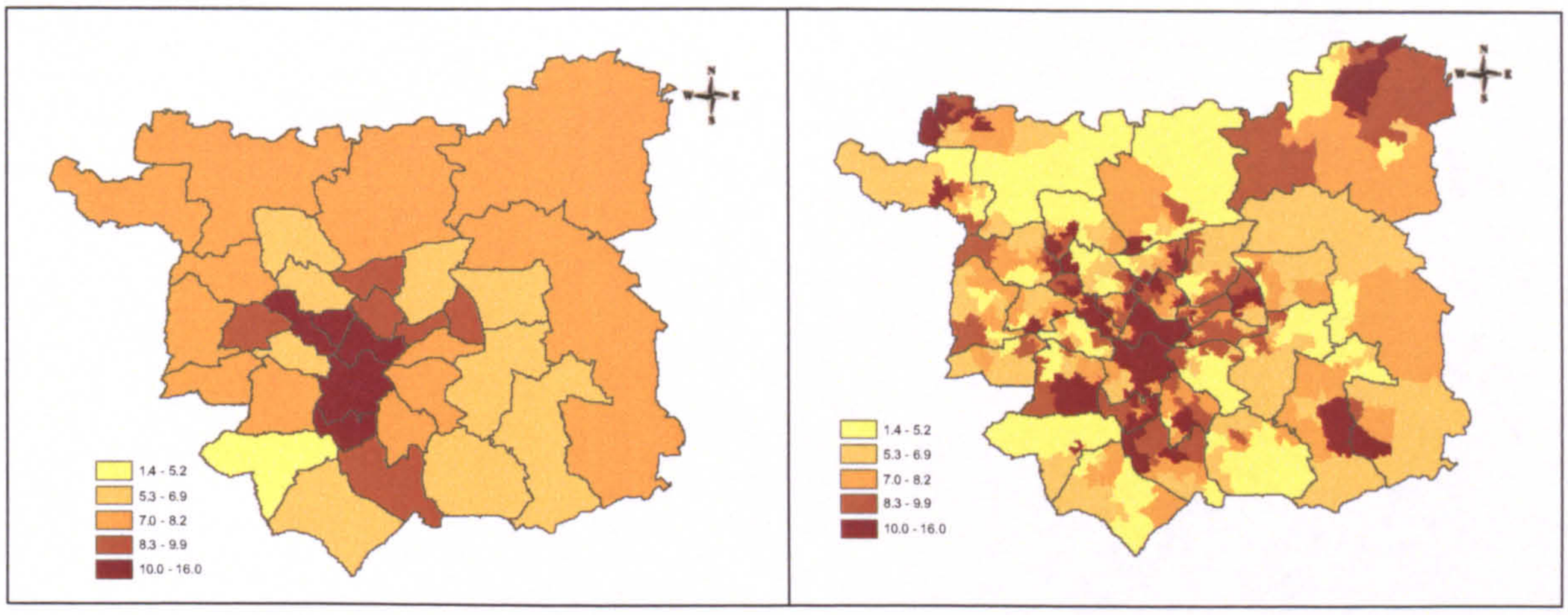
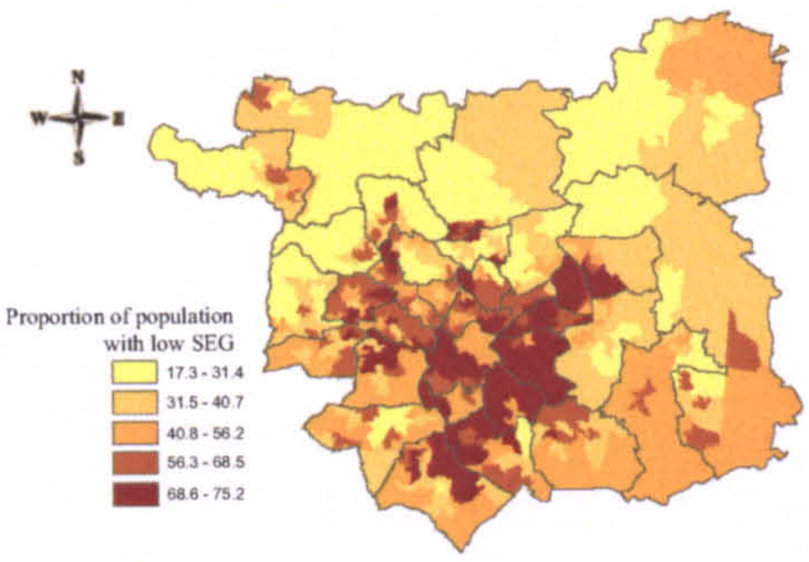


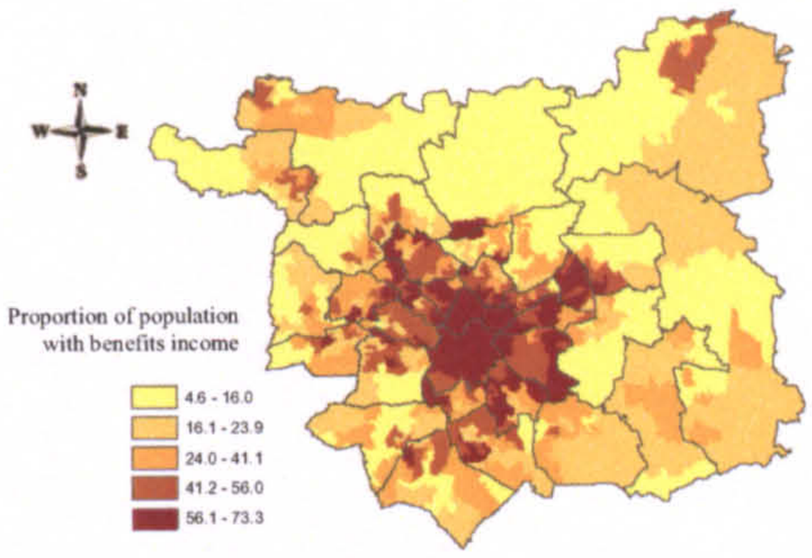
Figure 7.3. Map of childhood obesity across Leeds (proportion of measured children, in each area, who are obese). The darker the shading, the higher the prevalence of obesity. The left hand map is by ward, and the right hand map by lower super output area. Data were smoothed using Empirical Bayes and both maps use the same scale (based on a quintile scale for the lower map).

HSE variables	mean	min	max	EFS variables	mean	min	max
Obesogenic Environment Variables							
SEG_high	51.9	24.8	82.7	Household income low	33.1	7.1	67.8
SEG_low	48.1	17.3	75.2	Household income medium	52.0	30.8	64.3
urban	19.5	3.4	61.5	Household income high	14.8	1.2	44.3
suburban	59.0	37.3	65.9	Wage main source income	56.9	21.0	90.2
rural	21.5	1.3	36.7	Other main source income	7.6	1.8	19.3
Transport good	72.4	59.4	86.8	Benefits main source income	35.6	4.6	73.3
Transport bad	27.6	13.2	40.6	Own a computer	59.0	18.7	91.3
Leisure facilities good	62.4	51.5	72.6	Do not own a computer	41.0	8.7	81.3
Leisure facilities bad	37.6	27.4	48.5	One household TV	73.2	30.5	92.9
Easy access supermarket	93.1	88.1	97.3	More than one TV	26.8	7.1	69.5
Difficult access supermarket	6.9	2.7	11.9	Internet access	48.4	12.3	85.6
Teenagers are a problem	37.4	18.5	58.2	No internet access	51.6	14.4	87.7
No problem with teenagers	62.6	41.8	81.5				
Vandals are a problem	35.9	14.4	59.3				
No problem with vandals	64.1	40.7	85.6				
Obesogenic Behaviour Variables							
No fruit and vegetables eaten	9.2	2.3	17.8	Food Expenditure low	53.1	16.1	96.4
1-4 portions fruit and vegetables	70.5	65.3	74.4	Food Expenditure high	46.9	3.6	83.9
5 or more portions fruit and veg.	20.3	10.1	30.5	Food Expenditure £0-29	23.7	2.9	69.0
Child physical activity none	12.1	0.0	33.3	Food Expenditure £30-49	20.3	7.3	27.3
Child physical activity low	9.1	0.0	19.0	Food Expenditure £50-69	17.1	6.7	22.3
Child physical activity moderate	17.8	0.0	28.4	Food Expenditure £70-99	19.3	1.3	29.4
Child physical activity high	60.8	0.0	81.8	Food Expenditure £100+	19.7	0.9	48.3
Adult physical activity none	13.2	5.3	19.8	FE adusted for size - low	50.6	33.2	74.0
Adult physical activity low	33.8	24.6	42.5	FE adusted for size - high	49.4	26.0	66.8
Adult physical activity moderate	12.3	6.9	17.8	School meals eaten	14.3	0.0	34.4
Adult physical activity high	40.7	30.7	51.5	No school meals	85.7	65.6	100.0

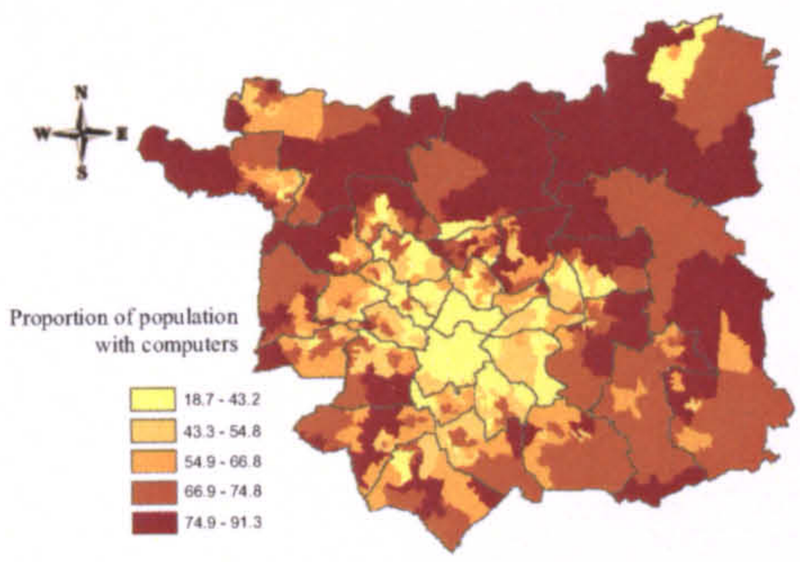
Table 7.2. Summary statistics for the obesogenic variables for the whole of Leeds. The left hand list come from the SimObesity simulation using the Health Survey for England 2002 (HSE) as the base population, and the right hand list from the simulation using the Expenditure and Food Survey 2005 (EFS). Both lists distinguish between the obesogenic environment variables, and the individual obesogenic behaviour variables. Each number represents the proportion of the population living in that SOA that display that attribute.



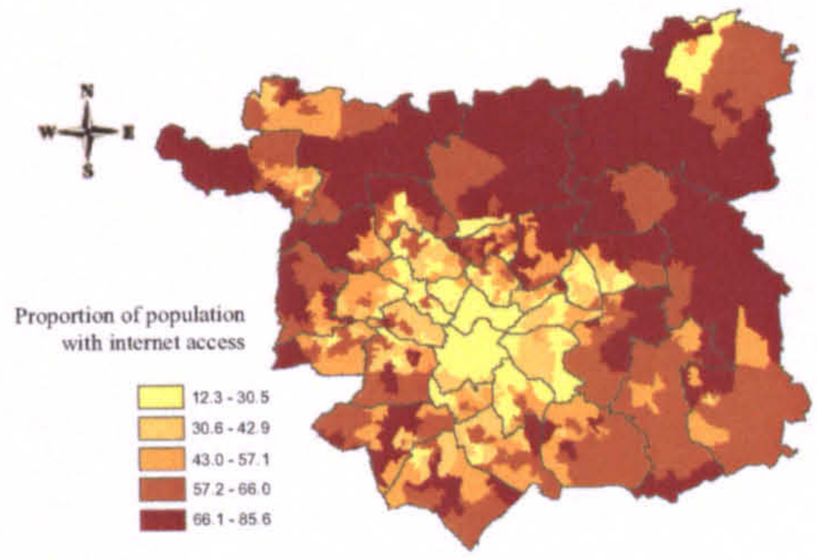
SEG low



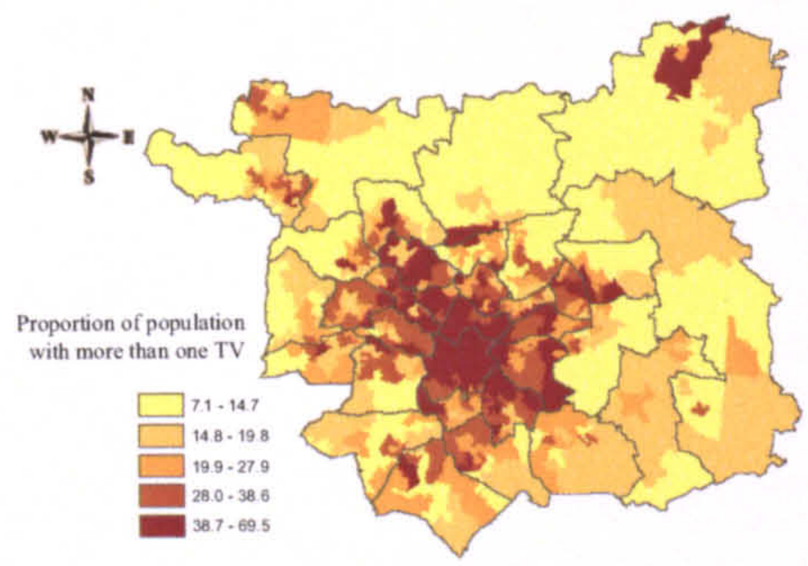
Main income from benefits



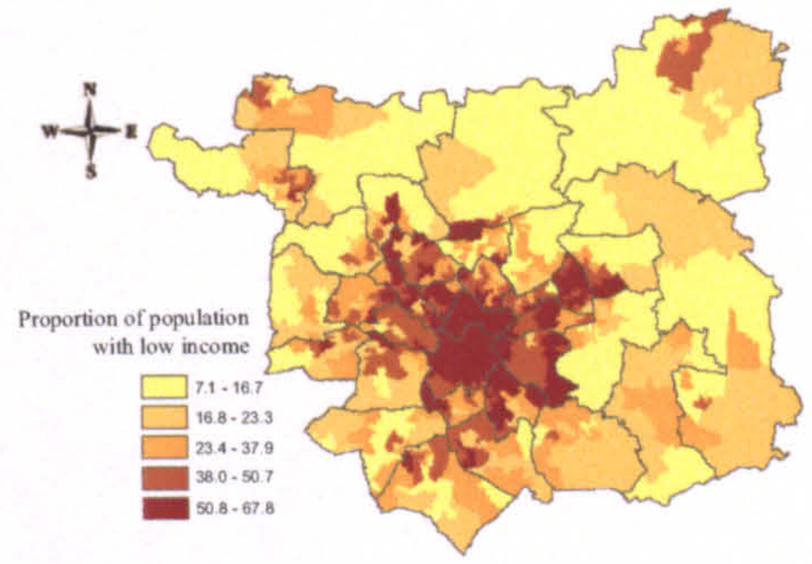
Own a computer



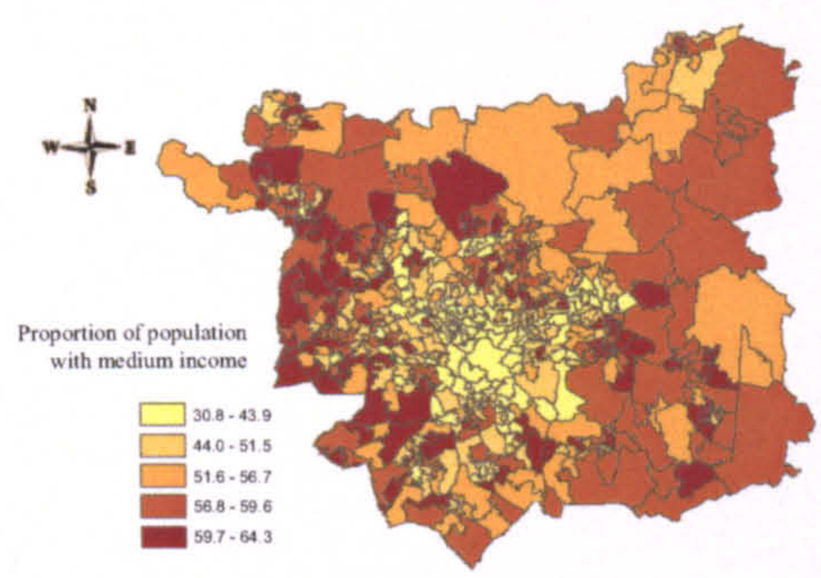
Have internet access



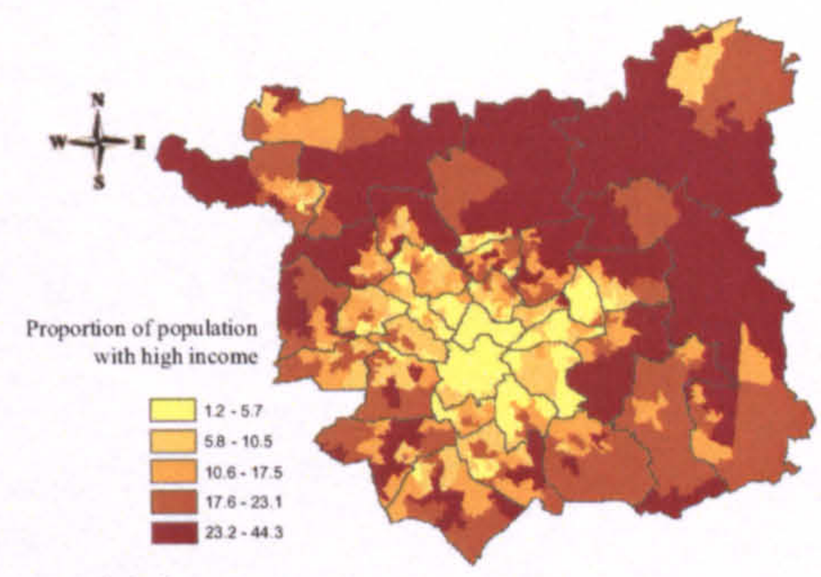
More than one TV per household



Income low



Income medium



Income high

Figure 7.4. Maps of obesogenic covariates – 1. The legend for each map indicates the colour (for each quintile) of the proportion of the population exhibiting the covariate of interest in each SOA. That is, the proportion with a low SEG; whose main source of income is from benefits; who do own a PC; who have internet access; of households that own more than one TV; and household income (low <£200pw, medium or high >£650pw).

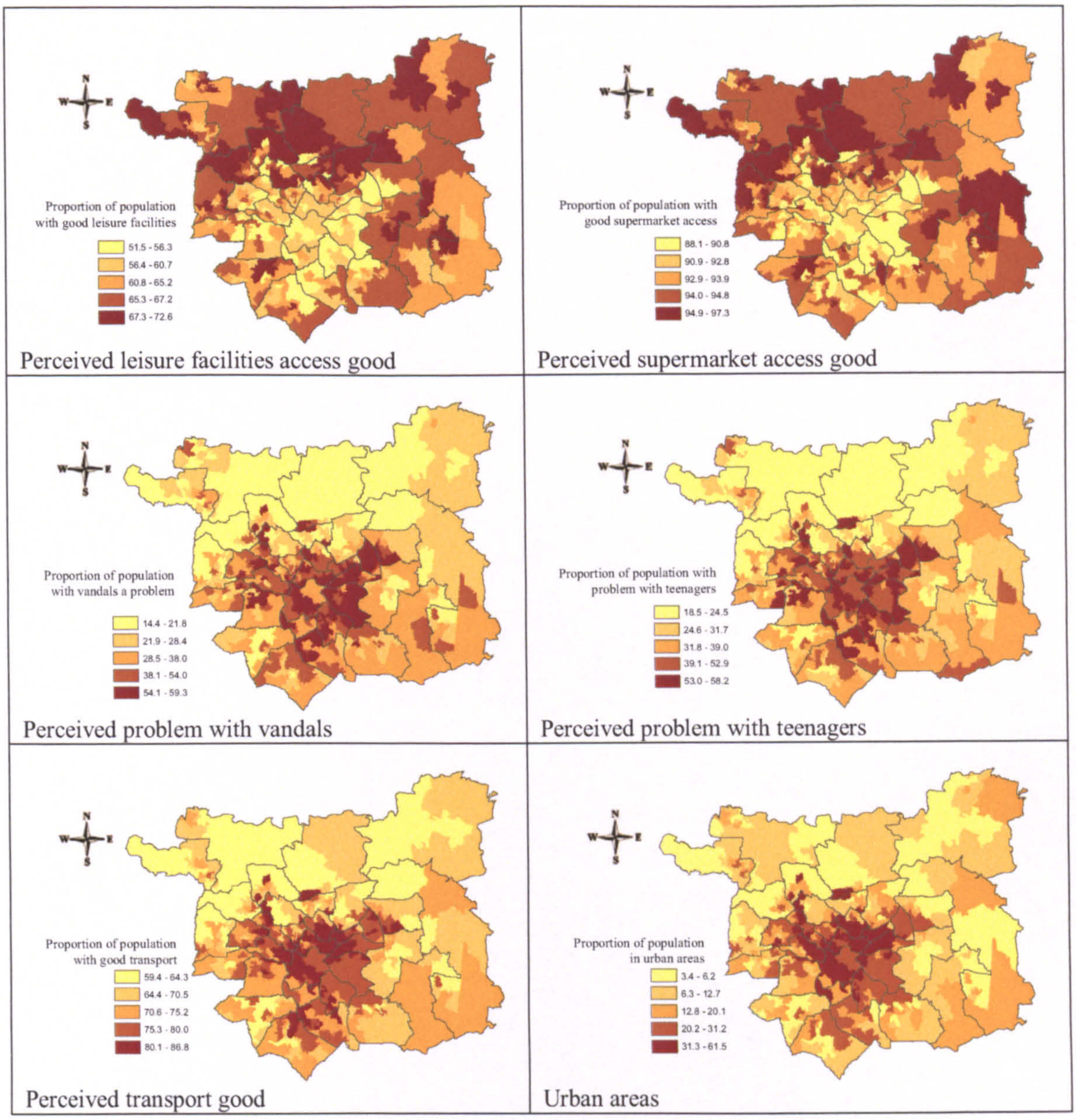
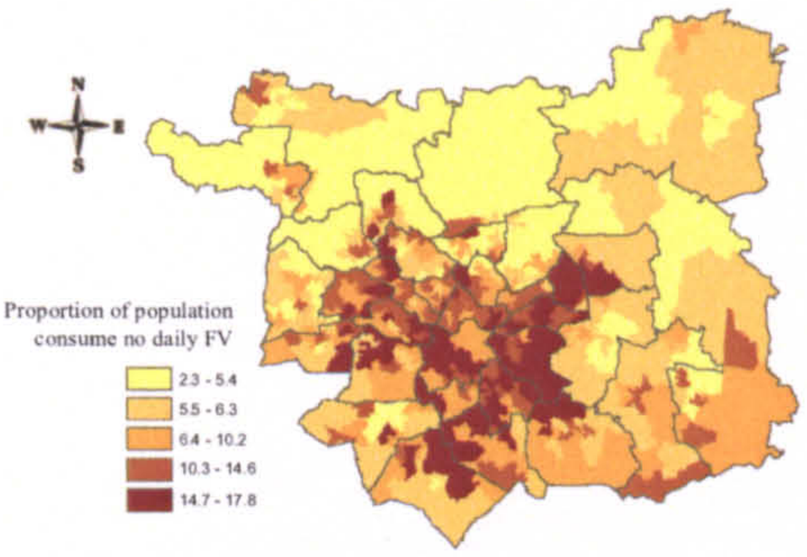
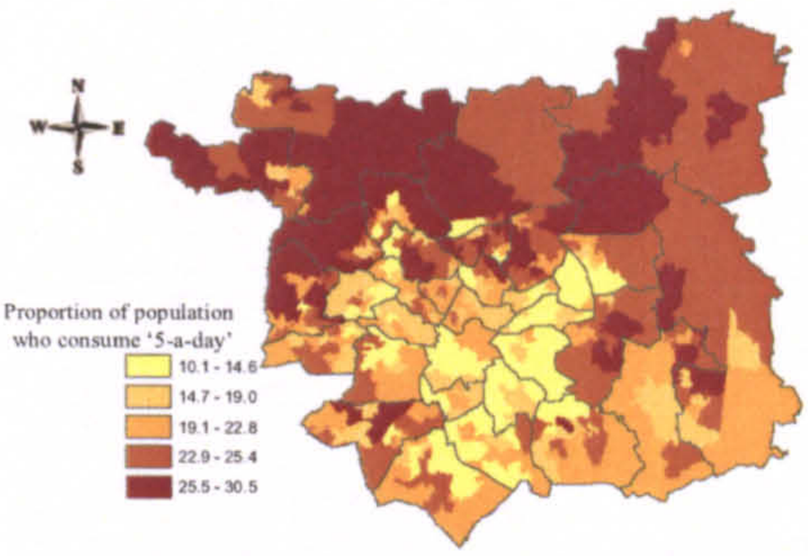


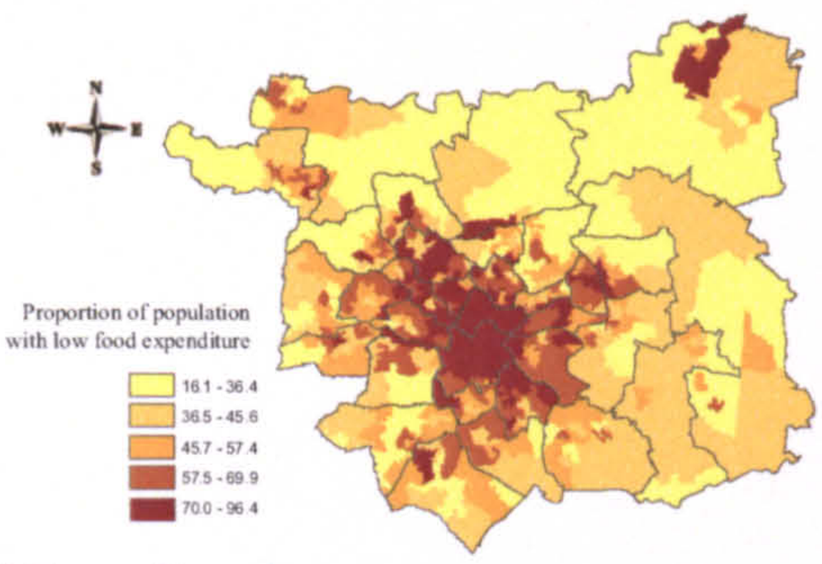
Figure 7.5. Maps of obesogenic covariates – 2. The legend for each map indicates the colour (for each quintile) of the proportion of the population exhibiting the covariate of interest in each SOA. That is, the proportion with perceived good access to leisure facilities; perceived good access to supermarkets; with a perceived problem with vandals or teenagers; perceived good local transport; the areas with a high degree of urbanisation.



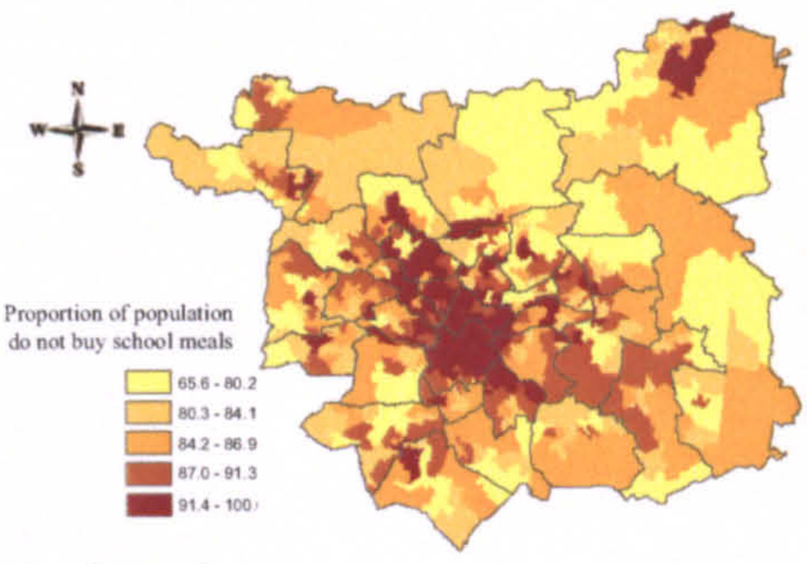
No daily fruit and vegetables



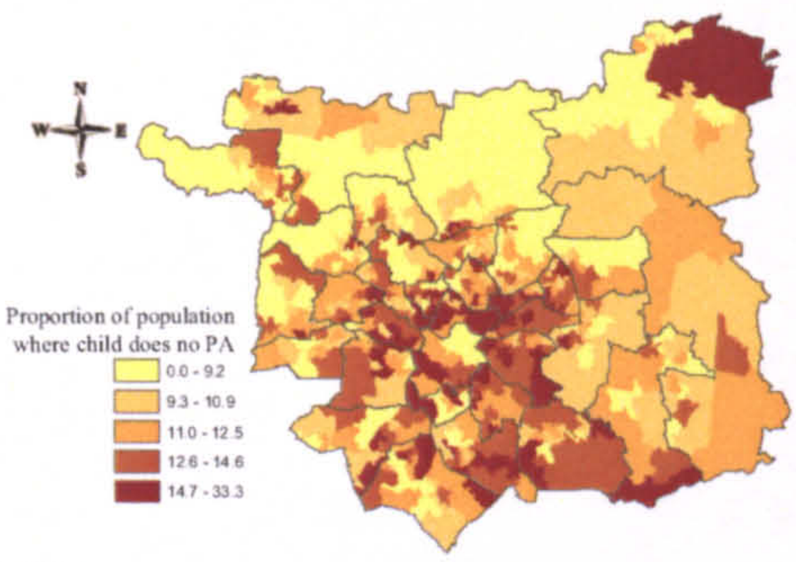
At least 5 portions fruit and vegetables per day



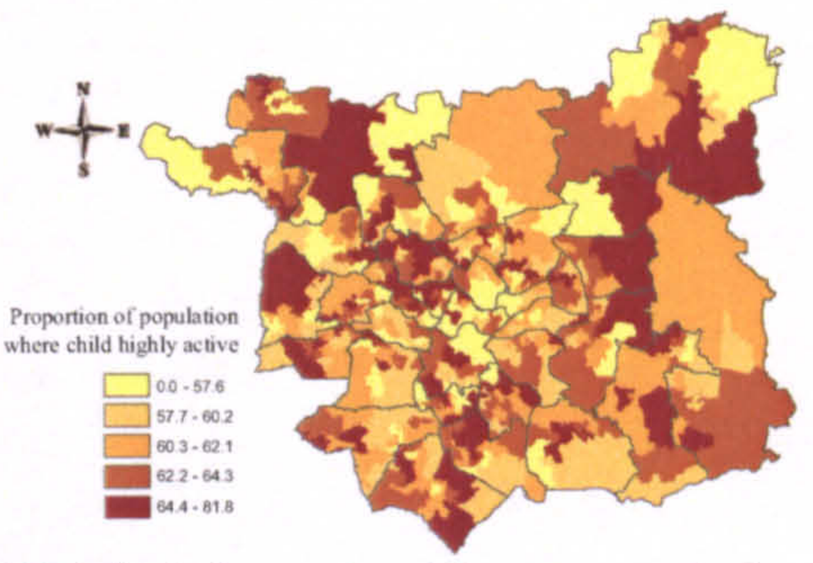
Food Expenditure low



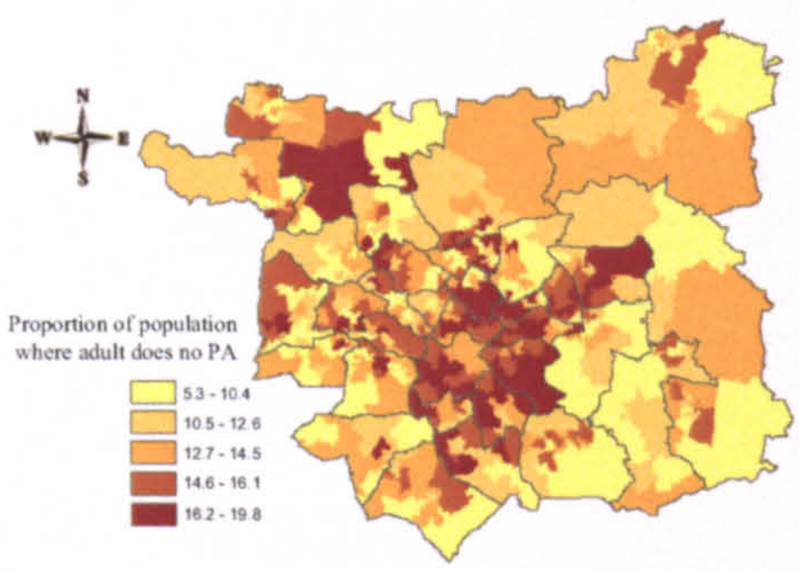
No school meals



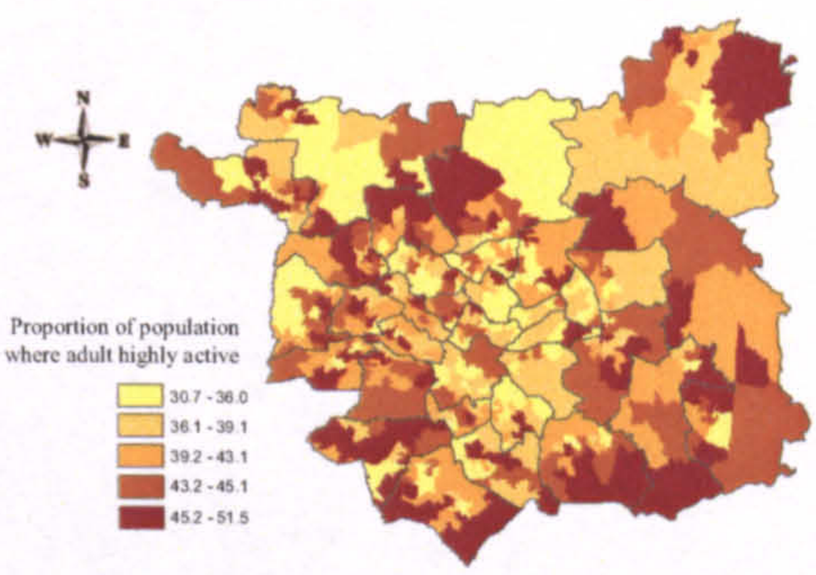
Child PA none



Child PA high (more than 7 hours pw active)



Adult PA none



Adult PA high (> 30 minutes ≥ 5 days per week)

Figure 7.6. Maps of obesogenic covariates – 3. The legend for each map indicates the colour (for each quintile) of the proportion of the population exhibiting the covariate of interest in each SOA. That is, the proportion that consume no daily fruit and vegetables; that consume at least 5 portions of fruit and vegetables per day; of households with low expenditure on food (less than £60 per two weeks); that do not buy school meals; where children undertake no physical activities; where children exhibit high levels of physical activity; where adults undertake no physical activities; where adults exhibit high levels of physical activity.

7.4.2 Global analyses

The results of the global regression analyses are summarised in Table 7.3, which categories the variables into four sub-groups: socio-economic indicators, dietary variables, physical activity variables, and social capital indicators. This shows whether each variables has a positive or negative relationship with childhood obesity. For example, obesity prevalence was lower in areas with a high income, and higher in areas where no of fruit and vegetables were consumed daily. Many of these variables are strongly correlated with income - see Table 7.4 (all $p < 0.001$). Car availability is also strongly positively correlated with income ($r > +/- 0.94$ for all categories, $p < 0.001$). Similarly there are strong positive correlations between fruit and vegetable consumption and expenditure on food (none eaten per day, low food expenditure per person, $r = 0.85$, $p < 0.001$; 5 or more portions per day, high food expenditure per person, $r = 0.84$, $p < 0.001$).

Similar results were obtained when other definitions of childhood obesity were used (results not shown); namely, the percentage of children who are overweight and obese (these correlations were very slightly weaker), and mean BMI SDS for the SOA (much weaker correlations for most variables) (results not shown). Using mean BMI SDS three variables produced different results to the percentage obese / overweight and obese results: degree of urbanisation (suburbia), child physical activity levels (none) and TV ownership. Suburban areas were positively associated ($r = 0.17$, $p < 0.001$) yet the obesity percentage data showed them equally negatively correlated. Thus the suburban data can be deemed inconclusive. A child undertaking no exercise had a negative correlation (i.e. protective!) with mean BMI SDS of the area, but this effect was small and insignificant ($r = -0.04$, $p = 0.362$), compared to a much stronger significant effect with percentage of obese children ($r = 0.22$, $p < 0.001$). Other child physical activity variables were in line with previous results. Similarly mean BMI SDS and ownership of more than one TV per household showed a very weak, insignificant negative correlation ($r = -0.02$, $p = 0.728$), whereas the obesity percentage data showed the opposite ($r = 0.32$, $p < 0.001$). These last two differences were not significant and thus may have occurred due to chance. The correlations with the percentage obesity data are stronger, significant and generally more robust.

In determining the covariates with the most important global relationship with childhood obesity, firstly, a high coefficient of determination was considered. The R^2 figures provide information regarding how well the fitted models replicate the observed datasets. The global R^2 figure says generally how well the model works, whereas the local R^2 figure gives information on how well the local model replicates the data around data point i weighted by distance from point i (i.e. the fit around different points). These figures represent the amount of variability in the obesity data explained by the predictor variable. Table 7.3 shows that many variables have a global R^2 above 10%, which is fairly low but adequate to enable the global model to be used

as a base to build a local model. Also important is having a large parameter estimate because for every unit change in the predictor variable, a change equal to the amount of this estimate induced in the obesity variable (e.g. for every unit change in the proportion of the child population who undertake no exercise, childhood obesity would increase by 0.16). Thus a high estimate is better, otherwise notwithstanding how significant the results, the impact on childhood obesity will only be small. Accordingly, the variables with the strongest global relationships with obesity are access to supermarket and leisure facilities, quality of public transport, fruit and vegetable consumption, sedentary behaviour, household income and urbanisation.

Independent variables	Pearson correlation coefficient	Global R^2	Sig. (2-tailed)	Global estimate	Local R^2
Socio-economic indicators:					
Income high (> £650 pw; > £33800 pa)	- 0.345	0.119	< 0.001	- 0.106	0.410
Income medium	- 0.303	0.092	< 0.001	- 0.099	0.093
Income low (< £200 pw; < £10400pa)	0.343	0.118	< 0.001	0.057	0.414
Live in rural area	-0.331	0.109	< 0.001	-0.079	0.425
Live in suburban area	-0.166	0.027	< 0.001	-0.105	0.401
Live in urban area	0.341	0.116	< 0.001	0.071	0.412
Do not / do own a home computer	+/- 0.337	0.114	< 0.001	+/- 0.057	0.411
Do not / do have internet access	+/- 0.342	0.117	< 0.001	+/- 0.054	0.414
Unemployed / employed	+/- 0.342	0.117	< 0.001	+/- 0.050	0.412
SEG low / high	+/- 0.290	0.084	< 0.001	+/- 0.045	0.438
Deprivation score	0.287	0.082	< 0.001	0.043	0.431
Social capital indicators:					
Supermarket access perceived as hard/easy	+/-0.275	0.075	< 0.001	+/-0.347	0.424
Leisure facilities access perceived bad/good	+/-0.272	0.074	< 0.001	+/-0.134	0.431
Public transport perceived as good/bad	+/-0.316	0.100	< 0.001	+/-0.120	0.419
Teenagers do / do not cause a problem	+/-0.306	0.094	< 0.001	+/-0.067	0.431
Vandals do / do not cause a problem	+/-0.301	0.091	< 0.001	+/-0.057	0.430
Dietary variables:					
5 or more portions of fruit and vegetables eaten per day	-0.247	0.061	< 0.001	-0.130	0.437
Low consumption fruit and vegetables (>0, <5 portions per day)	0.038	0.001	0.406	0.064	0.407
No fruit and vegetables eaten per day	0.284	0.080	< 0.001	0.180	0.436
Food expenditure low/ high	+/-0.319	0.102	< 0.001	+/-0.102	0.409
Food expenditure (adjusted for household size) low/high	+/-0.310	0.096	< 0.001	+/-0.102	0.417
Do / do not buy school meals	+/- 0.234	0.055	< 0.001	+/- 0.101	0.405
Physical activity variables:					
Child takes no physical exercise	0.217	0.047	< 0.001	0.156	0.398
Adult takes no physical exercise	0.209	0.044	0.006	0.181	0.416
Own more than one TV / only one (or no) TV per household	+/- 0.321	0.103	< 0.001	+/- 0.064	0.410
Child low activity (>0, <3 hours pw)	-0.052	0.003	0.261	-0.059	0.410
Child moderate active (3-7 hours pw)	-0.175	0.030	< 0.001	-0.131	0.404
Child highly active (takes 7 or more hours of exercise per week)	-0.068	0.005	0.141	-0.036	0.397
Adult low active (>0, <3 days pw)	0.125	0.016	< 0.001	0.096	0.386
Adult moderate active (3-5 days pw)	-0.100	0.010	0.029	-0.152	0.410
Adult highly active (exercise for at least 30 minutes on 5 or more days pw)	-0.197	0.039	< 0.001	-0.115	0.402

Table 7.3. Summary of the key coefficients of the global analyses, as well as the local coefficient of determination. Correlations with percentage of obese children in each SOA (obese data smoothed using Empirical Bayes) for Leeds (global regression model; N = 476). The R^2 figures provide information regarding how well the fitted models replicate the observed datasets. The global R^2 figure says generally how well the model works, whereas the local R^2 figure gives information on how well the local model

replicates the data around data point i weighted by distance from point i (i.e. the fit around different points). These figures represent the amount of variability in the obesity data explained by the predictor variable. The parameter global estimate figure represents the amount of change induced in the obesity variable for a unit change in the predictor variable; so a higher figure suggests the covariate has more impact on childhood obesity than a lower value. * Urban area is defined as inner city or other dense urban/town centre.

Independent variables - Pearson correlation coefficient	Deprivation score	Low income	Medium income	High income
Socio-economic indicators:				
Live in rural area	-0.92	-0.82	0.73	0.81
Live in suburban area	-0.32	-0.30	0.28	0.28
Live in urban area	0.90	0.81	-0.73	-0.80
Do own a home computer *	-0.70	-0.98	0.92	0.93
Do have internet access *	-0.76	-0.99	0.91	0.97
On benefits *	0.82	1.00	-0.93	-0.95
SEG low *	0.89	0.80	-0.69	-0.82
Deprivation score	1.00	0.79	-0.71	-0.78
Social capital indicators:				
Supermarket access perceived as easy *	-0.84	-0.82	0.73	0.82
Leisure facilities access perceived good *	-0.83	-0.79	0.69	0.79
Public transport perceived as good *	0.89	0.78	-0.67	-0.81
Teenagers do cause a problem *	0.92	0.83	-0.74	-0.83
Vandals do cause a problem *	0.92	0.82	-0.73	-0.82
Dietary variables:				
5 or more portions of fruit and vegetables eaten per day	-0.86	-0.75	0.65	0.76
Low consumption fruit and vegetables (>0, <5 portions per day)	0.43	0.23	-0.13	-0.30
No fruit and vegetables eaten per day	0.87	0.82	-0.74	-0.81
Food expenditure (adjusted for household size) low *	0.64	0.96	-0.90	-0.92
Do buy school meals *	-0.38	-0.79	0.75	0.75
Physical activity variables:				
Child takes no physical exercise	0.45	0.50	-0.43	-0.52
Adult takes no physical exercise	0.61	0.48	-0.47	-0.45
Own more than one TV per household *	0.59	0.94	-0.92	-0.86
Child low activity (>0, <3 hours pw)	-0.08	-0.17	0.12	0.19
Child moderate active (3-7 hours pw)	-0.26	-0.51	0.46	0.51
Child highly active (does ≥ 7 hours of exercise per week)	-0.12	-0.01	0.03	-0.01
Adult low active (>0, <3 days pw)	0.30	0.22	-0.23	-0.19
Adult moderate active (3-5 days pw)	-0.38	-0.26	0.30	0.21
Adult highly active (exercise ≥ 30 minutes on 5 or more days pw)	-0.49	-0.39	0.37	0.37

Table 7.4. Summary of results of Pearson correlation coefficient between the obesogenic variables and deprivation / income. All results were highly significant ($p<0.001$). Binary variables are marked with * - here, the opposite variable has the same but opposite Pearson correlation value e.g. the value of the Pearson correlation coefficient between households without a computer and deprivation is +0.70.

7.4.3 Geographically weighted regression analyses

Next the local GWR analyses were performed for those variables with strong global relationships with obesity. The key figures for each GWR analysis are summarised in Table 7.5. The AICc figure is a trade-off between goodness-of-fit and the number of parameters; accordingly it can be used to compare the global and local models of the same analysis (a lower figure represents a better fitting model), but it cannot be used to compare different datasets. The ANOVA analysis tests the null hypothesis that the local model does not represent an improvement over a global model. A significant F statistic suggests that the local model is better. Table 7.5 shows that for all independent variables the local model is a better fit than the

global model (using the results from both the AICc figure and the ANOVA analysis). Accordingly it is worthwhile perusing local level analyses.

Independent variables	Global AIC	Local AIC	ANOVA F	Local < global
Supermarket access perceived as easy	2288.3	2223.8	3.63	yes
Supermarket access perceived as hard	2288.3	2223.8	3.63	yes
Leisure facilities access perceived good	2289.0	2218.7	3.74	yes
Leisure facilities access perceived bad	2289.0	2218.7	3.74	yes
Public transport perceived as good	2275.6	2226.0	3.33	yes
Public transport perceived as bad	2275.6	2226.0	3.33	yes
5 or more portions of fruit and vegetables eaten pd	2295.7	2213.8	3.98	yes
Low consumption fruit and vegetables (>0, <5 portions pd)	2325.0	2244.6	3.93	yes
No fruit and vegetables eaten per day	2285.8	2214.0	3.77	yes
Child takes no physical exercise	2302.7	2250.4	3.38	Yes
Adult takes no physical exercise	2304.4	2241.6	3.57	Yes

Table 7.5. Summary of results of global and local GWR analyses using percentage of children who are obese as the dependent variable and the shortlisted covariates as predictor variables. This table shows whether the local model is an improvement on the global model: if local AICc < global AICc then suggests local model is a better fit of the data; similarly if the F statistic from the ANOVA analysis is significant this suggests the local model is an improvement on the global model.

For each relationship, a comparison of the range of the local parameter estimates with a confidence interval around the global estimate of the equivalent parameter determined if any relationships were non-stationary (Fotheringham et al, 2002, 2005). All of the simulated covariates showed a non-stationary relationship with obesity, which means that the same stimulus provokes a different response in some parts of Leeds. The results of these calculations for five key global determinants are given in Table 7.6. Perceived access to supermarkets ($p = 0.010$) and leisure facilities ($p = 0.010$), high or non-existence consumption of fruit and vegetables ($p < 0.001$ for both) and perceived quality of public transport services ($p < 0.001$) had statistically significant (using a Monte Carlo test) non-stationary processes. The insignificant results for sedentary behaviour in either adults or children are perhaps not surprising given the dispersed pattern of the physical activity maps seen in Figure 7.6 (above).

To list the parameter estimates for each location across Leeds would not be meaningful; accordingly these results have been mapped in order to visualise the spatial variation (see Figure 7.7). These maps show the areas where the local relationship between the covariates and childhood obesity is positive or negative, bearing in mind that the overall global relationship will be either positive or negative. The dark blue areas (cold spots) are areas where the predictor reduces the risk of childhood obesity, and the red areas (hot spots) are areas where the relationship is obesogenic: e.g. that a higher proportion of SOAs' population perceive supermarkets as hard to access is associated with less childhood obesity in blue areas (negative local relationship) and associated with more childhood obesity in red areas (positive). The

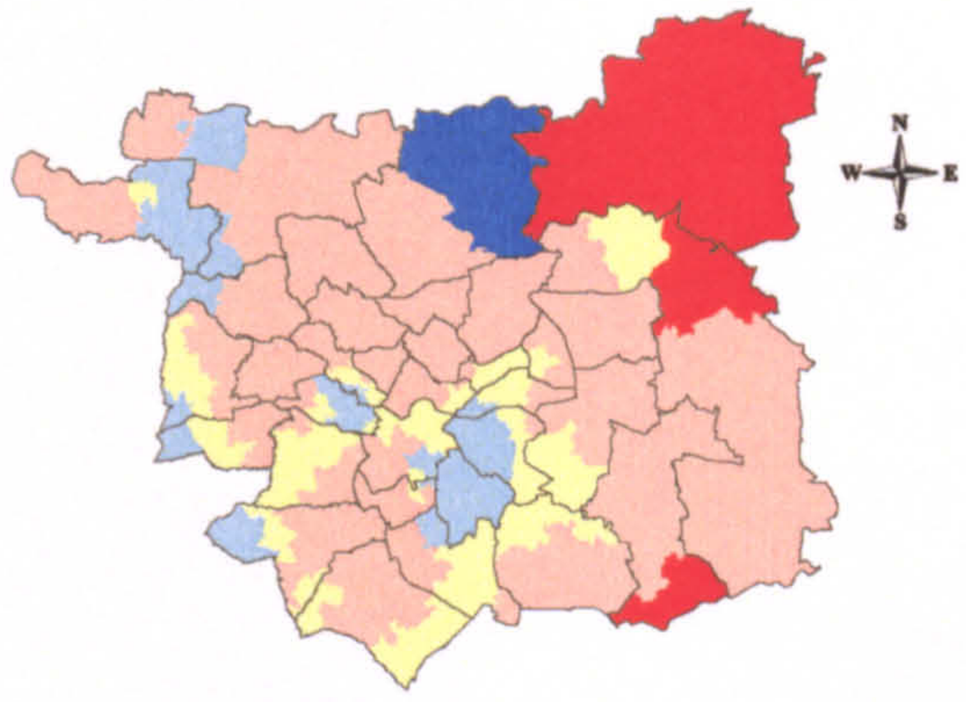
darkest blue and red areas are the areas with the strongest relationships, and these are the areas of specific interest; in particular the areas of red in a primarily blue map (and vice versa).

In practical terms, it is likely that any policies directed towards childhood obesity are likely to only go down to the ward level as the smallest unit. Accordingly in order to summarise this analysis by ward, rather than simply averaging the parameter estimates of all the SOAs in each ward which would serve to lose some of the detail of differences across a ward, the percentage of SOAs in each ward that show either a positive or negative relationship were calculated for each of the covariates (see Table 7.7). This information effectively provides a ward-level summary of Figure 7.7. For example, for the covariate that "supermarkets are difficult to access", in Moortown it can be see that all SOAs in this ward show a positive relationship between this variable and childhood obesity, which suggests this may be a variable to include in any future childhood obesity prevention policy. Whereas in Middleton the relationship is less clearly defined, which suggests this may not be a key variable to focus an intervention policy around. In order to facilitate identification of each of the 33 wards in Leeds, Figure 7.8 shows their locations.

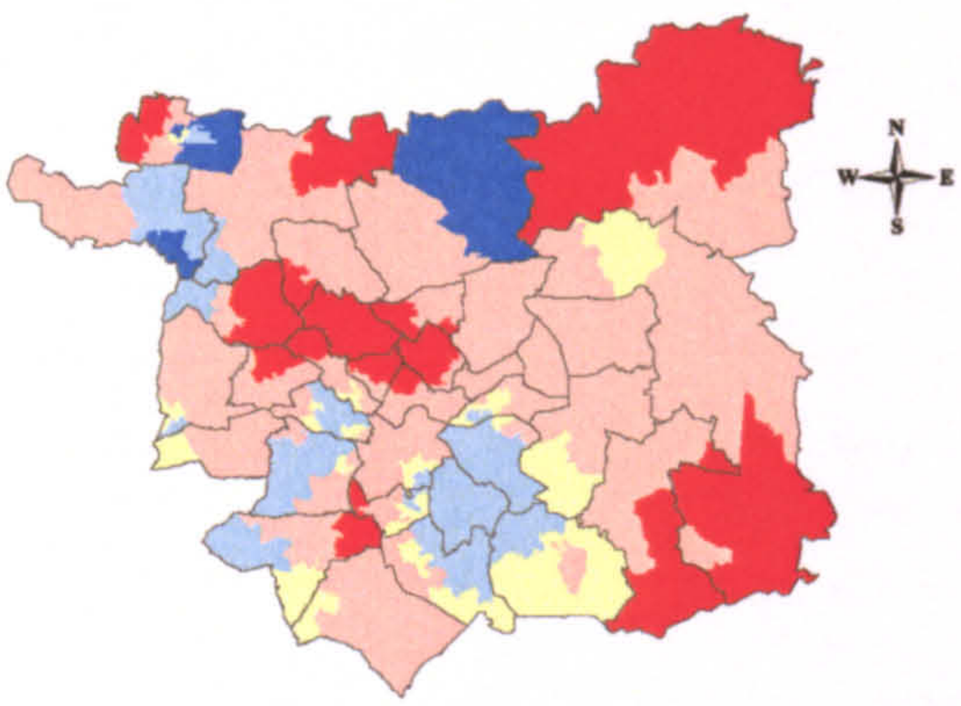
Independent Variables	Global SE	2.SE	LQ	UQ	IQR	Non stationary	P value	Local R^2
Supermarket access perceived as easy	0.06	0.11	-0.47	-0.07	0.39	Yes	0.010	-0.347
Supermarket access perceived as hard	0.06	0.11	0.07	0.46	0.39	Yes	0.010	0.347
Leisure facilities access perceived good	0.02	0.04	-0.18	-0.03	0.15	Yes	0.010	-0.134
Leisure facilities access perceived bad	0.02	0.04	0.03	0.18	0.15	Yes	0.010	0.134
Public transport perceived as bad	0.02	0.03	-0.14	-0.05	0.09	Yes	<0.001	-0.120
Public transport perceived as good	0.02	0.03	0.05	0.14	0.09	Yes	<0.001	0.120
5 or more portions of fruit and vegetables eaten per day	0.02	0.05	-0.18	-0.03	0.16	Yes	<0.001	-0.130
Low consumption fruit and vegetables	0.08	0.15	-0.19	0.32	0.51	Yes	0.180	0.064
No fruit and vegetables eaten per day	0.03	0.06	0.06	0.24	0.18	Yes	<0.001	0.180
Child takes no physical exercise	0.03	0.06	0.01	0.17	0.16	Yes	0.270	0.156
Adult takes no physical exercise	0.04	0.08	0.02	0.19	0.17	Yes	0.620	0.181

Table 7.6. Summary of local parameter estimate results for the relationship between childhood obesity and each of five of the key global determinants of obesity. This table shows the calculations of the extent of any spatial variability in order to determine whether any local relationship processes are non-stationary. That is, if the inter-quartile range is greater than twice the global standard error, then this suggests the local relationship might be non-stationary (Fotheringham et al, 2005). P values from the Monte-Carlo tests are given. Abbreviations used: SE: standard error; LQ: lower quartile; UQ: upper quartile; IQR: inter-quartile range.

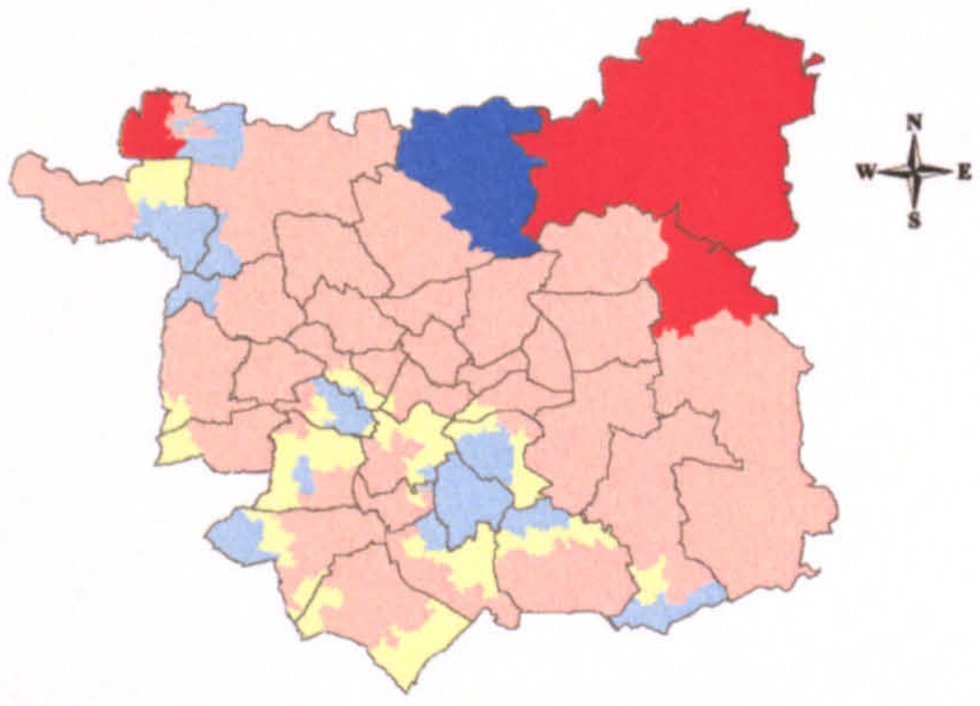
Public transport good. Min -.066 max +1.12



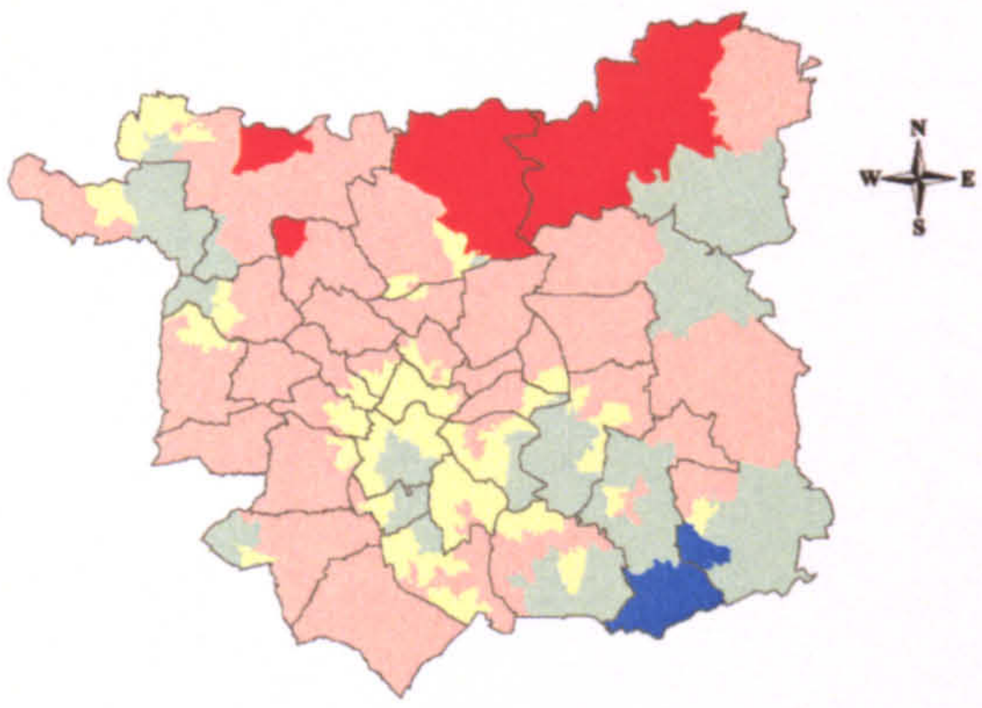
Leisure facilities poor. Min –0.79; max +1.00



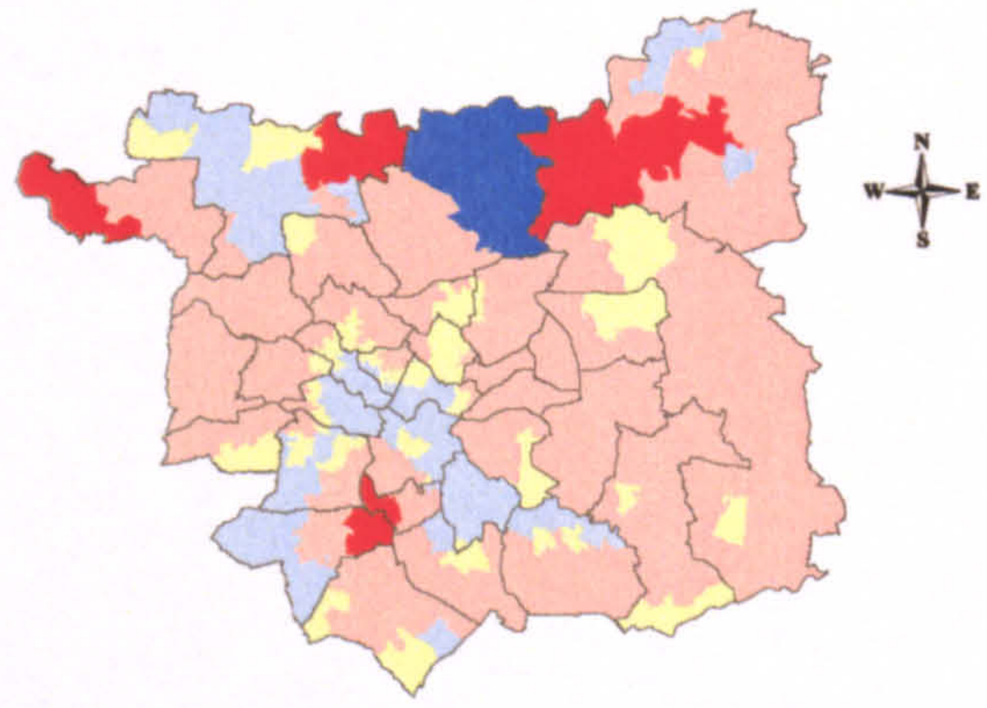
Supermarket access poor. Min –1.83; max +2.30



No daily fruit & vegetables. Min –1.36; max 2.31



Child no physical activity. Min –0.86; max 0.88



Adult no physical activity. Min –1.11; max 1.05

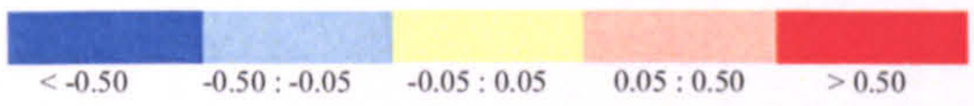


Figure 7.7. Map of the parameter coefficient for six of the key global predictor variables by SOA. The overall map colour (red or blue) illustrates the overall global relationship (increases or reduces risk of obesity respectively), although all were mapped from a global obesogenic perspective e.g. no daily fruit and vegetable intake, rather than consuming "5-a-day". The darkest blue and red areas are the areas with the strongest local relationship with obesity, and these are the areas of specific interest; in particular the areas of red in a primarily blue map (and vice versa). In order to enable comparison between maps, all were prepared using the same five-category manual scale.

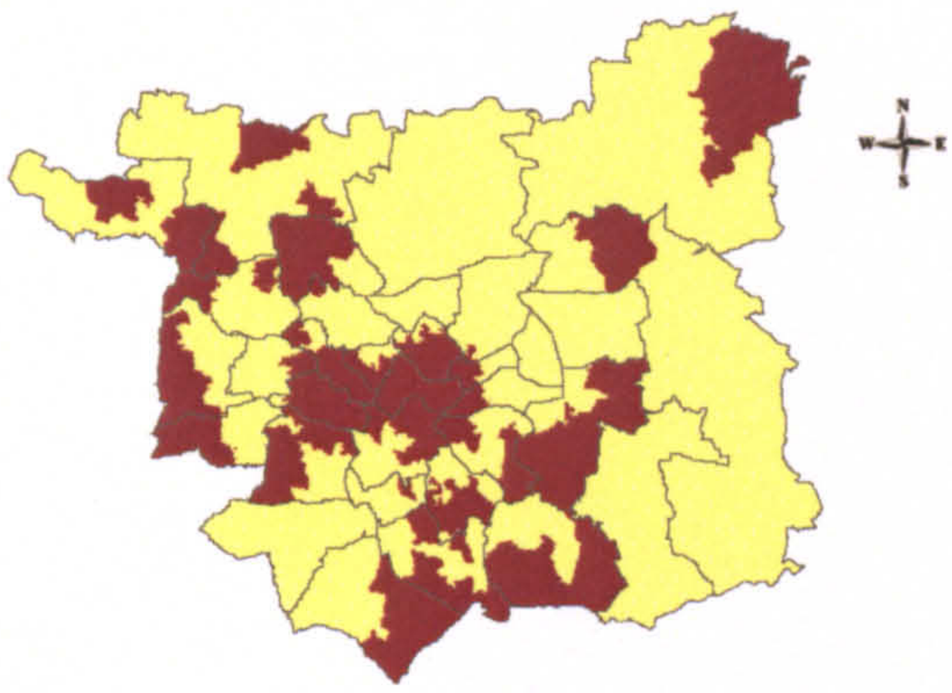
Blank Page

Ward	Shops hard		Leisure bad		Transport good		No fruit or vegetables		Child PA - none		Adult PA - none	
	negative	positive	negative	positive	negative	positive	negative	positive	negative	positive	negative	positive
Aireborough	88	12	82	12	76	12	76	12	88	6	0	100
Armley	33	40	40	33	0	33	40	33	0	67	60	20
Barwick and Kippax	0	94	0	94	6	88	0	100	31	56	0	88
Beeston	18	64	9	82	0	82	9	82	64	9	9	91
Bramley	0	100	0	100	0	100	0	100	0	100	0	93
Burmantofts	8	50	25	25	0	100	0	75	8	75	0	100
Chapel Allerton	0	100	0	100	0	83	0	100	0	75	0	25
City and Holbeck	21	57	21	43	14	36	14	43	71	0	36	50
Cookridge	0	100	0	100	0	100	0	100	0	100	0	93
Garforth and Swillington	0	100	0	100	0	100	6	88	31	63	0	88
Halton	0	93	0	93	0	40	0	100	53	27	0	100
Harehills	0	100	0	29	0	100	0	100	0	100	0	100
Headingly	0	100	0	100	0	94	0	94	0	41	29	24
Horsforth	7	93	7	93	7	93	7	93	7	86	0	100
Hunslet	100	0	100	0	36	0	100	0	0	9	91	0
Kirkstall	0	85	8	54	0	69	8	54	0	62	62	23
Middleton	36	36	7	57	0	50	7	57	7	43	7	79
Moortown	0	100	0	100	0	100	0	100	0	93	0	50
Morley North	24	65	18	65	6	71	18	65	12	82	47	53
Morley South	0	95	0	79	0	74	0	79	0	100	5	84
North	7	93	7	93	7	93	7	93	7	67	7	93
Otley and Wharfedale	31	63	31	69	25	63	25	75	25	31	44	31
Pudsey North	7	87	13	60	0	93	0	87	0	93	0	100
Pudsey South	0	93	7	60	0	100	0	93	0	100	0	80
Richmond Hill	67	25	58	17	0	75	25	33	17	42	0	92
Rothwell	36	14	0	50	0	29	7	36	7	79	21	50
Roundhay	0	100	0	100	0	100	0	100	0	100	0	87
Seacroft	0	100	0	75	0	100	0	100	0	67	0	100
University	0	100	0	86	0	43	0	93	0	14	50	14
Weetwood	0	100	0	100	0	100	0	100	0	100	0	86
Wetherby	0	100	0	100	6	94	0	100	11	89	39	56
Whinmoor	0	100	0	100	0	73	0	100	0	100	0	91
Wortley	67	27	0	33	0	33	7	20	0	80	33	33

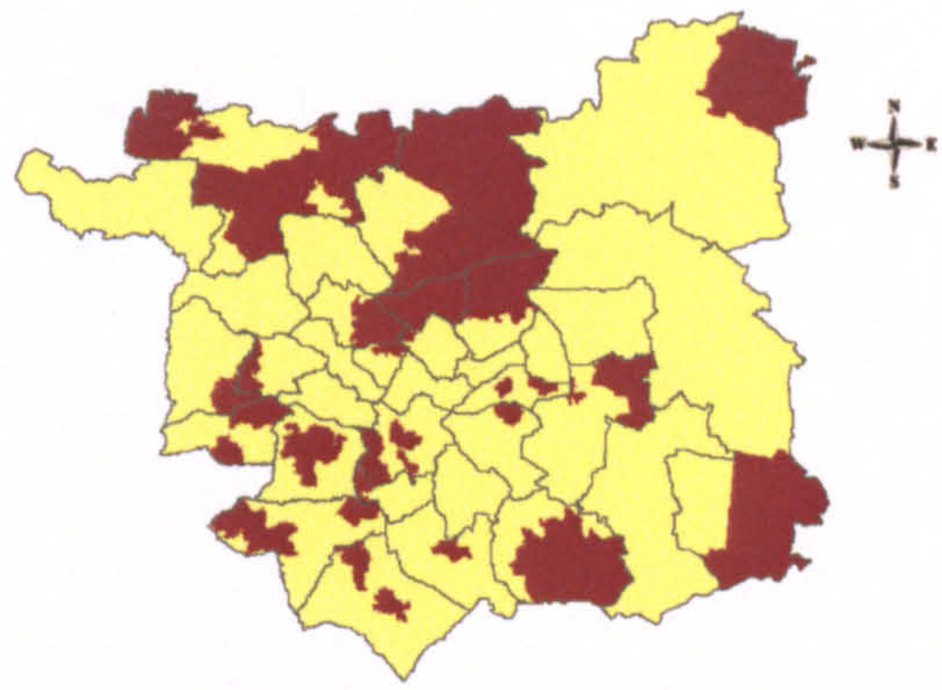
Table 7.7. Summary of variability in parameter estimates by ward. The variables are the ones that show the strongest global relationship with childhood obesity. The numbers relate to the percentage of SOAs in each ward that on a local level show either a positive or negative relationship with obesity (percentage neutral not shown). The higher percentage of the two will obviously concur with direction of the overall global relationship.

Figure 7.8. Map of Leeds showing the location of the wards. The grey shading corresponds to the grey shading in Tables 7.7 and 7.8 in order to facilitate locating a particular ward.

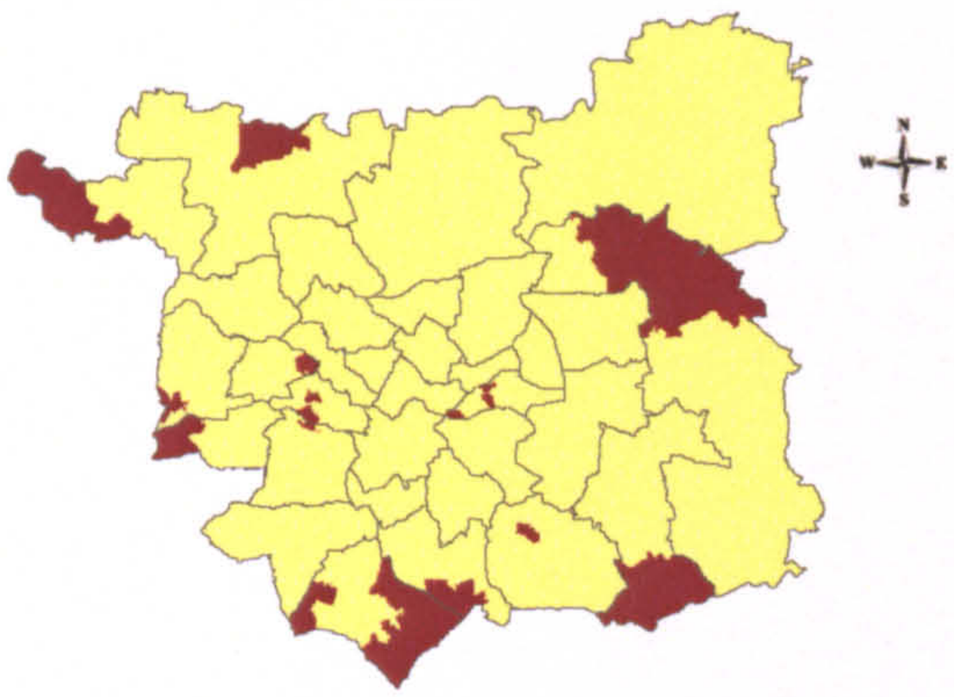
The local analysis also determined those areas where the most important local determinants of childhood obesity are different to the key global determinants by ranking the parameter estimates for all variables in each SOA – namely, expenditure on food, number of household televisions, problems with teenagers, internet access, school meals and children's levels of physical activity. The areas where these factors are influential drivers of childhood obesity are highlighted in Figure 7.9. Again these data have been summarised by ward with the percentage of SOAs in each ward where each covariates is an important determinant of childhood obesity in that locality detailed in Table 7.8. For example, in Armley (central Leeds) school meals is a local determinant of childhood obesity in most of that ward's SOAs (93%), as is moderate/high child physical activity levels (80%) and to a lesser degree households owning more than one television (13%).



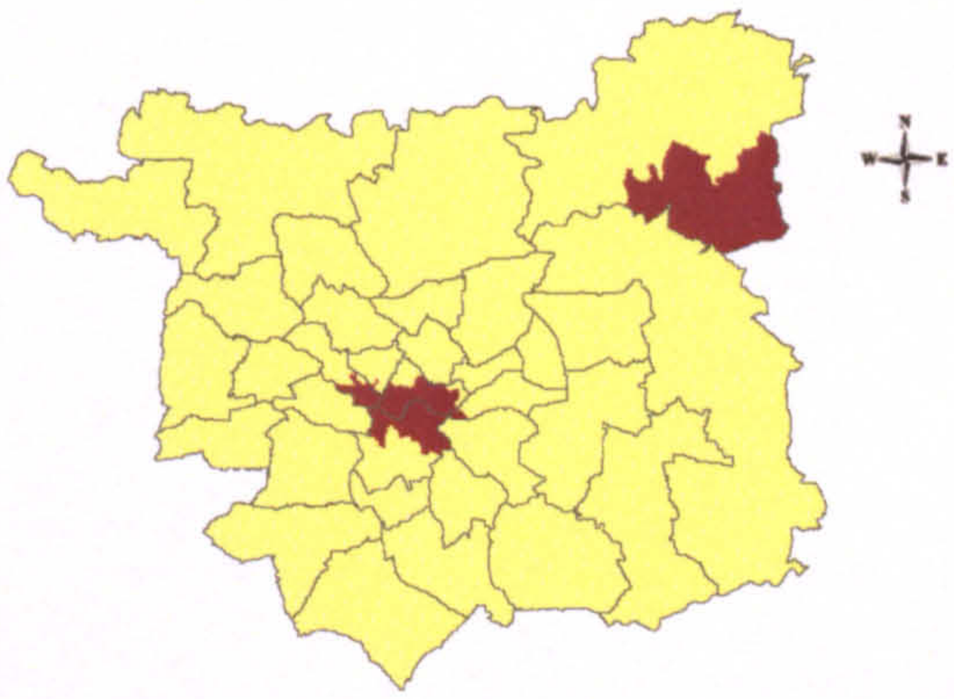
School meals are an important local determinant of childhood obesity (eating school meals lowers risk)



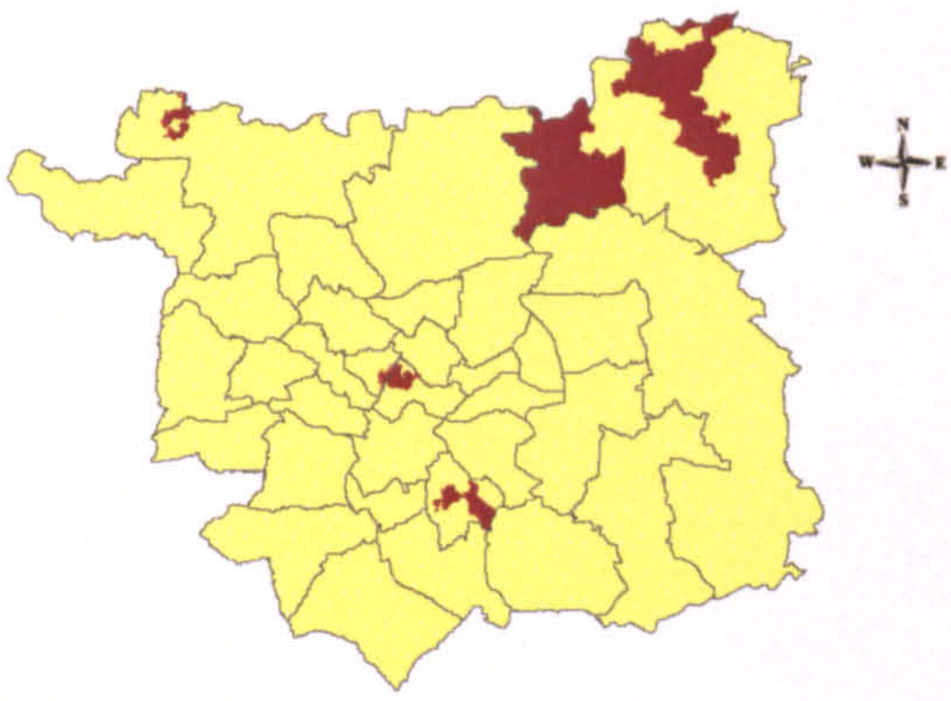
Similarly for food expenditure (high/low food expenditure lowers/increases risk respectively)



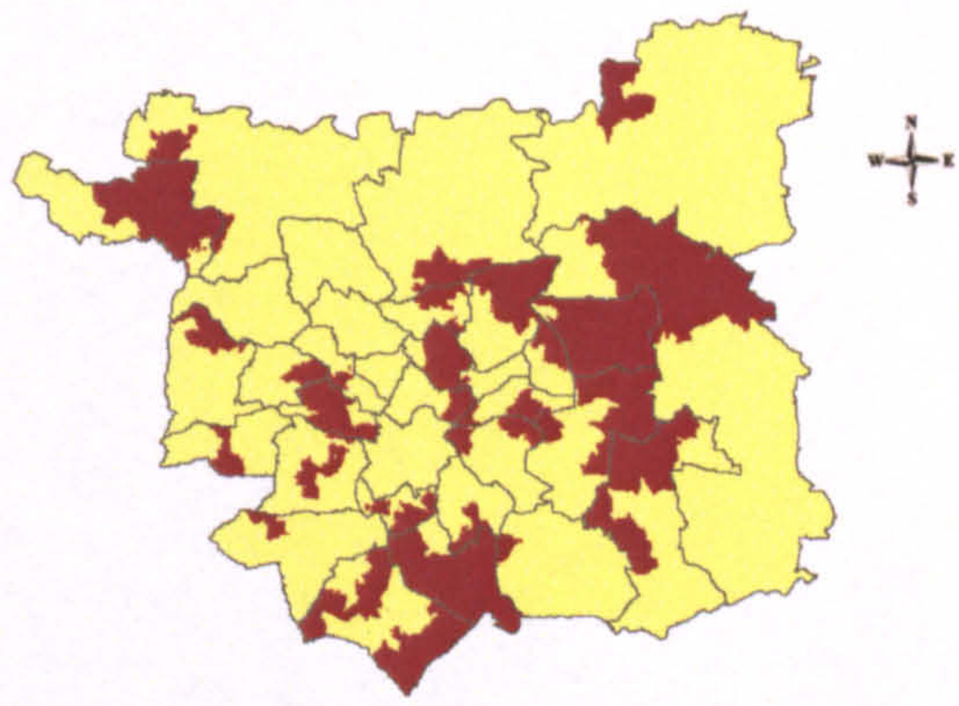
Households with more than one television are at increased risk of childhood obesity



Households that have internet access are at reduced risk of childhood obesity



Neighbourhoods with problems with teenagers hanging around are at increased risk (and vice versa)



Children who undertaken 3 hours or more of physical activity per week are at reduced risk

Figure 7.9. Maps showing where there are important local determinants of childhood obesity. For each map of each covariate shown (e.g. food expenditure) the shaded areas are those locations where that determinant is key. These key local determinants differ to the key global determinants (which are access to supermarkets and leisure facilities, public transport services, fruit and vegetable consumption, sedentary behaviour, household income and urbanisation). It also shows where multiple factors are important locally e.g. in the extreme north east corner of Leeds both purchasing school meals and household food expenditure are important local determinants of childhood obesity.

	School Meal	Food Expenditure	More than one TV	Internet access	Problem teenagers	Child PA mod/high
Aireborough	65		6			82
Armley	93		13			80
Barwick and Kippax	6	6	13			13
Beeston	9	9				36
Bramley	20	20	7			27
Burmantofts	42	17	17	8		33
Chapel Allerton	75	8				83
City and Holbeck	21	43		14		
Cookridge	79					
Garforth and Swillington			19			25
Halton	40	13				47
Harehills	64					7
Headingly	76	12		6	12	
Horsforth	14					
Hunslet	64				18	27
Kirkstall	85			31		8
Middleton	50	7	7			93
Moortown		100				36
Morley North		29				6
Morley South	26	37	26			37
North		87				47
Otley and Wharfedale	31	63	6		13	19
Pudsey North	47	27	7			7
Pudsey South	33	47	13			13
Richmond Hill	42	17				50
Rothwell	29	57	7			7
Roundhay	7	47				20
Seacroft						33
University	100			57	21	21
Weetwood	7	57				
Wetherby	22	6		6	50	6
Whinmoor						91
Wortley	67	40	7			27

Table 7.8. Summary by ward showing where there is an important local relationship with childhood obesity that is different to the key global relationships. The numbers represent the percentage of SOAs in that ward where this variable is an important determinant. For example, 26% of the SOAs in Morley South show a strong relationship between the number of household televisions and childhood obesity; similarly in Armley 93% of the SOAs show a strong relationship between buying school meals and childhood obesity.

In order to visualise whether the patterns of covariates' relationships with obesity are affected by income or deprivation, maps of low income, no car availability, low SEG and deprivation were prepared, which all show a clear pattern with the highest values in central Leeds (see Figure 7.10). Conversely, the strongest relationships between the covariates and obesity (as shown in Figure 7.7, above) are largely in the north east or south east of Leeds, and do not concur with these four patterns.

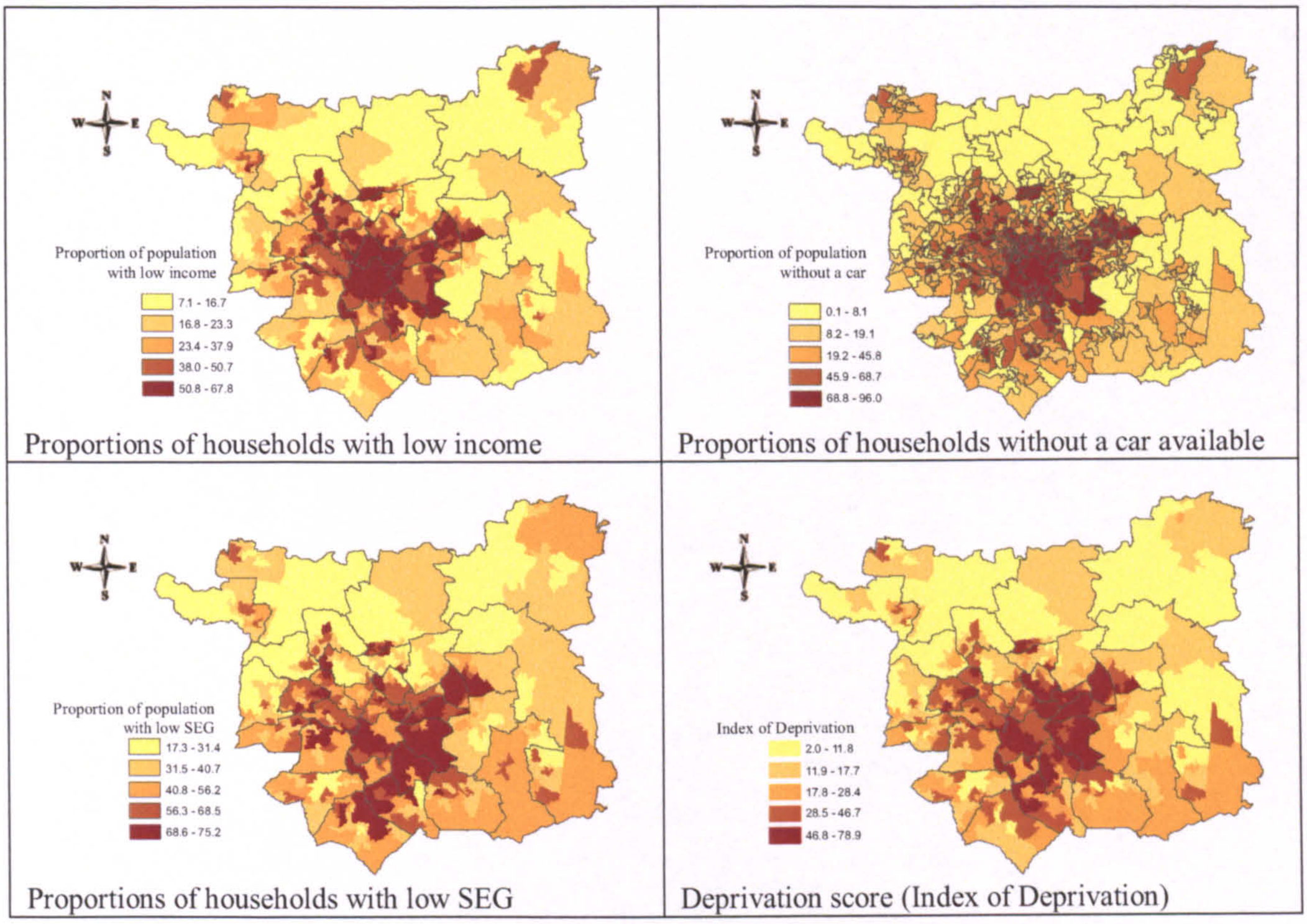


Figure 7.10. Maps of proportions of households with low income, without a car available, low SEG, as well as the deprivation score. All four are by SOA across Leeds, using simulated data, and the scales use quintiles.

The standarised residuals for each predictor variable were also mapped in order to see if any areas had particularly high or low values; a flat picture is desired (maps not shown). Negative residuals mean that the observed level of obesity is less that that predicted by the model, and positive ones mean that the observed values are greater. Very few SOAs (less than 2%) had a result less than -2.58 or greater than +2.58 areas, giving a uniform picture, suggesting the model is well fitting. The standard errors of the parameters at each regression point was also mapped to ensure rates were not too high (maps not shown). Higher rates were evident in the north east corner of Leeds, which might be expected because the model used a fixed kernel size in the analysis, so based on a set distance from the sample point rather than say the number of nearest neighbours, so SOAs with a larger area and/or fewer measurements will have higher standard errors. Overall standard errors were low.

7.5 Discussion

This chapter has taken two different approaches to consider the relationship between the obesogenic covariates and childhood obesity: using a global (for the whole of Leeds) and a local (for each SOA across Leeds) focus. The latter approach is novel in respect to obesity studies. The first approach enables the overall relationship between each covariate and childhood obesity in Leeds to be understood. This is the general stance that health planners work from; averaging data for the whole of a city in order to determine appropriate action to take. However, by also considering these relationships at the micro-level using local analyses, it highlights those micro-areas where the relationship may be at odds to the "norm". In these areas, it may be that different interventions or health policies are more appropriate to have the most effect on health (in this case, reducing prevalence of childhood obesity). Thus it is important that these geographical differences are understood, in order that they can be acted upon. This is a more informative position for a health planner to work from.

Whilst the obesity data were considered in more detail in Chapter 4, a brief summary of the global dataset is provided. The percentage of obese or overweight children in Leeds overall was 7.6% and 11.3% respectively. This compares to national UK prevalence of obesity and overweight amongst children of 16.7% (in 2002) (Sproston & Primatesta, 2002). The lower Leeds' obesity rates may be explained by differences in the definitions of obesity used; with the national dataset using the IOTF (International Obesity Task Force) definitions and this study using the British reference dataset definitions (with conservative, clinical, cut-offs of 91st / 98th centiles). Both datasets considered children of similar ages. The map of these data at ward and SOA level (see Figure 1) shows firstly how the prevalence of obesity varies across Leeds and secondly emphasises the increased detail obtained by analysing at the smaller scale – in the ward level map important areas of high and low prevalence (e.g. in the north east of Leeds) are not apparent.

7.5.1 Global analyses

Household income was shown to be a significant predictor of childhood obesity (inverse relationship) in line with other authors (Strauss & Knight, 1999; Stamatakis et al, 2005). This study also showed that SEG was negatively correlated with childhood obesity, but this relationship was fairly weak, perhaps indicating that SEG is not the best marker to typify the relationship between socio-economic status and childhood obesity. Other possibilities would include employment, tenure, education level and family size (not an exhaustive list). Other authors corroborate this result, although using other markers; for example, using education level as a proxy for socio-economic status, children from families with lower education levels had a higher risk of childhood obesity (Danielzik et al, 2004; Lamerz et al, 2005; Romon et al, 2005),

although this effect could be mediated by confounding factors, such as low income and lower levels of cognitive stimulation (Strauss and Knight, 1999). It may be that SEG is acting as a proxy for the effect of multiple adverse childhood circumstances, which are then manifesting as adult obesity in the long term (Power & Parsons, 2000). For example, it has been shown that there is a higher density of fast food outlets in poorer areas, which may (partially) explain the phenomenon (Reidpath et al, 2002). Computer ownership and access to the internet were shown to be negatively correlated with obesity; so in an area where a higher percentage of the population owned a computer (or internet access) then obesity prevalence was lower. It seems unlikely that simply owning a computer makes children thinner and this correlation does negate the increased sedentary behaviour theory somewhat. More likely the mechanism is to do with ownership acting as a marker for higher household income and/or lower deprivation.

Deprivation and living in urban areas were shown to be positively associated with obesity (and living in rural areas negatively associated). In relation to deprivation, studies using the Townsend Deprivation Score (an index score based on a combination of adult unemployment, household size, and car and home ownership) have shown that children from more deprived areas have higher risk of obesity (despite lower birth weights) (Kinra et al, 2000; Kinra et al, 2005). Also it has been shown that overcrowding, poverty, migration, pollution, housing, employment can all create environmental changes that may initiate the breakdown of community factors that adversely affect health, suggesting urbanization may have a role in the aetiology of obesity (Curtis et al, 2002). Furthermore urban sprawl, which reduces accessibility on foot, has been shown to be associated with obesity in America (Vandegrift & Yoked, 2004). Nevertheless despite the strong connection between urbanisation and health, there is a need to increase our understanding about its relationship with obesity.

Expenditure on food was shown to be a significant predictor of childhood obesity (inverse relationship). Household income has been shown to be associated with poor nutritional intake (Nelson, 2000) and low expenditure on food is strongly related to poor growth and health in children, after adjusting for confounders such as parental height or birth weight (Nelson, 2000). Also this study showed that household income and food expenditure were strongly positively correlated with each other, which concurs with the suggestion that the health consequences arise due to insufficient money to spend on food, rather than parents unwisely spending what money they have (Nelson, 2000).

Fruit and vegetable intake was negatively associated with obesity, which concurs with other authors (WHO, 2003; He et al, 2004; Sturm & Datar, 2005). There appeared to be a threshold effect of fruit and vegetable consumption before the protective effect was realised, as no daily fruit and vegetable consumption was positively associated with childhood obesity and eating

five or more portions was negatively correlated, with a weak, insignificant, positive correlation for 1-4 portions consumed daily. Also there is a strong positive correlation between income and fruit and vegetable consumption, and between expenditure on food and fruit and vegetable consumption. These indicate that the higher food expenditure is not about buying more food (more energy intake), but rather about buying more expensive, healthy (low fat, low sugar, low energy density) food such as fruit and vegetables, which is also suggested by Nelson (2000). Again, the collinearity between household income, food expenditure and fruit and vegetable consumption means it is impossible to isolate the variable that is having the effect with childhood obesity.

Consuming school meals was negatively correlated with obesity, suggesting a protective effect. Whilst this result was significant it was dichotomously categorised as the overwhelming majority of households did not spend any money on school meals, so the sample purchasing school meals was very small. Also "not spending on school meals" may also include children who consume free school meals, thus too much emphasis should not perhaps be placed upon this result. Furthermore there has been a lot of recent changes to the nutritional quality of school meals, which has affected both nutrient intake of children and the numbers of children taking school meals (many rejecting the forced healthy eating), and as this dataset is from 2005 or earlier it is contiguous with recent changes. It may be that changing to healthy snacks, such as free fruit at school (Nelson, 2000; Ransley at al, 2007) and reducing consumption of sugar-sweetened drinks (Gregory & Lowe, 2000; Wilson, 2000; Ludwig et al, 2001), has had an impact on the daily energy intake of children and thus on obesity.

In the relationship between childhood obesity and (children's or adults') physical exercise, a dose response is evident (i.e. more exercise, less obesity), with sedentary behaviour having a stronger effect than activity. These results coincide with other work, with a stronger link being found between lifestyles characterised by lack of physical activity and excessive inactivity with increased risk of obesity (Lowry et al, 2002; Matheson et al, 2004). These results suggest that it may be more beneficial to aim interventions towards those people undertaking no physical activity in order to get them to do some, rather than in trying to increase the amount of exercise that people who are already at least partly active do. The primary outlying result with these data was that the correlation with obesity prevention was lower in children who were highly active (more than 7 hours per week) than medium active (3-7 hours per week). It may be that either parents of obese children are reporting what they think they should be doing (whether or not their child is actually that active) or that obese children are encouraged to do more exercise in order to facilitate weight loss (or at least "growing into" their weight). Plus of course these data are self-reported, and one persons idea of "physical activity" is not necessarily the same as another's view.

Due to the strong effect of sedentary behaviour, many cross sectional and prospective studies have looked at the association between TV viewing and childhood obesity. In this study, it was found that the ownership of more than one television per household is positively correlated with the percentage of obese children. The mechanism between increased TV ownership and obesity is likely to be due to increased viewing (sedentary) time; a study in America (Wiecha et al, 2001) found that increased numbers of TV sets in the house, particularly if found in the child's bedroom, greatly increased TV viewing time. Also a prospective study by Gortmaker et al (1996) showed a strong positive dose-response relationship between time watching TV and prevalence of overweight, after adjusting for potential confounders (as measured at the end of the 4 year study). Whilst some studies have only found a weak association between TV viewing and childhood obesity (Robinson et al, 1993; Maffeis et al, 1998), most (including subsequent papers by Robinson (1999; 2001)) find a positive association (after adjusting for potential confounders, such as maternal overweight, previous overweight, family structure, ethnicity, socio-economic status and maternal and child aptitude test scores) in children all over the world; for example, in the UK (Reilly et al, 2005).

Recent debates on whether social capital affects health, suggest that variations in social capital between places may account for between-place variations in health (Mohan et al, 2005). Accordingly social capital was also considered, using data on residents' perception of different neighbourhood factors, and showed increased risk of obesity if there was a problem with teenagers or vandals, or with accessing supermarkets or leisure facilities. It has been shown that childhood obesity is associated with parents' perception of the safety of the neighbourhood (less safe, more obese) (Lumeng et al, 2006) and that fear is inversely associated with physical activity (Ross, 1993; Parkes & Kearns, 2006; Stafford et al, 2007). However, these relationships are with physical activity for leisure/pleasure and not with purposeful exercise, such as walking to get somewhere, which is not associated with safety ratings (Parkes & Kearns, 2006). In fact, in more deprived areas residents are more likely to walk than residents in affluent areas (Ross & Mirowsky, 2001), although this is likely to be due to necessity perhaps because of lack of car availability (plus it does not reflect overall activity levels). Whilst European evidence for an association between retail food access and fruit and vegetable consumption is mixed (Cummins & Macintyre, 2006), in the US the ease of access to supermarkets is associated with increased consumption of fruit and vegetables (Morland et al, 2002) and vice versa (Rose & Richards, 2004). This may be the mechanism by which the difficulty of accessing supermarkets, as found in this study, is associated with increased childhood obesity, particularly as a higher perceived problem with accessing supermarkets was linked with increased deprivation and low income. Physical activity levels have been found to be greater in locations with good local leisure facilities (Giles-Corti & Donovan, 2002) or with

designated play areas (Brownson et al, 2001). Accordingly it might be expected that populations in those SOAs with perceived good leisure facilities have higher physical activity levels, which concurs with finding that poor leisure facilities increases obesity risk. Conversely areas that reported perceived bad public transportation had reduced risk of obesity, lower deprivation scores and high income. As car availability is strongly positively correlated with income, most households with a high income also have a car available. Whether or not a car is available is likely to influence the perception of the quality of the public transport service available - if a car is available, the quality of the public transport would have to be very good before it is used in preference to the convenience of a car, thus leading to a rating of poor.

7.5.2 Local analyses

The local GWR analyses show how the relationship between the covariates and childhood obesity varies in each SOA across Leeds. All of the covariates showed a non-stationary relationship with childhood obesity, stressing the importance of a local level analysis because otherwise these local-level differences would be missed. Furthermore all of these non-stationary processes, with the exception of sedentary behaviour, were statistically significant. The maps of these results enable visualisation of the areas with the strongest negative (dark blue) or positive (red) relationship with obesity, bearing in mind that the overall global relationship will be either positive or negative – although in this instance all have been prepared from an obesogenic perspective (so using "good public transport" which has a positive relationship with childhood obesity, rather than "bad public transport" which has a negative relationship). Different interventions may be required in the SOAs with different (opposite) relationships.

All the covariates illustrated in Figure 7.7 (good public transport, poor leisure facilities, sedentary behaviour, poor supermarket access, and no fruit and vegetables) show a positive overall global relationship with childhood obesity (primarily red shading). Therefore, where a higher proportion of the population perceive the public transport as being good or the supermarket as being difficult to access, who think leisure facilities are poor, who do not eat any fruit and vegetables, and/or who are sedentary, then the higher the prevalence of childhood obesity. However the analyses show that some SOAs' local relationship with childhood obesity is opposite to Leeds' global relationship (blue, instead of red, shading); i.e. in these opposing areas the more people who, say, perceive leisure facilities in the neighbourhood as being poor, then the lower the prevalence of childhood obesity. The implications of this is that in these "blue" SOAs any intervention aimed at, say, improving access to leisure facilities may not be as effective at reducing childhood obesity as we might otherwise expect.

This study also highlights those areas where the key drivers of childhood obesity are different locally to globally. This yields food expenditure, television ownership and internet access, neighbourhood safety, school meals and children's physical activity levels as important determinants of obesity in certain parts of Leeds. Accordingly in these areas interventions focused on these covariates may be the most effective ways to prevent obesity in children. Tailoring interventions to specific localities may be key to their success. One size does not fit all.

These results facilitate greater understanding of local differences in obesity determinants. For example with fruit and vegetable consumption we saw globally, as we might expect given the literature, that increased consumption was associated with reduced levels of childhood obesity and no consumption associated with increased levels. However, some areas showed the converse relationship at the local level, yet it is unlikely that high consumption of fruits and vegetables causes children to become obese, or that no daily fruit and vegetables is protective. More likely is that other factors, peculiar to those specific SOAs showing the converse relationships, are the driving force. For instance, those areas that show a positive local relationship between fruit and vegetable consumption and childhood obesity also show that food expenditure, television ownership and the amount of exercise children undertaken are crucial local determinants of obesity, and it is likely that variations in these local determinants, rather than in the global determinants, that are key. It is necessary for public health officials to determine what these local factors are in order to be able to tailor solutions to each population's requirements for maximum effectiveness.

7.5.3 Relationship with income

The mean data (for the whole of Leeds) for the obesogenic variables does not tell us very much, as the range for most data are very wide indicating that there is a lot of variation across Leeds. This variation is highlighted in Figures 4-6, which map the different mean values for each SOA (as opposed to the mean for the whole of Leeds), enabling any patterns to be visually identified. These figures show that many covariates in isolation show a pattern around the centre of Leeds, with a few miscellaneous SOAs standing out. When these maps are compared to the income / deprivation maps in Figure 10, eyeballing the information suggests many covariates are strongly associated with deprivation and/or household income (as the patterns are the same/similar), which concurs with the correlation analysis. However the non-central hot/cold spots suggest that other relationships also exist and the local analyses delve into this more.

When the maps of the global relationships between the covariates and obesity (Figure 7) are compared to the income / deprivation maps (Figure 10), the patterns are not the same.

Conversely the strongest relationships between the covariates and obesity are largely in the north east or south east of Leeds, and does not concur with the deprivation patterns. That is, it seems that factors other than household income and deprivation are affecting the relationship between the covariates and the prevalence of childhood obesity in these outlying (i.e. not central Leeds) areas.

7.5.4 Limitations

This chapter has a number of limitations. These analyses are based on simulated covariate data, not "real" data, however the validation of the simulation was robust and these data do enable analysis at the micro-level, elucidating more about the relationship between obesity and covariates than would otherwise be the case. The covariate data are subject to all the data collection limitations of the underlying surveys, namely the HSE and EFS. In particular much of the information collected, such as the physical activity data, was by self report and may be subject to reporting bias, which both surveys have sort to minimise through their rigorous study designs. The actual obesity data are based on a sample of the population; the mean number of children measured in each SOA was 71 (range 2 to 156). This small number problem was adjusted for using Empirical Bayes to smooth the data, which works by borrowing strength from areas with higher (measured) populations. The choice of SOA as the unit of analysis was driven largely by the availability of data and scale of convenience and not due to a priori theoretical consideration. Using SOAs (n = 476) rather than the smaller scale output areas (n = approx. 2400) helped to reduce small number problems. There is no information about the exposure time for people; it is assumed that all individuals have lived in the same area their whole lives, yet migration in a small proportion of the population is likely, albeit possibly between areas with similar characteristics. Most of the covariates were strongly correlated with household income, which creates difficulties in interpreting the results. The problem of having two (or more) predictor variables that are correlated, is that it is not possible to identify which one is having an effect on the dependent variable (childhood obesity).

In order to be able to aggregate the data to SOAs it was necessary to categorise continuous data, the choice of categories being fairly arbitrary, although common sense options were chosen (e.g. the fruit and vegetable breakdowns were based around the fact many people consumed no daily fruit and vegetables and that the recommended daily intake is five portions per day). Limitations were driven by definitions given in the underlying datasets (i.e. either the Health Survey for England or Expenditure and Food Survey). For adults' physical activity, this was defined as the number of days that the person is active for more than 30 minutes per day. However a person who exercises vigorously for many hours in a day but only on 3 or 4 days per week (perhaps because they need a rest in-between!) would be categorised as moderate activity,

yet someone who does 30 minutes of walking every day would come out as a high exerciser. Total hours (and type) of exercise are very different and this categorisation (which stems from the underlying dataset) is poor.

With the GWR analyses the standarised residuals were considered and showed a flat picture, suggesting the model is well fitting. Similarly the standard errors were low in most areas. Higher rates were evident in the north east corner of Leeds (despite the smoothing) most likely because of the fixed kernel size in the GWR model. That is the local model used a set distance from the sample point (the geographic centre of each SOA) rather than number of nearest neighbours; so SOAs with a larger area and/or fewer measurements will have a larger standard error. Overall the standard errors were low, although arguably use of a floating kernel may have produced lower standard errors.

7.5.5 Conclusion

Individuals' behaviour was shown to be important in determining levels of obesity, including lack of physical activity, no or low consumption of fruit and vegetables, households that do not purchase school meals, and low expenditure on food. Additionally, this study adds to the increasing evidence of the existence of "obesogenic environments". Features of the local environment may affect childhood obesity by changing health behaviours, i.e. promoting a healthy diet and encouraging physical activity (and vice versa). Several local determinants of obesity were identified in this study; obesogenic environments are characterised by the following: scale of urban development, deprivation (high deprivation score, receives benefits, low income), low SEG, more than one television per household, poor leisure facilities, poor access to supermarkets, problems with teenagers hanging around or with vandals, and good public transport facilities. Consequently more effective interventions to reduce childhood obesity would be those targeted at multiple determinants, rather than focusing solely on the individual child.

Many of these determinants of childhood obesity are outside of the remit of the health care sector. This highlights the role of organisations outside the health sector (e.g. supermarkets, gyms, swimming pools, food manufacturers), as well as local and national government (e.g. decisions regarding public spending, agricultural policy) in contributing to the obesity problem. This emphasizes the responsibility that public and private sector organisations have in endorsing public health (Stafford et al, 2007). Thus in order to maximise childhood obesity prevention, health care professionals need to work with these organisations, otherwise it will not be possible to halt the flood of obesity. Multivariate analyses are constrained due to collinearity between the covariates. Further work considering combinations of local factors needs to be undertaken

to identify which combinations of risk factors are the most obesogenic in order to be able to stem the epidemic.

These determinants of childhood obesity are not uniform across Leeds; they vary such that there are some places with particularly high positive or negative associations, and where, consequently, there is a particular need for different types of public health action. Also, factors that are associated with obesity in one locality, may not be in another. This suggests that it is either a combination of risk factors that are important and/or perhaps a threshold of negative environmental factors needs to be attained, before the individual is tipped over into obesity and/or the environment becomes obesogenic. The ecological approach to health behaviour states that as well as being influenced by individual factors, factors outside the individual (i.e. social and cultural and the physical environment) also impact the choices people make in relation to health behaviour (see Chapter 2). Equally important is the interaction between these different factors - with different individuals being influenced by different factors, or in different ways by the same factors (see Chapter 2), which is the essence of non-stationary relationships whereby the same stimulus does not produce the same effect in all areas. This obviously complicates analysis and the ability to drill to the root of the problem. However with this in mind it is possible to identify differences between areas and to tailor interventions and health policies to embrace these differences to increase effectiveness.

These analyses highlight the different relationships between the obesogenic covariates and childhood obesity. The significant variation seen in these relationships across Leeds suggests that different interventions to reduce the prevalence of childhood obesity may be more effective in different areas. Accordingly there is considerable advantage to analysing health data at a small scale, otherwise this differences are simply "averaged" away and missed. The chapter clearly shows how micro-level spatial analysis can be used to enhance analysis, leading to more targeted decision making in health planning.

Chapter 8: The development of a childhood obesity prevention policy for Leeds using an ANGELO-style framework

8.1 Introduction
8.2 Background
8.2.1 Interventions to prevent obesity in children
8.2.2 The ANGELO model
8.2.3 Current obesity prevention policy in Leeds
8.3 Methodology
8.4 Results
8.4.1 ANGELO-style framework to prevent obesity
8.4.2 Identify three case study locations
8.4.3 Obesity variance across each area
8.4.4 Childhood obesity and deprivation / Open Area Classification super groups
8.4.5 Childhood obesity and the simulated covariates
8.4.6 Tailoring the framework to prevent obesity in Leeds
8.4.7 Comparison to Leeds policy
8.5 Discussion

8.1 Introduction

So far this thesis has considered how childhood obesity varies across Leeds, at both the residential level (chapter 4) and school environment (chapter 5). The variation in different groups of potential obesogenic variables, namely the distribution of populations with perceived low social capital, populations with low income, obesogenic behaviours of children and of people with low perceived neighbourhood safety in deprived areas, were summarised (chapter 6). It has also considered relationships between childhood obesity and different obesogenic factors, namely deprivation and OAC super group (chapter 4), and various simulated variables (chapter 7): fruit and vegetable consumption, school meal consumption, spending on food per household (eaten at home and externally), physical activity levels, number of televisions per household, perceived social capital, urbanisation, socio-economic group (SEG), size and source of household income, PC ownership, and internet access. This has shown that both childhood obesity varies across Leeds and also that the relationship of the obesogenic variables with childhood obesity varies across Leeds – the key global determinants of childhood obesity are not necessarily the same as the most important local determinants.

This chapter seeks to pull this work together into a practical application. Using the results from the previous chapters to identify the key obesogenic factors at the micro-level, three case studies of three different neighbourhoods in Leeds will be studied in depth – the most affluent, the most deprived, and a more average ward. An ANGELO-style framework (described below) will be used to develop and prioritise potential targeted interventions and health policies to facilitate the prevention of childhood obesity in these three areas, using factors that are amenable to measurement, intervention, and change. These suggestions will also be compared and

contrasted with actual local policy in Leeds with a discussion regarding what could, and perhaps should, be altered to facilitate the prevention of childhood obesity.

8.2 Background

8.2.1 Interventions to prevent obesity in children

Ecological evidence shows that the prevalence of childhood obesity is rising in the UK. Between 1995 and 2002 in children aged 2 to 15 years old, the prevalence of obesity in boys rose from 10% to 17% and in girls from 12% to 17% (Sproston & Primatesta, 2002). This coincides with secular decreases in habitual energy expenditure in children (DiGuiseppi et al, 1997; WHO, 2004b; Wareham, 2007) and paradoxically decreases in reported food intakes (Cavadini et al, 2000). It is interesting to note, however, food disappearance data suggest energy intakes have increased (Harnack et al, 2000).

The observational study evidence, where individuals' activity levels and changes in weight are measured, identifies the components of diet and physical activity that are key determinants of obesity. A review by Brown et al (2007) summarises the key determinants of obesity. The only factor that they concluded as having a "convincing" evidence base was increased (over time, not just high) total physical activity (reducing the risk). Other variables that are "probable" factors are breast feeding (reduced risk for the infant), diets rich in low energy-dense foods and high non-starch polysaccharides/dietary fibre foods (reduced risk), frequent large portions of energy-dense foods and/or sugary drinks (increases risk).

There are several papers (Lobstein et al, 2004; Summerbell et al, 2005; Flynn et al, 2006; Connelly et al, 2007) that thoroughly review interventions to prevent childhood obesity. There have been a small number of randomised controlled trials (RCTs) and of these very few have had an effect in preventing childhood obesity. Randomised and non-randomised controlled trials, with at least 30 participants, with a follow up of at least three months (short term) or at least 12 months (long term) are summarised in Table 8.1. The bulk of the interventions include both a diet and physical activity arm, with only two including only a dietary intervention and six undertaking only a physical activity intervention. Many were run in a school setting, some of which also had links to the home environment, and some were run in a community setting. Of these interventions, 11 were effective at reducing adiposity and 18 were not effective.

result	authors	NE	NS	PE	TVE	PA-c	PA-v	age	ethnicity	gender	Setting	term
Effective	Muller et al (2001a)	yes	yes	yes			yes	6-10 yrs	caucasian	both	home/school	LT
Effective	Robinson (1999)			yes	yes		yes	c.9yrs	caucasian	both	home/school	ST
Effective	Vandongen et al (1995)	yes		yes		yes		11-14 yrs	caucasian	both	home/school	ST
Effective	Fitzgibbon et al (2005)	yes		yes		yes		6-10 yrs	other	both	school	LT
Effective	Gortmaker et al (1999)			yes	yes		yes	c.12 yrs	cauc/minor	girls	school	LT
Effective	James et al (2004)	yes						6-10 yrs	caucasian	both	school	LT
Effective	Mo-Suwan et al (1998)			yes		yes		0-4 yrs	other	girls	school	LT
Effective	Sallis et al (2003)	yes		yes		yes		11-14 yrs	cauc/minor	boys	school	LT
Effective	Simonetti et al (1986)	yes		yes				3-8 yrs	caucasian	both	school	LT
Effective	Flores (1995)	yes		yes		yes		10-13 yrs	cauc/minor	girls	school	ST
Effective	Kain et al (2004)	yes		yes		yes		6-10 yrs	other	both	school	ST
Ineffective	Epstein et al (2001)	yes	yes	yes			yes	6-10 yrs	caucasian	both	community	LT
Ineffective	Beech et al (2003)	yes	yes	yes			yes	6-10 yrs	AfAmerican	girls	community	ST
Ineffective	Robinson et al (2003)			yes	yes		yes	6-10 yrs	AfAmerican	girls	community	ST
Ineffective	Stolley et al (1997)	yes	yes	yes			yes	7-12 yrs	AfAmerican	girls	community	ST
Ineffective	Warren et al (2003)	yes				yes		6-10 yrs	NatAmerican	both	home	LT
Ineffective	Caballero et al (2003)	yes	yes	yes			yes	6-10 yrs	NatAmerican	both	home/school	LT
Ineffective	Newmark-Sztainer et al	yes	yes	yes			yes	0-6 yrs	AfAmerican	girls	home/school	ST
Ineffective	Coleman et al (2005)	yes	yes	yes			yes	6-10 yrs	other	both	school	LT
Ineffective	Donnelly et al (1996)	yes	yes	yes			yes	6-10 yrs	caucasian	both	school	LT
Ineffective	Leupker et al (1996)	yes	yes	yes			yes	6-10 yrs	cauc/minor	both	school	LT
Ineffective	Sahota et al (2001)	yes	yes	yes			yes	6-10 yrs	cauc/minor	both	school	LT
Ineffective	Sallis et al (1993)			yes		yes		6-10 yrs	cauc/minor	both	school	LT
Ineffective	Baranowski et al (2003)	yes	yes	yes			yes	6-10 yrs	cauc/minor	both	school	ST
Ineffective	Dennison et al (2004)			yes	yes		yes	0-6 yrs	caucasian	both	school	ST
Ineffective	Harvey-Berino et al	yes	yes	yes			yes	14-17 yrs	cauc/minor	girls	school	ST
Ineffective	Pangrazi et al (2003)	yes	yes	yes			yes	6-10 yrs	AfAmerican	both	school	ST
Ineffective	Story et al (2003)	yes		yes	yes		Yes?	c.9yrs	?	girls	school	ST
Ineffective	Ramirez-Lopez et al	yes						6-10 yrs	other	both	school	?

Table 8.1. Summary of RCTs to prevent childhood obesity (adapted from Connelly et al (2007) and Summerbell et al (2005)). Key: cauc/minor: caucasian and black or minority ethnic group; AfAmerican: African American; NatAmerican: Native American; Community: community or other setting. NE: nutritional education; NS: nutritional skills education; PE: physical activity education information; TVE: education to reduce TV viewing (and/or media time); PA-c: Physical activity compulsory participation; PA-v: physical activity voluntary participation. ST: short term study (at least 3 months); LT: long term study (at least 12 months). All studies had at least 30 participants.

Positive outcomes on adiposity were shown with small-scale interventions in schools, such as, changing children's TV watching behaviour, nutritional / healthy lifestyle education, and using a price differential to promote the consumption of healthier foods. This type of intervention has demonstrable and significant favourable results, but these are small effects (given the scale of the obesity problem). Also, after interventions finish, any progress tends to regress (Lobstein et al, 2004; Summerbell et al, 2005): the interventions lack sustainability. This suggests that either interventions are required where the effects persist after they finish, or, more likely perhaps, we need permanent "interventions". That is, lifestyle changes and continuous encouragement and support to resist the obesogenic environment and prevent obesity – at least for the foreseeable future. The absence of long-term follow-up data makes it difficult to evaluate the efficacy of the interventions for population-wide effects on obesity prevalence and the effects of the interventions when the children become adults remain to be assessed.

Whilst it is important that a measure of childhood obesity is included as an outcome measure, otherwise it is not possible to tell if the intervention is having an impact on obesity (Lobstein, 2006), perhaps this is not the only way of assessing / measuring effectiveness. For example, there is an argument to consider positive behavioural changes as well (and perhaps the effect on emotional health). Many studies do see significant changes in diet and physical activity levels during the intervention. Whilst, ultimately, if there is never an effect on obesity, then the intervention is not effective at preventing obesity, but has enough follow up time been allowed to permit changes in behaviours to filter through to changes in obesity? Plus if physical activity is a component of the intervention then body fat and/or fitness tests may be appropriate outcome measures (Summerbell et al, 2005).

Lessons learnt from research to date suggests that future studies should also take care to ensure that they include the right number of participants to ensure adequate power, the follow up period is long term, and that appropriate and dependable outcome measures are included (Summerbell et al, 2005). It has not been possible to suitably assess many interventions because inadequate information is supplied (Lobstein, 2006). Other key problems have been in determining the most cost-effective interventions that are also generalisable to other situations (Summerbell et al, 2005). There is a lack of studies with interventions focused at pre-school children, with a gender specific component, or for immigrant (new to developed world, not simply ethnic minority populations) children, which needs to be addressed in the future research (Flynn et al, 2006).

It is necessary to understand the (potential) cause(s) of obesity in order to elucidate judicious public health strategies. But the equivocal evidence is far from compelling and does not

highlight one particular course of action as the best, thus failing to identify best practices to form the basis of an obesity prevention policy.

The traditional approach would be to undertake further good quality RCTs, which take time and are expensive. RCTs are regarded as the gold standard of evidence for the purpose of determining causality; other forms of evidence are considered less worthwhile. (Kroke et al, 2004). However, for public health purposes, RCT evidence may be "inappropriate, unachievable or irrelevant" (Swinburn et al, 2005). A RCT necessitates that a single or limited collection of factors are manipulated, and also that the distribution and context of the intervention are completely controlled (Lobstein et al, 2004). This level of control may make interventions too unrealistic and disruptive of the real world and the multifaceted approach required to affect population health may find these rigorous constraints "too artificial or unrealistic" (Swinburn et al, 2005). This is because obesity prevention interventions are necessarily complex public health approaches: dispensed in multifaceted circumstances over the long term (Lobstein et al, 2004), making it difficult to control all the variables, and to evaluate the effectiveness (Flynn et al, 2006). Further they are frequently targeting individuals and populations concurrently – to affect individual behaviours as well as the social and environment factors that affect population health (Lobstein et al, 2004). By definition there is continual danger of contaminating the control group, as a public health policy relies on its effects filtering through to the target community (Lobstein et al, 2004), which would serve to reduce the ostensible efficacy of an intervention. Accordingly, while the development and results of more RCTs are waited for, it is contended that policy should be based on the best of the existing evidence base (best available evidence as opposed to best possible evidence (Swinburn et al, 2005)). To sit on our (growing) laurels and wait for more evidence is likely to give sufficient time for the obesity time bomb to explode, and then it will be much harder to reverse the trend. We cannot afford to simply gather more evidence and wait.

So, based on the existing evidence base, what may be helpful in preventing obesity in children? Schools are a logical setting to roll out childhood obesity prevention interventions. But it is necessary to take a whole school approach, including cafeterias, PE classes, lunch and break activities, plus classroom teaching, and importantly it should also link back to the home and community (Summerbell et al, 2005). Programmes that impact several environments concurrently are required because of obesity's complex, multi-factorial aetiology, with the likely cause coming from combinations of variables with different levels of influence (Dietz & Gortmaker, 2001). Long-term benefits will only accrue from coordinating interventions with changing ingrained social and cultural beliefs (Lobstein et al, 2004). A multi-component school-based intervention addressing diet and/or physical activity is required. It is at least as important that environmental modifications, such as changes in school physical education, are

included in the programme as much as classroom-based educational interventions (Connelly et al, 2007; Wareham, 2007). More effective (at preventing child weight gain) interventions are those in the school directed at raising physical activity levels by PE classes and behaviour change rather than those directed towards decreasing sedentary time or those focused on the family environment (Wareham, 2007). In particular compulsory aerobic physical exercise seems to be important to success (Connelly et al, 2007), however it is not possible to strip out the effect from the rest of the intervention. That is, it may be that the provision of nutritional and nutritional skills education along side the compulsory physical activity component is essential for the intervention to be effective (Connelly et al, 2007). Additionally different age groups, genders and ethnic groups may need different approaches.

However, on top of this, a larger scale approach is advocated if the environmental drivers that promote obesity are to be confronted successfully. It is not sufficient to put in place interventions at simply the individual, family, school and/or community level, whilst contradictory messages continue to come from "higher" levels of media, industry and government (Lobstein, 2006). Yet to date, these macro levels (Swinburn et al, 1999) have been ignored. None of the studies in Table 8.1 dealt with population-wide policy-related issues such as food marketing, labelling or pricing, planning controls or transport policies (Lobstein, 2006), yet obesity is influenced by markets and governments (Lang & Rayner, 2007). There are many different stakeholders including the food industry, medical community, schools, employers, parents, advertisers and the media, recreation and sports planners, city planners, social and welfare services, manufacturers, retailers, transport, international trading and standard setting bodies, and local and central government – all of whom must recognise their role and responsibilities in this fight against obesity (Jebb et al, 2003). These macro-environment influences on obesity need to work in a co-ordinated fashion for maximum effect. For example, contradictory government policies, such as recommending individuals eat less fat whilst at the same time subsidising agriculture to churn out excess fat, are likely to have the effect of cancelling each other out (Lang & Rayner, 2007). Thus policy coherence is required to enable obesity to be successfully prevented.

8.2.2 The ANGELO model

Different models have been proposed to facilitate the understanding of the complex, multi-factorial aetiology of childhood obesity, including: Ecological Systems Theory (Davison & Birch, 2001); Epidemiological Traid (Swinburn & Egger, 2002; Egger et al, 2003); Ecological Model (Egger & Swinburn, 1997; Swinburn et al, 1999). All expand the energy balance equation to look at the broader environmental factors and their role in influencing energy balance in order to facilitate the identification of obesogenic factors and the prevention of

obesity. Similarly, Flodmark et al (2004) suggest, without the use of a model, that there are six levels that should each be considered when addressing a preventative programme for childhood obesity. Each of these models seeks to address the multifaceted nature of obesity. They all concur that the determinants of obesity sit at many different levels, and agree that successful prevention of obesity needs to work at all of these levels. However, it is how these levels are defined and summarised that varies between the models. Nevertheless each model is seeking to address the complex web of behaviours that impact a person's dietary and physical activity choices, including: the individual; family influences; neighbourhood and socio-cultural influences (including schools); industry, media and government.

One example is the framework that has been developed using the Ecological Model (Swinburn et al, 1999) that attempts to identify obesogenic factors in the environment, called the ANGELO (ANalysis Grid for Environments Linked to Obesity) framework. It is a grid made up of two sizes of environment (micro environmental "settings", such as schools and neighbourhoods, and macro environmental sectors, such as transport and health system) on one axis against four types of environment (physical, economical, political, socio-cultural), with food and activity as subcategories for each type, on the other axis (see Figure 8.1). These macro environmental "sectors" influence whole populations, including bodies such as the food industry and governments. The Physical environment looks at "what is available?" This includes food and physical activity options together with more intangible features such as technology, information and expertise, and the availability of training opportunities. The Economic environment looks at the "financial factors", covering both the consumers' costs and income as well as money spent on advertising/health promotion by industry and health departments and other government funding (such as on roads, public transport and recreation facilities). The Political environment explains the "rules". That is, the policies, regulations and laws, the institutional (including the home and school) rules that impact food and activity choices. Finally, the Sociocultural environment asks, "what are the attitudes, beliefs, perceptions and values?" At the micro level this covers the "culture" or "ethos" of a school, home, workplace or neighbourhood. At the macro environment level it would include things such as the influence of the media on influencing socio-cultural aspects of food and activity choice, particularly in relation to advertising and marketing. (Swinburn et al, 1999).

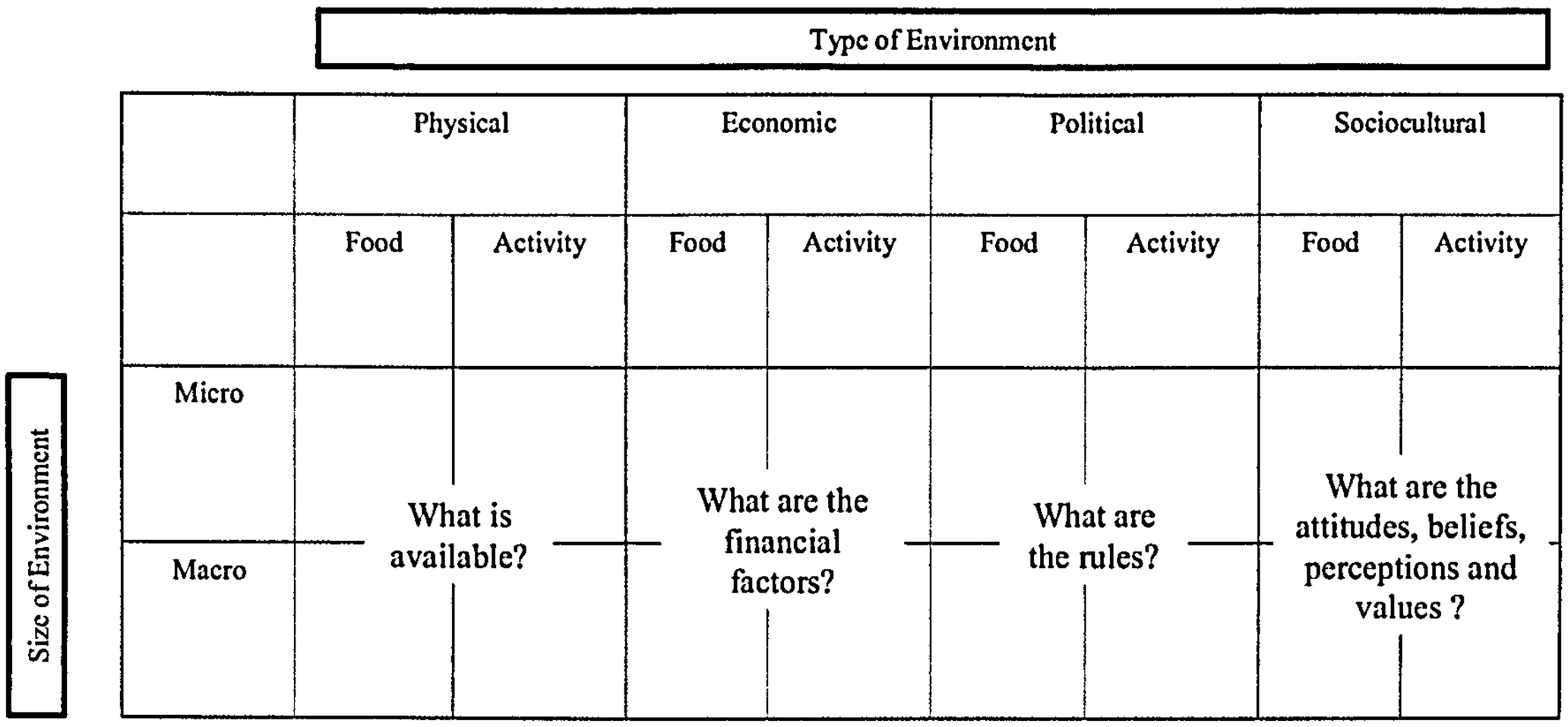


Figure 8.1. The ANGELO Framework. ANalysis Grid for Environments Linked to Obesity (ANGELO). This diagram is a pictorial representation of the ANGELO grid. The two axis are size of environment and type of environment. Each environment type is further subdivided into food and activity, to represent the two elements of the energy balance equation (Swinburn et al, 1999).

8.2.3 Current Obesity prevention policy in Leeds

The UK government has set a target for childhood obesity, namely "to halt, by 2010, the year-on-year increase in obesity among children under 11 years of age" (DH, 2004; Foster & Buttris, 2005), which was recently updated to "by 2020, we aim to reduce the proportion of overweight and obese children to 2000 levels" (DH, 2008). Three government departments are responsible for this target: the Department of Health (DH), Department for Culture, Media and Sport, and Department for Children, Schools and Families (formerly the Department for Education and Skills). Yet the UK (and Europe) does not have a comprehensive structure or policy model for what action to take to prevent obesity (Lang & Rayner, 2007). There are many government initiatives that have been detailed as priority measures by the DH. These are outlined in Table 8.2, with an indication of the age group the strategy is aimed at, together with an indication of the type of intervention.

To facilitate the achievement of the childhood obesity target, obesity prevention is being included on the agenda of local councils through Local Area Agreements (LAAs) and Local Public Service Agreements (LPSAs) (DH, 2007). These policies are a link between central and local government. The LPSAs are a mechanism to prioritise and agree the twelve most important targets for improvements in performance in the local area with central government. The LAAs provide a framework for service delivery in the local area, which can relate to national or local priorities.

Initiative	Target age group (years old)	Type of intervention
Healthy start	0-10	Nutrition
Breastfeeding promotion	0-2	
School food agenda	5-10	
Reformulation, portions sizes and signposting	2-10	
School fruit and vegetable scheme	4-6	
Food promotion to children	2-10	
Healthy schools programme	5-10	
PE in School s(PESSCL)	5-10	Physical activity
Children's play	0-10	
Active travel including school travel plans	5-10	
Obesity care pathway	2-10	Obesity
Obesity social marketing	2-10	

Table 8.2. Outline of various initiatives to tackle obesity set as priority measures by the DH (DH, 2006).

In Leeds the twelve LPSA targets cover topics such as fire, road safety, burglary, youth offending, homelessness and domestic violence. There is no specific target to reduce obesity prevalence to an agreed level, although three of these targets may impact obesity prevalence in Leeds – improving school children's health (through both tackling obesogenic behaviours and the healthy schools standard), improving physical activity, and cleaning up the local environment (see Table 8.3).

Target	Heading	Status	Notes	Target Value	Included in LAA block
4	Improving children's health in all Leeds schools (a) Tackling behaviours that contribute to obesity (b) Healthy Schools Standard	Agreed		½ 1	C&YP
7	Improving physical activity	Dropped	No further negotiation possible		
11	Improving the cleanliness of the local environment	Agreed		1	SSC

Table 8.3. A summary of three of the twelve Local Public Service Agreement (LPSA2) targets. These targets may impact obesity prevalence in Leeds.

The LAA for Leeds (Leeds City Council, 2007a) is very disappointing from a childhood obesity perspective. There are no specific outcomes or targets for obesity, although some, which are highlighted below, do consider factors that may be associated with obesity and thereby may affect obesity rates in Leeds. The Leeds LAA vision wants all Leeds' children "to be happy, [and] healthy", yet focuses "primarily, but not exclusively" on the 31 most deprived SOAs of Leeds. It also focuses on children in Year 9. There are four themes to the Leeds LAA: health and well being; housing and the environment; crime, safety and reassurance; employment,

skills, learning and enterprise. The specific outcomes and initiatives that may impact on childhood obesity for each of these themes are considered in turn.

Children and Young People

The only Mandatory Outcome remotely associated with obesity is to ensure all 284 schools in Leeds have an approved travel plan by March 2010 (milestones: 65% by 12/07 and 80% by 12/09). This strategy is to encourage children to walk and cycle to school, investing £40 million (nationally) to facilitate this (e.g. with increased cycle storage facilities, linking schools into the National Cycle Network, and to improve pedestrian safety through projects such as Kerbcraft) (Leeds City Council, 2007b).

Three additional outcomes (albeit not high enough up the agenda to be included in the outline of the policy – only readable in the small print) may be associated with childhood obesity. Firstly, to raise the percentage of 5-16 year olds undertaking at least two hours of "high quality PE and school sport" per week from 83% in 2006 to 90% by 8/08. Secondly to raise the percentage of schools that attain the National Healthy Schools Standard from 0.5% in 2/06 to 50% by 8/07 and 95% by 12/09. These standards include components on healthy eating and physical activity. Also, importantly, it involves the whole school community, from parents to school staff to governors (Healthy Schools, 2007). Thirdly to increase the percentage of schools that offer "the "core offer" of extended schools services" from 13% in 9/06 to 25% in 9/07 and 33% in 9/08. This core offer involves schools offering activities for the local community, including childcare from 8am to 6pm, two hours of extra-curriculum activities per week "for those who want it" (e.g. homework clubs, sport, music tuition), adult education facilities, referral to specialist support services (e.g. speech therapy, mental health services), and access to IT, sports and arts facilities (Education Leeds, 2007).

Of the other outcomes the bulk (9 out of 24) are to do with grades achieved by pupils. The remaining outcomes include topics such as teenage pregnancies, unemployment, child register/foster/adoption, attendance levels, number of excluded children (from school).

Healthier Communities and Older People

Mandatory Outcomes under this theme aim to reduce both rates of premature mortality and rates of inequalities in premature mortality, with an emphasis on reducing the risk factors for heart disease, stroke and related diseases (i.e. smoking, diet and physical activity). No specific references to diet (or smoking) were seen in the small print for this policy, however there were targets for increasing physical activity. That is, to raise the number of adults undertaking a

minimum of 30 minutes of "moderate intensity physical activity" on at least five days a week by 1% per annum (no baseline figures for this were given – to be established from the Leeds Physical Activity Survey in May 2007). Similarly, for a 1% per annum increase in the number of adults undertaking a minimum of 30 minutes of "moderate intensity sport and active recreation" on at least three days per week (baseline being established from Sport England Active People Survey in 10/06, results not published). Other Outcomes (Mandatory and Additional) focused largely on benefits and transportation.

Safer and Stronger Communities

There were a number of indicators under the Mandatory Outcome to "build respect in communities and reduce anti-social behaviour", largely around changing people's perception of such problems in their neighbourhood. For example, the people who feel anti-social behaviour is a very big problem in their area; those who feel it has got worse in the last year; problems with noisy neighbours, with teenagers hanging around, littered streets, abandoned or burnt out cars, vandalism, graffiti and other deliberate damage, or drug dealers and users. All of these examples have a target of a year on year reduction in the percentage, with no set percentage reduction given. Other examples include the percentage of residents who feel parents take responsibility for their children's behaviour, and the percentage of residents who feel treated with respect by their neighbours. Both of these have a target for year on year increases in the percentages found. Other outcomes under this theme included lots of focus on reducing crime and drink/drugs related problems, and some outcomes regarding quality of housing, cleaning up civic areas (particularly in deprived areas), promoting a sense of pride in neighbourhoods, accidental deaths and road traffic accidents, and homelessness.

Economic Development and Enterprise

Whilst the bulk of the outcomes for this theme were to support new businesses, create new jobs, reduce benefit claims and improve financial literacy support, there were a number of Mandatory Outcomes related to physical infrastructure. These focused on improving the transport network and investing in the public infrastructure to increase the percentage of non-car journeys into the city centre during peak times. The targets for this are very small, increasing the percentage of non-car journeys from 43% in 2005 to 43.9% in 2007/08 and 44.3% in 2008/09. This Mandatory Outcome also includes indicators to increase the number of visitors/users to the City Council's sports facilities and parks and countryside. The baseline figure for sports facilities is 4,105,506 (in 2005/06) and is targeted to actually decrease to 3.9 million in 2008/09, in order to allow for anticipated closures of facilities due to Private Finance Initiatives developments.

Similarly, the baseline figure for users of parks and countryside was 61,213,587 (in 2006/07), but this indicator will not be measured again and there are no targets.

8.3 Methodology

The first stage in developing a local childhood obesity prevention policy for Leeds was to place the results of existing evidence from interventions to prevent childhood obesity into an ANGELO-style framework. The ANGELO model was adapted slightly to clarify the micro- and macro-level environments considered: "micro" was broken down into individuals, schools and communities; "macro" was broken down into media, industry and government (in line with Lobstein et al, 2004).

The next stage was to calculate the mean and range of deprivation using the Index of Deprivation, 2004 (Communities and Local Government, 2004) for each SOA in each of the 33 wards in Leeds, in order to select three case study wards. Also, for the three case study areas, the proportions of populations in each of the National Statistics 2001 Open Area Classification of Output Areas (OAC) "super-groups" (Vickers & Rees, 2007) were calculated from the OAC data for the OAs in each ward, in order to give more demographic information about each area. A map showing the location of the three case study areas was prepared using ArcGIS v.9.0, as well as four maps of various combinations of the simulated covariates (as described in Chapter 6). The combinations were a distribution of perceived low social capital, less affluent households, obesogenic behaviours and unsafe/deprived areas so as to highlight the differences between the areas.

Following this, the variation in obesity across each case study area was examined. The obesity data used was that collected from routine measurements of children in the Leeds PCTs, data from the "Trends" study in Leeds, and the "RADs" study in Leeds (as described in Chapter 4). Obesity was considered in two ways. Firstly the percentage of obese children in each SOA, with "obese" classified as children above the 98th centile using the British reference dataset (Cole et al, 1995). Also using mean body mass index standard deviation score (BMI SDS) (standardised using the British 1990 growth reference dataset (Freeman et al, 1995; Cole et al, 1995)). The average individual (actual) childhood obesity data were tabulated for each study area: the percentage of obese children and the mean BMI SDS calculations for each case study ward are averages (mean) from the individual child data, which have then been smoothed (aggregated at ward level) using Empirical Bayes techniques to minimise small number problems (as described in chapter 4). The tabulated Minimum and Maximum BMI SDS figures are not smoothed but the actual range of individual child BMI SDS values. Next the SOA level data (i.e. the individual level data aggregated to SOA level) were mapped using ArcGIS v.9.0 to

show both prevalence of obesity and mean BMI SDS across each case study area. Similarly the location of the schools sampled in the Trends project in each of the study areas is shown, using the analysis from chapter 5 for whether the level of obesity at each school is in line with expectations (given deprivation and ethnic mix) or not. The hot spot (of childhood obesity) calculations (determined in chapter 4) were applied to each case study to highlight whether any hot spots exist in each case study area.

In order to be able to design targeted interventions for each of the case study areas, it is important to understand the relationship between childhood obesity and potential obesogenic factors in each area at the micro-level. Accordingly the work undertaken in Chapters 4 and 7 has been drawn on to ascertain the important local determinants of childhood obesity, and the direction of those relationships, for each case study area. Accordingly, next, the variance for the relationship between each covariate and childhood obesity was considered: firstly for deprivation and the OAC super groups, and then for the twelve simulated covariates highlighted in chapter 7 as having a strong relationship with obesity (namely, transportation, leisure facilities, supermarket access, child and adult sedentary behaviour, fruit and vegetable consumption, buying school meals, food expenditure, problems with teenagers loitering, television ownership, internet access, and levels of physical activity in children).

The mean BMI SDS for each of the seven OAC super groups was calculated using the individual (raw) data for the three case study areas cumulatively, and compared to the mean figures for the whole dataset for Leeds. The simulated variables were analysed using global and local geographically weighted regression (GWR) models in Chapter 7. The results of these models were summarised for each of the twelve principal determinants of obesity now being considered. A map of the relationship between each simulated covariate and childhood obesity was prepared for each case study area using ArcGIS V.9.0, and the obesogenic covariates in each area described and tabulated, thereby highlighting the differences in the relationships between the areas.

Finally the results from the analysis of the relationships between childhood obesity and its determinants in each case study area were applied to the framework in order to prevent childhood obesity in Leeds. This shows how policy can be tailored to the specific needs of each micro-area to prioritise the obesogenic factors for change and to guide the development of targeted interventions and perhaps how they are implemented. These recommendations were then compared and contrasted to the existing Leeds policy (the LPSAs and LAAs) in order to determine whether all the bases are being covered or if the policy is lacking in any respects.

8.4 Results

8.4.1 ANGELO-style framework to prevent obesity

The results of the ANGELO-style framework for the prevention of childhood obesity in Leeds is given in Table 8.4. This policy is based on literature both generalisable to the UK and internationally. This model is generic for Leeds. However, as is shown below (section 8.4.6), for maximum effect, it is likely to be beneficial to "customise" the policy for the area in question.

8.4.2 Identify three case study locations

Three wards were chosen for analysis based on a mixed demographic profile. The most affluent (Wetherby), the most deprived (City and Holbeck) and a mid-range ward (Morley South) were chosen (see Table 8.5). The Open Area Classification (OAC) "Super-group" was also considered: City and Holbeck's population is largely "multicultural" and "constrained by circumstances"; Morley South is "typical traits", "prospering suburbs" and "blue collar communities"; Wetherby is largely "prospering suburbs" and "countryside" (see Table 8.6). The location in Leeds of the three wards being used as case studies, City and Holbeck (CH), Morley South (MS), and Wetherby (W), are given in Figure 8.2.

The maps of combinations of the synthetic covariates highlight differences between each case study area. CH and to a lesser degree MS show a high distribution of low social capital, whereas the social capital in W is high (see Figure 8.3). Figure 8.4 illustrates some aspects of poverty, namely low food expenditure, unemployment, tenants without a car. Very few people in W or MS match these criteria, whereas most people in CH do. In relation to obesogenic behaviours of diet and activity (see Figure 8.5), many children in W do not eat fruit and vegetables daily and are very inactive. This is less true of the MS children and it seems that few children in CH are both eating a poor diet and inactive. Finally some areas of CH are both highly deprived and perceived as unsafe by the residents, as do a few areas in MS, but W is an affluent area (Figure 8.6).

Table 8.4. An ANGELO-style framework for childhood obesity prevention in Leeds.

MICRO ENVIRONMENT:

	Physical		Economic	
	Food	**Activity**	**Food**	**Activity**
Micro - individual	▪ Change diet composition to eat a healthy diet e.g. 5-a-day, replace sugary drinks with water ▪ Learn how to cook, and teach your children	▪ Take advantage of municipal spaces, rather than sedentary activities in the home ▪ Reduce dependence on motorized transport	▪ Take responsibility to buy healthy foods, rather than ready-made, highly processed foods or take-aways	▪ Invest in active toys for children (e.g. a football is much cheaper than a playstation)
Micro - schools	▪ Altering the composition of school meals and other foods/drinks available in schools (e.g. vending machines, tuck shops, breakfast clubs). ▪ Nutritional skills training /cooking lessons ▪ Nutritional / health education classes ▪ National School Fruit and Vegetable Scheme	▪ Introduce compulsory PE classes, preferably daily ▪ Introduce active after-school activities ▪ Walking buses ▪ Open sports facilities to members of the public to use ▪ PE education classes to encourage healthy, active lifestyles (e.g. curriculum to reduce TV viewing time) ▪ Supporting actions – monitoring, surveillance, programme evaluation	▪ Compulsory home economics classes for pupils ▪ Free school lunch (a set healthy menu) for all children	▪ Extra payments for the teachers required to run the compulsory PE classes (as may be extra to national curriculum) (ditto nutrition classes) ▪ Free after school activities
Micro - community	▪ Get collaboration of local retailers (e.g. within 0.5 mile of school) to limit sales of "unhealthy" foods during school day, including, if possible, before and after school.	▪ Develop and then encourage participation in community team exercises e.g. football league, cricket (for adults and children) ▪ Activities for adults e.g. walking clubs	▪ Establish local food cooperatives ▪ Community cooking classes (perhaps use school facilities) ▪ Fund raising / volunteers for schools nutrition classes	▪ Fund raise for schools to pay for extra school exercise classes or recruit volunteers from local community to run these classes ▪ Free / cheap activities for children e.g. youth clubs, scouts ▪ Build community spirit (increase sense of pride, reduce vandalisation)

	Political		Socio-cultural	
	Food	**Activity**	**Food**	**Activity**
Micro - individuals	▪ Eat 3 nutritionally balanced meals per day, at least one of which should be consumed as a family, at a table (no TV, no sugary drinks, no grazing).	▪ Lead by example (food & activity) – live a healthy, active lifestyle	▪ Acknowledge the need to eat less unless you exercise more ▪ Take parental responsibility to encourage healthy behaviours for the long term, rather than taking the short term easy option (whether in terms of meals or activities)	▪ Develop the "norm" of daily exercise ▪ Learn to choose physically active options to introduce more exercise into daily life e.g. stairs not lift, accept less car use ▪ Change the "default" lifestyle from sedentary to active
Micro - schools	▪ Children not allowed off school premises during the day (prevent purchase of foods outside of school) ▪ Monitor foods/drinks brought onto school premises; confiscate banned items (e.g. sugary drinks, confectionary) ▪ Set school lunch menu (2 options, both healthy) ▪ Remove or replace sugary drinks and confectionary from vending machines at schools and offer healthier choices (or remove machines all together)	▪ Discourage children being driven to school, e.g. move closest parking area to a distance away from the school, create walking bus groups ▪ Schools to develop coherent, coordinated food and activity policies	▪ Adapt school policy to encompass ethnic differences e.g. recipes in cooking classes	▪ Adapt school policy to encompass ethnic differences e.g. modest PE kit ▪ Require pupils to be physically active during the day. For example, in terms of locations of classes as well as structured PE lessons ▪ Devise strategies to assist pupils resistant to participating in normalised levels of physical activity to increase their activity levels
Micro - community	▪ Develop visible social norms that encourage people to be of "normal" weight (e.g. for adults BMI < 25)	▪ Accept the need for children to be more active and participate in school and community lead initiatives e.g. walking buses ▪ To facilitate safe neighbourhood play facilities	▪ Create a healthy eating ethos i.e. to be more discriminating about what, when and how much to eat ▪ Healthy eating education ▪ Increase acceptability of breastfeeding ▪ Transform the acceptability of working long hours (food and activity)	▪ Develop an environment where children are able to play in streets and parks ▪ Promote belief that fewer parking spaces for cars is healthy, in order to encourage more walking/cycling and increase the safety to do so

MACRO ENVIRONMENT:

	Physical		Economic	
	Food	**Activity**	**Food**	**Activity**
Macro – media	▪ Consistent and clear healthy eating messages	▪ Active playstations, i.e. which require movement	▪ Subsidise cost of healthy eating message	▪ Subsidise cost of active lifestyle messages
Macro - industry	▪ Stop increasing portion sizes / reduce portions sizes ▪ Decrease over supply of energy-dense foods ▪ Reduce amount of hidden fats and sugars in prepared foods ▪ Increase availability of nutrient dense foods, particularly for children ▪ Healthy options on restaurant menus (low fat, low sugar) ▪ Use food technology to make foods healthier e.g. ⇩ fat content of chips	▪ Urban planners to consider how to encourage more activity e.g. Locate car parks a (short) distance from the shops to increase walking; Install stairs in preference to lifts and escalators (or at least put in relatively more stairs) ▪ Industry to develop products to encourage physical activity in daily life ▪ Employers to develop more opportunities for physical activity ▪ Improve public transport to increase accessibility to leisure facilities	▪ Connect companies' profitability to healthier foods ▪ Price foods to encourage purchase of healthy options. ▪ Supermarkets provide extra points on reward schemes for purchase of healthy foods	▪ Tax breaks for companies who encourage physical activity during the working day and/or provide facilities ▪ Tax parking at work (or subsidise not driving to work)
Macro - government	▪ Nutritional restrictions on school meals / all food provided in schools ▪ Provide proper kitchens for all schools to cook wholesome, fresh, healthy food ▪ To ensure health information is widely available, easily understood and that the messages are clear, concise and consistent	▪ Local government staff to facilitate community and school based activities (food & activity) ▪ Ensure people feel safe to walk or bicycle to work/school/leisure and to play outside (e.g. policing, park wardens, protect urban spaces, more pedestrian zones/cycle paths) ▪ Provide affordable and accessible sports facilities e.g. gyms, swimming pools, skate parks ▪ Supporting actions – monitoring, programme evaluation	▪ Use subsidies/taxes, public procurement, and other fiscal measures to promote healthier foods / ⇩ demand for unhealthy foods (e.g. taxes on high fat/high sugar produces; tax companies' marketing spending based on a "healthy" rating for the food/drink item). ▪ Change prices so that they internalise currently externalised health costs. ▪ Subsidize school meals – free/subsidised, set, healthy menu for all children ▪ Remove agricultural subsidies for fat/sugar/meat/diary to stop unhealthy food mountains/surplus	▪ Develop subsidy/tax policy to promote exercise ▪ Congestion charging in city centers to indirectly encourage more walking ▪ Tax breaks / incentives for "healthy" businesses ▪ Public funding of quality physical education

	Political		Socio-cultural	
	Food	**Activity**	**Food**	**Activity**
Macro – media	▪ Self regulation to stop advertising unhealthy foods to children ▪ Healthy cooking programmes / Back to basics cooking programmes			▪ Develop social norm of a healthy, active lifestyle e.g. through soap operas
Macro - industry	▪ Develop self-regulation bodies ▪ Use Key Performance Indicators that work for health (food and exercise) ▪ Better food labeling (self regulated) ▪ Agree not to target children (particularly unhealthy foods) with advertising or other forms of marketing	▪ Harmonize consumers' health with companies' success (food and activity) ▪ Provide more opportunities for walking and cycling safely ▪ Employers to develop culture of active lifestyles	▪ Endorse wholesome food to everyone, especially low SEGs ▪ Provide truthful information for the consumer ▪ Advocate eating to excess as an exception (not the norm) ▪ Accept the spirit of encouraging healthy consumer lifestyles, rather than looking for ways round the law to ⇧ sales/profit ▪ Promote breastfeeding infants ▪ Health professionals to participate in the development of public health programmes	▪ Enable consumers to shop more often for less, thereby facilitating "person power" (not cars) (e.g. "two for one" discourages this; wheelie shopping baskets put purchases in, which also happen to be environmentally friendly)
Macro - government	▪ Regulation of food labeling (clear, consistent, messages) ▪ Regulation of food advertising to children and limit other forms of food marketing to children ▪ Restrict fast food outlets (e.g. based on outlet density, number per capita, or proximity to schools) ▪ Healthy Schools Initiatives ▪ Set nutritional guidelines for school lunches ▪ Place controls on the political contributions given by food industry ▪ Align government policies with obesity reduction targets: e.g. food industry sustainability strategy targets; PSLA/LAAs (food / activity)	▪ Employ planning functions to encourage habitual daily physical activity in the population (e.g. reduce distances between homes and shops to encourage person power; provide more opportunities for walking/cycling safely) ▪ Add PE as compulsory component of national school curriculum (ditto nutrition / cooking classes) ▪ Change planning laws to prevent overcrowding ▪ Set targets for population nutrient or food intakes, levels of physical activity/inactivity and obesity prevalence (food and activity) ▪ Set clear outcome measures to determine success of policies (food and activity)	▪ Education to ensure population have at least a minimum level of food choosing, sourcing and preparation skills ▪ Lead by example – establish high standards of public sector catering ▪ Consistent health messages ▪ Guidelines and support for maternal nutrition and to encourage breast feeding	▪ Education to develop the message of the importance of daily activity, and how to do it

Ward Name	No. of SOAs	Minimum Deprivation	Maximum Deprivation	Mean Deprivation
Wetherby	18	2.0	21.0	9.1
Horsforth	14	4.0	28.0	11.2
Otley and Wharfedale	16	5.0	29.0	11.7
Aireborough	17	5.0	32.0	14.5
Garforth and Swilling	16	6.0	26.0	14.7
Cookridge	14	4.0	38.0	14.9
Halton	15	7.0	22.0	15.3
Roundhay	15	5.0	28.0	15.4
Pudsey North	15	6.0	32.0	15.9
Barwick and Kippax	16	7.0	30.0	15.9
North	15	6.0	51.0	16.1
Moortown	14	9.0	48.0	18.9
Morley North	17	8.0	34.0	19.4
Weetwood	14	9.0	42.0	19.7
Rothwell	14	11.0	36.0	20.4
Pudsey South	15	10.0	39.0	22.2
Morley South	19	6.0	39.0	23.2
Headingly	17	15.0	40.0	24.8
Whinmoor	11	10.0	50.0	30.0
Wortley	15	15.0	54.0	31.3
Kirkstall	13	20.0	59.0	33.2
Bramley	15	16.0	67.0	35.7
Middleton	14	14.0	69.0	36.9
Armley	15	20.0	58.0	37.5
Beeston	11	23.0	72.0	39.3
University	14	22.0	73.0	42.9
Chapel Allerton	12	13.0	78.0	45.5
Hunslet	11	27.0	67.0	48.8
Richmond Hill	12	30.0	70.0	53.1
Harehills	14	19.0	76.0	53.9
Burmantofts	12	33.0	70.0	54.3
Seacroft	12	20.0	77.0	55.3
City and Holbeck	14	32.0	79.0	56.0

Table 8.5. Ranking of deprivation score (Index of Deprivation, 2004) for Leeds' wards, from most affluent to most deprived. The three grey shaded wards represent the thre case study areas.

OAC Supergroup	City and Holbeck	Morley South	Wetherby
1 – blue collar communities	6	20	11
2 – city living	6	1	0
3 – countryside	0	2	21
4 – prospering suburbs	0	22	45
5 – constrained by circumstances	33	15	5
6 – typical traits	15	39	18
7 – multicultural	40	0	0

Table 8.6. Details of the Open Area Classification (OAC) Super-group for each of the three wards selected for case study (based on the OAC data for actual OAs in each ward). Each number represents the percentage of population for each ward that are classified as each Super-group.

Figure 8.2. Map of location of the three case study wards.

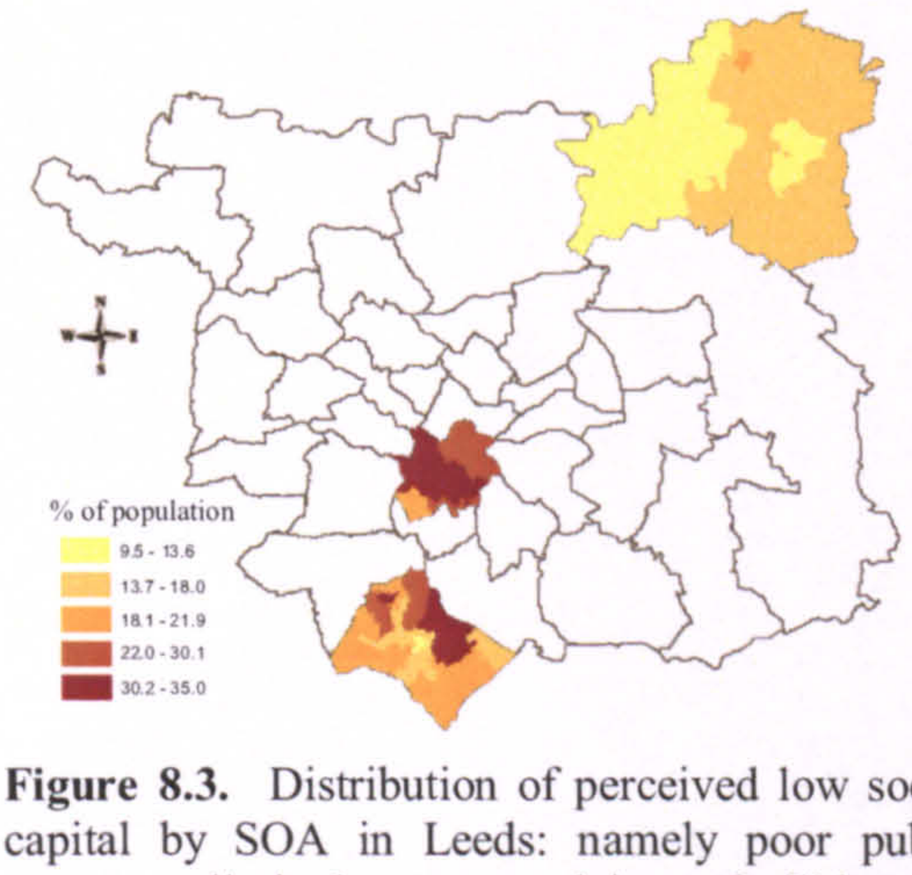


Figure 8.3. Distribution of perceived low social capital by SOA in Leeds: namely poor public transport, limited access to leisure facilities and supermarkets, problems with teenagers hanging around and vandalism.

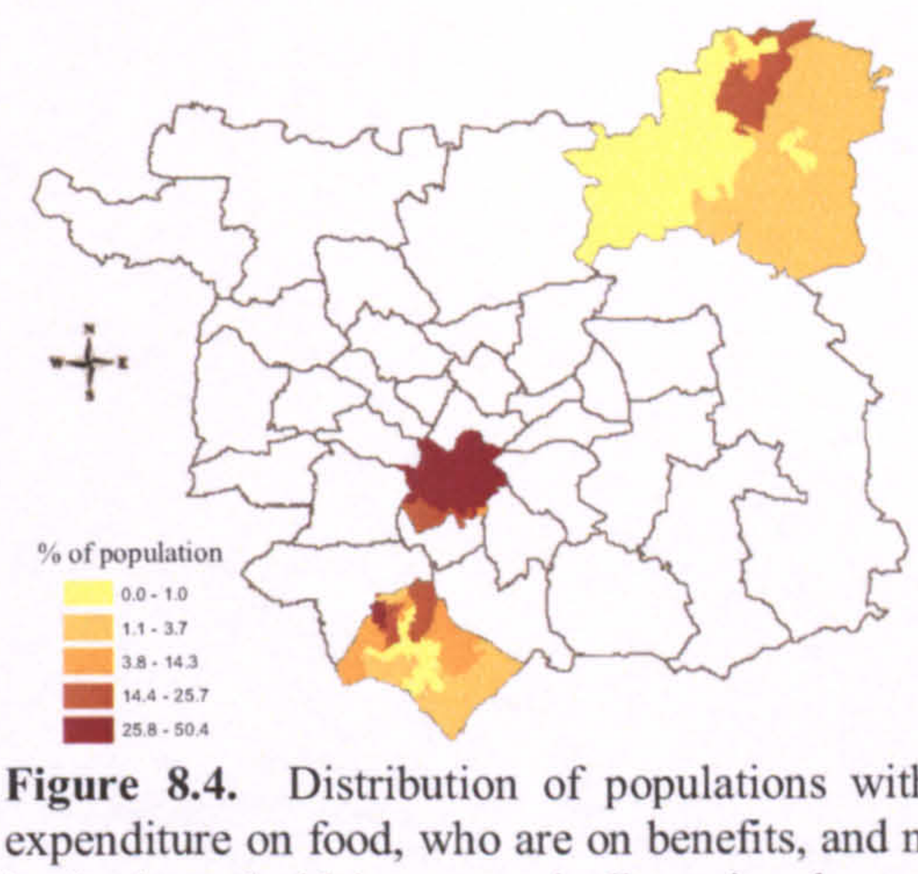


Figure 8.4. Distribution of populations with low expenditure on food, who are on benefits, and no one in the household is economically active, have a low income, no access to a car, and who do not own their home, by SOA in Leeds.

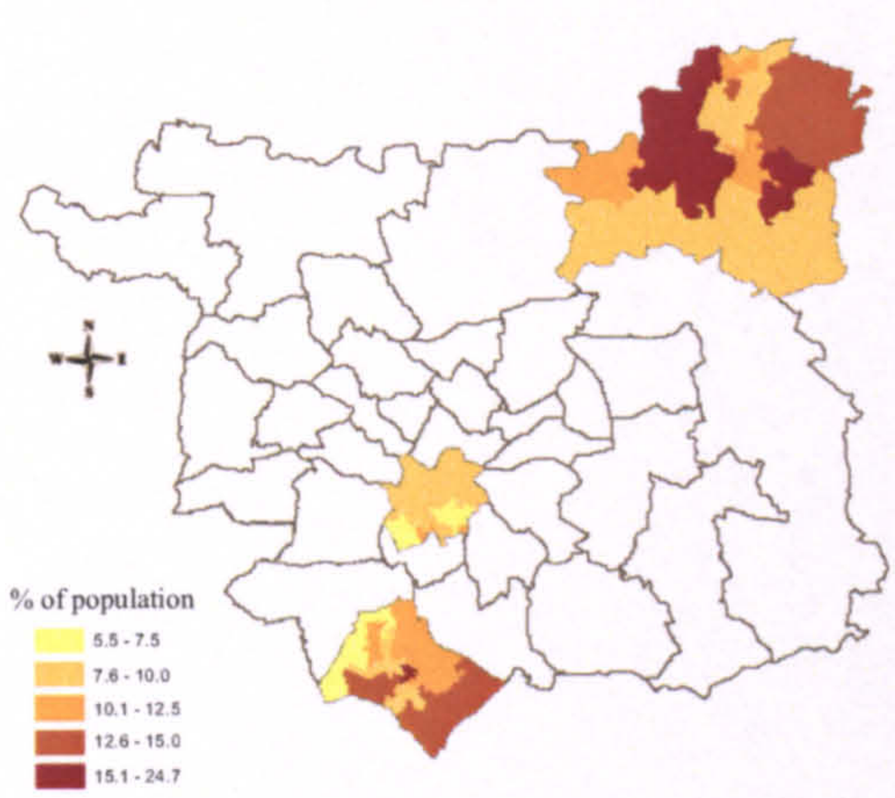


Figure 8.5. Distribution of obesogenic behaviours of children aged 7-15 years by SOA in Leeds: that is, with low daily fruit and vegetable consumption and low physical activity levels.

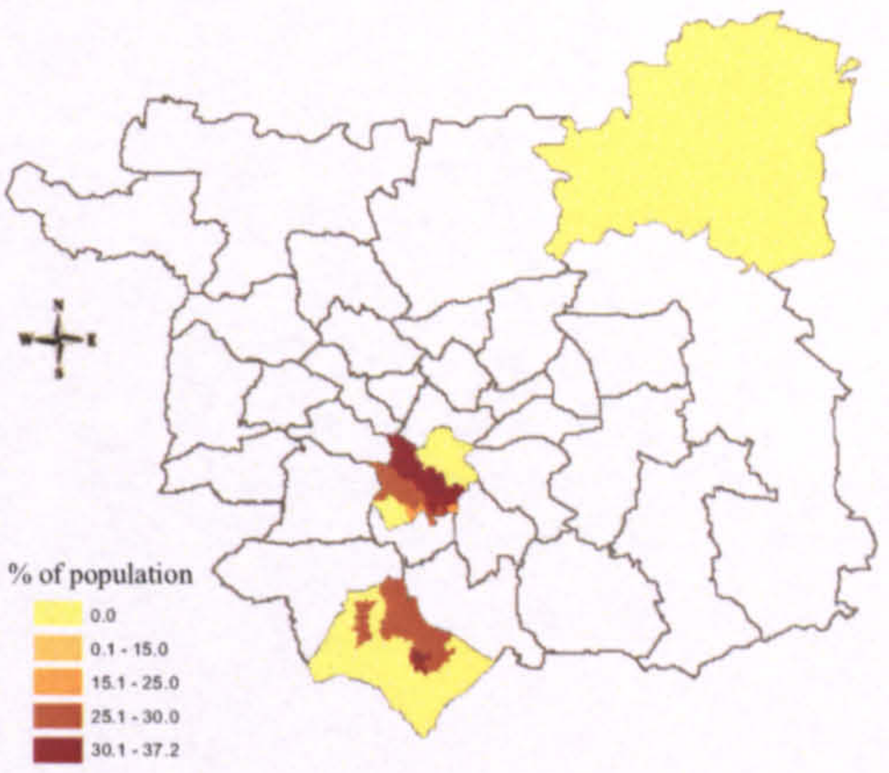


Figure 8.6. Distribution of people who perceive the safety of their neighbourhood as low and have a low income, who live in a highly deprived area, by SOA in Leeds.

8.4.3 Obesity variance across each area

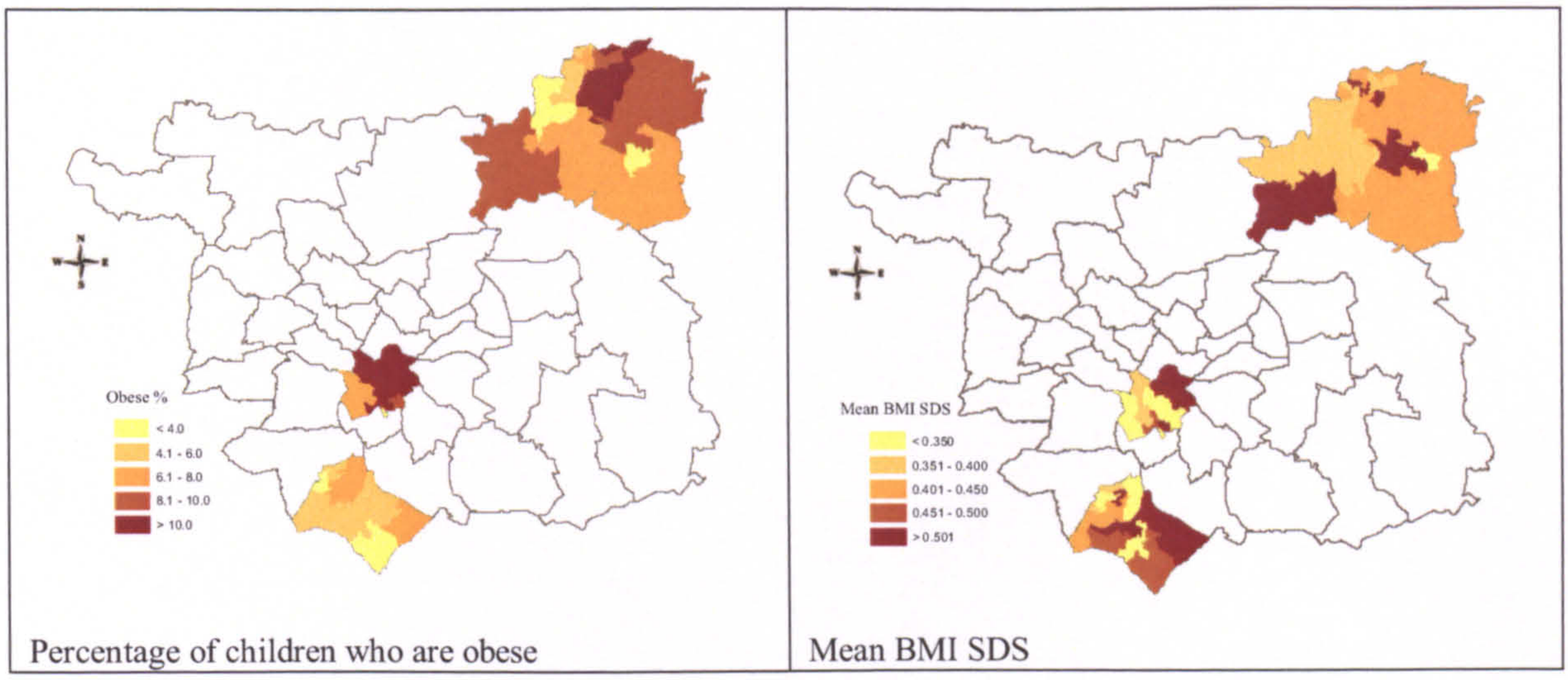
The prevalence of obesity is greatest in CH (10%) and lowest in MS (6%). Yet CH had the lowest mean BMI SDS (0.46) and Wetherby the highest (0.52) (see Table 8.7). The percentage of obese children is not uniform either within or across the three study areas (see Figure 8.7). This shows that CH has a high percentage of obese children, W has some areas with a high prevalence and MS has a much lower prevalence. To put it another way, both CH and W have higher prevalence of obesity than the global average figure for Leeds (7.6%), and MS has lower prevalence. By way of an alternative perspective using mean BMI SDS, CH is the area with SOAs with the lowest mean BMISDS. All three case study areas show SOAs with figures above and below the global Leeds mean BMI SDS (0.422).

We also have data on obesity in primary schools in each of the case study areas from the Trends study. The school in W shows obesity levels that we would expect given the deprivation and ethnicity mix, as does one school in CH. However the other school in CH and one school in MS display higher childhood obesity than would be expected (given the deprivation and ethnicity of the children) (see Figure 8.8), suggesting that these schools may be having a detrimental affect on the obesity of their pupils. It may be appropriate that a school is used as an environment for an intervention.

Hot spots were found in both deprived and affluent areas (in chapter 4), suggesting either a spread of obesity across socio-economic groups and/or something special about those areas which affects the aetiology of obesity. CH was in a hot spot for obesity prevalence (see Figure 4.13 in chapter 4). Children in this area are 1.5x more likely to be obese than children living in other parts of Leeds ($p = 0.001$). After adjusting for deprivation and OAC (see Figure 4.19 in chapter 4) it is apparent that there is a cold spot (cold spot 5) in CH and a hot spot (hot spot 4) in W.

Ward	Obese%	Minimum BMI SDS	Maximum BMI SDS	Mean BMI SDS
Wetherby	8.0	-2.440	3.566	0.525
Morley South	5.7	-3.480	3.849	0.467
City and Holbeck	10.0	-3.047	4.484	0.456

Table 8.7. Obesity data for the case study wards based on the individual level data. The percentage obese and mean BMI SDS figures were smoothed using Empirical Bayes (to minimise small number problems).



Percentage of children who are obese | Mean BMI SDS

Figure 8.7. Map of percentage of measured children who are obese (on left hand side) and of mean BMI SDS of measured children (smoothed), in each of the case study areas in Leeds. The global figures for percentage of obese children (7.6%) and mean BMI SDS (0.422) both lie within the central scale choice.

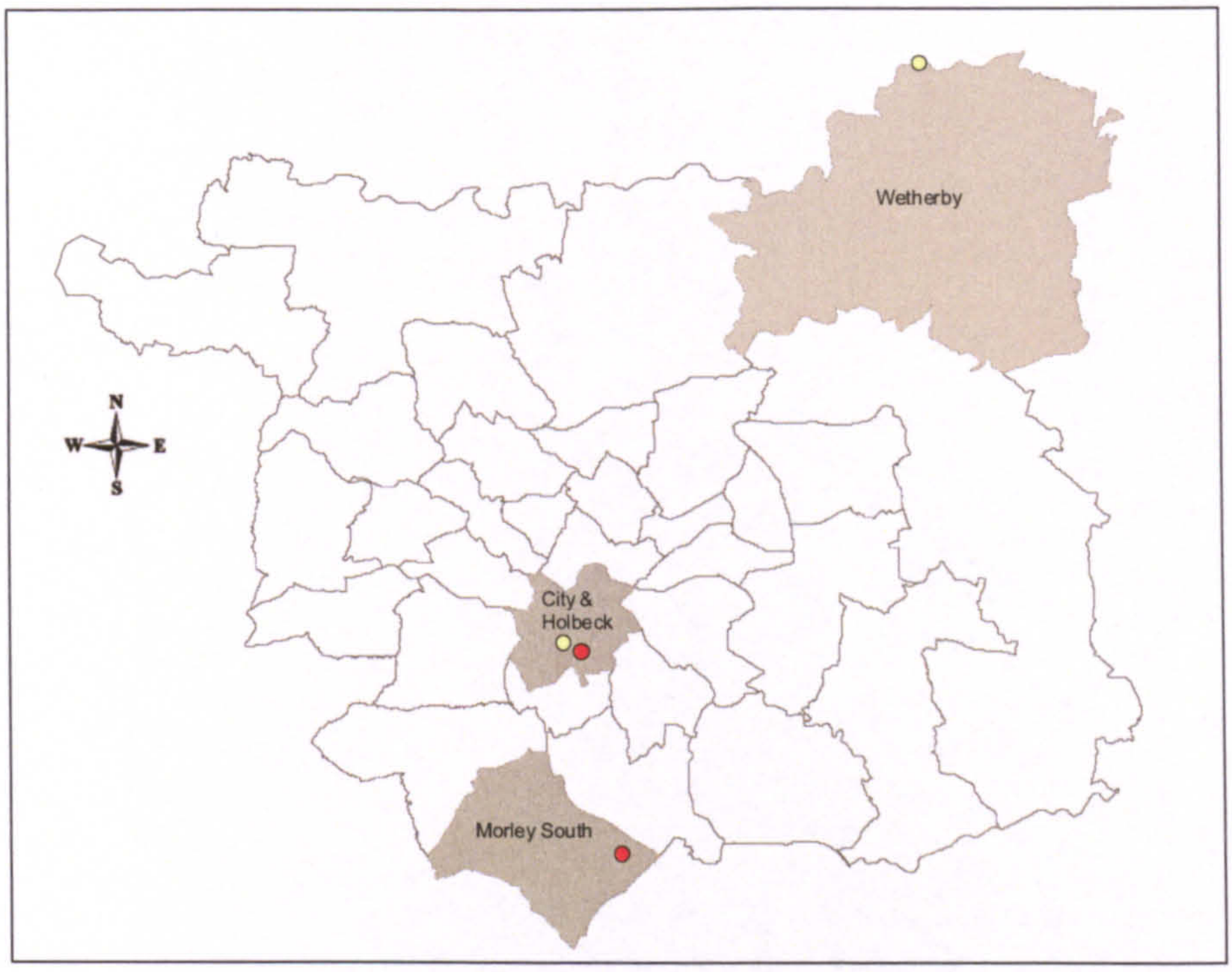


Figure 8.8. Map of location of primary schools, where children were measured as part of the Trends study. The red circles indicate the location of schools with higher obesity than expected given the deprivation and ethnicity mix, and the green circles show the location of schools where obesity is in line with expectations.

8.4.4 Childhood obesity and deprivation / OAC super group

There are high levels of obesity in each of the three case study areas, despite the differences in deprivation levels. In chapter 4 it was found that, based on mean BMI SDS for each super group, children in Leeds living in areas classified as OAC Super groups 1 (blue collar communities), 3 (countryside), and 5 (constrained by circumstances) were more obese than children living in OAC 2 (city living), 4 (prospering suburbs), and 7 (multicultural). This was not true of this subset (see Table 8.8). Super groups 4 and 7 still have a lower than average BMI SDS, and super groups 1 and 3 have higher than average BMI SDS. However in this sub-set, super group 2 has much higher than average obesity (but N is small). Both super groups 5 and 6 are close to the total average.

Super group	Leeds mean BMISDS	Leeds N	Case Mean BMISDS	Case N
1	0.488	6637	0.552	559
2	0.379	1002	1.215	14
3	0.492	607	0.555	196
4	0.392	7989	0.431	855
5	0.478	5655	0.462	357
6	0.423	7172	0.481	642
7	0.366	4532	0.333	495
Total	0.422	33594	0.462	3118

Table 8.8. A summary of the seven OAC super groups by mean BMI SDS firstly using the obesity dataset for the whole of Leeds and secondly using the sub-data for the case study areas only (based on individual-level raw data).

8.4.5 Childhood obesity and the simulated covariates

The twelve variables with the important relationships with childhood obesity (see chapter 7) were: quality of public transport, access to leisure facilities and supermarkets, sedentary behaviour of children and adults, fruit and vegetable consumption, buying school meals, expenditure on food, problems with teenagers hanging around, television ownership, having internet access and activity levels of children (as opposed to sedentary behaviour).

The results of the global and local GWR analysis (described in Chapter 7) showed the following (see Appendix G for definitions of how the covariates were assessed). Areas that reported perceived bad public transportation had reduced risk of obesity, lower deprivation scores and high income. The risk of obesity was increased if there was a problem with accessing supermarkets or leisure facilities. In the relationship between childhood obesity and (children's or adults') physical exercise, a dose response was evident (i.e. more exercise, less obesity), with sedentary behaviour having a stronger global effect than activity. Fruit and vegetable intake was negatively associated with obesity, and there appeared to be a threshold effect before the protective effect was realised. Buying school meals was negatively correlated with obesity.

Expenditure on food was shown to be a significant predictor of childhood obesity (inverse relationship) and the data suggested that the health consequences arise due to insufficient money to spend on food, rather than parents unwisely spending what money they have, and conversely that higher food expenditure is not about buying more food (more energy intake), but rather about buying more expensive, healthy food such as fruit and vegetables. There was increased risk of obesity if there was a problem with teenagers hanging around in the neighbourhood – potentially a proxy for neighbourhood safety. It was found that the ownership of more than one television per household was positively correlated with the percentage of obese children and access to the internet was shown to be negatively correlated.

In order to determine whether the relationships between these covariates and childhood obesity differ in the three case study areas, the relationships in each ward were considered separately. These are summarised in Table 8.9.

Wetherby

The consumption of no daily fruit and vegetables and the perception that leisure facilities in the neighbourhood are poor are strongly, positively associated with childhood obesity, without exception (Figure 8.9). A slightly weaker, but nevertheless consistent, positive relationship also exists if there is a perception that supermarkets are difficult to access. This may be because of the rural nature of this ward. Other variables that show areas of strong obesogenic relationship with childhood obesity are the perception of good public transport, children and adults with high levels of sedentary behaviour, and a perception of a problem with teenagers hanging around the neighbourhood. Variables that reduce the risk of childhood obesity in Wetherby (in decreasing order of strength) are children being at least moderately physically active (i.e. active for at least three hours per week), and children purchasing school meals. The relationship between childhood obesity and food expenditure, television ownership and internet access is weak across most of Wetherby.

City & Holbeck

Figure 8.10 shows that the parameter estimates for City & Holbeck are generally much weaker than for Wetherby. That is, for a unit change in a covariate, there is a smaller change (either positively or negatively) in childhood obesity in CH than in W. That aside, it shows that variables with a largely positive (obesogenic) relationship with childhood obesity (in decreasing strength order) include, the perception that supermarkets are difficult to access, low expenditure on food, no internet access, more than one television per household, no daily consumption of fruit and vegetables, the perception that leisure facilities in the area are poor, sedentary adult behaviour, and the perception that there is a problem with teenagers hanging around in the neighbourhood. Variables that seem to reduce the risk of childhood obesity in CH are the

purchasing of school meals, and sedentary children, the latter of which seems contrary to popular biological explanations of obesity and is presumably a marker for another behaviour/determinant.

Morley South

In Morley South the pattern of covariate relationships with childhood obesity differ once again (see Figure 8.11). Children undertaking no weekly physical activity (sedentary) show an across the board positive relationship (obesogenic) with childhood obesity, although none of the parameter estimates are above 0.5 (so a 1 unit change in sedentary behaviour of children will have a less than 0.5 unit change in obesity). The next most obesogenic factors in Morley South (in descending order) are: supermarkets perceived as difficult to access, more than one household television, no daily fruit and vegetables being consumed, perceived poor access to leisure facilities, perceived good public transport, a perception of a neighbourhood problem with teenagers hanging around, and sedentary adults. Conversely, purchasing school meals and moderately active children reduce the risk of obesity.

Area	Variables that increase the risk of obesity	Variables that reduce the risk of obesity
W	No daily fruit and vegetables Leisure facilities are poor Supermarkets perceived as difficult to access Good public transport Children with high levels of sedentary behaviour Adults with high levels of sedentary behaviour Problem with teenagers hanging around	Children being at least moderately physically active Children purchasing school meals
CH	Supermarkets perceived as difficult to access Low expenditure on food No internet access More than one television per household No daily consumption of fruit and vegetables Leisure facilities are poor Sedentary adult behaviour Problem with teenagers hanging around	Purchasing of school meals Sedentary children (!)
MS	Sedentary child behaviour Supermarkets perceived as difficult to access More than one household television No daily fruit and vegetables consumed Poor access to leisure facilities Good public transport Problem with teenagers hanging around Sedentary adults	Purchasing school meals Moderately active children

Table 8.9. Summary of variables that are associated with higher or lower obesity in each case study area

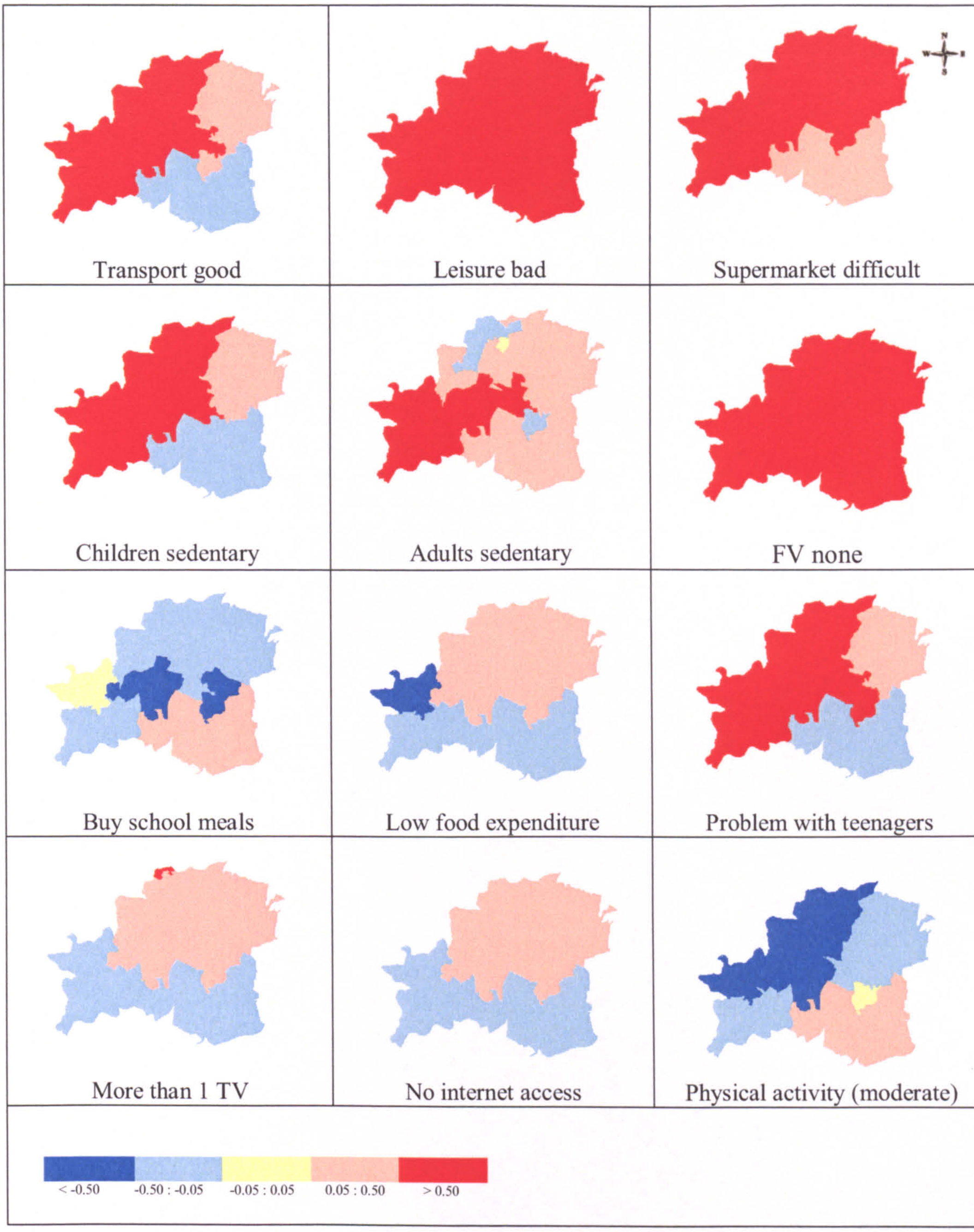


Figure 8.9. Map of covariates' relationships with childhood obesity for Wetherby, the most affluent ward in Leeds. The shading indicates the value of the parameter coefficient for each SOA (no boundaries shown). The overall colour of the map indicates the direction of the overall relationship (red: increases risk; blue: reduces risk of childhood obesity). The darkest blue and red areas are the areas with the strongest local relationship with obesity. In order to enable comparison between maps, all were prepared using the same five-category manual scale.

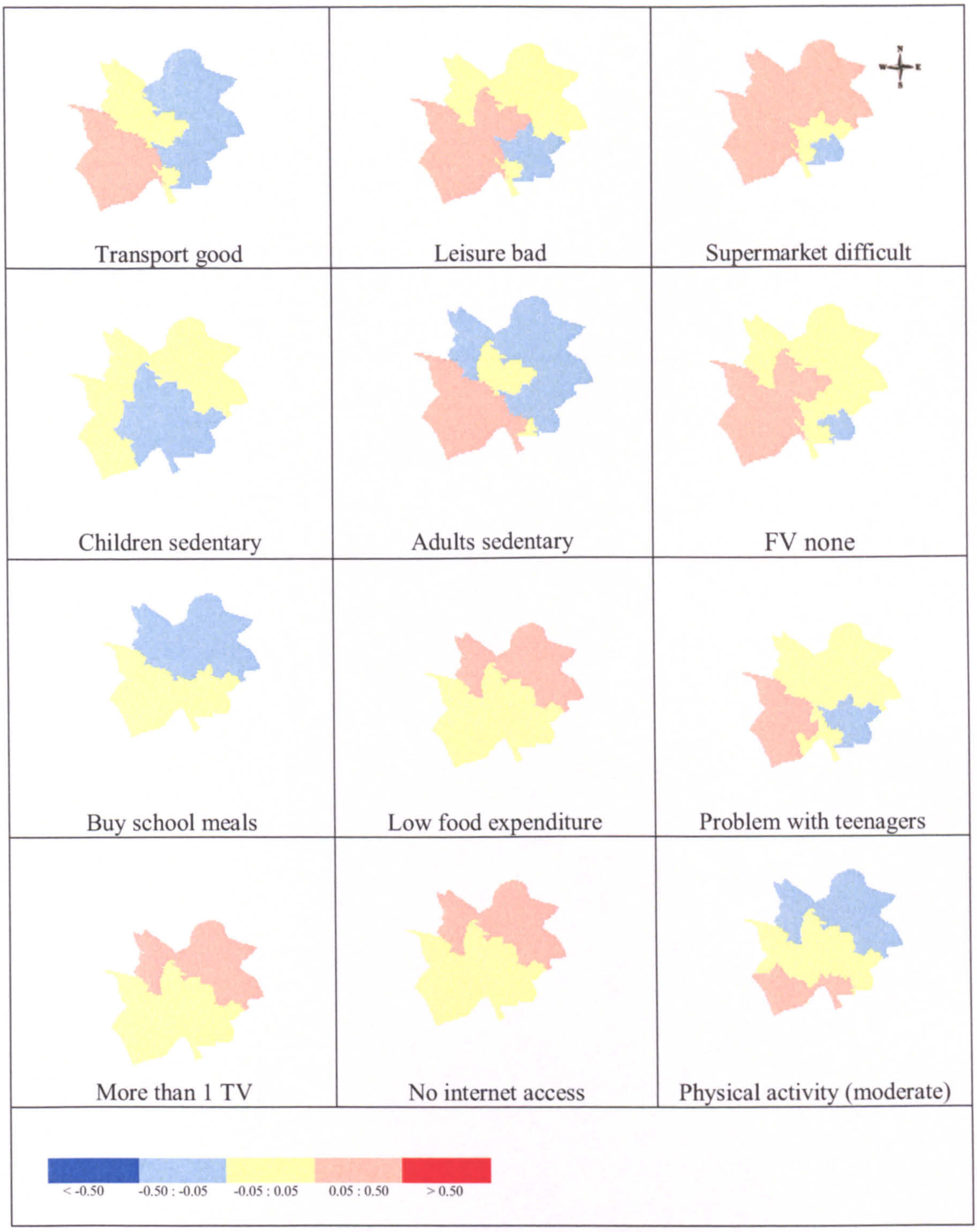


Figure 8.10. Map of covariates' relationships with childhood obesity for City & Holbeck, the most deprived ward in Leeds. The shading indicates the value of the parameter coefficient for each SOA (no boundaries shown). The overall colour of the map indicates the direction of the overall relationship (red: increases risk; blue: reduces risk of childhood obesity). The darkest blue and red areas are the areas with the strongest local relationship with obesity. In order to enable comparison between maps, all were prepared using the same five-category manual scale.

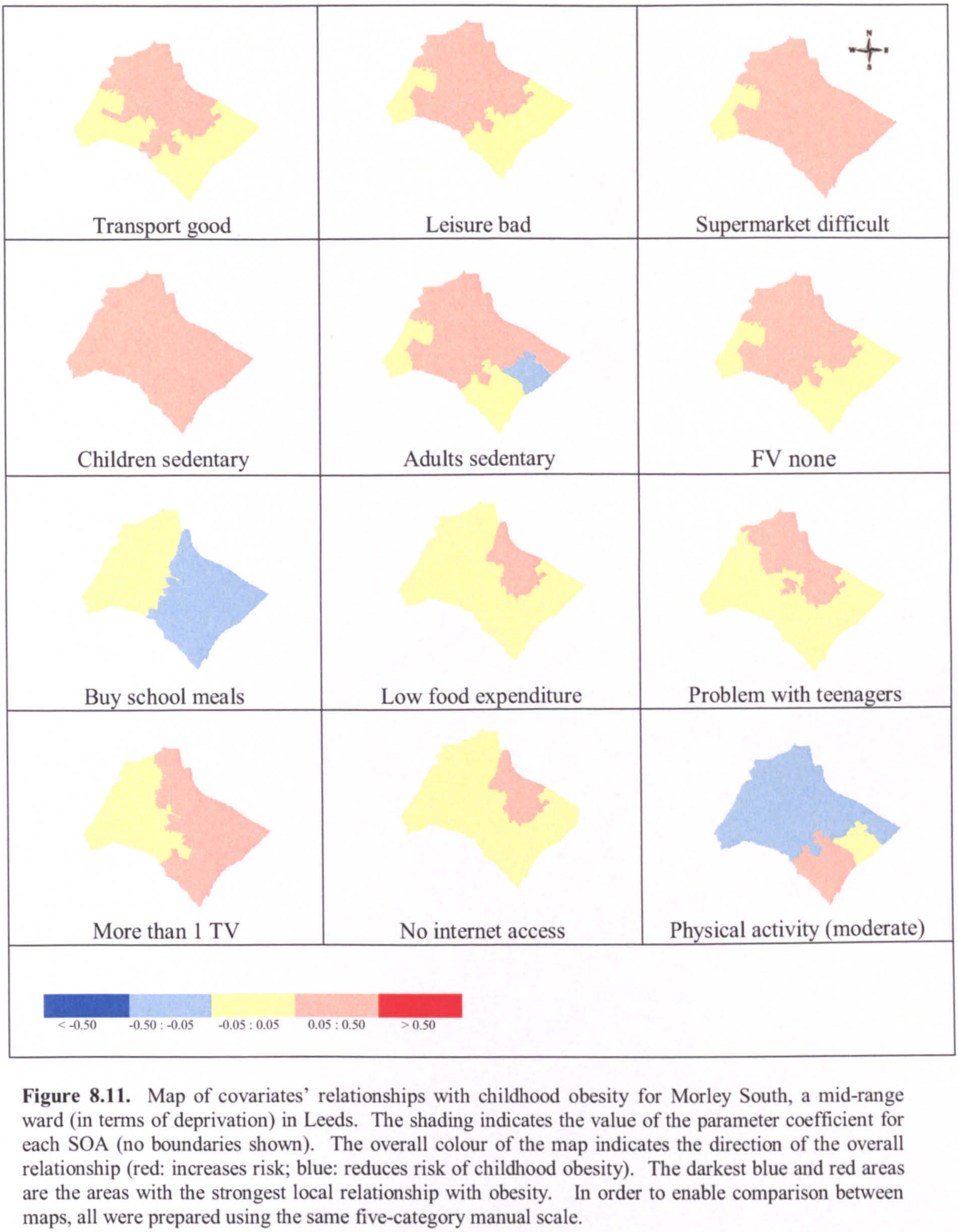


Figure 8.11. Map of covariates' relationships with childhood obesity for Morley South, a mid-range ward (in terms of deprivation) in Leeds. The shading indicates the value of the parameter coefficient for each SOA (no boundaries shown). The overall colour of the map indicates the direction of the overall relationship (red: increases risk; blue: reduces risk of childhood obesity). The darkest blue and red areas are the areas with the strongest local relationship with obesity. In order to enable comparison between maps, all were prepared using the same five-category manual scale.

8.4.6 Tailoring the framework to prevent obesity in Leeds

The analysis has shown that the key determinants of childhood obesity in each of the three case study wards in Leeds differ. Accordingly, the "micro" aspects of the ANGELO-style childhood obesity policy developed earlier in this chapter (see section 8.4.1) can be modified to take into account the nuances of each neighbourhood, both by the emphasis (in focus or funding) given to certain policies (the most important ones will vary), as well as how an intervention is implemented. Thus we can establish a location-specific policy for Leeds.

Across all three case study areas children purchasing school meals was a factor that reduced the risk of childhood obesity. Accordingly, the Leeds childhood obesity prevention policy should focus on the interventions that affect how healthy school meals are and to encourage children to eat them. Therefore, it is a priority that Leeds parents/schools/government ensure that meals provided in ALL schools for all age groups meet nutritional guidelines set by the School Food Trust to ensure they are healthy. This should be true for the first and the last child served. Anecdotal evidence suggests that some schools comply with healthy eating requirements by offering a small number of portions of the healthy option with large amounts of the unhealthy, popular, option (e.g. chips). A set healthy menu, provided free or subsidised, may be the best way to ensure all children eat a healthy meal at lunchtime.

Child activity as a compulsory element of the curriculum / school programme is also essential if school children are to be fit and healthy (and not obese) (Connelly et al, 2007). This was corroborated in the local analysis, with both Wetherby and Morley South showing a reduced risk of childhood obesity with more active children (data on compulsory versus voluntary was not available, but would be an interesting local analysis). Surprisingly City & Holbeck showed that sedentary behaviour in children was associated with a reduced prevalence of childhood obesity in this ward. This is not consistent with the biological explanation of energy balance causing obesity, nor with the experience of other authors and in the other areas in Leeds.

In relation to the obesogenic covariates, in Wetherby policy needs to focus on encouraging children (and adults) to eat more fruit and vegetables, improve access to leisure facilities and to supermarkets. In relation to fruit and vegetables, this may be by education to learn the importance of eating fruit and vegetables (although surely everyone has now heard of "5-a-day"?), as well as how to cook tasty, fresh, affordable meals that will appeal to children. Children must also be taught how to prepare and cook nutritionally balanced meals. Council housing needs to ensure that its tenants have adequate cooking facilities. Providing free fruit to school children (the School Fruit and Vegetable Scheme) is also a means to increase children's fruit and vegetable intake during the early years of schooling. Communities can also help to

increase fruit and vegetable intake, perhaps by establishing food cooperatives, or local farmers could provide fruit and vegetable boxes (obviously at a price, but Wetherby is an affluent area). Also there is a need to provide better leisure facilities – this may incorporate local councils and local transport companies improving access to the facilities, for example with regular, affordable buses, or by providing secure places to park bikes. Car parking is perhaps a double edged sword, being an energy-saving transportation method but if it enables people to go somewhere where they will enjoy being physically active then perhaps it is worthwhile. Also leisure facilities need to be clean, safe, and welcoming environments, which is down to the facility staff, local council and the community. Similarly issues arise in regard to supermarket access. Wetherby is a sparsely populated, largely rural area and it is likely that for many people the nearest supermarket is over ten miles away (which may be classified as difficult to access – depending how a question is worded). With a car this is not a problem, but without one then it is the bus or a taxi, so either time consuming or expensive.

For City & Holbeck, a highly deprived, inner-city area, the focus needs to be on supermarket access, low expenditure on food, television viewing and fruit and vegetable consumption. It is highly likely that internet access' association with childhood obesity is a marker for household income. Supermarket access can be improved by providing more local, nearby shops that sell healthy, affordable foods, as without access to a car or good public transport links out-of-town supermarkets can be inconvenient. Interventions in school (linked to home and the community) can help to reduce the amount of television children watch, as can parents monitoring the amount and removing televisions from children's bedrooms (Wiecha et al, 2001). However it is not as simple as stop watching television. Alternatives need to be provided, which is where support from the school and community can help, with increased after school clubs and community clubs (e.g. Scouts, youth clubs) as well as weekend activities in the neighbourhood. It is beyond the scope of this thesis to discuss the benefit system and incomes for unemployed families. Suffice to say that a third of people in CH are described as "constrained by circumstances", with higher than average unemployment, lone parent households and population densities. Nevertheless this will be a factor influencing the importance of food expenditure on this population. Energy dense foods are cheap and a healthy diet is more expensive (Baratt, 1997; Cade et al, 1999). It is important that local council and the community focus on educating parents (and children, due to the force of "pester power") regarding the importance of buying healthy foods and how this can be done in an affordable way if cooking skills are honed (rather than simply heating up highly processed, pre-prepared foods), as well as providing the ability to improve cooking skills (there is no point in telling people they need to be able to cook, if support to facilitate this is then not provided). Community cooking lessons and recipes need to take account of the demographics: given the high multi-cultural population (40%) it is important that ethnic meals are included in this intervention. Fruit and vegetable

consumption can be increased in line with the Wetherby suggestions, albeit with more ethnic and value-for-money options, to take account of the different demographics.

Morley South is similar to the previous two case studies in all aspects apart from sedentary child behaviour, which is highly associated with increased risk of childhood obesity. Reducing media time, such as television viewing, replacing this with more active home, school and neighbourhood activities, as well as increased (some?) compulsory exercise at schools, will all serve to limit this risk factor and so, hopefully, obesity in children.

8.4.7 Comparison to Leeds policy

How does this suggested childhood obesity prevention policy compare to actual policy in Leeds?

The lack of specific LPSA targets to reduce obesity prevalence to an agreed level suggests that this is perhaps too hard to achieve. The target to improve school children's health has two components: tackling obesogenic behaviours & healthy schools standard. However there are no clear "Outcomes" in the LAA that mention directly tackling obesity behaviours, so it is unclear how this LPSA target will be measured or achieved. Further, the emphasis given to tackling obesogenic behaviours is very low (i.e. a value of only half compared with a value of two for reducing youth offending). The target to improve physical activity does not have any target values, so measurement and achievement of this target is doubtful. This implies it is not a priority for the government. The target to clean up the city, which may improve social capital, is targeted and has clear "Outcomes" in the LAA.

The Leeds LAA does not include any specific targets regarding obesity prevalence. Also the focus on the 31 most deprived SOAs of Leeds includes six of the fourteen SOAs in CH ward, but would exclude Wetherby and Morley South wards – yet it has been shown that these areas also have obesity problems. It also focuses on children in Year 9, but all children can benefit from interventions to reduce obesity.

The Outcomes in the LAA that may be associated with obesity are as follows (taking each of the four themes in turn). To encourage children to walk and cycle to school, which is in line with the framework developed. To ensure more children do at least two hours of physical activity at school per week (which still seems very low if it is to have an effect – this would be classified as "low activity" in the analysis in this thesis). To increase the number of schools that attain the National Healthy Schools Standard, which is likely to be beneficial for the schools' pupils, plus the targets are high – from 50% of schools in 2007 to 95% by the end of 2009. To be more

beneficial the healthy eating and physical activity components in the schools need to link back to the home and community. To encourage more schools to offer extended schools services (but the targets are still pretty low – reaching only 33% by September 2008). Plus these services may or may not be beneficial from an obesity perspective depending upon what services are offered exactly. The healthy schools initiative and incorporating minimum levels of physical activity as a compulsory component of the school curriculum (although the current target levels are too low) are important targets. These cover many of the schools micro-environment interventions, such as altering the composition of school meals and other foods/drinks available in schools, nutritional and activity education classes, "walking buses", and active after-school activities. The bulk of the other outcomes in the Children and Young People theme are to do with grade achievement, which suggests the emphasis is misplaced. It is important that schools do not focus solely on grade attainment, particularly at the exclusion of all else (e.g. exam classes rather than PE lessons), as to do so teaches children that a rounded, balanced lifestyle is not important, and perhaps that a healthy diet and regular exercise is not important.

The key Healthier Communities outcome was to reduce the risk factors for heart disease, stroke and related diseases (i.e. smoking, diet and physical activity). But there was no mention of how diet would be changed, nor any targets to measure any change. Further the targets to increase physical activity were tiny – just 1% increases each year. So of the 545,427 adults that live in Leeds that would be an increase in physical activity levels in approximately 5-6000 people per year.

The indicators for Safer Communities centered largely on changing people's perception of anti-social behaviour problems in their neighbourhood. These indicators may affect a person's social capital or feelings of safety in their community. But the targets are simply for reductions/increases (as appropriate) each year, so even a 0.01% change would be classified as a success, yet really what impact would this have? Further, the baseline figures for this are not stratified across Leeds, with a simple average showing in the report. Thus it is unclear how changes in micro-areas (for example the 31 highly deprived SOAs the LAA highlights as being a priority for change) will be accurately monitored. Obviously an average decrease in a percentage for the whole of Leeds may be masking an increase in any or all of these target SOAs.

The Economic Development outcomes were related to physical infrastructure. Firstly to increase the percentage of non-car journeys – but the targets for this are miniscule with a 1.3% increase over five years. Accordingly this is likely to have minimal impact of physical activity levels. Also to increase the number of visitors/users to the City Council's sports facilities, yet

the target is a decrease in the number of users. The rationale is that facilities are expected to close, but this is irrelevant. If the outcome is to increase, then the target needs to be to increase – so this is an indicator that is targeted to not meet the lexis of the outcome. Similarly the number of visitors to Leeds parks does not have any target figures and is not being measured after baseline, so it is unclear how the council will monitor whether this goal has been met or not. Plus the number of users is greater than the population of Leeds (approximately 720,000 people) so it includes people from surrounding areas and potentially, depending how it is measured, includes individuals more than once if they make multiple visits. Thus an increase could be seen because the existing users become more active, not because the population of Leeds is generally more active. Further who is to say a visitor is "active"? They could be visiting to watch a sport or simply have an ice cream.

It is clear that this policy is not as comprehensive as that suggested in the framework in this chapter. Obviously time and money are going to be practical limitations in the "real world", but childhood obesity is a major problem and targets for obesity prevalence in Leeds should be agreed and prioritized in the top twelve LPSAs. There are too many missing policies to illustrate one by one, but the following illustrates the point. Whilst improving the safety/cleanliness of a neighbourhood is being addressed in Leeds, the policies do not address other local community level factors, such as to encourage participation in community team games and activities, to establish food coops, to encourage activities for children (youth clubs, scouts), creating a healthy eating ethos, increasing the acceptability of breast feeding, and reducing the long-working hours culture. The macro environment is hardly addressed at all, with the outcomes for public transport, cycling/walking, healthy schools initiatives, and physical activity curriculum being the only areas that are covered. Perhaps this is appropriate in a local government's policy, as these areas are more suitably addressed by regulation by central government. But some aspects could be addressed locally. For example, urban planners considering how to affect more activity, local planning laws to prevent overcrowding and restrict fast food outlets close to schools, provide reasonably priced, healthy foods in public sector facilities, encourage local restaurants to provide healthy options and local employers to develop more opportunities for physical activity, and to ensure health information is widely available. So more could be done.

8.5 Discussion

The prevalence of obesity is greatest in CH (10%) and lowest in MS (6%). The "deprivation theory" would predict that Wetherby had the lowest prevalence, which is not the case. The obesity data for each case study area suggests that as well as many children with a high enough BMI SDS to be classified as obese, CH also has many children with low BMI SDS, thereby pulling the average mean BMI SDS down. Whereas in W it seems that there are more children with a high BMI SDS, but not quite high enough to be classified as obese. Arguably this could be storing up more problems for the future if the children continue to gain weight, rather than gaining height and growing out of the problem. It has been suggested that higher BMI SDS in more deprived areas are due to shorter stature (White et al, 1995).

The significant variation seen in the relationships of the covariates with childhood obesity across Leeds suggests that different interventions to reduce the prevalence of childhood obesity may be more effective in different areas. It is necessary for public health officials to determine what these local factors are in order to be able to tailor solutions to each population's requirements for maximum effectiveness. One size does not fit all. Further messages should be modified to suit individuals at different stages of decision making: from the "pre-contemplation individual" who has not yet decided to change their lifestyle to reduce the risks of obesity so needs motivational guidance, to the "action individual" who has decided to change and needs advice on how to do that (Jebb et al, 2003). Given the complex interplay between different covariates, it is likely that the most effective interventions to prevent childhood obesity will be a policy incorporating interventions targeting multiple determinants.

The individualistic model is outdated. Few obesity experts accept that individuals "choose" to become overweight – the question of what constitutes free choice being an important one. Only individuals with greater amounts of "social power" are to be expected to be capable of consistently making healthy choices (Lang & Rayner, 2007). How can children choose? Accordingly the solution to the ideological framework within which to sit an obesity prevention policy should not focus on individual level interventions. The optimum model and most realistic way forward is to use a population level approach, changing obesogenic environments towards becoming leptogenic, or at least obeso-resistant, environments.

There are clear benefits to taking a public health perspective to childhood obesity prevention policy. Universal prevention, aiming at everyone in the community, seeks to stabilize or reduce the mean rates of obesity in the population (Muller et al, 2001b). Selective interventions that focus on high-risk children do help to target finite resources, but can lead to stigmatisation, especially if individual children are singled out for action (Lobstein et al, 2004). Furthermore,

as genetic studies show that most children are at risk of weight gain, population focused obesity prevention strategies will benefit the health of all children not just those who are (or would be) obese. The real strength of this approach is that significant population benefits can result from even fairly small effects if a large number of people are exposed to that environment (Swinburn & Egger, 2002).

However, public health policies for obesity prevention are not without difficulties: at the end of the day "all" they have to do is persuade the population to eat a healthy, balanced diet and to be physically active. However obesity is a chronic condition due to positive energy balance over time, probably many years (Jebb et al, 2003). As such, the impact of any prevention policy is likely to be equally gradual (Lang & Rayner, 2007), which limits the visibility and measurability of its success (or otherwise) (Heller & Page, 2002), which likely relies on average measures of population health indicators, such as BMI. Further, the complex aetiology of obesity means the policy message is also going to be complex. It is not as simple as "stop smoking" as we all have to eat – and it is both what and how much we eat that is important (Jebb et al, 2003) and similarly for physical activity. This complexity increases the likelihood that the message given is interpreted as contradictory in nature and also the perception that authorities cannot agree on what the best action to take is (Jebb et al, 2003). Given this complexity, it is not sufficient to put in place interventions at simply the individual, school or community level, whilst contradictory messages continue to come from "higher" levels of media, industry and government (Lobstein, 2006). Yet to date, these macro levels (Swinburn et al, 1999) have been ignored; unsurprisingly perhaps given the difficulties in influencing strategies at this level. It feels way beyond the influencing circle of the "average researcher", particularly against the might of the powerful and influential food industry. This influence is further hampered by the fact that there is insufficient evidence for the obesity professionals to prescribe what "best practice" should be. Notwithstanding that this task is extremely daunting, we should still seek to make changes at these levels.

It should be asked whether the policy is appropriate, adequate, and equitable. It is not sufficient for an obesity prevention policy to only focus on urging individuals to live less unwholesome lives and/or promoting public health messages through public organizations (the traditional approach) – for maximum leverage it is also necessary to engage other levels of environments, such as communities, media and industry. The importance of the environment in controlling obesity is widely acknowledged. A WHO report (2003) states that major social and environmental changes to make healthier choices more accessible and preferable are required to prevent obesity. Ecological interventions that consider multiple levels of influence on behaviour may be more effective. For example, a programme that increases preferences for fruit and vegetables without changing availability of these foods is probably going to produce

smaller increases in consumption than a programme that addresses both these factors. Also they can change health behaviours without the need to make changes at the individual level. For instance, reducing the amount of fat in foods served in schools cafeterias can reduce overall fat and energy consumption without altering children's beliefs or attitudes about low fat foods. However it can be difficult to understand all levels of ecological influence on health behaviours, plus levels of ecological influence interact, making the model even more complex. In addition intervening at more than one level of the ecosystem can be difficult and costly.

Some of these determinants of childhood obesity are outside of the remit of the health care sector. This highlights the role of organisations outside the health sector (e.g. supermarkets, gyms, swimming pools, food manufacturers and to some extent schools), as well as local and national government (e.g. decisions regarding public spending, agricultural policy) in contributing to the obesity problem. This emphasizes the responsibility that public and private sector organisations have in endorsing public health (Stafford et al, 2007). Thus in order to maximise childhood obesity prevention, health care professionals need to work with these organisations, otherwise it will not be possible to halt the flood of obesity.

This study is not without its limitations. The relationships between childhood obesity and the obesogenic covariates are based on small-area estimates for Leeds using SimObesity (a spatial microsimulation model) (see chapter 6). Whilst the validation of the inputs and outputs from the model were robust, it is nevertheless not "real" data. Also care must be taken in interpreting the maps because whilst each SOA has approximately the same population size, they vary in size geographically, thus more sparsely populated areas (and thus larger SOAs in geographical size) may be visually given more weight. To minimise this problem the interpretations of the maps was undertaken both visually and with the use of the underlying numbers (not provided). Further the cross-sectional nature of the data means that this work cannot identify causality, but may just provide some clues about what interventions may work best in Leeds.

This theoretical framework does not consider real world limitations of budgets and finite resources. But it is better to work from the ideal scenario, rating the elements in the framework for validity, relevance and potential for change in order to then prioritise action. This testing should be undertaken with a mixed committee (not solely government staff), including senior representatives from different stakeholders, which would be an important mechanism for gaining the commitment of the stakeholders to common goals related to childhood obesity prevention.

Conclusion

This chapter has outlined a framework for a childhood obesity prevention policy for Leeds. This has been compared to current Leeds policies, which have been shown to be lacking in many areas, although some important steps, such as the Healthy Schools Initiative, have been made. There is significant variation in obesity across Leeds and hot spots of problem areas are not restricted to the most highly deprived areas, which is where the focus of local policies are. There is also significant variation across Leeds in the relationship of childhood obesity with key determinants, and three case studies were used to highlight some of the differences. This then supports the debate that solutions need to be tailored to the locality. Interventions need to focus on the different levels of environment that influence obesity – from the individual, schools/communities, through to media/industry and government. Changes at all these different levels need to be coordinated, to enable benefits to occur and to be sustained over the long-term. Small individual programmes are unlikely to make a difference to the obesity epidemic. For maximum benefit an obesity prevention policy needs to take a coordinated, multi-component, multi-sectorial public health approach and overall policy unity and coherence is required, with buy-in of all stakeholders.

Chapter 9: Conclusion

9.1 Introduction
9.2 Summary of research findings
9.3 Limitations of the research
9.4 Possibilities for future research
9.5 Conclusion

9.1 Introduction

This final chapter draws the thesis to a close by reflecting on the study's results, and summarising the key limitations of the research. Possibilities for future work are proposed before finishing with some concluding statements regarding the key findings.

9.2 Summary of research findings

The principal aim of this thesis was to investigate the micro-level variability in childhood obesity and obesogenic environments/behaviours. To achieve this 14 research objectives were established and needed to be undertaken. Each of these objectives will be addressed to evaluate whether they were successfully attained.

The first objective was to *review the aetiology and geography of childhood obesity, and interventions to prevent childhood obesity.* This was covered in two literature review chapters (2 and 3) and also in chapter 8, where a table summarising the key interventions that have taken place was provided. This gave a solid knowledge of essential background information. Further background information was provided through the second objective, namely *to evaluate small area population estimation methods and spatial analysis techniques.* This was covered gradually in several chapters, building up the information base. Chapter 6 considered different small area population estimation methods, establishing that spatial microsimulation using a deterministic re-weighting algorithm was a suitable method with which to estimate obesogenic covariates. Chapter 4 covered the practicalities of spatial analysis of health data including information about spatial scan statistics, with chapters 5 and 7 going into more detail about specific techniques, namely multi-level modelling and geographically weighted regression respectively. The third objective *was to describe obesogenic environments and the ANGELO model*, which were covered in the background sections of chapters 7 and 8. The last objective focused on providing background information was the fourth one, which was to *examine current obesity prevention policy in Leeds.* Chapter 8 addressed this point, outlining the key government strategies to prevent/reduce childhood obesity that are taking place and then focusing on the Local Public Service Agreements and Local Area Agreements for Leeds.

Objective 5, *to identify and obtain the available data for childhood obesity in Leeds*, lead to the identification of three possible sources of childhood obesity data. The first source to be identified was the routinely collected data by the Leeds Primary Care Trusts. As these data were only available for young children (aged 3-6 years), other sources of data were sought. This work then lead to the procurement of the height and weight data for the Trends and RADs studies, which are ongoing studies in Leeds. This information is included in chapter 4, however the text does not reflect the massive amount of work that went into obtaining these data. Similarly objective 6, *to investigate and determine sources of data for obesogenic covariates* was also a large piece of work. This was covered in chapter 6 and involved three stages. Firstly to establish a "wish list" of childhood obesity obesogenic covariates to be simulated based upon those variables identified as having a significant relationship with obesity from a review of the literature (chapter 2). However the ability to simulate a variable depends upon being able to find a reliable dataset that includes that variable and is as disaggregated as possible. Accordingly the University of Essex data archives were searched for suitable datasets and the obesogenic covariate list shortened accordingly. The third step is to ensure that the obesogenic covariates (output variables) have a strong correlation with the input variables for the spatial microsimulation model (within a chosen dataset, such as the Health Survey for England), thereby establishing the final list of obesogenic covariates for simulation. That is, from the Health Survey for England: fruit and vegetable consumption; physical activity levels; degree of urbanisation; socio-economic group; perception of quality of public transport, access to leisure facilities and supermarkets, and whether there is a problem with teenagers hanging around or vandals in the locality. From the Expenditure and Food Survey, the covariates were: household expenditure on food; ownership of a PC and more than one television; having internet access; household income and its main source; spending on school meals.

Chapter 4 principally covered objective 7, *to explore the temporal and micro-level spatial variations in childhood obesity in Leeds at residential level, using spatial analysis techniques, and identify "hot" and "cold" spots of childhood obesity*. The results showed that obesity in 3 to 13 year olds in Leeds has risen since 1998 and that prevalence of obesity increases with age. These results serve as a base from which future trends in Leeds can be monitored and also for comparison with trends elsewhere in the UK. Also childhood obesity was significantly associated with deprivation, with children living in highly deprived areas being twice as likely to be obese (or overweight) than children living in more affluent areas. Nevertheless some hot and cold spots were found in affluent and deprived areas, suggesting either a spread of obesity across socio-economic groups and/or something special about those areas affecting the aetiology of obesity. Some homogenous geo-demographic groups were more likely to be obese than others. Furthermore some such groups displayed obesity prevalence contrary to the "deprivation theory"; that is, groups that were less likely to be obese were located in deprived

parts of Leeds, and vice versa. These are important demographic differences between areas of high and low prevalence of obesity. Spatial analysis techniques used at the micro-level can help us to understand the variations in obesity within an area more thoroughly, allowing us to gain a better understanding of the driving factors and to identify key problem areas, rather than relying on averages for the whole of an area, which may mean that health professionals would miss small problem areas.

Objective 8 was covered by Chapter 5, *to explore the variation in childhood obesity at the school level, using spatial analysis techniques, and identify the schools that are under or over performing in respect to their levels of pupils' obesity.* This chapter demonstrated that there is considerable variation between schools in terms of their pupils' obesity, but that only part of this variation was explained by differences in the socio-demographic backgrounds of the pupils. Using a cross classified multi-level modelling process ethnicity, deprivation and "school effect" were taken into account and so allowed identification of schools that deviate from their expected levels of obesity. The strength of these results is that they provide an indication of how well, or otherwise, schools are functioning in an obesogenic environment.

Chapter 6 focused on objectives 9 and 10, namely *to build and validate SimObesity, a static spatial microsimulation model for obesogenic the environment in Leeds* and *to use SimObesity to estimate obesogenic covariates at the individual level, and aggregate to low super output area, in Leeds.* The key finding of this work was that the input (constraint) variables must be selected carefully, as they must be correlated with the output variables that are being synthesised if these output variables are to be simulated accurately (i.e. to closely match the characteristics of the actual population). Accordingly validation of the output variables stemmed from the correlation analysis undertaken in choosing the input variables, and then in determining whether these input variables simulate well; if yes, then it can be assumed that the correlated output variables also simulated well. Four methods of calibration were used and drew the same conclusions, that is, that the estimates for the obesogenic covariates were robust. SimObesity is a novel application of spatial microsimulation modelling.

To fulfil object 11, *to study the relationship between childhood obesity and obesogenic covariates in Leeds, using spatial analysis techniques, to identify those covariates with the strongest associations with childhood obesity and also to determine how the relationships vary at the micro-level*, chapter 7 took two different approaches, using both global and local analyses. The latter approach is novel in respect to obesity studies. The first approach enables the overall relationship between each covariate and childhood obesity in Leeds to be understood, and the second approach highlighted the differences in the covariates' relationships with obesity in the different micro-areas, showing where the relationships were weaker or

stronger. This significant variation across Leeds suggests that different interventions to reduce the prevalence of childhood obesity may be more effective in different areas in order to have the most effect on reducing prevalence of childhood obesity. This study showed that individuals' behaviours are important in determining prevalence of childhood obesity: namely, lack of physical activity, no or low consumption of fruit and vegetables, not purchasing school meals, and low expenditure on food were all globally associated with an increased risk of obesity (with these relationships varying across Leeds at the micro-level). This study also adds to the increasing evidence of the existence of "obesogenic environments". Features of the local environment may affect childhood obesity by changing health behaviours, i.e. promoting a healthy diet and encouraging physical activity (or vice versa). Several local determinants of obesity were identified in this study, with different covariates being more important in different areas. However to generalise for Leeds, obesogenic environments are characterised by the following: scale of urban development, deprivation (high deprivation score, receives benefits, low income), low SEG, more than one television per household, poor leisure facilities, poor access to supermarkets, problems with teenagers hanging around or with vandals, and good public transport facilities. Consequently more effective interventions to reduce childhood obesity would be those targeted at multiple determinants, and also take a coordinated approach with stakeholders.

Chapter 8 was concerned with objectives 12 and 13. That is, *to apply the knowledge gained in objectives 1-11 to develop a framework for childhood obesity prevention policy, and tailor it for different population groups* and to *compare and contrast the theoretical childhood obesity prevention policy with the actual obesity policy in Leeds*. It outlined a generic framework for a childhood obesity prevention policy, which was then tailored for each case study area in Leeds addressing the requirements of different populations. This policy was then compared to the current Leeds policies, which have been shown to be lacking in many areas, although some important steps, such as the Healthy Schools Initiative, have been made. There is significant variation in obesity across Leeds and hot spots of problem areas are not restricted to the most highly deprived areas, which is where the focus of local policies are. Interventions need to focus on the different levels of environment that influence obesity – from the individual, schools/communities, through to media/industry and government. Changes at all these different levels need to be coordinated, to enable benefits to occur and to be sustained over the long-term. Small individual programmes are unlikely to make a difference to the obesity epidemic. For maximum benefit an obesity prevention policy needs to take a coordinated, multi-component, multi-sectorial public health approach and overall policy unity and coherence is required, with buy-in of all stakeholders.

9.3 Limitations of the research

This study has achieved its principle aim of investigate the micro-level variability in childhood obesity and obesogenic environments/behaviours. However all studies have their limitations and this one is no exception. The key limitations were as follows:

- These analyses are based on simulated covariate data, not "real" data, however the validation of the simulation was robust and these data do enable analysis at the micro-level, elucidating more about the relationship between obesity and covariates than would otherwise be the case.
- The covariate data are subject to all the data collection limitations of the underlying surveys, namely the HSE and EFS. In particular much of the information collected, such as the physical activity data, was by self report and may be subject to reporting bias, which both surveys have endeavoured to minimise through their rigorous study designs.
- The obesity data were based on a sample and thus suffered from some small number problems, although a number of steps were taken to minimise this, including aggregating the data to SOA (rather than a smaller geography), aggregating data across time and age groups to enlarge the dataset, and smoothing the data using Empirical Bayes techniques.
- The use of SOA as unit of analysis was decided based on the availability of data and as a scale of convenience. It was not based on any a priori theoretical considerations.
- In order to be able to aggregate the data to SOAs it was necessary to categorise continuous data, the choice of categories being fairly arbitrary, although common sense options were chosen (e.g. the fruit and vegetable breakdowns were based around the fact many people consumed no daily fruit and vegetables and that the recommended daily intake is five portions per day).
- Running a separate simulation for each covariate may also have reduced errors, but would have been very time intensive.
- The size of the base populations for SimObesity may have been too small to generate a synthetic population that truly matches the actual population and amalgamation of parent files across years (e.g. using the HSE 2002 and 2003) may have improved the results, although this would have prevented the analysis of some key variables.
- There is no information about the exposure time for people; it is assumed that all individuals have lived in the same area their whole lives, yet migration in a small proportion of the population is likely, albeit possibly between areas with similar characteristics. Also it is likely that migration is less of an issue for children than for adults (purely due to age).
- Most of the covariates were strongly correlated with household income, which creates difficulties in interpreting the results. The problem of having two (or more) predictor

variables that are correlated, is that it is not possible to identify which one is having the effect on the dependent variable (childhood obesity).

- This study uses cross-sectional, not longitudinal, data (for both obesity and the covariates) and thus the results cannot indicate causality.

9.4 Possibilities for future work

They say a PhD thesis is never finished so much as abandoned. Accordingly details of potential future work are detailed below:

The existing simulated obesogenic covariates data could be used to address other research questions: for example, to facilitate the development of an obesity prevention initiative for parents of pre-school children in specific children's centres in Leeds. Similarly SimObesity could be used to generate other synthetic data to answer other research questions, depending on the data available in the various national level surveys, such as National Diet and Nutrition Survey or the British Household Panel Survey, that are available in the UK.

This work could be extended by using qualitative research methods to drill down into residential or school level hot or cold spots of obesity, in order to explore why, and maybe how, areas/schools differ from each other in this regard; studying the exceptional areas/schools in depth may elucidate why some areas/schools manage to resist the obesogenic environment to the benefit of its residents/pupils while others do not.

Furthermore it may be possible to use SimObesity to undertake "what if" scenario analysis to theoretically evaluate the potential impact of potential interventions on the prevalence of childhood obesity, particularly if SimObesity were upgraded to a dynamic model. Then it would be able to represent future population structures and be able to evaluate the future impact of health policies on disease prevalence, which is cheaper and much quicker than running a pilot study.

Multivariate analyses are constrained due to collinearity between the covariates. Further work considering combinations of local factors needs to be undertaken to identify the combination that it would be necessary to determine which combinations of risk factors are the most obesogenic in order to be able to stem the epidemic.

This analysis serves to highlight that there is a need to extend the current research base in order to build a well-founded framework to form the basis of a strategy for the prevention of

childhood obesity. At the current time there is no coherent obesity prevention policy in the UK and the evidence base does not indicate any best practices.

9.5 Conclusion

This thesis has added to the literature in the following ways. The methodologies used in this thesis are generalisable to other locations and chronic conditions.

In terms of methodology the key lessons are as follows. Spatial microsimulation modelling was shown to be a robust method to estimate obesogenic covariates at the micro-level. In the design of a spatial microsimulation model using a deterministic re-weighting algorithm, the input variables must be strongly correlated with the output variables to be able to accuracy simulate micro-area estimates. Also this thesis has highlighted that there is considerable advantage to analysing health data at a small scale, otherwise micro-level differences are simply "averaged" away and missed.

In relation to the application, the important results are firstly that deprivation was shown to be associated with childhood obesity in Leeds but that hot spots also exist in affluent areas, suggesting a spread of childhood obesity across the socio-economic groups. Secondly this study adds to the increasing evidence of the existence of "obesogenic environments": features of the local environment in Leeds may affect childhood obesity by changing health behaviours giving some evidence for contextual effects, however individual effects (compositional) are also important. Finally, relationships between obesogenic covariates and childhood obesity vary across Leeds, highlighting the need for tailored public health policies that are based on locally relevant evidence.

References

Agostoni C, Scaglioni S, Ghisleni D, Verduci E, Giovannini M, Riva E (2005). How much protein is safe? *International Journal of Obesity*, 29 (Suppl 2): S8-13

Agras WS, Hammer LD, McNicholas F, Kraemer HC (2004). Risk factors for childhood overweight: a prospective study from birth to 9.5 years. *The Journal of Pediatrics*, 145 (1): 20-25

Agras WS, Kraemer HC, Berkowitz RI, Hammer LD (1990). Influence of early feeding style on adiposity at 6 years of age. *The Journal of Pediatrics*, 116: 805–809

Allison DB, Matz PE, Pietrobelli A, Zannolli R, Faith MS (2001). Genetic and environmental influences on obesity. In: Bendich A, Deckelbaum RJ (eds). Primary and Secondary Preventive Nutrition. Humana Press: Totowa, New Jersey, pp 147–164

Arenz S, Rucker R, Koletzko B, von Kries R (2004). Breast-feeding and childhood obesity: a systematic review. *International Journal of Obesity*, 28: 1247–1256

Armstrong J, Reilly JJ, Child Health Information Team (2002). Breastfeeding and lowering the risk of childhood obesity. *Lancet*, 359 (9322): 2003–2004

Asthana S, Curtis S, Duncan C, Gould M (2002). Themes in British health geography at the end of the century: a review of published research 1998-2000. *Social Science and Medicine*, 55: 167-173

Baird J, Fisher D, Lucas P, Kleijnen J, Roberts H, Law C (2005). Being big or growing fast: systematic review of size and growth in infancy and later obesity. *BMJ*, 331: 929-935

Ballas D & Clarke G (2001). Modelling the local impacts of national social policies: a spatial microsimulation approach. *Environment and Planning C: Government and Policy*, 19: 587-606

Ballas D (2004). Simulating trends in poverty and income inequality on the basis of 1991 and 2001 census data: a tale of two cities. Area, 36 (2): 146–163

Ballas D, Clarke G, Dorling D, Rigby J, Wheeler B (2006). Using geographical information systems and spatial microsimulation for the analysis of health inequalities. Health Informatics Jounal, 12(1): 65-79

Ballas D, Rossiter D, Thomas B, Clarke G, Dorling D (2005). Geography matters: simulating the local impacts of national social policies. Joseph Rowntree Foundation, York, UK

Banerjee S, Morgan RJ, Rees SA, Latif AH (2003). Height screening at school: ineffective without high standards and adequate resources. *Archives of Disease in Childhood,* 88 (6): 477-481

Baranowski T, Baranowski JC, Cullen KW, Thompson DI, Nicklas T, Zakeri IE, Rochon J (2003). The fun, food, and fitness project (FFFP): the Baylor GEMS pilot study. *Ethnicity and Disease*, 13 (Suppl 1): S30–39

Barkeling B, Andersson I, Lindroos AK, Birkhed D, Rossner S (2001). Intake of sweet foods and counts of cariogenic microorganisms in obese and normal-weight women. *European Journal of Clinical Nutrition,* 55 (10): 850–855

Barker M, Robinson S, Osmond C, Barker DJ (1997). Birth weight and body fat distribution in adolescent girls. *Archives of Disease in Childhood*, 77: 381–383

Barlow SE & Dietz WH (1998). Obesity evaluation and treatment: expert committee recommendations. *Pediatrics*, 102: e29

Barratt J (1997). The cost and availability of healthy food choices in southern Derbyshire. *Journal of Human Nutrition and Dietetics,* 10 (1), 63–69

Baum F & Ziersch A (2003). Social capital. *Journal of Epidemiology and Community Health,* 57: 320-323

Baum FE, Bush RA, Modra CC, Murray CJ, Cox EM, Alexander KM, Potter RC (2000). Epidemiology of participation: An Australian community study. *Journal of Epidemiology and Community Health*, 54: 414-423

Beech BM, Klesges RC, Kumanyika SK, Murray DM, Klesges L, McClanahan B, Slawson D, Nunnally C, Rochon J, McLain-Allen B, Pree-Cary J (2003). Child- and parent-targeted interventions: the Memphis GEMS pilot study. *Ethnicity and Disease*, 13 (Suppl 1): S40–53

Belderson P, Harvey I, Kimbell R, O'Neill J, Russell J, Barker ME (2003). Does breakfast-club attendance affect schoolchildren's nutrient intake? A study of dietary intake at three schools. *The British Journal of Nutrition*, 90 (6): 1003-6

Bellizzi MC & Dietz WH (1999). Workshop on childhood obesity: summary of the discussion. *The American Journal of Clinical Nutrition*, 70: 173S-175S

Benigni, R; Giuliani, A. Cancer incidence and socioeconomic geography of Finland: A correlation study (2002). *Journal of Environmental Science and Health Part C - Environmental Carcinogenesis and Ecotoxicology Reviews*, 20 (1): 29-43

Bergmann KE, Bergmann RL, Von Kries R, Bohm O, Richter R, Dudenhausen JW, Wahn U (2003). Early determinants of childhood overweight and adiposity in a birth cohort study: role of breastfeeding. *International Journal of Obesity,* 27: 162–172

Berteus Forslund H, Torgerson JS, Sjostrom L, Lindroos AK (2005). Snacking frequency in relation to energy intake and food choices in obese men and women compared to a reference population. *International Journal of Obesity*, 29 (6): 711-719

Betts P (2003). Commentary. *Archives of Disease in Childhood*, 88: 480–481

Bhave S, Baydekar A, Otiv M, 2004. IAP national task force for childhood prevention of adult diseases: childhood obesity. *Indian Pediatrics*, 41 (6): 559-575

Billson H, Pryer JA, Nichols R (1999). Variation in fruit and vegetable consumption among adults in Britain. An analysis from the dietary and nutritional survey of British adults. *European Journal of Clinical Nutrition*, 53 (12): 946-952

Biro FM, McMahon RP, Striegel-Moore R, Crawford PB, Obarzanek E, Morrison JA, Barton BA, Falkner F (2001). Impact of timing of pubertal maturation on growth in black and white female adolescents: The National Heart, Lung, and Blood Institute Growth and Health Study. *The Journal of Pediatrics*, 138: 636–643

Black D, Morris J, Smith C, Townsend P (1980). Inequalities in health: report of a Research Working Group. London: Department of Health and Social Security

Blair AJ (1991). When are calories most fattening? *Appetite*, 17 (2): 161

Booth ML, Wake M, Armstrong T, Chey T, Hesketh K, Mathur S (2001). The epidemiology of overweight and obesity among Australian children and adolescents, 1995–97. *Australian and New Zealand Journal of Public Health,* 25: 162–169

Borzekowski DL & Robinson TN (2001). The 30-second effect: an experiment revealing the impact of television commercials on food preferences of preschoolers. *Journal of the American Dietetic Association*, 101 (1): 42–46

Boseley S (2004). United States accused of sabotaging obesity strategy. *International Journal of Health Services*, 34(3):553-554

Bouchard C & Tremblay A (1997). Genetic influences on the response of body fat and fat distribution to positive and negative energy balances in human identical twins. *The Journal of Nutrition,* 127 (5): 943S–947S

Bourdieu P (1986). The forms of capital. In Richardson JG (ed), The handbook of theory and research for the sociology of education. New York: Greenwood Press

Bowman SA, Gortmaker SL, Ebbeling CB, Pereira MA, Ludwig DS (2004). Effects of fast-food consumption on energy intake and diet quality among children in a national household survey. *Pediatrics,* 113: 112-118

Brown DG & Xie Y (2006). Spatial agent-based modelling. *International Journal of Geographical Information Science*, 20 (9): 941-943

Brown T, Kelly S, Summerbell C (2007). Prevention of obesity: a review of interventions. *Obesity Reviews*, 8 (suppl 1):127-30

Brownson RC, BakerEA, Housemann RA, Brennan LK, Bacak SJ (2001). Environmental and policy determinants of physical activity in the United States. *American Journal of Public Health*, 91 (12): 1995-2003

Brunsdon CA, Fotheringham AS, Charlton ME (1998). Geographically weighted regression - modelling spatial nonstationary. Statistician, 47, 431–43

Bundred P, Kitchiner D, Buchan I (2001). Prevalence of overweight and obese children between 1989 and 1998: population based series of cross sectional studies. BMJ; 322: 1-4

Burdette HL and Whitaker RC, 2005. A national study of neighbourhood safety, outdoor play, television viewing and obesity in preschool children. *Pediatrics*, 116 (3): 657-662

Butler D (2004). Health experts find obesity measures too lightweight. *Nature*, 428(6980):244

Caballero B, Clay T, Davis SM, Ethelbah B, Rock BH, Lohman T, Norman J, Story M, Stone EJ, Stephenson L, Stevnes J (2003). Pathways Study Research Group. Pathways: a school-based, randomized controlled trial for the prevention of obesity in American Indian schoolchildren. *The American Journal of Clinical Nutrition*, 78:1030–1038

Cade J, Upmeier H, Calvert C, Greenwood D (1999). Costs of a healthy diet: analysis from the UK Women's Cohort Study. *Public Health Nutrition*, 2: 505-512

Cade JE, Frear L, Greenwood DC (2006). Assessment of diet in young children with an emphasis on fruit and vegetable intake: using CADET--Child and Diet Evaluation Tool. *Public Health Nutrition, 9(4):501-508*

Carstairs V (1981). Small area analysis and health service research. Community Medicine, 3 (2); 131-39

Cavadini C, Siega-Riz AM, Popkin BM (2000). US adolescent food intake trends from 1965 to 1996. *Archives of Disease in Childhood*, 83: 18–24

Cawley JH (2006). Markets and childhood obesity policy. *Future of Children*, 16(1):69-88

Cecil JE, Watt P, Murrie IS, Wrieden W, Wallis DJ, Hetherington MM, Bolton-Smith C, Palmer CN (2005). Childhood obesity and socioeconomic status: a novel role for height growth limitation. *International Journal of Obesity*, 29 (10): 1199-1203

Charlton M, Fotheringham S, Brunsdon C (2003), GWR 3: software for geographically weighted regression. Spatial analysis research group, Department of Geography, University of Newcastle Upon Tyne, Version 3.0.1

Chinn S & Rona RJ (2001). Prevalence and trends in overweight and obesity in three cross sectional studies of British children, 1974-94. *BMJ,* 322: 24-26

Chinn S & Rona RJ (2002). International definitions of overweight and obesity for children: a lasting solution? *Annals of Human Biology*, 29 (3): 306-313

Clarke G (1996). Microsimulation for urban and regional policy analysis. Pion, London, UK

Clarke G, Eyre H, Guy C (2002). Deriving Indicators of Access to Food Retail Provision in British Cities: Studies of Cardiff, Leeds and Bradford. *Urban Studies*, 39 (11): 2041–2060

Cliff AD, Haggett P, Smallman-Raynor M (1998). Deciphering global epidemics: analytical approaches to the disease records of world cities 1888-1912. Cambridge University Press, Cambridge, UK

Clifford TJ (2003). Breast feeding and obesity. *BMJ*, 327 (7420): 879–880

Coen SE & Ross NA (2006). Exploring the material basis for health: Characteristics of parks in Montreal neighbourhoods with contrasting health outcomes. *Health and Place*, 12: 361-71

Cohan, J (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159

Cole TJ & Green PJ (1992). Smoothing reference centile curves: the LMS method and penalized likelihood. *Statistics in Medicine*, 11: 1305-1319

Cole TJ, Freeman JV, Preece MA (1995) Body Mass Index reference curves for the UK, 1990. *Archives of Disease in Childhood*, 73: 25-29.

Cole TJ, Bellizzi MC, Flegal KM, Dietz WH (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*, 320: 1240–1243

Coleman KJ, Tiller CL, Sanchez J, Heath EM, Sy O, Milliken G, Dzewaltowski DA (2005). Prevention of the epidemic increase in child risk of overweight in low-income schools. The El Paso coordinated approach to child health. *Archives for pediatrics and adolescent medicine*, 159: 217–224

Communities and Local Government (formerly called the Office of the Deputy Prime Minister, Neighbourhood Renewal Unit). The English Indices of Deprivation 2004. http://www.communities.gov.uk/index.asp?id=1128449 Accessed September 2007

Comuzzie AG & Allison DB (1998). The search for human obesity genes. *Science*, 280 (5368): 1374-1377

Connelly JB, Duaso MJ, Butler G (2007). A systematic review of controlled trials of interventions to prevent childhood obesity and overweight: A realistic synthesis of the evidence. *Public Health*, 121: 510-517

Connor SM (2006). Food-related advertising on preschool television: building brand recognition in young viewers. *Pediatrics*, 118(4):1478-1485

Cooper R, Cutler J, Desvigne-Nickens P, Fortmann SP, Friedman L, Havlik R, Hogelin G, Marler J, McGovern P, Morosco G, Mosca L, Pearson T, Stamler J, Stryer D, Thom T (2000). Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States - Findings of the National Conference on Cardiovascular Disease Prevention. *Circulation*, 102 (25): 3137-47

Coulston AM & Johnson RK (2002). Sugar and sugars: myths and realities. *Journal of the American Dietetic Association*, 102(3):351-353

Cowles MK (2004). Review of WinBUGS 1.4. *The American Statistician*, 58 (4): 330-336

Cummins S & Macintyre S (2006). Food environments and obesity – neighbourhood or nation? *International Journal of Epidemiology*, 35 (1): 101-104

Cummins, C., Winter, H., Cheng, K.K., Maric, R., Silcocks, P., Varghese, C. (1999). An assessment of the Nam Pehchan computer program for the identification of names of south Asian ethnic origin. *Journal of Public Health Medicine,* 21(4), 401-406.

Curhan GC, Chertow GM, Willett WC, Spiegelman D, Colditz GA, Manson JE, Speizer FE, Stampfer MJ (1996a). Birth weight and adult hypertension and obesity in women. *Circulation*, 94 (6): 1310–1315

Curhan GC, Willett WC, Rimm EB, Spiegelman D, Ascherio AL, Stampfer MJ (1996b). Birth weight and adult hypertension, diabetes mellitus, and obesity in US men. *Circulation*, 94 (12): 3246–3250

Curtis S, Cave B, Coutts A (2002). Is urban regeneration good for health? Perceptions and theories of the health impacts of urban change. *Environment and Planning C - Government and Policy*, 20 (4): 517-34

Dabelea D, Hanson RL, Lindsay RS, Pettitt DJ, Imperatore G, Gabir MM, Roumain J, Bennett PH, Knowler WC (2000). Intrauterine exposure to diabetes conveys risks for type 2 diabetes and obesity: a study of discordant sibships. *Diabetes* 49: 2208–2211

Daniels SR, Khoury PR, Morrison JA (1997). The utility of body mass index as a measure of body fatness in children and adolescents: differences by race and gender. *Pediatrics*, 99 (6): 804-807

Danielzik S, Czerwinski-Mast M, Langnase K, Dilba B, Muller MJ (2004). Parental overweight, socioeconomic status and high birth weight are the major determinants of overweight and obesity in 5–7 y-old children: baseline data of the Kiel Obesity Prevention Study (KOPS). *International Journal of Obesity*, 28: 1494–1502

Davey RC (2004). The obesity epidemic: too much food for thought? *British Journal of Sports Medicine*, 38(3):360-363

Davison KK & Birch LL (2001). Childhood overweight: a contextual model and recommendations for future research. *Obesity Reviews*, 2: 159-171

Davison KK & Birch LL (2002). Obesogenic families: parents' physical activity and dietary intake patterns predict girls' risk of overweight. *International Journal of Obesity*, 26: 1186–1193

Daynard RA, Howard PT, Wilking CL (2004). Private enforcement: litigation as a tool to prevent obesity. *Journal of Public Health Policy*, 25(3-4):408-417

Deckelbaum RJ & Williams CL (2001). Childhood obesity: the health issue. *Obesity Research*, 9 (S4): 239S-243S

Dennison BA, Russo TJ, Burdick PA, Jenkins PL (2004). An intervention to reduce television viewing by preschool children. *Archives for pediatrics and adolescent medicine*, 158: 170–176

Department of Health (2004) Choosing Health. Making Healthy Choices Easier. London: Department of Health

Department of Health (2006). Obesity delivery framework Leeds PCTs. Powerpoint presentation slides

Department of Health (2007). Local Area Agreements and Local Public Service Agreements. http://www.dh.gov.uk/en/AboutUs/MinistersAndDepartmentLeaders/ChiefMedicalOfficer/Features/DH_4116435 Accessed October 2007

Dietz WH & Bellizzi MC (1999). Introduction: the use of body mass index to assess obesity in children. *The American Journal of Clinical Nutrition*, 70: 123S-125S

Dietz WH & Gortmaker SL (1985). Do we fatten our children at the television set? Obesity and television viewing in children and adolescents. *Pediatrics*, 75 (5): 807–812

Dietz WH & Gortmaker SL (2001). Preventing obesity in children and adolescents. *Annual Review of Public Health*, 22: 337-353

Dietz WH (2001a). The obesity epidemic in young children: Reduce television viewing and promote playing. *BMJ*, 322 (7282): 313–314

Dietz WH (2001b). Breast feeding may help prevent childhood overweight. *The Journal of the American Medical Association*, 285 (19): 2506-2507

DiGuiseppi C, Roberts I, Li L (1997). Influence of changing travel patterns on child death rates from injury: trend analysis, *BMJ*, 314 (7082): 710–713

Diliberti N, Bordi P, Conklin M, Roe L, Rolls B (2004). Increased portion size leads to increased energy intake in restaurant meal. *Obesity Research.*12(3):562-568

Donnelly JE, Jacobsen DJ, Whatley JE, Hill JO, Swift LL, Cherrington A, Polk B, Tran ZV, Reed G (1996). Nutrition and physical activity program to attenuate obesity and promote physical and metabolic fitness in elementary school children. *Obesity Research*, 4: 229-243

Dorling D (1995). The visualization of local urban change across Britain, *Environment and Planning B,* 22 (3): 269-290

Dorling D, Mitchell R, Shaw M, Orford S, Smith GD (2000). The Ghost of Christmas Past: health effects of poverty in London in 1896 and 1991. *BMJ*, 321 (7276): 1547-51

Dorosty AR, Emmett PM, Cowin S, Reilly JJ (2000). Factors associated with early adiposity rebound. ALSPAC Study Team. *Pediatrics*, 105: 1115–1118

Drewnowski A & Popkin BM (1997). The nutrition transition: new trends in the global diet. *Nutrition Reviews,* 55: 31–43

Drewnowski A (2004). Obesity and the Food Environment: Dietary Energy Density and Diet Costs. *American Journal of Preventative Medicine*, 27(3S):154–162

Dummer TJ, Gibbon MA, Hackett AF, Stratton G, Taylor SR (2005). Is overweight and obesity in 9-10-year-old children in Liverpool related to deprivation and/or electoral ward when based on school attended?. *Public Health Nutrition*, 8 (6): 636-41

Dwyer JT & Ouyang CM (2000). What can industry do to facilitate dietary and behavioural changes?. *British Journal of Nutrition*, 83 Suppl 1:S173-80

Ebbeling CB, Pawlak DB, Ludwig DS (2002). Childhood obesity: public-health crisis, common sense cure. *Lancet,* 360: 473-482

Ebbeling CB, Sinclair KB, Pereira MA, Garcia-Lago E, Feldman HA, Ludwig DS (2004). Compensation for energy intake from fast food among overweight and lean adolescents. *The Journal of the American Medical Association*, 291 (23): 2828-2833

Education Leeds (2007). Extended schools and services. http://www.educationleeds.co.uk/DisplayText.aspx?section=8&pageno=316 Accessed October 2007

Egger G & Swinburn B (1997). An 'ecological' approach to the obesity pandemic. *BMJ*, 315 (7106): 477–480

Egger G, Swinburn B, Rossner S (2003). Dusting off the epidemiological triad: could it work with obesity? *Obesity Reviews*, 4 (2): 115-119

Ellaway A & Macintyre S (1999). Social capital and self rated health: Support for a contextual mechanism (letter). *American Journal of Public Health*, 90: 988

Elliot A (2003). US food industry ensures that consumers are not told to eat less. BMJ, 327(7423):1067

Epstein LH, Gardy CC, Raynor HA, Beddome M, Kilanowski CK, Paluch R (2001). Increasing fruit and vegetable intake and decreasing fat and sugar intake in families at risk for childhood obesity. *Obesity Research*, 9: 171-178

Epstein LH, Handley EA, Dearing KK, Cho DD, Roemmich JN, Paluch RA, Raja SP, Youngju SB (2006). Purchases of food in youth. Influence of price and income. *Psychological Science*, 17(1):82-89

Etelson D, Brand DA, Patrick PA, Shirali A (2003). Childhood obesity: do parents recognize this health risk? *Obesity Reviews*, 11 (11): 1362-1368

Fall CH, Osmond C, Barker DJ, Clark PM, Hales CN, Stirling Y, Meade TW (1995). Fetal and infant growth and cardiovascular risk factors in women. *BMJ*, 310 (6977): 428–432

Finkelstein E, French S, Variyam JN, Haines PS (2004). Pros and cons of proposed interventions to promote healthy eating. *American Journal of Preventive Medicine*, 27(3 Suppl):163-171

Finkelstein EA, Ruhm CJ, Kosa KM (2005). Economic causes and consequences of obesity. *Annual Review of Public Health*, 26:239-257

Fischer M, Scholten HJ, Unwin D (eds) (1996). Spatial analytical perspectives of GIS. Taylor & Francis, London, UK

Fitzgibbon ML, Stolley MR, Schiffer L, Van Horn L, Kaufer Christoffel K, Dyer A (2005). Two-year follow-up results for hip-hop to health jr:a randomized controlled trial for overweight prevention in preschool minority children. *Journal of Pediatrics*, 146: 618–625

Flodmark C-E, Lissau I, Moreno LA, Pietrobelli A, Widhalm K (2004). New insights into the field of children and adolescents' obesity: the European perspective. *International Journal of Obesity*, 28: 1189-1196

Florencio TM, Ferreira HS, de Franca AP, Cavalcante JC, Sawaya AL (2001). Obesity and undernutrition in a very-low-income population in the city of Maceio, northeastern Brazil. *The British Journal of Nutrition,* 86 (2): 277–284

Flores R (1995). Dance for health: improving fitness in African American nad Hispanic adolescents. *Public Health Reports*, 110: 189-193

Flynn MAT, McNil DA, Maloff B, Mutasingwa D, Wu M, Ford C, Tough SC (2006). Reducing obesity and related chronic disease risk in children and youth: a synthesis of evidence with 'best practice' recommendations. *Obesity Reviews*, 7 (suppl.1): 7-66

Foresight Report (2007). Tackling Obesities: future choices http://www.foresight.gov.uk/Obesity/obesity_final/20.pdf Accessed October 2007

Foster RK & Buttriss J (2005) A review of the Public Health White Paper. Nutrition Bulletin 30, 70-75

Fotheringham AS, Charlton M, Brunsdon C (2002). Geographically weighted regression: The Analysis of Spatially Varying Relationships. Wiley: Chichester

Fotheringham AS, Charlton M, Brunsdon C (2005). Geographically weighted regression. ESRC National Centre for Research Methods course lecture notes

Franks PW, Ravussin E, Hanson RL, Harper IT, Allison DB, Knowler WC, Tataranni PA, Salbe AD (2005). Habitual physical activity in children: the role of genes and the environment. *The American Journal of Clinical Nutrition,* 82 (4): 901-908

Freedman DS, Khan LK, Mei Z, Dietz WH, Srinivasan SR, Berenson GS (2002). Relation of childhood height to obesity among adults: the Bogalusa Heart Study. *Pediatrics*, 109: E23

Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS (2004). Inter-relationships among childhood BMI, childhood height, and adult obesity: the Bogalusa Heart Study. *International Journal of Obesity*, 28: 10-16

Freeman JV, Cole TJ, Chinn S, Jones PRM, White EM, Preece MA (1995) Cross-sectional stature and weight reference curves for the UK, 1990. *Archives of Disease in Childhood*, 73: 17-24

Frontini MG, Bao W, Elkasabany A, Srinivasan SR, Berenson G (2001). Comparison of weight-for-height indices as a measure of adiposity and cardiovascular risk from childhood to young adulthood: the Bogalusa heart study. *Journal of Clinical Epidemiology*, 54 (8): 817-822

Frumkin H (2003). Healthy places: Exploring the evidence. *American Journal of Public Health*, 2003; 93 (9): 1451-6

Garn SM, Bailey SM, Higgins IT (1976). Fatness similarities in adopted pairs. *The American Journal of Clinical Nutrition,* 29: 1067–1068

Garner P, Panpanich R, Logan S (2000). Is routine growth monitoring effective? A systematic review of trials. *Archives of Disease in Childhood*, 82:197–201

Gatrell A & Löytönen M (eds) (1998). GIS and health. Taylor & Francis, London, UK

Gatrell, AC (2002). Geographies of health: an introduction. Blackwell publishers, Oxford, UK

Gattrell A, Thomas C, Bennett S, Bostock L, Popay J, Williams G, Shahtahmasebi S (2000). Understanding health inequalities: Locating people in geographical and social spaces. In Graham H (ed), Understanding health inequalities (pp. 156-169). Buckingham: Open University Press

Gattrell AC (2005). Complexity theory and geographies of health: a critical assessment. *Social Science and Medicine*, 60: 2661–71

Geier AB, Rozin P, Doros G. Unit bias: A new heuristic that helps explain the effect of portion size on food intake. *Psychological Science.* 2006;17:521-525

Gibson SA (2000). Associations between energy density and macronutrient composition in the diets of pre-school children: sugars vs starch. *International Journal of Obesity*, 24 (5): 633–638

Giles-Corti B & Donovan RJ (2002). The relative influence of individual, social and physical environment determinants of physical activity. *Social Science and Medicine*, 54 (12): 1793-1812

Gillman MW, Rifas-Shiman SL, Camargo CA Jr, Berkey CS, Frazier AL, Rockett HR, Field AE, Colditz GA (2001). Risk of overweight among adolescents who were breastfed as infants. *The Journal of the American Medical Association*, 285: 2461–2467

Gilmer, M.J., Harrell, J.S., Miles, M.S., Hepworth, J.T. (2003). Youth characteristics and contextual variables influencing physical activity in young adolescents of parents with premature coronary heart disease. *Journal of Pediatric Nursing*. 18(3), 159-168

Gortmaker SL, Must A, Sobol AM, Peterson K, Colditz GA, Dietz WH (1996). Television viewing as a cause of increasing obesity among children in the United States, 1986–90. *Archives of Pediatrics and Adolescent Medicine*, 150: 356–362

Gortmaker SL, Peterson K, Wiecha J, Sobol AM, Dixit S, Fox MK, Laird N (1999). Reducing obesity via a school-based interdisciplinary intervention among youth. *Archives of Pediatrics and Adolescent Medicine*, 153: 409-419

Gould R, Russell J, Barker ME (2006). School lunch menus and 11 to 12 year old children's food choice in three secondary schools in England-are the nutritional standards being met? *Appetite*, 46 (1): 86-92

Gregory J & Lowe S (2000). National Diet and Nutritional Survey: Young People Aged 4–18 Years, Vol. 1: Report of the Diet and Nutrition Survey. The Stationery Office: London

Grove KL, Sekhon HS, Brogan RS, Keller JA, Smith MS, Spindel ER (2001). Chronic maternal nicotine exposure alters neuronal systems in the arcuate nucleus that regulate feeding behaviour in the newborn rhesus macaque. *The Journal of Clinical Endocrinology and Metabolism,* 86: 5420-5426

Guillaume M & Lissau I (2002). Epidemiology. In: Burniat W, Cole T, Lissau I, Poskitt EME (eds). Child and Adolescent Obesity. Causes and Consequences; Prevention and Management. Cambridge University Press: Cambridge, pp 28–49

Guillaume M, Lapidus L, Lambert A (1998). Obesity and nutrition in children. The Belgian Luxembourg Child Study IV. *European Journal of Clinical Nutrition*, 52 (5): 323–328

Guo SS, Chumlea WC, Roche AF, Siervogel RM (1998). Age- and maturity-related changes in body composition during adolescence into adulthood: the Fels longitudinal study. *Applied Radiation and Isotopes*, 49 (5-6): 581-585

Guo SS, Roche AF, Chumlea WC, Gardner JD, Siervogel RM (1994). The predictive value of childhood body mass index values for overweight at age 35 year. *The American Journal of Clinical Nutrition,* 59 (4): 810-819

Hakeem R (2001). Socio-economic differences in height and body mass index of children and adults living in urban areas of Karachi, Pakistan. *European Journal of Clinical Nutrition*, 55(5):400-406

Hall DM (2000). Growth monitoring. *Archives of Disease in Childhood,* 82:10–15

Hanley AJ, Harris SB, Gittelsohn J, Wolever TM, Saksvig B, Zinman B (2000). Overweight among children and adolescents in a Native Canadian community: prevalence and associated factors. *The American Journal of Clinical Nutrition*, 71 (3): 693–700

Hardy R, Wadsworth M, Kuh D (2000). The influence of childhood weight and socioeconomic status on change in adult body mass index in a British national birth cohort. *International Journal of Obes*ity, 24 (6): 725-34

Harnack LJ, Jeffery RW, Boutelle KN (2000). Temporal trends in energy intake in the United States: an ecologic perspective. *The American Journal of Clinical Nutrition*, 71 (6): 1478–1484

Harvey-Berino J, Rouke J (2003). Obesity prevention in preschool Native-American children: a pilot study using home visiting. *Obesity Research*, 11: 606–611

Hawe P & Shiell A (2000). Social capital and health promotion: a review. *Social Science and Medicine*, 51: 871-885

Hayne CL, Moran PA, Ford MM (2004). Regulating environments to reduce obesity. Journal of Public Health Policy, 25(3-4):391-407

Haynes R (1991). Inequalities in health and health service use: evidence from the General Household Survey. *Social Science and Medicine*, 33 (4), 361-8

He K, Hu FB, Colditz GA, Manson JE, Willett WC, Liu S (2004). Changes in intake of fruits and vegetables in relation ot risk of obesity and weight gain among middle-aged women. *International Journal of Obesity*, 28(12): 1569-1574

Healthy Schools, 2007. Healthy Schools National Programme. www.healthyschools.gov.uk Accessed October 2007

Heaney RP, Davies KM, Barger-Lux MJ (2002). Calcium and weight: clinical studies. *Journal of the American College of Nutrition*, 21:152S–155S

Hediger ML, Overpeck MD, Kuczmarski RJ, Ruan WJ (2001). Association between infant breastfeeding and overweight in young children. *The Journal of the American Medical Association*, 285 (19): 2453–2460

Heller RF & Page J (2002). A population perspective to evidence based medicine: "evidence for population health". *Journal of Epidemiology and Community Health*, 56: 45-47

Hernandez B, Gortmaker SL, Colditz GA, Peterson KE, Laird NM, Parra-Cabrera S (1999). Association of obesity with physical activity, television programmes and other forms of video viewing among children in Mexico city. *International Journal of Obesity*, 23 (8): 845–854

Hill J & Peters J (1998). Environmental contributions to the obesity epidemic. *Science*, 280:1371-1374

Hoek J & Gendall P (2006). Advertising and obesity: a behavioral perspective. *Journal of Health Communication*, 11(4):409-423

Hood MY, Moore LL, Sundarajan-Ramamurti A, Singer M, Cupples LA, Ellison RC (2000). Parental eating attitudes and the development of obesity in children. The Framingham Children's Study. *International Journal of Obesity*, 24: 1319–1325

Horne BD, Muhlestein JB, Lappé DL, Renlund GD, Bair TL, Bunch TJ, Anderson JL, for the intermountain heart collaborative study group (2004). Less affluent area of residence and

lesser-insured status predict an increased risk of death or myocardial infarction after angiographic diagnosis of coronary disease. *Annals of Epidemiology*, 14 (2): 143-50

Huang Z and Williamson P (2001). A comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. Working paper 2001/02, Department of Geography, University of Liverpool [online] http://pcwww.liv.ac.uk/~william/microdata/workingpapers/hw_wp_2001_2.pdf, accessed 13/08/07

Hyppa M & Maki J (2001). Individual-level relationships between social capital and self-rated health in a bilingual community. *Preventive Medicine*, 32: 148-155

Independent Television Commission (2005), *Television: the Public's View, 2003* [computer file]. Colchester, Essex: UK Data Archive [distributor], March 2005. SN: 5138.

Ipsos-RSL & Office for National Statistics (2003). *United Kingdom Time Use Survey, 2000* [computer file]. *3rd Edition.* Colchester, Essex: UK Data Archive [distributor], September 2003. SN: 4504

Jahns L, Siega-Riz AM, Popkin BM (2001). The increasing prevalence of snacking among US children from 1977 to 1996. *The Journal of Pediatrics*, 138: 493–498

James J, Thomas P, Cavan D, Kerr D (2004). Preventing childhood obesity by reducing consumption of carbonated drinks: cluster randomised controlled trial. *BMJ*, 328: 1237

Jebb SA (2005). Dietary strategies for the prevention of obesity. *Proceedings of the Nutrition Society*, 64: 217–227

Jebb SA, Lang R, Penrose A (2003). Improving communication to tackle obesity in the UK. *Proceedings of the Nutrition Society*, 6: 577-581

Jeffery AN, Voss LD, Metcalf BS, Alba S, Wilkin TJ (2005). Parents' awareness of overweight in themselves and their children: cross sectional study within a cohort (EarlyBird 21). *BMJ*, 330: 23-24

Johnstone AM, Shannon E, Whybrow S, Reid CA, Stubbs RJ (2000). Altering the temporal distribution of energy intake with isoenergetically dense foods given as snacks does not affect total daily energy intake in normal-weight men. *The British Journal of Nutrition*, 83 (1): 7–14

Jones SE, James-Ellison M, Young S, Gravenor MB, Williams R (2005). Monitoring trends in obesity in South Wales using routine data. *Archives of Disease in Childhood*, 90: 464-467

Joshi H, Wiggins RD, Bartley M, Mitchell R, Gleave S, Lynch K (2000). Putting health inequalities on the map: Does where you live matter, and why? In Graham H (ed), Understanding health inequalities (pp. 156-169). Buckingham: Open University Press

Kain J, Uauy R, Albala, Vio F, Cerda R, Leyton B (2004). Schoolbased obesity prevention in Chilean primary school children: methodology and evaluation of a controlled study. *International Journal of Obesity*, 28: 483–493

Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D(1997). Social capital, income inequality, and mortality. *American Journal of Public Health*, 87: 1491-498

Kelley B & Smith JA (2004). Legal approaches to the obesity epidemic: an introduction. *Journal of Public Health Policy*, 25(3-4):346-352

Kinra S, Baumer JH, Davey Smith G (2005). Early growth and childhood obesity: a historical cohort study. *Archives of Disease in Childhood*, 90 (11): 1122-1127

Kinra S, Nelder RP, Lewendon GJ (2000). Deprivation and childhood obesity: a cross sectional study of 20,973 children in Plymouth, United Kingdom. *Journal of Epidemiology & Community Health,* 54 (6): 456-460

Klesges RC, Shelton ML, Klesges LM (1993). Effects of television on metabolic rate: potential implications for childhood obesity. *Pediatrics*, 91 (2): 281-286

Kongmuang C (2007). Modelling Crime: A Spatial Microsimulation Approach. PhD thesis (available from the University of Leeds library).

Kroke A, Boeing H, Rossnagel K, Willich SN (2004). History of the concept of 'levels of evidence' and their current status in relation to primary prevention through lifestyle interventions. *Public Health Nutrition*, 7: 279-284

Krukowski RA, Harvey-Berino J, Kolodinsky J, Narsana RT, Desisto TP (2006). Consumers may not use or understand calorie labeling in restaurants. *Journal of the American Dietetic Association*, 106(6):917-920

Kulldorff M (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods*, 26: 1481-1496

Kulldorff M (2006). SaTScan User Guide for version 7.0. http://www.satscan.org/

Kulldorff M & Information Management Services Inc (2005). SaTScan™ v6.0: Software for the spatial and space-time scan statistics. http://www.satscan.org/

Kumanyika S, Jeffery RW, Morabia A, Ritenbaugh C, Antipatis VJ (2002). Public health approaches to the prevention of obesity. Working Group of the International Obesity TaskForce.Obesity prevention: the case for action. *International Journal of Obesity*, 26: 425–436

Kumanyika S, Jeffery RW, Morabia A, Ritenbaugh C, Antipatis VJ (2002). Public health approaches to the prevention of obesity. Working Group of the International Obesity

TaskForce.Obesity prevention: the case for action. *International Journal of Obesity*, 26: 425–436

Lake JK, Power C, Cole TJ (1997). Child to adult body mass index in the 1958 British birth cohort: Associations with parental obesity. *Archives of Disease in Childhood*, 77: 376–381

Lamerz A, Kuepper-Nybelen J, Wehle C, Bruning N, Trost-Brinkhues G, Brenner H, Hebebrand J, Herpertz-Dahlmann B (2005). Social class, parental education, and obesity prevalence in a study of six-year-old children in Germany. *International Journal of Obesity*, 29: 373–380

Lang T & Rayner G (2007). Overcoming policy cacophony on obesity: an ecological public health framework for policymakers. *Obesity reviews*, 8 (suppl 1): 165-181

Lara-Esqueda A, Aguilar-Salinas CA, Velazquez-Monroy O, Gomez-Perez, Rosas-Peralta M, Mehta R, Tapia-Conyer R (2004). The body mass index is a less-sensitive tool for detecting cases with obesity-associated co-morbidities in short stature subjects. *International Journal of Obesity*, 28: 1443-1450

Lazarus R, Baur L, Webb K, Blyth F (1996). Adiposity and body mass indices in children: Benn's index and other weight for height indices as measures of relative adiposity. *International Journal of Obesity*, 20 (5): 406-412

Leeds City Council (2007a). The Leeds initiative: A Local Area Agreement for Leeds 2006-2009. www.leedsinitiative.org/initiativeDocuments/2007712_18168276.pdf Accessed October 2007

Leeds City Council (2007b). Sustainable School Travel Strategy, consultation draft (Aug 2007). www.educationleeds.co.uk/documents/download.aspx?fid=23916) Accessed October 2007

Leupker RV, Perry CL, McKinlay SM, Nader PR, Parcel GS, Stone EJ, Webber LS, Elder JP, Feldman HA, Johnson CC, Kelder SH, Wu M for the CATCH Collaborative Group (1996). Outcomes of a field trial to improve children's dietary patterns and physical activity. The Child and Adolescent Trial for Cardiovascular Health (CATCH). *The Journal of the American Medical Association*, 275(10): 768-776

Levine RS, Feltbower RG, Connor AM, Robinson M, MCJ Rudolf (2007). Monitoring trends in childhood obesity : A simple school-based model. *Public Health*, in press, doi: 10.1016/j.puhe.2007.04.019

Lewis MK & Hill AJ (1998). Food advertising on British children's television: a content analysis and experimental study with nine-year olds. *International Journal of Obesity*, 22 (3): 206-214

Leyland AH & Davies CA (2005). Empirical Bayes methods for disease mapping. Statistical Methods in Medical Research, 14 (1): 17-34. Emperical Bayes software (add-in for ArcView 3x) is available from: http://arcscripts.esri.com/details.asp?dbid=13900

Li L, Parsons TJ, Power C (2003). Breast feeding and obesity in childhood: cross sectional study. *BMJ*, 327 (7420): 904–905

Linn SE (2004). Food marketing to children in the context of a marketing maelstrom. *Journal of Public Health Policy*, 25 (3-4): 367-37

Lissau I & Sorensen TI (1994). Parental neglect during childhood and increased risk of obesity in young adulthood. *Lancet*, 343 (8893): 324-327

Liu GC, Cunningham C, Downs SM, Marrero DG, Fineberg N (2002). A spatial analysis of obesogenic environments for children. *Proceedings of the AMIA Annual Symposium*, 459-63

Liu YJ, Araujo S, Recker RR, Deng HW (2003). Molecular and genetic mechanisms of obesity: implications for future management. *Current Molecular Medicine*, 3 (4): 325-340

Livesey G (2005). Low-glycaemic diets and health: implications for obesity. *Proceedings of the Nutrition Society*, 64(1):105-13

Lobstein T & Dibb S (2005). Evidence of a possible link between obesogenic food advertising and child overweight. *Obesity Reviews*, 6(3):203-208

Lobstein T (2006). Comment: Preventing child obesity – an art and a science. *Obesity Reviews*, 7 (suppl.1): 1-5

Lobstein T, Baur L, Uauy R for the IASO International Obesity TaskForce (2004). Obesity in children and young people: a crisis in public health, *Obesity Reviews*, 5 (Suppl. 1): 4–104

Lostao L, Regidor E, Aiach P, Dominguez V. (2001). Social inequalities in ischaemic heart and cerebrovascular disease mortality in men:Spain and France, 1980–1982 and 1988–1990. *Social Science and Medicine*, 52: 1879–87

Lowry R, Wechsler H, Galuska DA, Fulton JE, Kann L (2002). Television viewing and its associations with overweight, sedentary lifestyle, and insufficient consumption of fruits and vegetables among US high school students: differences by race, ethnicity, and gender. *Journal of School Health,* 72 (10): 413-421

Ludwig DS, Peterson KE, Gortmaker SL (2001). Relation between consumption of sugar-sweetened drinks and childhood obesity: a prospective, observational analysis. *Lancet*, 357 (9255): 505–508

Lumeng JC, Appugliese D, Cabral HJ, Bradley RH, Zuckerman B (2006). Neighborhood safety and overweight status in children. *Archinves of Pediatric and Adolescent Medicine*, 160 (1): 25-31

Macintyre S, Ellaway A, Cummins S (2002). Place effects on health: how can we conceptualise, operationalise and measure them? *Social Science and Medicine*, 55: 125–39

Maffeis C, Talamini G, Tato L (1998). Influence of diet, physical activity and parent's obesity on children's adiposity: a four-year longitudinal study. *International Journal of Obesity*, 22 (8): 758–764

Magarey AM, Daniels LA, Bouton TJC, Cockington RA (2001). Does fat intake predict adiposity in healthy children and adolescents aged 2–15y? A longitudinal analysis. *European Journal of Clinical Nutrition*, 55 (6): 471–481

Maheswaran R, Strachan DP, Dodgeon B, Best NG (2002). A population-based case-control study for examining early life influences on geographical variation in adult mortality in England and Wales using stomach cancer and stroke as examples. *International Journal of Epidemiology*, 31 (2): 375-82

Mallam KM, Metcalf BS, Kirkby J, Voss LD, Wilkin TJ (2003). Contribution of timetabled physical education to total physical activity in primary school children: cross sectional study. *BMJ*, 327 (7415): 592-593

Marmonier C, Chapelot D, Louis-Sylvestre J (2000). Effect of macronutrient content and energy density of snacks consumed in a satiety state on the onset of the next meal. *Appetite*, 34 (2): 161–168

Matheson DM, Killen JD, Wang Y, Varady A, Robinson TN (2004). Children's food consumption during television viewing. *American Journal of Clinical Nutrition*, 79 (6): 1088-1094

McCabe, M.P., Ricciardelli, L.A., Finemore, J. (2002). The role of puberty, media and popularity with peers on strategies to increase weight, decrease weight and increase muscle tone among adolescent boys and girls. *Journal of psychosomatic research*, 52(3), 145-153

McConahy KL, Smiciklas-Wright H, Birch LL, Mitchell DC, Picciano MF (2002). Food portions are positively related to energy intake and body weight in early childhood. *The Journal of Pediatrics*, 140: 340–347

McDermott R, O'Dea K, Rowley K, Knight S, Burgess P (1998). Beneficial impact of the homelands movement on health outcomes in central Australian aborigines. *Australian and New Zealand Journal of Public Health*, 22: 653–658

McGloin AF, Livingstone MBE, Greene LC, Webb SE, Gibson JMA, Jebb SA, Cole TJ, Coward WA, Wright A, Prentice AM (2002). Energy and fat intake in obese and lean children at varying risk of obesity. *International Journal of Obesity*, 26 (2): 200–207

Mensah GA (2005). Eliminating disparities in cardiovascular health - Six strategic imperatives and a framework for action. *Circulation*, 111 (10): 1332-6

Metcalf B, Voss L, Jeffery A, Perkins J, Wilkin T (2004). Physical activity cost of the school run: impact on schoolchildren of being driven to school (EarlyBird 22). *BMJ*, 329 (7470): 832-833

Metcalf BS, Voss LD, Wilkin TJ (2002). Accelerometers identify inactive and potentially obese children (EarlyBird 3). *Archives of Disease in Childhood*, 87: 166-167

Michielutte R, Diseler RA, Corbett WT, Schey HM, Ureda JR (1984). The relationship between weight-height indices and the triceps skinfold measure among children age 5 to 12. *American Journal of Public Health*, 74 (6): 604-606

Mitchell R, Dorling D, Shaw M (2000). Inequalities in life and death: what if Britain were more equal? Joseeph Rowntree Foundation, York

Mohan J, Twigg L, Barnard S, Jones K (2005). Social capital, geography and health: a small-area analysis for England. *Social Science and Medicine,* 60: 1267-1283

Monden CWS, van Lenthe FJ, Mackenbach JP (2006). A simultaneous analysis of neighbourhood and childhood socio-economic environment with self-assessed health and health-related behaviours. *Health and Place*, 12(4): 394-403

Monge-Rojas, R., Nunez, H.P., Garita, C., Chen-Mok, M. (2002). Psychosocial aspects of Costa Rican adolescents' eating and physical activity patterns. *The Journal of Adolescent Health*, 31(2), 212-219

Monmonier M (1996). How to lie with maps. University of Chicago Press, Chicago, 2nd edition

Morland K, Wing S, Diez Roux A (2002). The contextual effect of the local food environment on reidents' diets: The Atherosclerosis Risk in Communities study. *American Journal of Public Health*, 92 (11): 1761-1767

Mo-Suwan L, Lebel L, Puetpaiboon A, Junjana C (1999). School performance and weight status of children and young adolescents in a transitional society in Thailand. *International Journal of Obesity*, 23 (3): 272–277

Mo-Suwan L, Pongprapai S, Junjana C, Puetpaiboon A (1998). Effects of a controlled trial of a school-based exercise program o nthe obesity indexes of preschool children. *American Journal of Clinical Nutrition*, 68: 1006-1011

Muller MJ, Asbeck I, Mast M, Langnase K, Grund A (2001a). Prevention of obesity – more than an intention. Concept and first results of the Kiel Obesity Prevention Study (KOPS). *International Journal of Obesity*, 25 (suppl 1): 566-574

Muller MJ, Mast M, Asbeck I, LangnaseK, Grund A (2001b). Prevention of obesity – is it possible? *Obesity Reviews*, 2: 15-28

National Centre for Social Research and University College London, Department of Epidemiology and Public Health (2004), Health Survey for England, 2002 [computer file]. Colchester, Essex: UK Data Archive [distributor]. SN: 4912

National Social Marketing Centre (NSMC), 2006. It's our health! Realising the potential of effective social marketing. Summary independent review, findings and recommendations. http://www.nsms.org.uk/images/CoreFiles/NCCSUMMARYItsOurHealthJune2006.pdf (accessed 30 April 2007

National Statistics (2006). United Kingdom Input-Output Analyses, 2006 Edition. *http://www.statistics.gov.uk/inputoutput*

Nelson M (2000). Childhood nutrition and poverty. Proceedings of the Nutrition Society. 59(2):307-15

Nelson MC, Gordon-Larsen P, Song Y, Popkin BM (2006). Built and Social Environments: Associations with Adolescent Overweight and Activity. *American Journal of Preventive Medicine,* 31 (2): 109-17

Nelson R (2004). "Cheeseburger Bill" protects food industry. *The Lancet*, 363: 954

Neovius M, Linné Y, Barkeling B, Rössner S (2004). Discrepancies between classification systems of childhood obesity. *Obesity Reviews*, 5: 105-114

Nestle M (2003). The ironic politics of obesity. *Science*, 299(5608):781

Neumark-Sztainer D, French SA, Hannan PJ, Story M, Fulkerson JA (2005). School lunch and snacking patterns among high school students: associations with school food environment and policies. *The International Journal of Behavioral Nutrition and Physical Activity*, 2 (1): 14

Neumark-Sztainer D, Story M, Hannan PJ, Rex J (2003). New Moves: a school-based obesity prevention program for adolescent girls. Preventive Medicine, 37:41–51

Neumark-Sztainer DR, Wall MM, Haines JI, Story MT, Sherwood NE, van den Berg PA (2007). Shared risk and protective factors for overweight and disordered eating in adolescents. American Journal of Preventive Medicine, 33(5):359-369

Neville L, Thomas M, Bauman A (2005). Food advertising on Australian television: the extent of children's exposure. *Health Promotion International*, 20(2):105-112

New SA & Livingstone MB (2003). An investigation of the association between vending machine confectionery purchase frequency by schoolchildren in the UK and other dietary and lifestyle factors. *Public Health Nutrition*, 6 (5): 497-504

Nielsen BM, Bjornsbo KS, Tetens I, Heitmann BL (2005). Dietary glycaemic index and glycaemic load in Danish children in relation to body fatness. British Journal of Nutrition, 94(6):992-7

Noor MI (2002). The nutrition and health transition in Malaysia. *Public Health Nutrition*, 5(1A):191-195

Norman P, 1999. Putting iterative proportional fitting on the researcher's desk. Working paper 99/03, School of Geography, University of Leeds

O'Connor TM, Yang S-J, Nicklas TA (2006). Beverage intake among preschool children and its effect on weight status. *Pediatrics*, 118(4):1010-1018

Office for National Statistics, 2001a Census: Standard Area Statistics (England and Wales) [computer file]. ESRC/JISC Census Programme, Census Dissemination Unit, MIMAS (University of Manchester)

Office for National Statistics, 2001b Census: Digitised Boundary Data (England and Wales) [computer file]. ESRC/JISC Census Programme, Census Geography Data Unit (UKBORDERS), EDINA (University of Edinburgh)

Office for National Statistics and Department for Environment, Food and Rural Affairs (2006), *Expenditure and Food Survey, 2004-2005* [computer file]. *2nd Edition.* Colchester, Essex: UK Data Archive [distributor]. SN: 5375

Office for National Statistics. Social and Vital Statistics Division and Food Standards Agency (2005). *National Diet and Nutrition Survey : Adults Aged 19 to 64 Years, 2000-2001* [computer file]. Colchester, Essex: UK Data Archive [distributor], May 2005. SN: 5140.

Ogden CL, Flegal KM, Carroll MD, Johnson CL (2002). Prevalence and trends in overweight among US children and adolescents, 1999–2000. *The Journal of the American Medical Association,* 288 (14): 1728–1732

Okasha M, McCarron P, McEwen J, Durnin J, Davey Smith G (2003). Childhood social class and adulthood obesity: findings from the Glasgow Alumni Cohort. *Journal of Epidemiology and Community Health*, 57: 508–9

Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DB (2000). Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ*, 320 (7236): 967–971

Openshaw S (1983). The modifiable areal unitproblem. Concepts and techniques in model geography. Environmental publications, Norwich, UK

Orcutt GH (1957). A new type of socio-economic system. *The Review of Economics and Statistics*, 39(2): 116-123

Padez C, Mourao I, Moreira P, Rosado V (2005). Prevalence and risk factors for overweight and obesity in Portuguese children. *Acta paediatrica*, 94 (11): 1550-1557

Pangrazi RP, Beighle A, Vehige T, Vack C (2003). Impact of promoting lifestyle activity for youth (PLAY) on children's physical activity. *The Journal of School Health*, 73: 317–321

Parkes A & Kearns A (2006). The multi-dimensional neighbourhood and health: A cross-sectional analysis of the Scottish Housing Survey, 2001. *Health and Place*, 12 (1), 1-18

Parsons TJ, Power C, Logan S, Summerbell CD (1999). Childhood predictors of adult obesity: a systematic review. *International Journal of Obesity*, 23 (Suppl. 8): S1–107

Parsons TJ, Power C, Manor O (2001). Fetal and early life growth and body mass index from birth to early adulthood in 1958 British cohort: longitudinal study. *BMJ*, 323 (7325): 1331–1335

Pearson T, Russell J, Campbell MJ, Barker ME (2005). Do 'food deserts' influence fruit and vegetable consumption? A cross-sectional study. *Appetite*, 45 (2): 195-7

Pelletier AL, Chang WW, Delzell JE Jr, McCall JW (2004). Patients' understanding and use of snack food package nutrition labels. *Journal of the American Board of Family Practice.* 17(5):319-323

Phillips SM, Bandini LG, Naumova EN, Cyr H, Colclough S, Dietz WH, Must A (2004). Energy-Dense Snack Food Intake in Adolescence: Longitudinal Relationship to Weight and Fatness. *Obesity Research*, 12: 461-472

Pickett K & Pearl M (2001). Multilevel analyses of neighbourhood socio-economic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*, 55: 111–22

Pickle LW, Waller LA, Lawson AB (2005). Current practices in cancer spatial data analysis: a call for guidance. *International Journal of Health Geographics*, 4:3

Popkin BM & Udry JR (1998). Adolescent obesity increases significantly in second and third generation US immigrants: the National Longitudinal Study of Adolescent Health. *The Journal of Nutrition*, 128 (4): 701-706

Popkin BM (1998). The nutrition transition and its health implications in lower-income countries. *Public Health Nutrition,* 1: 5–21

Post GB & Kemper HC (1993). Nutrient intake and biological maturation during adolescence. The Amsterdam growth and health longitudinal study. *European Journal of Clinical Nutrition,* 47: 400–408

Poulton R & Williams S (2001). Breastfeeding and risk of overweight. *The Journal of the American Medical Association,* 286 (12): 1449–1450

Power C & Jefferis BJ (2002). Fetal environment and subsequent obesity: a study of maternal smoking. *International Journal of Epidemiology,* 31: 413– 419

Power C & Parsons T (2000). Nutritional and other influences in childhood as predictors of adult obesity. *The Proceedings of the Nutrition Society*, 59 (2): 267-72

Power C, Lake JK, Cole TJ (1997). Body mass index and height from childhood to adulthood in the 1958 British born cohort. *The American Journal of Clinical Nutrition*, 66 (5): 1094–1101

Pratt M, Macera CA, Sallis JF, O'Donnell M, Frank LD (2004). Economic interventions to promote physical activity: Application of the SLOTH model. *American Journal of Preventive Medicine*, 27(Suppl 3): 136-145

Prentice AM & Jebb SA (2003). Fast foods, energy density and obesity: a possible mechanistic link. *Obesity Reviews,* 4 (4): 187-194

Prentice AM (1998). Body mass index standards for children: Are useful for clinicians but not yet for epidemiologists. *BMJ,* 317: 1401-1402

Propper C, Jones K, Bolster A, Burgess S, Johnson R, Sarker R (2005). Local neighbourhood and mental health: Evidence from the UK. *Social Science and Medicine*, 61: 2065–83

Putman RD (1995). Bowling alone: American's declining social capital. *Journal of Democracy*, 6: 65-78

Ramirez-Lopez E, Grijalva-Haro MI, Valencia ME, Antonio Ponce J, Artalejo E (2005). Effect of a school breakfast program on the prevalence of obesity and cardiovascular risk factors in children. *Salud pública de México*, 47: 126–133 [in Spanish]

Ransley JK, Greenwood DC, Cade JE, Blenkinsop S, Schagen I, Teeman D, Scott E, White G, Schagen S (2007). Does the school fruit and vegetable scheme improve children's diet? A non-randomised controlled trial. Journal of Epidemiology and Community Health, 61: 699-703

Rasbash J, Steele F, Browne W, Prosser B (2004). A user's guide to MlwiN Version 2.0. Centre for multilevel modelling, Institute of Education, London, UK

Ravelli AC, van Der M, Osmond C, Barker DJ, Bleker OP (1999). Obesity at the age of 50 y in men and women exposed to famine prenatally. *The American Journal of Clinical Nutrition*, 70 (5): 811–816

Ravussin E, Valencia ME, Esparza J, Bennett PH, Schulz LO (1994). Effects of a traditional lifestyle on obesity in Pima Indians. *Diabetes Care* 17: 1067–1074

Reading R, Langford IH, Haynes R, Lovett A (1999). Accidents to preschool children: Comparing family and neighbourhood characteristics. *Social Science and Medicine*, 48: 321–30

Rees P, Norman P, Brown D (2004). A framework for progressively improving small area population estimates. Journal of the Royal Statistical Society Series A – Statistics in Society, 167: 5-36

Reidpath DD, Burns C, Garrard J, Mahoney M, Townsend M (2002). An ecological study of the relationship between social and environmental determinants of obesity. *Health and Place*, 8: 141–145

Reilly JJ & McDowell ZC (2003). Physical activity interventions in the prevention and treatment of paediatric obesity: systematic review and critical appraisal. *Proceedings of the Nutrition Society*, 62 (3): 611-619

Reilly JJ (2002). Assessment of childhood obesity: national reference data or international approach? *Obesity Review*, 10 (8): 838-840

Reilly JJ, Armstrong J, Dorosty AR, Emmett PM, Ness A, Rogers K, Steer C, Sherriff A for the Avon Longitudinal Study of Parents and Children Study Team (2005). Early life risk factors for obesity in childhood: cohort study. *BMJ*, 330 (7504): 1357-1363

Rietschlin J (1998). Voluntary association membership and psychological distress. *Journal of Health and Social Behaviour*, 39: 348-355

Rigby JE & Gatrell AC (2000). Spatial patterns in breast cancer incidence in NW Lancashire, Area; 32: 71-78

Rigby NJ, Kumanyika S, James WPT (2004). Confronting the Epidemic: The Need for Global Solutions. *Journal of Public Health Policy*, 25: 418-434

Robert SA (1999). Socioeconomic position and health: the independent contribution of community socio-economic context. *Annual Review of Sociology*, 25: 489–516

Robertson SM, Cullen KW, Baronowski J, Baronowski T, Shaouhua H, de Moor C (1999). Factors related to adiposity among children aged 3–7y. *Journal of the American Dietetic Association*, 99 (8): 939–943

Robinson TN (1999). Reducing children's television viewing to prevent obesity: a randomised controlled trial. *The Journal of the American Medical Association*, 282: 1561-1567

Robinson TN (2001). Television viewing and childhood obesity. *Pediatric Clinics of North America*, 48 (4): 1017-1025

Robinson TN, Hammer LD, Killen JD, Kraemer HC, Wilson DM, Hayward C, Taylor CB (1993). Does television viewing increase obesity and reduce physical activity? Cross-sectional and longitudinal analyses among adolescent girls. *Pediatrics*, 91: 273–280

Robinson TN, Killen JD, Kraemer HC, Wilson DM, Mathesan DM, Haskell WL, Pruitt LA, Powell TM, Owens AS, Thompson NS, Flint-Moore NM, Davis GJ, Emig KA, Brown RT, Rochon J, Green S, Varady A (2003). Dance and reducing television viewing to prevent weight gain in African-American girls: the Stanford GEMS pilot study. *Ethnicity and Disease*, 13 (suppl 1): 565-577

Roefs A & Jansen A (2004). The effect of information about fat content on food consumption in overweight/obese and lean people. Appetite, 43(3):319-322

Rolls B, Roe L, Meengs J, Wall D (2004). Increasing the portion size of a sandwich increases energy intake. *Journal of the American Dietetic Association.*;104(3):367-372

Rolls BJ, Engell D, Birch LL (2000). Serving portion size influences 5-year-old but not 3-year-old children's food intake. *Journal of the American Dietetic Association,* 100 (2): 232-234

Romon M, Duhamel A, Collinet N, Weill J (2005). Influence of social class on time trends in BMI distribution in 5-year-old French children from 1989 to 1999. *International Journal of Obesity*, 29: 54–59

Rose D & Richards R (2004). Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutrition*, 7(8), 1081-1088

Ross CE & Mirowsky J (2001). Neighborhood disadvantage, disorder, and health. *Journal of Health and Social Behavior*, 42(3): 258-276

Ross CE (1993). Fear of victimization and health. *Journal of Quantitative Criminology*, 9(2): 159-175

Ross CE (2000). Walking, exercising, and smoking: does neighbourhood matter? *Social Science and Medicine*, 51: 265-74

Rothman KJ and Greenland S (1998). Modern epidemiology. Second edition, Lippincott-Raven publishers, Philadelphia, USA

Rudolf MCJ, Cole TJ, Krom AJ, Sahota, P, Walker J (1999). Growth of primary school children: a validation of the 1990 standards and their use in growth monitoring. *Archives Disease in Childhood* 83: 298-301

Rudolf MCJ, Greenwood DC, Cole TJ, Levine R, Sahota P, Walker J, Holland PC, Cade J, Truscott J (2003) Rising obesity and expanding waistlines in school children: a cohort study. *Archives Disease in Childhood;* 89:235-237

Rudolf MCJ, Hochberg Z, Speiser P (2005a). Perspectives on the development of an international consensus on childhood obesity. *Archives of Disease in Childhood*, 90: 994-996

Rudolf MC, Levine R, Feltbower RG, Connor A, Robinson M (2005b). The trends project: Development of a methodology to reliably monitor the obesity epidemic in childhood. *Archives of Disease in Childhood*, published online 14/12/05: doi:10.1136/adc.2005.078915

Rudolf MCJ, Levine R, Feltbower RG, Connor A, Robinson M (2006). The Trends project: development of a methodology to reliably monitor the obesity epidemic in childhood. *Archives of Disease in Childhood*; 91: 309-311

Sahota P, Rudolf MC, Dixey R, Hill AJ, Barth JH, Cade J (2001a). Evaluation of implementation and effect of primary school based intervention to reduce risk factors for obesity. *BMJ*, 323: 1027

Sahota P, Rudolf MCJ, Dixey R, Hill AJ, Barth JH, Cade J (2001b). Randomised controlled trial of primary school based intervention to reduce risk factors for obesity. *BMJ*, 323: 1029

Sallis JF, McKenzie TL, Alcaraz JE, Kolody B, Hovell MF, Nader PR (1993). Project SPARK. Effects of physical education on adiposity in children. *Annals of the New York Academy of Sciences*, 699: 127-136

Sallis JF, McKenzie TL, Conway TL, Elder JP, Prochaska JJ, Brown M, Zive MM, Marshall SJ, Alcaraz JE (2003). Environmental interventions for eating and physical activity. A randomized controlled trial in middle school. *American Journal of Preventive Medicine*, 24: 209-217

Salvy, S.J., Romero, N., Paluch, R., Epstein, L.H. (2007). Peer influence on pre-adolescent girls' snack intake: effects of weight status. *Appetite*, 49(1), 177-182

Samples of Anonymised Records (SARs) website at the Cathie Marsh Centre for Census and Survey Research from Census 2001. http://www.ccsr.ac.uk/sars/ Accessed February 2008

Saxena S, Ambler G, Cole TJ, Majeed A (2004). Ethnic group differences in overweight and obese children and young people in England: cross sectional survey. *Archives of Disease in Childhood*, 89(1): 30-36

Schlundt DG, Hargreaves MK, McClellan L (2006). Geographic clustering of obesity, diabetes, and hypertension in Nashville, Tennessee. *Journal of Ambulatory Care Management*, 29 (2): 125-32 (p.127)

Schwartz J & Byrd-Bredbenner C (2006). Portion distortion: typical portion sizes selected by young adults. *Journal of the American Dietetic Association*. 106(9):1412-1418

Sekine M, Yamagami T, Handa K, Saito T, Nanri S, Kawaminami K, Tokui N, Yoshida K, Kagamimori S (2002). A dose-response relationship between short sleeping hours and

childhood obesity: results of the Toyama Birth Cohort Study. *Child Care, Health and Development*, 28 (2): 163-170

Select Committee on Health – Third Report (Obesity) (2004). *Health Committee Publications*, http://www.parliament.thestationeryoffice.co.uk/pa/cm200304/cmselect/cmhealth/23/2302.htm Accessed December 2005

Senior M, Williams H, Higgs G (2000). Urban-rural mortality differentials: controlling for material deprivation. *Social Science and Medicine*, 51: 289-305

Shang L, Xu YY, Du XH, Chen CS (2004). Geographical differences in blood pressure of male youth aged 17-21 years in China. *Blood Pressure*, 13 (3): 169-75

Shi H, Laurent EJ, LeBouton J, Racevskis L, Hall KR, Donovan M, Doepker RV, Walters MB, Lupi F, Liu J (2006). Local spatial modeling of white-tailed deer distribution. Ecological Modelling, 190 (1-2): 171-89

Shohaimi S, Welch A, Bingham S, Luben R, Day N, Wareham N, Khaw KT (2004). Residential area deprivation predicts fruit and vegetable consumption independently of individual educational level and occupational social class: a cross sectional population study in the Norfolk cohort of the European Prospective Investigation into Cancer (EPIC-Norfolk). *Journal of Epidemiology and Community Health*, 58 (8): 686-691

Shouls S, Congdon P, Curtis S (1996). Geographic variation in illness and mortality: the development of a relevant area typology for SAR districts. *Health and Place*, 2 (3): 139-55

Siega-Riz AM, Popkin BM, Carson T (1998). Trends in breakfast consumption for children in the United States from 1965 to 1991. *The American Journal of Clinical Nutrition*, 67 (4): 748S-756S

Simonetti D'Arca A, Tarsitani G, Cairella M, Siani V, De Filippis S, Mancinelli S, Marazzi MC, Palombi L (1986). Prevention of obesity in elementary and nursery school children. *Public Health*, 100: 166-173

Simpson L & Tranmer M (2005). Combining sample and census data in small area estimates: Iterative proportional fitting with standard software. Professional Geographer, 57 (20): 222-234

Sorensen HT, Sabroe S, Rothman KJ, Gillman M, Fischer P, Sorensen TI (1997). Relation between weight and length at birth and body mass index in young adulthood: cohort study. *BMJ*, 315 (7116): 1137

Sorensen TI, Holst C, Stunkard AJ (1998). Adoption study of environmental modifications of the genetic influences on obesity. *International Journal of Obesity*, 22(1):73-81

Speiser PW, Rudolf MC, Anhalt H, Camacho-Hubner C, Chiarelli F, Eliakim A, Freemark M, Gruters A, Hershkovitz E, Iughetti L, Krude H, Latzer Y, Lustig RH, Pescovitz OH, Pinhas-Hamiel O, Rogol AD, Shalitin S, Sultan C, Stein D, Vardi P, Werther GA, Zadik Z, Zuckerman-Levin N, Hochberg Z; Obesity Consensus Working Group (2005). Childhood Obesity, *The Journal of Clinical Endocrinology and Metabolism*, 90 (3):1871-87

Spiegelhalter DJ, Thomas A, Best NG, Lunn D (2002). WinBUGS Version 1.4 User Manual. MRC Biostatistics Unit, Cambridge. http://www.mrc-bsu.cam.ac.uk/bugs 2002. Software available free from http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

Sproston K & Primatesta P (2002). Health survey for England 2002, The Health of Children and Young People. *Joint Health Surveys Unit - National Centre for Social Research, Department of Epidemiology and Public Health at the Royal Free and University College Medical School,* The Stationery Office http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsStatistics/DH_4078027 Accessed 17 Sep 2007

Stafford M, Cummins S, Ellaway A, Sacker A, Wiggins RD, Macintyre S (2007). Pathways to obesity: Identifying local, modifiable determinants of physical activity and diet. *Social Science & Medicine*, doi:10.1016/j.socscimed.2001.05.042

Stamatakis E, Primatesta P, Chinn S, Rona R, Falascheti E (2005). Overweight and obesity trends from 1974 to 2003 in English children: what is the role of socioeconomic factors? *Archives of Disease in Childhood*, 90: 999-1004

Stettler N, Zemel BS, Kumanyika S, Stallings VA (2002). Infant weight gain and childhood overweight status in a multicenter, cohort study. *Pediatrics*, 109 (2): 194–199

Stolley MR, Fitzgibbon ML (1997). Effects of an obesity prevention program on the eating behaviour of African American mothers and daughters. *Health Education and Behavior*, 24: 152-164

St-Onge MP, Keller KL, Heymsfield SB (2003). Changes in childhood food consumption patterns: a cause for concern in light of increasing body weights. *The American Journal of Clinical Nutrition*, 78 (6): 1068–1073

Story M, Sherwood NE, Himes JH, Davis M, Jacobs DR Jr, Cartwright Y, Smyth M, Rochon J (2003). An after-school obesity prevention program for African-American girls: the Minnesota GEMS pilot study. *Ethnicity and Disease*, 13 (suppl 1): S54-64

Strauss RS & Knight J (1999). Influence of the home environment on the development of obesity in children. *Pediatrics*, 103 (6): e85

Strauss RS & Pollack HA (2001). Epidemic increase in childhood overweight 1986–1998. *The Journal of the American Medical Association,* 286 (22): 2845–2848

Strauss RS (2000). Childhood obesity and self-esteem. *Pediatrics*, 105 (1): e15

Sturm R & Datar A (2005). Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health*, 119(12): 1059-1068

Sturm R (2005). Childhood obesity -- what we can learn from existing data on societal trends, part 2. *Preventing Chronic Disease*, 2 (2): http://www.cdc.gov/pcd/issues/2005/apr/04_0039.htm

Subramanian SV & Smith GD (2006). Patterns, distribution, and determinants of under- and overnutrition: a population-based study of women in India. *American Journal of Clinical Nutrition*, 84(3): 633-640

Subramanian SV, Kawachi I, Kennedy BP (2001). Does the state you live in make a difference? Multilevel analysis of self-rated health in the US. *Social Science and Medicine*, 53: 9-19

Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ (2005). Interventions for preventing obesity in children. Cochrane Database of Systematic Reviews 2005, Issue 3. Art. No: CD001871. DOI. 10.1002/14651858. CD001871.pub2

Swinburn B & Egger G (2002). Preventive strategies against weight gain and obesity. *Obesity Reviews,* 3: 289-301

Swinburn B, Egger G, Raza F (1999). Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine*, 29: 563-570

Swinburn B, Gill T, Kumanyika S (2005). Obesity prevention: a proposed framework for tanslating evidence into action. *Obesity Reviews*, 6: 23-33

Takahashi E, Yoshida K, Sugimori H, Miyakawa M, Izuno T, Yamagami T, Kagamimori S (1999). Influence factors on the development of obesity in 3-year-old children based on the Toyama study. *Preventive Medicine*, 28 (3): 293-296

Takahashi K, Yokoyama T, Tango T (2005). FleXScan v1.1: Software for the Flexible Scan Statistic. National Institute of Public Health, Japan. http://www.niph.go.jp/soshiki/gijutsu/index_e.html

Tango T & Takahashi K (2005). A flexibly shaped spatial scan statistic for detecting clusters. International Journal of Health Geographics, 4: 11 (http://www.ij-healthgeographics.com/content/4/1/11

Thompson OM, Ballew C, Resnicow K, Gillespie C, Must A, Bandini LG, Cyr H, Dietz WH (2006). Dietary pattern as a predictor of change in BMI z-score among girls. *International Journal of Obesity*, 30 (1):176-82

Tickle M (2002). The 80:20 phenomenon: help or hindrance to planning caries prevention programmes? *Community Dental Health*, 19(1), 39-42

Tillotson JE (2004). America's obesity: conflicting public policies, industrial economic development, and unintended human consequences. *Annual Review of Nutrition*, 24:617-43

Timperio A, Salmon J, Telford A, Crawford D (2005). Perceptions of local neighbourhood environments and their relationship to childhood overweight and obesity *International Journal of Obesity*, 29: 170–175

Toh CM, Cutter J, Chew SK (2002). School based intervention has reduced obesity in Singapore. *BMJ*, 324 (7334): 427

Tseng J-H, Merchant E, Tseng M-Y (2006). Effects of socioeconomic and geographic variations on survival for adult glioma in England and Wales. *Surgical Neurology*, 66 (3): 258-63

Tucker LA, Seljaas GT, Hager RL (1997). Body fat percentage of children varies according to their diet. *Journal of the American Dietetic Association*, 97 (9): 981–986

Twigg L, Barnard S, Mohan J (2006). Developing and evaluating small-area indicators of the neighbourhood social environment. Environment and Planning A, 38: 2173-2192

Twigg L, Moon G, Jones K (2000). Predicting small-area health-related behaviour: a comparison of smoking and drinking indicators. Social Science and Medicine, 50: 1109-1120

University of Essex, Institute for Social and Economic Research (2006). *British Household Panel Survey; Waves 1-14, 1991-2005* [computer file]. *2nd Edition.* Colchester, Essex: UK Data Archive [distributor], June 2006. SN: 5151

Unsworth, R. & Stillwell, J. (2004). Twenty-first Century Leeds: Geographies of a regional City. Leeds University Press, Table 3.1, p.51

Vandegrift D & Yoked T (2004). Obesity rates, income, and suburban sprawl: an analysis of US states. *Health and Place*, 10: 221–9

Vandongen R, Jenner DA, Thompson C, Taggart AC, Spickett EE, Burke V, Beilin LJ, Milligan RA, Dunbar DL (1995). A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10 to 12 year old children. *Preventive Medicine*, 24: 9-22

Veenstra G (2000). Social capital, SES and health: An individual-level analysis. *Social Science and Medicine*, 50: 619-629

Vickers D & Rees P (2007). Creating the National Statistics 2001 Output Area Classification. *Journal of the Royal Statistical Society*, Series A 170(2)

Vickers D (2007). Open Geodemographics: At the Real Life Methods (part of the ESRC National Centre for Research Methods) Geodemographics and the Social Sciences conference. http://www.reallifemethods.ac.uk/training/workshops/geodemographics/documents/vickers-geodemographics.pdf

Victoria CG, Barros FC, Lima RC, Horta BL, Wells J (2003). Anthropometry and body composition of 18 year old men according to duration of breast feeding. birth cohort study from Brazil. *BMJ*, 327: 901–904

Von Kries R, Koletzko B, Sauerwald T, von Mutius E, Barnert D, Grunert V, von Voss H (1999). Breast feeding and obesity: cross sectional study. *BMJ,* 319 (7203): 147–150

Von Kries R, Toschke AM, Koletzko B, Slikker W Jr (2002). Maternal smoking during pregnancy and childhood obesity. *American Journal of Epidemiology*, 156: 954-961

Voss LD, Metcalf BS, Jeffery AN, Wilkin TJ (2005). IOTF thresholds for overweight and obesity and their relation to metabolic risk in children (EarlyBird 20). *International Journal of Obesity*, 29 (11): advance online publication: doi:10.1038/sj.ijo.0803187

Wadsworth M, Marshall S, Hardy R, Paul A (1999). Breast feeding and obesity. Relation may be accounted for by social factors. *BMJ*, 319 (7224): 1576

Wake M, Hesketh K, Waters E (2003). Television computer use and body mass index in Australian primary school children. *Journal of Paediatrics and Child Health*, 39 (2): 130–134

Wallace R & Wallace D (1997). Resilience and persistence of the synergism of plagues: stochastic resonance and the ecology of disease, disorder and disinvestment in US urban neighborhoods. *Environment and planning A*, 29: 789–804

Wang H, Sekine M, Chen X, Kanayama H, Yamagami T, Kagamimori S (2006). Sib-size, birth order and risk of overweight in junior high school students in Japan: Results of the Toyama Birth Cohort Study. *Preventive Medicine*, In Press, Corrected Proof

Wang Y, Ge K, Popkin BM (2000). Tracking of body mass index from childhood to adolescence: a 6-y follow-up study in China. *The American Journal of Clinical Nutrition,* 72 (4): 1018–1024

Wang Y, Monteiro C, Popkin BM (2002). Trends of obesity and underweight in older children and adolescents in the United States, Brazil, China, and Russia. *The American Journal of Clinical Nutrition,* 75 (6): 971–977

Wang Y & Lobstein T (2006). Worldwide trends in childhood overweight and obesity. *International Journal of Pediatric Obesity*, 1(1): 11-25

Wardle J, Guthrie C, Sanderson S, Birch L, Plomin R (2001). Food and activity preferences in children of lean and obese parents. *International Journal of Obesity*, 25 (7): 971-977

Wardle J, Henning Brodersen N, Cole TJ, Jarvis MJ, Boniface DR (2006). Development of adiposity in adolescences: five yaer longitudinal study of an ethnically and socioeconomically diverse sample of young people in Britain. *BMJ*, 332: 1130-1135

Wardle J, Sanderson S, Guthrie CA, Rapoport L, Plomin R (2002). Parental feeding style and the inter-generational transmission of obesity risk. *Obesity Research*, 10 (6): 453-462

Wareham N (2007). Physical activity and obesity prevention. Obesity Reviews, 8 (suppl.1): 109-114

Warren JM, Henry CJ, Lightowler HJ, Bradshaw SM, Perwaiz S (2003). Evaluation of a pilot school programme aimed at the prevention of obesity in children. *Health Promotion International*, 18: 287–296

Weiss RI & Smith JA (2004). Legislative approaches to the obesity epidemic. *Journal of Public Health Policy*, 25(3-4):379-390

Wells JC, Stanley M, Laidlaw AS, Day JM, Stafford M, Davies PS (1997). Investigation of the relationship between infant temperament and later body composition. *International Journal of Obesity*, 21 (5): 400–406

Whelan A, Wrigley N, Warm D, Cannings E (2002). Life in a 'food desert'. *Urban Studies*, 39(11): 2083-100

Whincup PH, Gilg JA, Papacosta O, Seymour C, Miller GJ, Alberti KG, Cook DG (2002). Early evidence of ethnic differences in cardiovascular risk: cross sectional comparison of British SouthAsian and white children. *BMJ*, 324 (7338): 635-641

Whincup PH, Owen CG, Sattar N, Cook DG (2005). School dinners and markers of cardiovascular health and type 2 diabetes in 13-16 year olds: a cross sectional study. *BMJ*, 331: 1060-1061

Whitaker RC & Dietz WH (1998). Role of the prenatal environment in the development of obesity. *The Journal of Pediatrics*, 132: 768–776

Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH (1997). Predicting obesity in young adulthood from childhood and parental obesity. *The New England Journal of Medicine*, 337: 869–873

White E, Wilson A, Greene SA, Berry W, McCowan C, Cairns A, Ricketts I (1995). Growth screening and urban deprivation. *Journal of medical screening*, 2(3): 140-144

Wiecha JL, Sobol AM, Peterson GE, Gortmaker SL (2001). Household television access: Associations with screen time, reading and homework among youth. *Ambulatory Pediatrics*, 1: 244–251

Wilkinson P, Grundy C, Landon M, Stevenson S (1998). Chapter 11 in Gatrell A & Löytönen M (eds) (1998). GIS and health. Taylor & Francis, London, UK

Wilson JF (2000). Lunch eating behavior of preschool children. Effects of age, gender, and type of beverage served. *Physiology and Behavior*, 70: 27–33

Winkler E, Turrell G, Patterson C (2006). Does living in a disadvantaged area mean fewer opportunities to purchase fresh fruit and vegetables in the area? Findings from the Brisbane food study. *Health and Place*, 12: 306–19

Wolfe WS, Campbell CC, Frongillo EA Jr, Haas JD, Melnik TA (1994). Overweight schoolchildren in New York State: prevalence and characteristics. *American Journal of Public Health*, 84 (5): 807-813

Woods LM, Rachet B, Riga M, Stone N, Shah A, Coleman MP (2005). Geographical variation in life expectancy at birth in England and Wales is largely explained by deprivation. *Journal of Epidemiology and Community Health*, 59: 115-20

Wootan MG & Osborn M (2006). Availability of nutrition information from chain restaurants in the United States. *American Journal of Preventive Medicine*, 30(3):266-268

World Health Organisation (2003). Report of a joint WHO/FAO Expert Consultation on diet, nutrition and the prevention of chronic disease. *WHO technical report series*; 916. Geneva

World Health Organisation (2004a). Report of a WHO Consultation on Obesity: preventing and managing the global epidemic. WHO Technical Report Series; 894. Geneva. http://www.who.int/nutrition/publications/obesity/en/index.html (accessed Oct 2007)

World Health Organization (2004b). Global strategy on diet, physical activity and health. http://www.who.int/nutrition/publications/obesity/en/index.html (accessed Oct 2007)

World Health Organization Expert Consultation (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet*, 363 (9403): 157-163

World Health Organisation (2006). Fact Sheet No. 311: Obesity and overweight. http://www.who.int/mediacentre/factsheets/fs311/en/ (accessed Oct 2007)

Wrigley N (2002). 'Food Deserts' in British Cities: Policy Context and Research Priorities. *Urban Studies*, 39 (11): 2029–2040

Wrigley N, Warm D, Margetts B (2003). Deprivation, diet and food retail access: findings from the Leeds 'food desert' survey. *Environment and Planning A*, 35, 151-188

Wrigley N, Warm D, Margetts B, Whelan A (2002). Assessing the Impact of Improved Retail Access on Diet in a 'Food Desert': A preliminary Report. *Urban Studies*, 39 (11): 2061–2082

Yiannakouliasa N, Rowe BH, Svenson LW, Schopflocher DP, Kelly K, Voaklander DC (2003). Zones of prevention: the geography of fall injuries in the elderly. *Social Science and Medicine*, 57: 2065–73

Young LR & Nestle M (2002). The contribution of expanding portion sizes to the US obesity epidemic. *American Journal of Public Health*, 92: 246–249

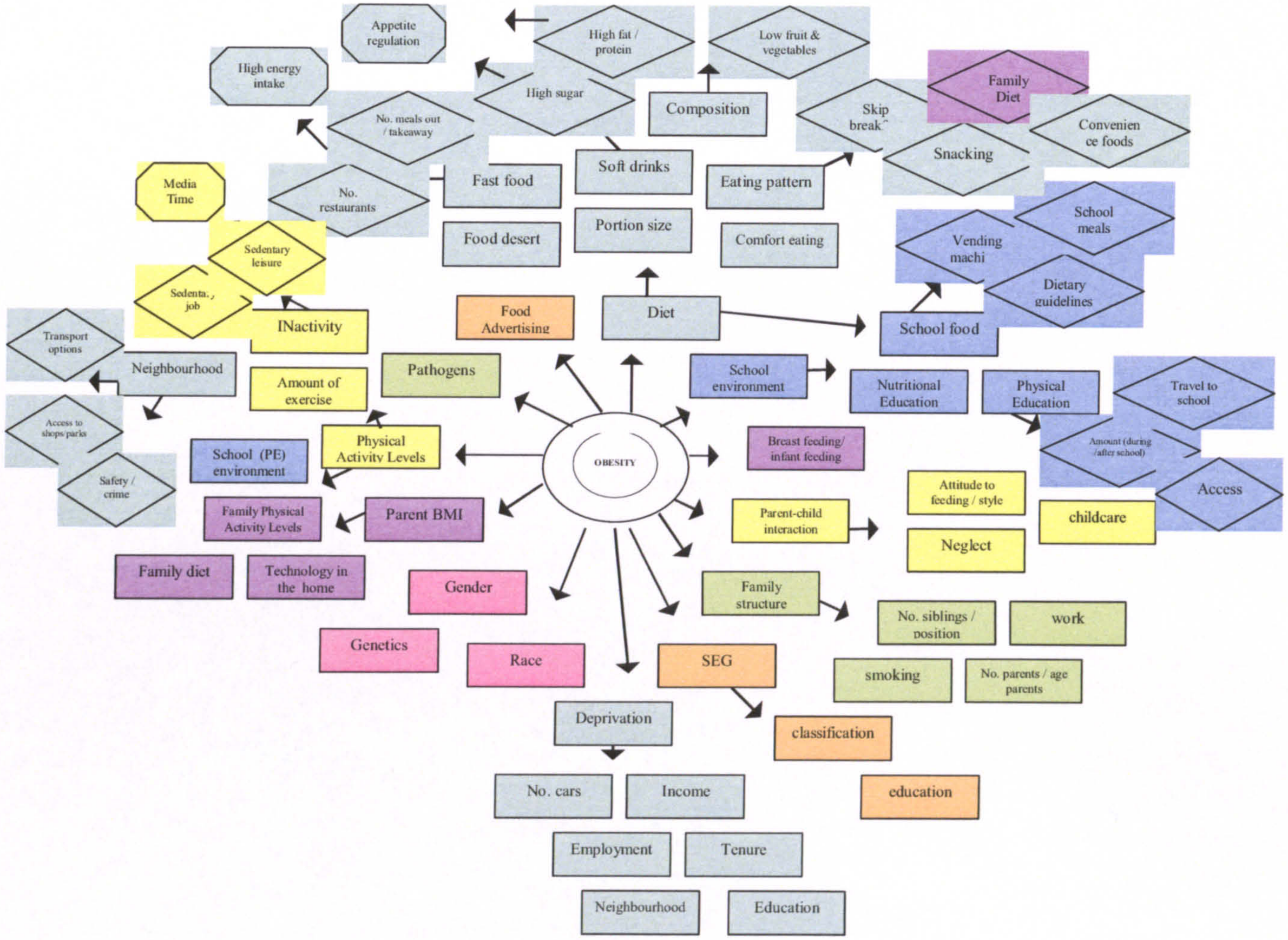
Young LR & Nestle M (2003). Expanding portion sizes in the US marketplace: Implications for nutrition counselling. *Journal of the American Dietetic Association,* 103(2):231-234

Zabinski MF, Daly T, Norman GJ, Rupp JW, Calfas KJ, Sallis JF, Patrick K (2006). Psychosocial correlates of fruit, vegetable, and dietary fat intake among adolescent boys and girls. *Journal of the American Dietetic Association*, 106(6), 814-821

Zhang LC & Chambers RL (2004). Small area estimates for cross-classifications. Journal of the Royal Statistical Society Series B – Statistical Methodology, 66: 479-496

Ziersch AM, Baum FE, MacDougall C, Putland C (2005). Neghbourhood life and social capital: the implications for health. *Social Science and Medicine*, 60: 71-86

Appendix A - Diagram of the causes of childhood obesity



APPENDIX B – SimObesity: structure of java code

Microsimulation package:

Class name	Function	Change
Dataloader	Loads data	Yes
Menu	Main entrance for the framework package	
Menulistener		
MicroModel	Runs the bulk of the code (i.e. the microsimulation model). Put in so it is easy to call the reweighter class.	Only if change reweighter class
MicroMonitor	Keeps an eye on what is happening	
ModelData	Stores values	Only if > 6 constraints
Reweighter	Algorithms 1 (reweighter) and 2 (integisation)	Yes
SaveResults	Creates output files (either postal sectors or wards)	Yes
TaskLoadData	Simply how to run tasks in different threads	
TaskRunModel	Simply how to run tasks in different threads	
TaskSaveData	Simply how to run tasks in different threads	

Database Management Package:

Class name	Function	Change
ConnectionManager	All are about getting data into and out of the database /	
DAL	/ managing data	
FindDatabase		
SQLTypeConverter	Converting from java type to SQL type	

Framework Package:

Class name	Function	Change
AppInitialisation	The basis for the whole idea.	
ApplicationInformation	Where the nuts and bolts are.	
ApplicationMenu	Allows us to load up packages without them	
ApplicationListeners	knowing about each other	
DisplaySettings		
FileManager		
GenericClassLoader		
MainMenuLoader		
MainScreen		
MainScreenListener		
MenuClassListener		
ResourceExtractor		
StartUp		

SharedObjects Package:

Class name	Function	Change
IApplicationInformation	Provides communications and interfaces	
IDAL	that are promised throughout the system.	
Imenu	Easy way to transport information around.	
ISQLTYpeConverter		
ITask		
ITaskThread		
Menu		
Messagebox		
Model		
NewTaskThread		

Relevant classes if want to change the code slightly:

- *How change from 476 SOAs to 2400 OAs or 100 postal sectors (output)?* SaveResults (or change the way the ZoneCode is coded to swap between OAs and SOAs)
- *How change number of constraints used?* Reweighter and Dataloader
- *Any impact (to code) from number of categories per constraint?* No. Array size is dynamic.
- *Where is number of iterations?* Reweigher
- *Where is break clause (i.e. if difference < 0.001 then stop)?* Reweighter

APPENDIX C - SimObesity: populations files

For HSE 2002 simulation:

weight	sex	AgeFull	nimd	SimTen	SimEth	SimTQ	ID
1.00	2	7	4	2	1	1	10101101
4.50	2	1	4	1	1	4	10107103
1.00	2	7	4	1	2	2	10111105
1.00	1	7	4	2	1	3	10115101
1.00	2	7	4	1	1	2	10120101
1.00	1	7	4	2	1	2	10127101
1.00	2	7	4	2	1	1	10127102
...							
1.00	2	7	5	1	1	2	82038103

17517 individuals (rows) in HSE population file.

For EFS 2005 simulation:

weight	SimHH	SimCars	Ten	hse2	Nofolks	ID
1.0	2	2	2	2	3	1
1.0	2	1	2	3	3	2
1.0	1	1	2	3	1	3
1.0	1	1	2	3	1	4
1.0	2	1	2	2	3	5
1.0	1	1	1	3	1	6
1.0	1	1	2	3	1	7
...						
1.0	2	1	2	2	2	6798

6798 individuals (rows) in EFS population file.

APPENDIX D – SimObesity: constraint files

For HSE 2002 simulation:

ZoneCode	male	female
00DAFAE01011264	701	737
00DAFAE01011265	689	802
00DAFAE01011266	700	757
00DAFAE01011267	730	806
...		
00DAGKE01011739	781	836

ZoneCode	0_2yrs	3_6yrs	7_9yrs	10_12yrs	13_15yrs	16_18yrs	adult
00DAFAE01011264	32.02	59.04	51.04	64.04	66.05	46.03	1119.78
00DAFAE01011265	38.95	49.93	37.95	37.95	47.94	59.92	1218.37
00DAFAE01011266	31.98	74.95	63.96	72.95	58.96	56.96	1097.25
00DAFAE01011267	50.00	63.00	56.00	68.00	62.00	55.00	1182.00
...							
00DAGKE01011739	48.00	83.00	79.00	78.00	74.00	64.00	1191.00

ZoneCode	IMD1_leastdep	IMD2	IMD3	IMD4	IMD5_mostdep
00DAFAE01011264	0	0	1438	0	0
00DAFAE01011265	0	1491	0	0	0
00DAFAE01011266	0	1457	0	0	0
00DAFAE01011267	0	0	0	1536	0
...					
00DAGKE01011739	0	0	0	0	1617

ZoneCode	Own	Rent_other
00DAFAE01011264	1066.26	371.74
00DAFAE01011265	1142.09	348.91
00DAFAE01011266	1381.90	75.10
00DAFAE01011267	922.63	613.37
...		
00DAGKE01011739	613.38	1003.62

ZoneCode	White	NotWhite
00DAFAE01011264	1433.99	4.01
00DAFAE01011265	1457.98	33.02
00DAFAE01011266	1391.91	65.09
00DAFAE01011267	1502.11	33.89
...		
00DAGKE01011739	1584.96	32.04

ZoneCode	Degree	AO	None_other	FT_student
00DAFAE01011264	122.08	455.32	588.41	272.19
00DAFAE01011265	339.54	477.36	461.38	212.71
00DAFAE01011266	340.77	458.69	354.76	302.79
00DAFAE01011267	154.00	441.00	642.00	299.00
...				
00DAGKE01011739	64.00	459.00	732.00	362.00

For EFS 2005 simulation:

ZoneCode	no_kids	with_kids
00DAFAE01011264	804.0	634.0
00DAFAE01011265	902.8	588.2
00DAFAE01011266	681.4	775.6
00DAFAE01011267	828.3	707.7
...		
00DAGKE01011739	817.0	800.0

ZoneCode	Own	Rent_other
00DAFAE01011264	1066.26	371.74
00DAFAE01011265	1142.09	348.91
00DAFAE01011266	1381.90	75.10
00DAFAE01011267	922.63	613.37
...		
00DAGKE01011739	613.38	1003.62

ZoneCode	No_car	Car_avail
00DAFAE01011264	345.41	1092.59
00DAFAE01011265	412.96	1078.04
00DAFAE01011266	123.82	1333.18
00DAFAE01011267	494.40	1041.60
...		
00DAGKE01011739	802.98	814.02

ZoneCode	detached	semi_terr	flat_oth
00DAFAE01011264	62.8	1348.6	26.6
00DAFAE01011265	143.0	1005.2	342.8
00DAFAE01011266	656.4	724.3	76.4
00DAFAE01011267	197.9	1192.0	146.1
...			
00DAGKE01011739	30.0	1080.9	506.1

ZoneCode	1 person	2 people	3_4people	5ormore
00DAFAE01011264	340.5	541.4	472.8	83.3
00DAFAE01011265	583.0	487.6	355.4	65.0
00DAFAE01011266	242.8	511.7	598.4	104.1
00DAFAE01011267	456.7	531.2	461.5	86.5
...				
00DAGKE01011739	582.4	525.0	410.3	99.3

476 SOAs (rows) in each file.

APPENDIX E – SimObesity: summary files

For HSE 2002:

area	male	female					
1	8260	9257					
area	0_2yrs	3_6yrs	7_9yrs	10_12yrs	13_15yrs	16_18yrs	19+
1	1424	1946	1545	1565	1530	1376	8131
area	least	Dep2	Dep3	Dep4	most		
1	2822	2638	3153	3817	5087		
area	own	rent					
1	12185	5332					
area	white	NotWhite					
1	15595	1922					
area	Degree	AO	None	Student			
1	1283	4512	1934	9788			

For EFS 2005:

area	NoKids	Kids		
1	4637	2161		
area	Own	Rent		
1	4773	2025		
area	NoCar	CarAvail		
1	1684	5114		
area	detached	semi_terr	flat_other	
1	1539	4100	1159	
area	1person	2people	34people	5ormore
1	1872	2472	1972	482

APPENDIX F – SimObesity: micro_config files

For HSE 2002 simulation:

Array	Identifier	ItemName
CONST1	ID	ZoneCode
CONST2	ID	ZoneCode
CONST3	ID	ZoneCode
CONST4	ID	ZoneCode
CONST5	ID	ZoneCode
CONST6	ID	ZoneCode
SUMMARY1	ID	Area
SUMMARY2	ID	Area
SUMMARY3	ID	Area
SUMMARY4	ID	Area
SUMMARY5	ID	Area
SUMMARY6	ID	Area
CONST1	TABLE	Const_sex
CONST2	TABLE	Const_agef
CONST3	TABLE	Const_IMD
CONST4	TABLE	Const_ten
CONST5	TABLE	Const_eth
CONST6	TABLE	Const_TQ
POPULATION	TABLE	HSE02_pop4
SUMMARY1	TABLE	Sum_sex
SUMMARY2	TABLE	Sum_agef
SUMMARY3	TABLE	Sum_IMD
SUMMARY4	TABLE	Sum_ten
SUMMARY5	TABLE	Sum_eth
SUMMARY6	TABLE	Sum_TQ

For EFS 2005 simulation:

Array	Identifier	ItemName
CONST1	ID	ZoneCode
CONST2	ID	ZoneCode
CONST3	ID	ZoneCode
CONST4	ID	ZoneCode
CONST5	ID	ZoneCode
SUMMARY1	ID	Area
SUMMARY2	ID	Area
SUMMARY3	ID	Area
SUMMARY4	ID	Area
SUMMARY5	ID	Area
CONST1	TABLE	Const_HH
CONST2	TABLE	Const_car
CONST3	TABLE	Const_ten
CONST4	TABLE	Const_hse
CONST5	TABLE	Const_NoFolks
POPULATION	TABLE	EFS_pop3
SUMMARY1	TABLE	Sum_HH
SUMMARY2	TABLE	Sum_car
SUMMARY3	TABLE	Sum_ten
SUMMARY4	TABLE	Sum_hse
SUMMARY5	TABLE	Sum_NoFolks

APPENDIX G - Definitions and categorisation of the simulated obesogenic variables

HSE variables	Definition	Categories
SEG	Social class of household reference person	High (I, II, IIIN); Low (IIIM, IV, V)
urbanisation	Degree of urbanisation of residential location. Urban: inner city / other dense urban/town centre; suburban: suburban residential (city/large town outskirts); rural: rural residential / village centre, or rural agricultural with isolated dwellings or small hamlets	Urban; Suburban; Rural
Transport	This area has good local transport	Agree (good transport); Disagree (bad transport)
Leisure facilities	This area has good leisure things for people	Agree (good leisure facilities); Disagree (bad leisure facilities)
Access supermarket	Ease of getting to supermarket	Easy; Difficult
Teenagers	Problem of teenagers hanging around	Problem; Not a problem
Vandals	Problem of vandalism, graffiti or deliberate damage	Problem; Not a problem
Fruit and vegetables	Total number of portions of fruit and vegetables eaten yesterday	None; 1-4 portions; 5 or more portions
Child physical activity	Time spent last week in activities (children only)	None; <3 hours (low); 4-7 hours (moderate); 7 or more hours (high)
Adult physical activity	Total days per week active for at least 30 minutes (adults only)	None; <2 days pw (low); 3-4 days pw (moderate); 5 or more days pw (high)
EFS variables		
Food Expenditure	Amount household spent on food in last two weeks	< £60 (low); > £60 (high)
Food Expenditure adjusted for size	Amount household spent on food in last two weeks divided by number of household members	< £25 (low); > £25 (high)
Computer	Home computer in household	Yes; No
TV	Number of TVs in household	0 or 1; 2 or more
Internet access	Internet connection in household	Yes; No
Source income	Main source of household income	Wage, self employed, pension, investments, other; benefits
School meals	Amount of money spent on school meals in the last week	None; Some
Household income	Amount of household income	Low <£200 pw (equiv £10400pa); Medium; High >£650 pw (equiv £33800pa)