Statistical and Policy Analysis of Large-scale Public Health Interventions

Nicholas Jonathan Preval

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin,
New Zealand.

18 August 2014

ABSTRACT

Background: The outcome evaluation of large-scale public health interventions is an important part of the rational policy cycle. Timely outcome evaluation can inform decisions by policy makers on whether to continue, modify or discontinue an intervention. Cost-benefit analysis-based outcome evaluations are particularly valuable, producing results that allow easy comparisons of interventions that have multiple benefits and/or externalities. It is important that the evaluation of benefits and costs be based on sound statistical principles, and correct characterisation of the key features of an intervention is vital if results are to be generalised in order to inform future interventions.

The key element of cost-benefit analyses of large-scale public health interventions is the evaluation of health benefits. Health data present conceptual and analytical challenges due to the characteristics of cost data and the complexities of assigning dollar values to changes in the rates of events such as mortality. Health data from New Zealand's Warm Up New Zealand: Heat Smart (WUNZ:HS) evaluation, which retrofitted homes with improved insulation and space heating, provided the case study that is the basis of this thesis, enabling the exploration of suitable statistical and econometric approaches to the evaluation of health benefits. A related comparison of the WUNZ:HS programme with Australia's controversial Home Insulation Programme (HIP) was an opportunity to consider the key characteristics of a successful large-scale public health intervention of this type, an evaluation which complements the outcome evaluation of health cost data.

Aims: The primary aim of this thesis was to examine the methodological issues around the analysis of administrative data in large-scale public health intervention evaluations, using data available from the WUNZ:HS evaluation as the basis of a case study. The methodologies examined aimed to accurately quantify the health benefits or costs of a programme in dollar terms.

Secondary to this aim was a policy analysis-based comparison of the WUNZ:HS programme with the Australian Home Insulation Programme (HIP). This comparison informed retrofit programme specific recommendations and also a discussion about the generalisability of the results of the primary analysis and the importance of intervention design and implementation.

Methods: Mortality data were analysed using Cox Proportional Hazard models. Statistically significant changes in mortality were then valued using a "Value of a Statistical Life"-based methodology. Changes in hospitalisation costs and pharmaceutical use costs were modelled at the household level using difference-in-difference-based fixed

effects models. The WUNZ:HS HIP policy comparison was conducted using a combination of a literature review and a key informant interview.

Results: Individuals aged 65 and over who occupied a household that received a WUNZ:HS insulation retrofit and who had had a baseline circulatory hospitalisation had a Hazard Ratio for mortality of 0.673 (95% CI 0.535 - 0.847) ($p<0.001$) relative to comparable control group individuals. Individuals aged 65 and over who occupied a household that received a WUNZ:HS insulation retrofit and who had had a baseline respiratory hospitalisation had a Hazard Ratio of 0.8 (95% CI 0.635 – 1.008), ($p=0.058$) relative to comparable control group individuals. Annual household level mortality-related benefits of insulation can be valued at approximately $1,120, although this figure varies greatly depending on modelling assumptions. Heating did not have a statistically significant Hazard Ratio for vulnerable treatment vs. control individuals.

Households that received insulation had reductions in monthly all-cause hospitalisation costs relative to their matched controls of $3.18 per month ($p<0.001$) (95% CI 1.83 - 4.54) in the favoured model. There was no evidence of an impact of heating retrofits on hospitalisation costs.

Households that received insulation had reductions in monthly all-cause pharmaceutical use costs relative to matched control group households of $2.01 ($p<0.001$) (95% CI 1.52 – 2.50) in the favoured model. Households that received heating demonstrated reductions in monthly all-cause pharmaceutical use costs of $3.01 ($p<0.001$) (95% CI 1.43 – 4.59).

The central estimate of the combined annual household health benefits of insulation was $859 using a ($p<0.05$) threshold for inclusion of benefits and $1,173 using a ($p<0.1$) threshold. The combined annual household level health benefits of heating were approximately $41 using either threshold. When health results were incorporated into total cost-benefit figures the benefit cost ratio for the programme was estimated at 6.4:1, driven largely by health benefits, indicating strong evidence for the value of the WUNZ:HS programme. This result was highly sensitive to the assumptions used to estimate mortality-related benefits.

Comparison of the WUNZ:HS and HIP programmes suggested that differences in programme outcomes were driven by differences in the characteristics of retrofitting organisations, the financial contribution of households towards retrofits, the experience of implementing agencies, the relative speeds of programme roll-outs, and differences in policy officials' commitment to economic stimulus vs. other potential programme co-benefits. Some of these differences can be linked to the wider policy contexts, particularly the impact of housing research in the New Zealand context.

Conclusion: Exploration of statistical and econometric methods for the evaluation of large-scale public health interventions, using the WUNZ:HS evaluation dataset as a case study, produced plausible and consistent results, and demonstrated the importance of modelling assumptions. Policy analysis-based comparison of WUNZ:HS and Australia's HIP demonstrated the importance of characterising key elements of an intervention if results are to be generalisable.

ACKNOWLEDGEMENTS

I would first and foremost like to acknowledge my beautiful life-partner Mary-Lou Harris. The past few years have been difficult and I don't think I would have completed this thesis without your support. Thanks too to our families and friends, who have given us a great deal of financial and moral support.

Thanks to the University of Otago. The University of Otago Postgraduate Scholarship that funded my first three years of study was incredibly valuable and I could not have completed my thesis without it. Thanks also to the administrative staff of the University of Otago, Wellington who ensure students have access to every available resource.

I would like to thank my supervisors, in alphabetical order: Arthur Grimes, Philippa Howden-Chapman and Michael Keall. It was a pleasure to share the room with such active and powerful minds.

Thanks are due to the Ministry of Economic Development (now Ministry of Business Innovation and Employment), who funded the Warm Up New Zealand: Heat Smart evaluation and my extension of the evaluation. Thanks to the Energy Efficiency and Conservation Authority, who provided data and expertise during the evaluation and particularly to Jenny Lackey.

Thanks are due to the Ministry of Health, Quotable Value, and PHARMAC (Pharmaceutical Management Agency) for their data and expertise. Particular thanks to Chris Lewis of the Ministry of Health and Chris Peck of PHARMAC.

Thanks are due to the members of the Warm Up New Zealand: Heat Smart evaluation team: Lucy Telfar-Barnard (my Stata mentor), Tim Denne, Steve Bond-Smith, Prof Richard Arnold and Chris Young.

Thanks are due to the whole *He Kainga Oranga* team, particularly, Helen Viggers, Nevil Pierce and Michael Baker. Thanks also to Ralph Chapman who first introduced me to the team back in 2007.

Thanks to my fellow PhD travellers Marie Russell, Ramona Tiatia, Caroline Shaw, and all the rest – you were all shining examples and kept my eyes on the prize.

Thanks also to Idea Services, my employer during the last 13 years, Matthew and Brendan who I have supported during many of those years, their families, and of course my fellow staff members Teresa Trask, Regan Austin, and a rotating cast of many more. Idea has been an excellent place to work as a student and I am grateful for the financial flexibility that this job has provided.

Finally, I must also thank my haematologist Professor John Carter for his kind support and proactive approach, which ensured that I received the best treatment currently available when I was first diagnosed with chronic myeloid leukaemia in early 2013. I must also acknowledge the researchers who developed Imatinib, a targeted anti-cancer drug that transformed the prognosis of my illness. Naturally, as a student of health economics, I also thank PHARMAC and the New Zealand healthcare system.

STATEMENT OF AUTHOR CONTRIBUTION

This thesis presents some paragraphs in Subsections 3.3.8 and 3.3.3.1 that are direct quotes from one of the four reports that I directly contributed to as part of my role in the Warm Up New Zealand (WUNZ:HS) evaluation team (the Group). Such paragraphs are clearly identified with italics and appropriate referencing, and are my own work. Methods that I employed when working on the WUNZ:HS evaluation are also replicated in this thesis, typically in a modified or expanded form. I carried out and wrote up these analyses alone when working on the original reports but discussion and consultation with other Group members naturally occurred both before and after analysis. Table 1 sets out the WUNZ:HS evaluation reports, the authors, states my role in each report and notes the material from the report that contributes to this thesis.

Table 1: Statement of author contribution

Report citation	Candidate contribution to report	Contribution of material in report to this thesis	Year	Peer review
Denne, T. and S. Bond-Smith (2011). Impacts of the NZ Insulation Fund on Industry & Employment.	• N/A	• Results contribute to benefit:cost ratios presented in Grimes, Denne et al. (2011): these results were used in Chapter Seven of this thesis.	2011	Reviewed by EECA and MED prior to completion, No external peer review.
Grimes, A., C. Young, R. Arnold, T. Denne, P. Howden-Chapman, N. Preval and L. Telfar-Barnard (2011). "Warming Up New Zealand: Impacts of the New Zealand Insulation Fund on Metered Household Energy Use."	• Participated in regular meetings to discuss method etc.	• Fixed effects model used to analyse energy use was initial inspiration for models used to analyse health cost data in other reports (Telfar Barnard, Preval et al. 2011, Preval 2012) and later in this thesis (Chapters Five and Six)	2011	Reviewed by EECA and MED prior to completion, No external peer review.
Grimes, A., T. Denne, P. Howden-Chapman, R. Arnold, L. Telfar-Barnard, N. Preval and C. Young (2011). Cost Benefit Analysis of the Warm Up New Zealand: Heat Smart Programme.	• Contributed analyses of health data and imputed health benefits presented in Telfar Barnard, Preval et al. (2011)	• Benefit:cost ratios and calculations were used as basis for analysis in Chapter Seven.	2011	Reviewed by EECA and MED prior to completion, No external peer review.

Telfar Barnard, L., N. Preval, P. Howden-Chapman, R. Arnold, C. Young, A. Grimes and T. Denne (2011). The impact of retrofitted insulation and new heaters on health services utilisation and costs, pharmaceutical costs and mortality: Evaluation of Warm Up New Zealand: Heat Smart.	• Assigned costs to mortality model output. • Developed approach to assessing Pharms data. • Analysed household level hospitalisation and pharmaceutical use costs. • Calculated imputed health benefits from previous studies. • Summarised benefits at household level.	• Datasets used are basis of thesis. Dataset design and initial construction were work of other team-members primarily Lucy Telfar-Barnard prior to my involvement. • Mortality *cost* modelling in this thesis was based on my approach here (Section 4.3). • Household level hosp cost (Chapter Five) and pharms use cost (Chapter Six) models derived from my work here.	2011	Reviewed by EECA and MED prior to completion, No external peer review.
Preval, N. (2012). An extended analysis of the impacts of retrofitted insulation and new heaters on health services utilisation and costs, and mortality: Evaluation of Warm Up New Zealand: Heat Smart Extension.	• Re-analysed mortality data with more suitable model and additional year of data. • Reran household level hospitalisation cost models with additional year of data.	• I obtained an additional year of mortality and hospitalisation data which are used in this thesis. • Mortality model used in thesis (Sections 4.2 and 4.3) was derived from approach in this report.	2012	Reviewed by EECA and MED prior to completion, No external peer review.

PREFACE

I re-entered the academic world, almost by accident, at the age of 27, having completed an undergraduate philosophy degree many years previously and then spent several years working with people with an intellectual disability. During this time I became interested in the subject of peak oil. This lead me to study Environmental Studies at Victoria University of Wellington, and then to complete a Master of Environmental Studies under the tutelage of Associate Professor Ralph Chapman. My thesis topic, a cost-benefit analysis of a major randomised controlled trial of heating retrofits, brought me into contact with He Kainga Oranga, the Housing and Health Programme of the Public Health Department of the University of Otago, Wellington. I worked with Nevil Pierse and Philippa Howden-Chapman and many others, had my first taste of statistical analysis and contracted Philippa's contagious fascination with the connections between housing and health.

When I completed my thesis I briefly worked as a policy analyst for a private research organisation and then returned to the fold at the University of Otago, Wellington as a research assistant for the newly formed Centre for Sustainable Cities. While there I studied epidemiology as part of the Masters of Public Health Programme, which confirmed my interest in the topic and encouraged me to pursue a PhD with a statistical focus. The opportunity arose to be part of a group which was carrying out a government commissioned evaluation of a large-scale public health intervention, the $347 million Warm Up New Zealand: Heat Smart programme. This seemed like a logical extension of my Master's thesis and a great opportunity to be involved in high impact research with real policy consequences, so I came aboard as part of the group charged with assessing any health benefits or changes in health costs resulting from the programme.

The WUNZ:HS evaluation was a challenging and interesting process to be part of. Many serious hours were spent in deep concentration as data were collected and cleaned, analyses planned and reports written. The key conclusion resulting from the group's analyses was that the programme had a highly favourable ratio of benefits to costs. A subsequently commissioned evaluation that I carried out confirmed the key results driving this conclusion. Gratifyingly, our results appeared to have real influence on the policy process, both in modifications of the programme and then in the extension of the programme in the form of the Warm Up New Zealand: Healthy Homes programme.

In the one and a half years since these evaluations were completed I have extended the work that I carried out as part of the group, treating it as a case study of large-scale public health intervention outcome evaluation, attempting a more thorough analysis of the available health data and exploring the generalisability of the results of the evaluation by contrasting the WUNZ:HS programme with a much less successful programme, the Australian Home Insulation Programme. During the same period of time Mary-Lou and I purchased our first home, learned that I had chronic myeloid leukaemia (which potentially explained much previous ill-health) and adopted two poodles, refugees

from the Christchurch earthquake and thus part of the larger New Zealand housing and health story.

It has been a busy, challenging and interesting four years, and I will take with me many new skills and valuable experiences. Having completed my thesis I look forward to the next chapter ☺

Table of Contents

ABSTRACT iii

ACKNOWLEDGEMENTS vi

STATEMENT OF AUTHOR CONTRIBUTION viii

PREFACE x

TABLE OF FIGURES xix

TABLE OF TABLES xx

LIST OF ABBREVIATIONS xxv

CHAPTER 1. INTRODUCTION 1

1.1. What is a large-scale public health intervention? 1

1.2. What is meant by evaluation? 2

1.3. Why should the outcomes of large-scale public health interventions be evaluated? 3

1.4. The New Zealand context 5

1.5. Aim of thesis 6

1.6. Structure of thesis 7

CHAPTER 2. THE CASE STUDY: WARM UP NEW ZEALAND: HEAT SMART 8

2.1. Key connections between housing, insulation, heating and health 9

2.1.1. Insulation 9

2.1.2. Heating 10

2.1.3. Insulation, heating and the rebound effect 10

2.1.4. Insulation, heating and health 11

2.2. Home insulation and heating in New Zealand: a brief background 15

2.2.1. EECA and housing: small scale public health interventions 16

2.2.2. WUNZ:HS: programme details 19

2.3. WUNZ:HS evaluation 20

CHAPTER 3. WUNZ:HS EVALUATION: DESIGN, DATASET CONSTRUCTION AND DESCRIPTION 22

3.1. WUNZ:HS study design 23

3.1.1. Administrative data 24

3.1.2. Quasi-experimental design 25

3.1.3. Limitations of the WUNZ:HS evaluation design 26
3.2. Dataset creation 28
3.2.1. Ministry of Health Data 31
3.2.2. The MED extension 32
3.2.3. Data summary 35
3.3. Dataset description 36
3.3.1. NHI 36
3.3.1.1. Age 38
3.3.1.2. Number of occupants 40
3.3.1.3. Ethnicity 42
3.3.1.4. Sex 44
3.3.1.5. All causes mortality 45
3.3.1.6. Accuracy of NHI address field 46
3.3.2. NMDS data 48
3.3.2.1. Hospitalisation categories and exclusions 49
3.2.2.2. Cost variables in the NMDS 52
3.3.2.3. Possible issues for NMDS data 53
3.3.3. Pharms data 53
3.3.3.1. Pharmaceutical cost data 54
3.3.3.2. Pharmaceutical categories 56
3.3.3.3. Possible issues for Pharms data 57
3.3.4. QV data 59
3.3.4.1. Possible issues for QV data 63
3.3.5. EECA data 63
3.3.6. Metered Energy Use Data 64
3.3.7. Climate data 66
3.3.8. Imputed benefits: GP data, days off work data and days off school data 68
CHAPTER 4. MODELLING AND VALUING CHANGES IN MORTALITY 75
4.1. Approach to modelling in this thesis 76
4.1.1. Initial modelling decisions 76

4.1.2. Model specification 77
4.1.3. Exposures of interest 77
4.1.3.1. Timing of exposures of interest 78
4.1.4. Selection of potential confounders and model construction 79
4.1.4.1. Transformation of covariates 81
4.1.4.2. Interaction terms 82
4.1.4.3. Model assumptions, leverage and goodness-of-fit 83
4.1.5. Accounting for clustering and matching 85
4.1.6. Approach to statistical significance 86
4.1.7. Policy relevant stratification 87
4.1.8. Adjustment of costs and benefits 87
4.2. Modelling changes in mortality 88
4.2.1. Model specification 89
4.2.2.Confounding and mortality 90
4.2.2.1. Age and mortality 91
4.2.2.2.Sex and mortality 92
4.2.2.3. Ethnicity and mortality 92
4.2.2.4. Socio-economic status and mortality 93
4.2.2.5. Health status and mortality 94
4.2.2.6. Housing characteristics 96
4.2.2.7. Climate, location and mortality 97
4.2.3. Initial multiple regression models 98
4.2.4. Transformation of continuous variables 98
4.2.5. Interaction terms 100
4.2.6. Finalising model 102
4.2.6.1. Testing proportional hazards assumption 103
4.2.6.2. Assessing leverage, influence, outliers 104
4.2.6.3. Assessment of global goodness-of-fit 106
4.2.6.4. Stratification by Community Services Card status 108
4.2.6.5.Summary 111

4.3. Valuing changes in mortality 113
4.3.1. The value of a statistical life in New Zealand 113
4.3.2. Value of a statistical life year: age and health 117
4.3.3. Combining the value of a statistical life year with change in mortality hazard 118
4.3.3.1. Valuing change in mortality in the circulatory cohort 118
4.3.3.2. Valuing changes in mortality in the respiratory cohort 125
4.3.3.3. Accounting for individuals in both cohorts 126
4.3.4. Summary 127
CHAPTER 5. MODELLING CHANGES IN HOSPITALISATION COSTS 128
5.1. Individual level analysis 130
5.2. Household level analysis: Original model 133
5.2.1. Model specification 133
5.2.2. Model construction and approach to extreme values 134
5.2.3. Confounding 138
5.2.3.1. Occupant characteristics 138
5.2.4. Initial multiple regression model 140
5.2.5. Transformation of continuous variables and interaction terms 143
5.2.6. Assessing the impact of extreme value removal 143
5.2.7. Model specification and global goodness-of-fit measures 145
5.2.8. Stratification by Community Services Card status 146
5.2.9. Summary 148
5.3. Household level analysis: Extended model 149
5.3.1. Model specification 149
5.3.2. Model construction 150
5.3.3. Confounding 150
5.3.3.1. Occupant characteristics 151
5.3.3.2. Season and location 151
5.3.4. Multiple regression model 152
5.3.5. Assessing the impact of extreme value removal 154
5.3.6. Model specification and global goodness-of-fit measures 156

5.3.7. Stratification by Community Services Card Status 156
5.3.8. Summary 157
5.4. Chapter summary 159
CHAPTER 6. MODELLING CHANGES IN PHARMACEUTICAL USE COSTS 160
6.1. Original model 161
6.1.1. Model specification 161
6.1.2. Model construction and approach to extreme values 162
6.1.3. Confounding 164
6.1.3.1. Occupant characteristics 164
6.1.4. Initial multiple regression model 164
6.1.5. Transformation of continuous variables and interaction terms 166
6.1.6. Assessing the impact of extreme value removal 166
6.1.7. Model specification and global goodness-of-fit measures 168
6.1.8. Stratification by Community Services Card status 168
6.1.9. Summary 170
6.2. Household level analysis: Extended model 171
6.2.1. Model specification 171
6.2.2. Model construction 171
6.2.3. Confounding 172
6.2.3.1. Occupant characteristics 172
6.2.3.2. Climate and location 173
6.2.4. Multiple regression model 174
6.2.5. Assessing the impact of extreme value removal 178
6.2.6. Model specification and global goodness-of-fit measures 180
6.2.7. Stratification by Community Services Card Status 181
6.2.8. Summary 182
6.3. Chapter summary 183
CHAPTER 7. COMBINING HEALTH BENEFITS 184
7.1. Comparison of health benefits 185
7.2. Annual household health benefits 188

7.2.1. Cost-benefit analysis 191
7.2.2. Summary 192
CHAPTER 8. POLICY COMPARISON OF WUNZ:HS AND AUSTRALIA'S HIP 194
8.1. HIP: Context and overview 196
8.2. Comparison of WUNZ:HS and HIP 201
8.2.1. Agenda-setting, policy formation and decision making 203
8.2.1.1. Entry to programmes and installer characteristics 203
8.2.1.2. Rebate structure 206
8.2.1.3. Product inclusion and installation methods 207
8.2.1.4. Differences in policy context 209
8.2.1.5. Government structure 211
8.2.1.6. Differences in programme formation and organisational structure 212
8.2.1.7. Consultation, risk assessment and independent policy advice 213
8.2.2. Policy implementation 214
8.2.3. Policy evaluation 215
8.2.4. Conclusion and recommendations 216
CHAPTER 9. DISCUSSION AND CONCLUSION 218
9.1. Discussion 219
9.1.1. Key results in the context of housing, heating and health literature 219
9.1.2. Thesis in broader context of large-scale public health intervention evaluation research 221
9.1.3. Strengths of the WUNZ:HS evaluation 221
9.1.4. Limitations of the WUNZ:HS evaluation 222
9.1.5. Lessons of from the WUNZ:HS HIP policy comparison 227
9.1.6. Future research directions 228
9.2. Conclusion 231
REFERENCES 233
APPENDIX 1. Ethics Committee Approvals 244
1.1 . Ethics approval: WUNZ:HS health evaluation 244
1.2 . Ethics approval: Jenny Lackey interview 246

1.2.1. Information form 248
1.2.2. Consent form 250
APPENDIX 2. Additional Tables 251
APPENDIX 3. Evaluation of Warm Up New Zealand: Heat Smart Extension 262

TABLE OF FIGURES

Figure 1. Dataset creation 35
Figure 2: Removal of people aged 101 or more as at Jan 2008 39
Figure 3: Age structure of treatment and control group as at 2008 and 2006 Census population 39
Figure 4: Removal of households that had 16 or more occupants as at Jan 2008 42
Figure 5: Pharmaceutical Data Pathway 58
Figure 6: Climate Zones 66
Figure 7: Model building process 84
Figure 8: Cumulative hazard function graphed against Cox Snell residuals 108
Figure 9: Kernel density plots, hosp_cost_diff, with and without removal of extreme values 136
Figure 10: Dataset construction, original hospitalisation costs models 137
Figure 11: Histograms showing distribution of occ_diff and age_diff, "after" period 139
Figure 12: Dataset construction, extended hospitalisation costs model 150
Figure 13: Dataset construction, original pharmaceutical cost difference models 163

TABLE OF TABLES

Table 1: Statement of author contribution viii
Table 2 Funding available under WUNZ:HS 19
Table 3 QV Variables 29
Table 4. Match of treatment and control dwellings, by count and percentage 30
Table 5. Administrative health data 32
Table 6. Matching based on MoH demographic data 34
Table 7: Age structure of treatment group and control group as at Jan 2008 and 2006 Census Population 39
Table 8: Ethnic composition of treatment group and control group as at Jan 2008 and 2006 Census Population 44
Table 9: Sex composition of treatment group and control group as at Jan 2008 and 2006 Census Population 45
Table 10: Proportion of NMDS records that are mismatched based on comparison of NMDS Domicile Code and QV matched Domicile Code by year, treatment group, ethnicity, sex and age-group 47
Table 11: Key Hospitalisation Exclusions 51
Table 12: Hospitalisation Categories. 52
Table 13: Pharmaceutical Categories 57
Table 14. QV data available for each study home 59
Table 15: Distribution of "dwelling health risk" variable for individuals from treatment and control group households 61
Table 16: Datasets linkable using address data 62
Table 17: Key EECA variables 64
Table 18: *Statistically significant health-related savings documented in previous H&HRP analyses* 70
Table 19: *Demographic profile of treatment households as at July 2009* 71
Table 20: *Additional imputed yearly benefits per 1000 households assuming 3.61 occupants per household and age structure of treatment group as at July 2009* 73
Table 21: Summary of insulation and heating retrofits for treatment homes with suitable health data 77
Table 22: Initial multiple regression mortality models, circulatory and respiratory cohorts 98
Table 23: Fractional polynomial transformation of initial circulatory and respiratory models 99
Table 24: Testing interaction terms for the circulatory and respiratory models 101
Table 25: Preliminary final models, circulatory and respiratory cohorts 102

Table 26: Preliminary final models, circulatory and respiratory cohorts, addition of time-varying covariates 104
Table 27: Dfbetas for coefficients of exposures of interest, circulatory and respiratory cohorts 105
Table 28: Quantiles of risk, circulatory and respiratory cohorts 107
Table 29: Results for circulatory model stratified by Community Services Card status 108
Table 30: Results for respiratory model stratified by Community Services Card status 109
Table 31: VSLY values for various combinations of VSL, discount rate and time period 115
Table 32: Values of change in mortality per treatment group individual who avoided mortality as a result of receiving insulation (values of J): circulatory cohort 120
Table 33: Total value of change in mortality per treatment group household due to mortality avoided during first year post-treatment (values of K): circulatory cohort 121
Table 34: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): circulatory cohort 123
Table 35: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): circulatory cohort: results by CSC status of household 124
Table 36: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): respiratory cohort 125
Table 37: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): respiratory cohort: results by CSC status of household 126
Table 38: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): respiratory cohort: adjusted to avoid double counting 127
Table 39: Summary measures, all study individuals' post-treatment average annual hospitalisation cost data 130
Table 40: Summary statistics, dependent cost variables, with and without exclusions 135
Table 41: Simple regression analyses, occupant-diff and age_diff variables, treatment/control pairings with values below 1st and above 99th percentiles excluded 139
Table 42: Initial multiple regression models, hospitalisation cost difference models, treatment/control pairings with values below 1st and above 99th percentiles excluded 141
Table 43: Hospitalisation cost diff models, no exclusions and treatment/control pairings with values below the 5th or above the 95th percentiles excluded 144
Table 44 All-cause hospitalisation cost difference model with robust clustered standard errors and bootstrapped standard errors, treatment/control pairings with values below 1st and above 99th percentiles excluded 145

Table 45: Hospitalisation cost models with treatment stratified by Community Services Card status, treatment/control pairings with values below 1st and above 99th percentiles excluded 146
Table 46: Simple regression analyses, occupant-diff and age_diff variables, extended model, months with values below 1st and above 99th percentiles excluded 151
Table 47: Simple regression analyses: winter and monthly average temperature, extended model, months with values below 1st and above 99th percentiles excluded 152
Table 48: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded 153
Table 49: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded, winter interaction terms added 154
Table 50: All-cause hospitalisation cost difference, extended model, no exclusions 155
Table 51: All-cause hospitalisation cost difference extended model with robust clustered standard errors and bootstrapped standard errors, months with values below 1st and above 99th percentiles excluded 156
Table 52: All-cause hospitalisation cost difference, extended model, treatment stratified by Community Services Card status, months with values below 1st and above 99th percentiles excluded 157
Table 53: Summary statistics, dependent pharmaceutical use cost difference variables, with and without exclusions 162
Table 54: Simple regression analyses, occupant_diff and age_diff variables, treatment/control pairings with values below 1st and above 99th percentiles excluded 164
Table 55: Initial multiple regression models, original pharmaceutical cost difference models, , treatment/control pairings with values below 1st and above 99th percentiles excluded 165
Table 56: Cost difference models, no exclusions and , treatment/control pairings with values below the 5th or above the 95th percentiles excluded 167
Table 57 All-cause pharmaceutical use cost difference model with robust clustered standard errors and bootstrapped standard errors, treatment/control pairings with values below 1st and above 99th percentiles excluded 168
Table 58: Hospitalisation cost models stratified by Community Services Card status, treatment/control pairings with values below 1st and above 99th percentiles excluded 169
Table 59: Simple regression analyses, occupant-diff and age_diff variables, extended model, months with values below 1st and above 99th percentiles excluded 173
Table 60: Simple regression analyses: winter and monthly average temperature, extended mode, months with values below 1st and above 99th percentiles excluded l 173
Table 61: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded 174

Table 62: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded, winter interaction terms added 176
Table 63: All-cause pharmaceutical use cost difference, extended model, no exclusions and months above 5th or below 95th percentiles excluded 178
Table 64: Extended model with robust clustered standard errors and bootstrapped standard errors, months with values below 1st and above 99th percentiles excluded..... 180
Table 65: All-cause hospitalisation cost difference, extended model, stratified by Community Services Card status, months with values below 1st and above 99th percentiles excluded..... 181
Table 66 Summary table of model coefficients ordered by treatment type 185
Table 67 Annual health-related benefit per household of WUNZ:HS insulation and heating retrofits for all households 189
Table 68: Annual health-related benefit per household of WUNZ:HS insulation and heating retrofits for CSC households and non-CSC households 190
Table 69: Combined annual health benefits of insulation, all households, for a variety of mortality cost modelling assumptions 191
Table 70 WUNZ:HS evaluation benefit:cost ratios for a variety of modelling assumptions 192
Table 71: WUNZ:HS and HIP key features 202
Table 72: Age structure, circulatory and respiratory cohorts 251
Table 73: Estimated Hazard Ratios for Time to Death for one year increase in age, circulatory and respiratory cohorts 251
Table 74: Sex structure, circulatory and respiratory cohorts..... 252
Table 75: Estimated Hazard Ratios for Time to Death by sex, circulatory and respiratory cohorts..... 252
Table 76: Ethnicity structure, circulatory and respiratory cohorts, as at first month *after*. 253
Table 77: Estimated Hazard Ratios for Time to Death for ethnicity, circulatory and respiratory cohorts 253
Table 78: NZdep2006 quintiles, circulatory and respiratory cohorts..... 254
Table 79: Estimated Hazard Ratios for Time to Death for 10 unit increase in NZDep2006, circulatory and respiratory cohorts 254
Table 80: Baseline hospitalisation costs per year, circulatory and respiratory cohorts..... 255
Table 81: Estimated Hazard Ratios for time to death for hospitalisation measures, circulatory and respiratory cohorts 255
Table 82: Housing risk category, circulatory and respiratory cohorts, as at first month *after* 256
Table 83: Estimated Hazard Ratios for time to death for "dwelling health risk" dummy variables, circulatory and respiratory cohorts 256

Table 84: Climate Zone, circulatory and respiratory cohorts 257
Table 85: Estimated Hazard Ratios for time to death for "Climate Zone" dummy variables, circulatory and respiratory cohorts 257
Table 86: Initial multiple regression models, circulatory and respiratory cohorts 258
Table 87: Full regression output for model presented in Table 48 259
Table 88: Full regression output for model presented in Table 49 259
Table 89: Full regression output for models presented in Table 61 260
Table 90: Full regression output for models presented in Table 62 261

LIST OF ABBREVIATIONS

CAU:	Census Area Unit
CBA	Cost Benefit Analysis
CSC	Community Services Card
DALY	Disability Adjusted Life Year
EECA	Energy Efficiency and Conservation Authority
ESF	Energy Saver Fund
EWHG	Energy Wise Home Grants programme
GDP	Gross Domestic Product
GLM	Generalised Linear Model
HIP	Home Insulation Programme (Australian)
kWh	Kilowatt hour
MED	Ministry of Economic Development
MoH	Ministry of Health
NHI	National Health Index
NIWA	National Institute of Water and Atmospheric Research
NMDS	National Minimum Dataset
OECD	Organisation for Economic Co-operation and Development
PHARMAC	Pharmaceutical Management Agency
QALY	Quality Adjusted Life Year
QV	Quotable Value
RCT	Randomised Controlled Trial
VSL	Value of a Statistical Life
VSLY	Value of a Statistical Life Year
WHO	World Health Organisation
WUNZ:HS	Warm Up New Zealand: Heat Smart

CHAPTER 1. INTRODUCTION

This thesis is focused on methodological issues surrounding the quantification of the health benefits of large-scale public health interventions. Accurate evaluation of large-scale public health interventions is important from a policy perspective, and this thesis explores some of the potential tools that can be used to carry out such analyses.

1.1. What is a large-scale public health intervention?

The term "public health" refers not only to the health status of a population but also to "a philosophy of intervention aimed at protecting and promoting the health of the population" (Baggot, 2000, p1). What then is a public health intervention? A useful definition is provided by Rychetnik, Hawe et al. (2004, p 356):

> "*[a]n intervention comprises an action or programme that aims to bring about identifiable outcomes. A public health intervention can be defined by the fact that it is applied to many, most, or all members in a community, with the aim of delivering a net benefit to the community or population as well as benefits to individuals.*"

Under such a definition any change in the way things are done, be it a new policy, law, project or programme applied at a group level with an intended community or population health outcome can be seen as a public health intervention. It is important to note that interventions that are applied to many, most or all members of a group are typically preventative, e.g. immunisation rather than curative medicine. Preventative interventions can be further categorised using Geoffrey Rose's famous distinction between preventative interventions that target high risk individuals (who gain a great deal of benefit from such interventions) and population level preventative interventions that may only offer a small reduction in risk for most individuals, but that may more effectively reduce the societal burden of disease by addressing the root cause of a given problem (Rose 2001). Public health interventions fall into the latter category.

Large-scale is a more subjective term and will mean different things in countries with different sized populations. In the context of this thesis, the significance of the term is that it implies that enough people will potentially be affected by an intervention that administrative data will be an optimal tool for analysing the outcomes of the intervention.

Having defined a large-scale public health intervention loosely as any action, change in law or policy or new programme, which is intended to impact community or population health and which is large enough in size to be optimally analysed using administrative data, it is then important to ask what is meant by evaluation.

1.2. What is meant by evaluation?

Rychetnik and Hawe et al. (2004, p 540) quote the following definition of evaluation, *''a process that attempts to determine as systematically and objectively as possible the relevance, effectiveness, and impact of activities in the light of their objectives"* (Last 2001). This broad definition captures the key elements of the term, and is a useful starting point. Public health interventions are amenable to evaluation from a number of perspectives: their design and implementation can be evaluated as well as their outcomes; and both quantitative and qualitative tools can be used to evaluate them. This thesis focuses primarily on outcome evaluation, assessing the impact of an intervention on the target population. Outcome evaluation as a term encompasses both the evaluation of the outcomes of pilot interventions, and evaluations of the outcomes of full interventions either during their period of operation or after it.

A logical starting point for carrying out an outcome evaluation is to ask whether an intervention has met its stated goals. An example of such an evaluation would be comparing immunisation rates before and after an advertising campaign aimed at encouraging immunisation by new parents. Even if a causal connection could be established an evaluation such as this would raise other significant questions such as: could immunisation rates have been raised more cost-effectively using an alternative approach? Could the money spent on the programme have been better spent trying to address a different health problem (or even in a different sector) and achieved better results (and how would we define better)? Were there additional benefits or costs that resulted from the programme (for example other changes in parenting behaviours following the advertising campaign)?

Economic evaluations of public health interventions address these questions by allowing policy makers to quantify outcomes in terms of a common unit of measurement, such as dollars per additional immunisation, or quality adjusted life years (QALYs) gained, which means that interventions can be compared. The key methods of health economic evaluation include effectiveness analysis, cost utility analysis and cost-benefit analysis (Drummond, Sculpher et al. 2005). While each method has its advantages and disadvantages, cost-benefit analysis, which attempts to comprehensively quantify the costs and benefits of an intervention in dollar terms, has the advantage of producing results which allow for the comparison of projects that have a range of intended outcomes and/or significant externalities that are borne by society as a whole. This thesis focuses on exploring methods for quantifying the health impacts of large-scale public health interventions in dollars terms using a whole-of-government approach, thus informing future cost-benefit analysis-based evaluations. In order for such quantification to constitute high quality evidence sound statistical and/or econometric techniques must be used in the design of evaluations.

1.3. Why should the outcomes of large-scale public health interventions be evaluated?

Accurate evaluation of public health interventions is important for a number of reasons:

Firstly, there is has always been a public interest in ensuring that the taxpayer or ratepayer money, which typically funds public health interventions, is spent effectively (however effectiveness is measured). If it is found that an intervention is ineffective policy makers will ideally then make informed decisions about whether it should be redesigned or terminated, as part of a rational public policy cycle. This principle is a core justification for any analysis of the outcomes of any government-funded intervention. In addition, at the time of writing the world is still recovering from the global financial crisis, to which afflicted Western governments responded with a mixture of policies ranging from stimulus to austerity. During this period it has become ever more necessary to justify government spending, particularly new areas of spending, given that the adoption of austerity has meant that core government services in many nations are being cut back, with public sector job losses and pay cuts being a common theme throughout much of the West.

Secondly, total spending on healthcare has been increasing as a proportion of gross domestic product (GDP) for the majority of Organisation for Economic Co-operation and Development (OECD) nations over the last 50 years (OECD 2011). Such increases are often particularly marked during recessions, as the majority of healthcare spending in OECD nations comes from government, and it is relatively difficult for governments to decrease healthcare spending in the short term (OECD 2011). The reasons for the long-term increases in healthcare spending seen in the OECD, both as a proportion of GDP and per-capita vary from country to country, and may include demographic changes (aging populations and longer lifespans with attendant greater medical needs), costly new medical technologies, and ratcheting public demand for services (OECD 2011). In this context, claims that public health interventions can improve the health of a community or population, leading to reductions in the numbers of expensive medical procedures or pharmaceuticals prescribed and thus achieve net cost savings are particularly appealing.

Despite the importance of evaluation, government spending, including public health interventions, is often under-evaluated (Wanless 2004). While public health interventions may, increasingly, be designed based on high quality evidence, for example evidence from randomised controlled trials (RCTs), more usually their outcomes are not evaluated. In some cases it simply may not be possible to evaluate the outcomes of an intervention, for example the target population may be too small to produce data that can be meaningfully analysed, predicted outcomes may only be

realised many years in the future, a suitable control group may not exist, or ethical issues may preclude the randomisation necessary to create a suitable control group (Wanless 2004). In other cases, the organisations implementing or funding the programme may lack the expertise or financial resources to evaluate the outcomes of a programme (Wanless 2004).

Even when public health interventions are based on evidence derived from RCTs it is important to note that there may be questions about the degree of generalisability of such evidence (Victora, Habicht et al. 2004, McMahon and Chopra 2012); "real world" implementation brings with it many potential complications, such as organisational problems, lower levels of compliance, differences in implementation, corruption or unanticipated changes in circumstances. This issue is exacerbated in the case of complex interventions which contain multiple specific interventions or stages (Craig, Dieppe et al. 2008).

In this context the potential value of quasi-experimental[1] outcome evaluations of large-scale public health interventions becomes clear. Quasi-experimental outcome evaluations, sometime described in the literature as "plausibility studies" (Victora, Habicht et al. 2004), may either strengthen the evidence of RCTs by demonstrating generalisability or they may provide the best evidence possible if RCTs have not occurred, are not possible, or are not ethically justifiable (Victora, Habicht et al. 2004). Quasi-experimental outcome evaluations of programmes that incorporate some element of randomisation in their design in order to facilitate evaluation provide better quality evidence because in the absence of randomisation concerns about selection bias will limit the strength of the conclusions that can be drawn from an outcome evaluation.

Large-scale public health interventions are an ideal candidate for outcome evaluation. Their higher cost means that it is more important that they are evaluated, their scale suggests potential resources to fund evaluation, and a large target population increases the potential for statistically valid evaluation, given appropriate study design. Administrative data, when available, are an ideal data source for the evaluation of the outcomes of large-scale public health interventions, with lower data collection costs than carrying out surveys, or accessing individual medical records, although they entail additional concerns about accuracy and privacy.

[1] Quasi-experimental designs attempt to approximate RCTs by creating a suitable control group.

1.4. The New Zealand context

New Zealand has the same need for evaluation of its large-scale public health interventions as other nations. Policy makers in New Zealand have become increasingly aware of the importance of scientific evidence as a guide to policy choice and design and also the need for scientifically valid outcome evaluation (Gluckman 2011). New Zealand's Cabinet adopted the Gateway Review Process in 2007, a systematic approach to intervention design and evaluation, signalling the importance of outcome evaluation (State Services Commission 2013).

The policy imperative to accurately evaluate the outcomes of large-scale public health interventions informed the funding of the outcome evaluation of a major housing intervention, the Warm Up New Zealand: Heat Smart (WUNZ:HS) programme (Grimes, Denne et al. 2011). This cost-benefit-based analysis required the quantification of any health benefits of the programme in dollar terms.

The author's involvement with the WUNZ:HS evaluation included co-authoring a report which attempted to quantify the health benefits of the programme in dollar terms (Telfar Barnard, Preval et al. 2011) and authoring another (Preval 2012)[2] . The data gathered for these evaluations provide the opportunity to carry out a case study in the application of statistical techniques to quantify health outcomes in dollar terms. The WUNZ:HS case study thus forms the primary basis of this thesis.

[2] See Appendix Three

1.5. Aim of thesis

The primary aim of this thesis is to examine the methodological issues around the analysis of administrative data in large-scale public health intervention evaluations, using data available from the WUNZ:HS evaluation as the basis of a case study. The methodologies examined aim to accurately quantify the health benefits or costs of a programme in dollar terms.

Secondary to this aim is a policy analysis-based comparison of the WUNZ:HS programme with the Australian Home Insulation Programme (HIP). This comparison informs a discussion about the generalisability of the results of the primary analysis and the importance of intervention design and implementation.

1.6. Structure of thesis

Chapter Two describes the connections between housing and health, explains the origins of the WUNZ:HS programme and its intended outcomes and introduces the WUNZ:HS evaluation.

Chapter Three describes the quasi-experimental design underlying this thesis and the WUNZ:HS evaluation, describes dataset creation, characterises datasets, and highlights potential limitations of those datasets.

Chapter Four sets out the approach taken to model specification and fitting in this thesis and describes and carries out an analysis of the mortality-related benefits of WUNZ:HS. Chapter Five describes and carries out the analysis of hospitalisation costs. Chapter Six describes and carries out the analysis of pharmaceutical use costs.

Chapter Seven presents the combined results of the analyses described in Chapters Four to Six in order to consider the consistency of these results, calculate the value of combined annual health benefits per household and to assess the impact of different modelling assumptions on annual health benefits. The impact of different modelling choices on the benefit cost ratios calculated in the original cost-benefit analysis presented in (Grimes, Denne et al. 2011) is also briefly considered.

Chapter Eight compares the design and implementation of the WUNZ:HS programme with the unsuccessful Australian Home Insulation programme (HIP) from a policy analysis perspective. This comparison also provides some additional insight into the interpretation of the results presented in Chapter Seven.

Chapter Nine is a discussion chapter, placing the results of this thesis in the context of the literature on large-scale public health evaluations and previous housing and health research and considering the implication of results for future research. Finally, the strengths and weaknesses of the WUNZ:HS evaluation design are considered including reflections on future methodological choices for the design of similar evaluations.

Chapter Ten is a brief conclusion, summarising the key finding of the thesis as a whole, presenting the public health implications of this research and making.

CHAPTER 2. THE CASE STUDY: WARM UP NEW ZEALAND: HEAT SMART

WUNZ:HS was a $347 million four year programme funded by the New Zealand government that provided part-funding for insulation retrofits, draught-proofing, moisture barriers and clean, efficient heating for New Zealand households. WUNZ:HS began on July the 1st 2009 and was originally intended to run for four years and retrofit 188,500 homes, although this target was slightly extended, reflecting the quickness and cost efficiency with which the target number of installations was reached in November 2012, and final WUNZ:HS retrofits were completed in Sept 2013. In the 2013 Budget a more focused version of WUNZ:HS called Warm Up New Zealand: Healthy Homes was announced. Beginning on July the 1st 2013, Warm Up New Zealand: Healthy Homes will operate for three years, has a budget of $100 million and will subsidise insulation for low-income households (EECA 2013).

WUNZ:HS was administered by the Energy Efficiency and Conservation Authority (EECA), a Crown Entity whose core mandate is improving the way that New Zealand uses energy, which includes any "changes in energy-use that result in an increase in net-benefits", which clearly encompasses health improvements. WUNZ:HS was predicted to produce energy savings, efficiency gains and also health improvements for participants and was available to any home-owner whose home was built prior to 2000 and as such qualifies as a public health intervention. With its large budget and number of retrofits WUNZ:HS can thus clearly be viewed as a large-scale public health intervention.

In order to understand why WUNZ:HS was created and what it was hoped it would achieve, it is necessary to briefly explore the evidence for connections between housing and health, and the history of home design and home design policy in New Zealand.

2.1. Key connections between housing, insulation, heating and health

The following discussion focuses on how to make homes warmer, a common problem in temperate countries such as New Zealand. When the temperature outside is too cold for people to be comfortable it is necessary to create a "net temperature difference" between outdoor and indoor temperatures in order for people to be comfortable indoors (Lloyd, Bishop et al. 2007, p7). Such a net temperature difference can only be created and maintained by using an external energy source to heat the interior. There are two sources of external energy which can be used: the sun, or a space heating device of some type (including the human body). However, any net temperature difference will diminish over time (or require further heating to maintain), with excess heat "flowing" from the building either through the fabric of the building via conduction, convection or radiant heat transfer, or by air infiltration via cracks, windows or gaps.

2.1.1. Insulation

Insulation is one of the key ways of slowing such heat loss in buildings. There are two ways that insulation can be used to reduce heat loss. Firstly, bulk materials such as fibreglass, foam or wool can be installed in ceilings, walls or floors to reduce heat lost via convection and conduction. These materials contain pockets of trapped still air which does not conduct heat well. These materials may also be used to reduce heat loss from hot water cylinders and pipes. The effectiveness of such insulation materials is influenced by their thickness, correct installation and dryness.

Secondly, reflective insulation such as aluminium foil works by limiting the transfer of radiant heat (Lloyd, Bishop et al. 2007), and has often been used as under-floor insulation in New Zealand. The thermal resistance of insulation (or any material) is measured in m_2°C/W and is known as its R value: the higher the R value, the greater the thermal resistance (McChesney, Cox-Smith et al. 2008).

Improvements in insulation are easy to define: the insulation of a home has been improved if the installation or upgrading of an insulation product causes the building's net thermal resistance to increase.

2.1.2. Heating

Heat, either from the sun or from a space heater of some type, creates the net temperature differential that insulation helps maintain. Space heaters comprise a wide variety of sources from the human body to devices such as fan heaters or wood burners. While insulation can be expected to have a somewhat predictable effect on indoor temperature for a given level of heating, the effect of a given space heating *device* on temperature is dependent on the quantity of energy consumed by that device, which is dictated by the preferences, habits and financial constraints of occupants. Improvements in heating are more difficult to define than improvements in insulation, but typically will include some combination of increases in energy efficiency, reduction in harmful emissions (see 2.1.4. below) and increased potential heat output measured in kW.

2.1.3. Insulation, heating and the rebound effect

Improvements in the insulation and/or space heating devices installed in a home (retrofitting) improve energy efficiency and thus could potentially allow occupants to maintain the same temperature post-retrofit while using less energy (assuming that they have some form of space heating). However occupants typically "spend" some of this potential energy saving in the form of higher temperatures: this is an example of the direct rebound effect. Occupants have a range of potential responses: at one extreme they may choose (assuming they have control over the amount of energy used to heat the home and that that amount is not zero) to remain at the same indoor temperature by reducing energy use and take all of the gains from the retrofit as a financial saving via reduced heating costs. At the other extreme, occupants may spend 100% of energy efficiency gains in the form of increased temperature. In addition, as described in Howden-Chapman, Viggers et al. (2009), a subgroup of occupants may respond to improved insulation by changing their space heating behaviour to take advantage of improved heat retention or heating efficiency and actually spend more.

Influential U.K. research suggests that at indoor temperatures of 16.5°C 30% of potential energy savings are expected to be taken as improved comfort i.e. temperature increases (Milne and Boardman 2000). It is likely that households that experienced relatively comfortable temperatures prior to a retrofit will take a higher proportion of the benefit of a retrofit as energy use savings, while colder households will take a higher proportion of efficiency gains as increased temperature (Milne and Boardman 2000). This means that the relative impact of a retrofit on energy use and temperature gains will be influenced by pre-retrofit indoor temperature, which will in turn be directly related to factors such as household income.

A number of studies have demonstrated temperature increases following insulation and/or heating retrofits (Howden-Chapman, Matheson et al. 2007, Shortt and Rugkåsa 2007, Howden-Chapman, Pierse et al. 2008, Lloyd, Callau et al. 2008). The Insulation,

Heating and Health Study carried out by *He Kainga Oranga* demonstrated both statistically significant increases in indoor temperature (Howden-Chapman, Matheson et al. 2007), and statistically significant reductions in combined measured and self-reported energy use (Chapman, Howden-Chapman et al. 2009).

2.1.4. Insulation, heating and health

There are a number of hypothesised causal pathways by which improved insulation and space heating can potentially have an impact on health in a temperate country such as New Zealand.

The first causal pathway by which improved insulation and space heating can potentially have an impact on health is increased temperatures. Low temperatures can cause bronchoconstriction, particularly in combination with exercise. Improvement in indoor temperatures resulting from improved insulation or heating will reduce exposure to cold and thus potentially reduce the incidence of bronchoconstriction-related events. Statistically significant reductions in respiratory symptoms (e.g. wheeze) following an insulation/ heating retrofit are reported in a number of insulation/heating intervention studies summarised in a systematic review by Thomson, Thomas et al. (2009) and in a more recent Cochrane review (Thomson, Thomas et al. 2013), although some studies did not report any statistically significant change in the outcomes of interest.

Excess winter mortality is reported in many temperate countries, driven primarily by higher rates of cardiovascular and respiratory mortality amongst the elderly (Keatinge, Donaldson et al. 1997, Telfar-Barnard 2010). Davie, Baker et al. (2007) found that during the period 1980-2000 there were 1,600 excess winter deaths per year in New Zealand (relative to what would be predicted if summer and winter death rates were identical). Similarly, Telfar-Barnard found that there were 7,166 excess winter hospitalisations per year during the period February 2000 – and January 2006 (Telfar-Barnard 2010). A recent New Zealand RCT reported non-statistically significant reductions in all-cause and respiratory hospitalisations following an insulation retrofit (Howden-Chapman, Matheson et al. 2007), and a New Zealand study which utilised a within-person crossover design found a statistically significant reduction in acute hospitalisations for individuals aged 5- 34 following participation in the Healthy Housing Programme (which provided a suite of advice and physical housing upgrades including insulation retrofits) (Jackson, Thornley et al. 2011).

While researchers have debated the underlying causes of excess winter mortality (with outdoor cold exposure an alternative hypothesis to indoor exposure) the indoor exposure hypothesis has strong plausibility; a recent model based on previous epidemiological analyses of indoor cold-related mortality Wilkinson, Smith et al. (2009) assumed a 2% increase in cardiovascular disease-related winter mortality for each 1°C

reduction in standardised indoor temperature. Interestingly, however, one New Zealand paper reviewed did not find any statistically significant associations between measures of low daily temperatures and mortality rates (Hales, Salmond et al. 2000).

It is speculated that exposure to colder temperatures may cause cardiovascular events via thrombosis due to haemoconcentration or other cardiovascular reflexes that are induced by low temperature (Keatinge, Donaldson et al. 1997). Excess winter respiratory mortality and morbidity may be explained by a combination of factors including the impact of cold on immune resistance, and increased bacterial survival time in droplets at colder temperatures (Keatinge, Donaldson et al. 1997).

In addition to the increased mortality risk resulting from cold ambient temperatures there is a considerable body of literature linking increased cardiovascular and cerebrovascular mortality with heat waves and higher temperatures summarised in reviews by Basu and Samet, and Basu (2002, 2009), including in temperate New Zealand (Hales, Salmond et al. 2000). It is plausible that insulation, by reducing extremes of indoor temperature may contribute to a reduction in the risk of such mortality for vulnerable sub-groups such as the elderly, or those with pre-existing disease.

The second causal pathway by which improved insulation and space heating can potentially impact health is that warmer homes are less damp, and dampness is correlated with concentrations of mould and dust mite faeces, both of which can trigger allergic reactions in sensitised asthmatics (Carrer, Maroni et al. 2001). Moulds can also cause non-immune specific inflammation (Carrer, Maroni et al. 2001, Howden-Chapman 2004). Mudarri and Fisk (2007) carried out a meta-analysis of damp and mould studies which and found statistically significant associations between damp and mould and upper respiratory tract symptoms, cough, wheeze and current asthma. An additional connection between improved heating and dampness is that unflued gas heaters (common in New Zealand) release water vapour into the indoor environment as a product of combustion which contributes to dampness, thus replacing unflued gas heaters can potentially improve health via reduced dampness even if there is no change in temperature.

The third causal pathway by which improved insulation and space heating can potentially have an impact on health is that some space heating devices emit harmful chemicals and particulate matter into the indoor environment. For example, unflued gas heaters emit nitrogen dioxide as a product of combustion. A randomised controlled trial carried out by Howden-Chapman, Pierse et al. (2008) demonstrated that homes using unflued gas heaters had more than three times the level of NO_2 in their living rooms than homes that did not and that replacing unflued gas heaters with healthier heaters reduced indoor NO_2 levels by two thirds relative to households that continued to have an unflued gas heater as their primary space heater (Gillespie-Bennett, Pierse et

al. 2008). Indoor nitrogen dioxide levels have been correlated with asthma symptoms in a number of studies (Chauhan, Inskip et al. 2003, Pilotto, Nitschke et al. 2004), which may be due to an interaction between nitrogen dioxide and other allergens which increases their effect on sensitised asthmatics (Tunnicliffe, Burge et al. 1994, Barck, Sandström et al. 2002).

The fourth causal pathway by which improved insulation and space heating can potentially have an impact on health is that indoor temperatures and heating behaviours (for example heating either one room or multiple rooms) may influence the probability of cross infection within a household via the degree of indoor crowding (i.e. if the entire home is warm then household members will not have to crowd together in a single heated room, a common heating pattern in New Zealand). A meta-analysis of 345 studies found that 55% showed a statistically significant association between higher levels of crowding and higher rates of close contact infectious diseases while only 1% showed a statistically significant association between higher levels of crowding and lower rates of close contact infectious diseases (Baker, McDonald et al. 2013).

Finally, there is some evidence of a causal connection between housing condition and mental health. Of the warmth and energy efficiency improvement intervention studies reviewed by Thomson, Thomas et al. (2009) one study found an odds ratio of 0.560 ($p < 0.001$) for intervention group members for low happiness based on the SF (Short Form)-36 (Howden-Chapman, Matheson et al. 2007) and another 0.510 ($p<0.001$) for low vitality (Howden-Chapman, Pierse et al. 2008) but other intervention studies did not find statistically significant changes in mental health-related measures. More recently Howden-Chapman, Chandola et al. (2011) found that participants in the Whitehall II Study who had "some to great problems with house" had higher modelled GHQ (General Health Questionnaire)-30 scores (poorer mental health) relative to individuals who had "slight to no problems" after controlling for plausible confounders. Other previous research has found a correlation between self-reported poor housing quality (including damp and mould) and poor self-reported mental health (Hunt 1990, Hopton and Hunt 1996), although the subjectivity of self-reported measures of housing quality used in these two papers raises the possibility that these results may partially reflect a tendency for those with poorer mental health to assess their surroundings more negatively.

Speculatively, other mental health-related causal pathways could include increased comfort, reduced social embarrassment, reduced worry about risks relating to health, and financial savings from improved home energy efficiency which could reduce stress on home occupants.

New Zealanders spend a large proportion of their time inside their own homes (69.7%), which is slightly higher than comparable groups such as UK and US residents (Baker,

Keall et al. 2007). The impact of insulation and heating on health is also exacerbated because those who are most vulnerable (young children, the elderly, the unemployed and the disabled or unwell) typically spend more time in their homes (Baker, Keall et al. 2007, O'Sullivan, Howden-Chapman et al. 2011).

Connections between indoor temperature and health have long been recognised by the international medical establishment, with World Health Organisation (WHO) recommendations for a minimum indoor temperature of 18°C first being made in 1987 (World Health Organization 1987). New Zealand homes do not currently meet this requirement, with average winter temperatures in New Zealand living rooms recently reported to be only 17.8 °C from 5-11 pm, and bedrooms typically much colder (French, Camilleri et al. 2007). Consistent with this pattern, New Zealand has patterns of excess winter mortality (Davie, Baker et al. 2007) and morbidity (Telfar-Barnard 2010) which may partially be explained by some physical housing characteristics (Telfar-Barnard 2010).

Having discussed connections between home insulation, heating and health it is natural to turn to the question of why temperatures in New Zealand homes are low, which policy interventions have influenced New Zealand home temperatures and energy efficiency over the years and how WUNZ:HS came into being. Answering these questions will inform analytical choices made in the substantive chapters of this thesis.

2.2. Home insulation and heating in New Zealand: a brief background

New Zealand homes built in the last 100-120 years typically have solid construction; however, this durability is a mixed blessing: with an estimated average lifespan of 90 years (Chapman, Howden-Chapman et al. 2009) the consequences of construction techniques, design decisions and materials utilised in the earlier part of the twentieth century continue to have an impact on New Zealanders today. In particular, New Zealand homes built before 1978 were not required to be insulated and were often not insulated, which means that unless later retrofitted with insulation they are difficult to heat and often damp. French and Camilleri et al. (2007) found that, for a representative sample of New Zealand homes, an increase in the age of a home by a decade resulted in a 0.2±0.05°C drop in temperature ($p=0.00004$). In addition, New Zealand homes are typically larger than European homes (comparable in size to Australian or American homes), making them more difficult to heat than European homes (Howden-Chapman, Viggers et al. 2009). Victorian villas constructed at the turn of the 20th century were often closed off from the sun for reasons of design and were oriented to the street rather than to the sun: such houses are not able to efficiently use solar energy (Ministry for the Environment 2005). As construction techniques and housing preferences changed in the following fifty years newer homes were more able to utilise solar radiation via larger windows and more open design but due to the absence of insulation the heat was not retained (Ministry for the Environment 2005). Other trends were less positive, with Bastings (1958) reporting a reduction in 1950s era roof and wall section R-values in comparison with 1920's values due to changes in materials and construction.

New Zealand policy makers' awareness of home temperature issues developed in the 1950s, as New Zealand research addressed concerns around condensation and mould and thermal insulation began to be imported and utilised for the first time (Isaacs 1993). Insulation was first manufactured locally in the 1960s and was actively promoted, however the 1973/74 oil shock is generally credited with the first major increase in the proportion of residential buildings built with insulation in New Zealand (Isaacs 1993, Ministry for the Environment 2005, McChesney, Cox-Smith et al. 2008).

New Zealand's central government responded to the oil shock and other related energy concerns by instituting an interest free loans scheme for ceiling insulation in 1975 (McChesney, Cox-Smith et al. 2008) and eventually by introducing the first mandatory minimum insulation standards for new buildings and alterations to existing buildings, NZS 4218P: 1977, which set out a variety of acceptable combinations of minimum ceiling, wall and floor R-values (Isaacs 1993). Similar minimum insulation standards were being set in other developed nations around this time (Lah 2009).

Other than efforts by the government-run Housing Corporation of New Zealand to retrofit ceiling insulation in many of its South Island properties in the early 1980s (McChesney, Cox-Smith et al. 2008), these measures represented the only significant policy interventions aimed at improving insulation by the central government until the early 1990s. This contrasts strongly with nations such as Germany, which regularly increased minimum insulation standards during this time (Lah 2009). In order to understand this difference it is necessary to place New Zealand housing in the broader energy supply context. New Zealand has historically had extremely cheap electricity, thanks to its hydroelectric dams, large coal deposits and the development of resources such as the Maui gas field in 1969. It is speculated that this historical surplus created a culture of energy inefficiency that was only briefly interrupted by the oil shocks of the 1970s: however, steeply declining oil prices in the 1980s and 1990s combined with a central government focus on reforming and restructuring the electricity sector moved energy efficiency from the top of the policy agenda (Lah 2009), with most consumer-focused home energy efficiency policies during the 1980s and early 1990s limited to the provision of information and advice (McChesney, Cox-Smith et al. 2008).

2.2.1. EECA and housing: small scale public health interventions

In 1992, the Rio Earth Summit took place and New Zealand signed the Framework Convention on Climate Change. In the same year poor rainfall meant that southern hydro-lakes were very low leading to an electricity crisis and to serious attempts to improve electricity conservation (PCE 2000). The Electricity Shortage Review Committee created to assess and report on the crisis highlighted the potential to reduce consumer demand for energy using both financial incentives and education, and the need to promote awareness and understanding of energy efficiency more generally (PCE 2000). Later that year the Government set out its Energy Policy framework which included the following outcome:

> […] the efficient and effective provision of energy services through well-functioning commercial systems with competitive incentives, operating within:
>
> - an effective and stable regulatory environment; and
> - an environment of **effective energy conservation** from the promotion and adoption of efficient practices and technology[.]

This policy framework established what would eventually be named the Energy Efficiency and Conservation Authority (EECA), an entity whose functions included developing a long-term energy efficiency strategy (PCE 2000).

Following a pilot study, in 1995 EECA began central government's first residential energy programme, the Energy Saver Fund (ESF), an intervention which funded residential energy efficiency upgrade projects and was intended to target sectors of the community with high barriers to purchasing efficiency upgrades such as beneficiaries.

The ESF was a $13 million programme over a six-year period (McChesney, Cox-Smith et al. 2008) and funded a wide variety of energy efficiency upgrades, a small percentage of which were ceiling and floor insulation retrofits. The ESF was later criticised for its focus on kWh savings, which ignored the potential health and environmental benefits of improved insulation (EECA, 2001).

In 2000 EECA became a Crown Entity following the passing of the Energy Efficiency and Conservation Act, and in 2001 it produced the first National Energy Efficiency and Conservation Strategy, which defined energy efficiency as "a change in energy-use that results in an increase in net-benefits", clearly indicating a broader focus than that which underpinned the ESF.

Following the completion of the ESF in 2001 EECA developed the Energywise Home Grants (EWHG) programme. The EWHG part-funded the insulation of pre-1978 houses occupied by low-income households, with whole house retrofit packages typically including ceiling and underfloor insulation, cylinder wraps, pipe lagging and draught-proofing of doors as required. The retrofits were delivered by 'service providers', commercial entities or not-for-profit community trusts who won large-scale contracts for projects with EECA (EECA 2007). A variety of other schemes operated on a small scale during this period including schemes targeting landlords and funding clean heat projects.

The logic underlying EWHG was that barriers such as limited access to capital and imperfect information were preventing homeowners and landlords from making an economically rational investment in improved home energy efficiency. In addition, the problem of split incentives meant that tenants were unlikely to purchase insulation as they would not gain the benefits of such improvements beyond the period of their tenancy, while landlords do not typically pay energy bills or gain the health benefits from living in a more comfortable home and so will be reluctant to invest in insulation unless the investment is rewarded via higher rental returns (IEA 2007). It was believed that improved awareness of the value of insulation and the encouragement of a nascent retrofit industry would eventually create a self-sustaining market for insulation that did not require further subsidisation.

In 2007, EECA conducted a major review of the EWHG programme. Recommendations included extending the scope of the programme beyond pre-1978 households to include pre-2000 households (in 2000 minimum insulation values had been raised slightly by standard NZS 4218:1996: it was also widely acknowledged that insulation installed in the years directly following 1978 was not installed well or thoroughly), incorporating funding for clean heat retrofits and raising minimum required R-values for ceiling and floor insulation retrofits. Some of these recommendations were incorporated into the operation of the programme. In late 2008, a one billion dollar 15-year extension to the EWHG programme was announced by the Labour led

government, to be funded by profits generated by a recently passed emissions trading scheme (Science Media Centre 2008); however, following the November 8th election, the newly elected National led government announced the replacement of the EWHG programme with the $NZ 323 million WUNZ:HS programme in its 2009 Budget as part of an arrangement with the Green Party (programme funding increased to $347 million in November 2009). The WUNZ:HS programme built on the EWHG programme, extending its availability to all households. The programme was explicitly expected to produce a combination of energy efficiency gains and reduced health risks (Brownlee 2009).

The evidence that informed WUNZ:HS (and the proposed EWHG extension before it) included research derived from randomised controlled trials carried out by *He Kainga Oranga*[3], the Housing and Health Programme of the Public Health Department of the University of Otago, Wellington on the health and energy use impacts of retrofitted insulation (Howden-Chapman, Matheson et al. 2007), and improved home heating (Howden-Chapman, Pierse et al. 2008). In particular, a cost-benefit analysis of floor and ceiling insulation based on the results of the insulation retrofit study suggested a favourable ratio of benefits to costs of 1.7:1 driven largely by improvements in health (Chapman, Howden-Chapman et al. 2009). Early versions of this paper and other as-yet unpublished research findings were widely circulated within government (Carroll, Blewden et al. 2008), and it is likely that both the design of the WUNZ:HS programme and indeed the decision to implement the programme were informed by these findings. The relationship between the research of *He Kainga Oranga* and housing policy in New Zealand is explored in greater depth in Chapter Eight.

[3] www.healthyhousing.org.nz

2.2.2. WUNZ:HS: programme details

WUNZ:HS operated from the 1st of July 2009 to the end of funding in June 2014 (J. Lackey, personal communication, 2014). WUNZ:HS was universally available to all New Zealand homeowners, whether landlords or owner-occupiers, whose homes were built before 2000, and provided part-funding for a range of retrofits. Funding could be used to retrofit underfloor and ceiling insulation, conditional on both being insulated if they were accessible. Participants whose floor and ceiling insulation met minimum R values could receive funding toward clean and energy efficient heating devices such as a heat pump or pellet burner, and finally funding was available for hot water cylinder wraps, pipe lagging, draught-stopping, and a ground moisture barrier. Clean heating funding was limited, from Sept 2012 onwards, to the replacement of inefficient wood burners and open fires in areas of poor air quality (EECA 2012). Table 2 sets out details of funding initially available.

Table 2 Funding available under WUNZ:HS

	Insulation [includes all non-heating retrofits]	**Clean heating**
All eligible houses	33% of the total cost up to $1300 (including GST)	$500 (including GST)
Homeowners who hold Community Services Cards[4]	60% of the total cost, or more than 60% if you qualify for a special project.	$1200 (including GST)
Landlords with tenants who hold Community Services Cards	60% of the total cost	$500 (including GST)

(EECA 2009a)

Complex arrangements were made to cover the proportion of costs that were not covered by WUNZ:HS funding. Non-profit organisations sometimes fully covered or further subsidised costs for households that they identified as particularly likely to benefit from the programme, and a number of councils provided upfront funding for retrofits to be paid back over time in the form of higher rates.

WUNZ:HS operated with approximately 60 'service providers' who carried out retrofits and were then reimbursed by the government for their work. These service providers included both private companies and non-profit organisations, which could sub-contract work to other organisations. Service providers tendered to be part of the programme and amongst other conditions had to be able to demonstrate capacity to

[4] Community Services Cards (CSCs) are available to low and middle-income adults who are New Zealand residents or citizens. CSCs entitle bearers to reductions in charges for healthcare products and services such as prescription fees.

deliver retrofits to agreed quality and safety standards (J. Lackey, personal communication, 2010). Various compliance measures were enforced by EECA including a comprehensive series of checks filled out by installers and occasional inspections carried out by EECA (J. Lackey, personal communication, 2010).

2.3. WUNZ:HS evaluation

In late 2009, the then Ministry of Economic Development (MED), EECA's parent organisation, requested proposals for an economic evaluation of the WUNZ:HS programme. The MED request for proposals included requirements that the evaluation would assess the impact of WUNZ:HS on energy-use, health, and employment within the insulation industry, and combine the results of these assessments in the form of a cost-benefit analysis.

The contract was won by a group including academics from the University of Otago, Wellington (including me) and Victoria University of Wellington, and from consultancy firms Motu Economic and Public Policy Research, who were the head contractors, and Covec: (collectively referred to in this thesis as "the Group"). The winning proposal was based on an innovative linking of administrative health data with EECA data via a commercially-owned dataset.

From the official start of the evaluation in November 2009 to its completion in October 2011 the Group produced four reports, assessing the industry impact of WUNZ:HS (Denne and Bond-Smith 2011), changes in energy use (Grimes, Young et al. 2011), health impacts (Telfar Barnard, Preval et al. 2011) and finally a cost-benefit analysis (Grimes, Denne et al. 2011). The cost-benefit analysis suggested a highly favourable benefit cost ratio of 4.3:1 which was driven primarily by a reduction in all-causes mortality for people aged 65 and over who had recently been hospitalised with a circulatory condition (Grimes, Denne et al. 2011). Smaller, but significant, benefits from the programme included a reduction in metered energy use (Grimes, Young et al. 2011) and industry benefits (Denne and Bond-Smith 2011). In 2012 the MED commissioned an additional report on the health impacts of WUNZ:HS (Preval 2012). Various modelling changes were made for the health extension but this did not materially alter the benefit:cost ratio presented in Grimes, Denne et al. (2011).

Amongst the policy recommendations made by the Group were suggestions that WUNZ:HS be more targeted, and that funding insulation should be prioritised over funding heating (Grimes, Denne et al. 2011,p 28). In late 2012 funding for heating retrofits under the WUNZ:HS programme was greatly limited, as noted in Section 2.2. In addition, the Warm Up New Zealand: Healthy Homes programme which succeeded WUNZ:HS is limited to the provision of insulation for low-income homes. It appears likely that these decisions were influenced by the work of the Group.

The data gathered for the WUNZ:HS evaluations provided the opportunity to carry out a case study in the application of statistical techniques to quantify the health outcomes of large-scale public interventions in dollar terms and forms the primary basis for this thesis. The thesis presents some of the models used in Preval (2012) and Telfar Barnard, Preval et al. (2011) as well as additional analyses carried out using the same basic datasets, although changes in dataset composition, modelling approach and timeframe mean that the results presented in this thesis in some cases differ from previously presented results. I have indicated where work presented is not exclusively my own (further details are presented in Section 1.7). In Chapter Seven the cost-benefit calculations presented in (Grimes, Denne et al. 2011) are used as a basis for crude estimates of the impact of different approaches to modelling benefits and costs on the ratio of benefits-to-costs.

In the following chapter the WUNZ:HS evaluation is described in detail, the design of the study is discussed, the characteristics of administrative data are considered, the dataset creation process is thoroughly described and summarised, and finally the available data is described in detail.

CHAPTER 3. WUNZ:HS EVALUATION: DESIGN, DATASET CONSTRUCTION AND DESCRIPTION

The WUNZ:HS evaluation used a quasi-experimental design and was primarily based on administrative or publicly available data. The starting point of the evaluation design was the assumption that, given EECA's relationship with its parent organisation MED, administrative data from the WUNZ:HS programme would be available, and that there was some potential to link this data with other administrative datasets. Designing an evaluation based on this data thus required knowledge of the specific datasets that could be linked to address data from EECA, a suitable study design that could potentially capture key benefits of participating the programme and an awareness of the limitations of such a design and the administrative data it utilised.

3.1. WUNZ:HS study design

The starting point of the WUNZ:HS evaluation study design was the availability of data pertaining to each of the retrofits carried out by EECA, particularly the address and the details of the work done. In order to evaluate the programme it was necessary to obtain relevant health outcome and demographic information about the occupants of these retrofitted homes and to compare the health outcomes of these individuals with those of a suitable control group. Such a quasi-experimental design would be greatly strengthened if information about the health outcomes of individuals prior to the intervention was able to be obtained.

The key to the study design was the existence of Quotable Value (QV), a state-owned housing valuation organisation possessing highly detailed data on the condition, age and characteristics of New Zealand's housing stock. Previous work by Lucy Telfar-Barnard established that Ministry of Health held administrative data which could be linked with QV data[5]. The additional innovation proposed for the WUNZ:HS evaluation was utilising QV's data and skilled technicians to match EECA provided treatment addresses with QV's dataset and to then identify suitable control group homes for each treatment home. Having identified these treatment and control group homes using the QV dataset it would then be possible to obtain, via suitably anonymised processes, administrative health data and energy use data associated with both treatment and control individuals and/or addresses as Telfar-Barnard (2010) had previously demonstrated with regard to health data.

The benefits of such a design would include the construction of a large study cohort without the need for expensive and time consuming surveys or individual consent processes (in addition, EECA had signalled that they did not want WUNZ:HS participants to be contacted for research purposes). Potential disadvantages would include the limitations of quasi-experimental study designs, the limitations of administrative data in general, data loss at each stage of the data-matching process, and the quality of the specific health data that could be linked (discussed in detail in Section 3.3).

The final design balanced issues of validity with timeliness and cost, and successfully met the criteria set by MED. Once the contract was won, ethics approval was sought. Data collection began on the 2nd of March 2010, following Multi-Regional Ethics Committee approval (See Appendix A).

[5] Lucy Telfar-Barnard's PhD thesis (Telfar-Barnard 2010) linked anonymised administrative morbidity data (NMDS) with Quotable Value data using the address field in the National Health Index (NHI) data (with an extremely high success rate of 68% given the complexities of address matching). It was then possible to calculate the relative rate ratio for excess winter hospitalisation for cohort members living in housing with a certain characteristic (for example decade of construction) relative to a baseline characteristic.

Because the WUNZ:HS evaluation relies so heavily on administrative data, the strengths and weaknesses of administrative data are explored in detail in Subsection 3.1.1. The quasi-experimental design of the evaluation is further explored in Subsection 3.1.2 below and the full detail of the data collection process described in Section 3.2.

3.1.1. Administrative data

The WUNZ:HS evaluation was designed to almost exclusively utilise administrative data, linking various administrative datasets with addresses from EECA's administrative data. When data from multiple administrative datasets are linked by researchers, particularly when some pertain to rare events, the possibility that individuals may be identified increases. In order to address this possibility various anonymity preserving measures had to be considered for the WUNZ:HS design. Making a successful proposal required addressing privacy concerns, as did subsequently obtaining ethics approval (see Appendix A).

Administrative data, generally defined as operational data collected by a medium- or large-scale organisation while carrying out its core functions, have increasingly become a key part of public health intervention evaluations. Administrative data have a great deal of benefit from a research perspective. Administrative data have already been collected as part of their primary function; if administrative data can be accessed by researchers then this will greatly reduce costs relative to collection using more intensive means such as surveys, interviews or directly accessing medical records. Access to administrative data makes studies with large cohorts much more feasible, which increases the potential to evaluate rare events such as hospitalisation.

Administrative data also have a number of weaknesses from a research perspective. They may be subject to quality issues, for example missing data and data entry errors, additionally quality may vary by region or over time for a given dataset, potentially introducing systematic biases to data that researchers need to be aware of. By contrast, a study using primary data sources may incorporate a number of checks, for example double entry of data, in order to limit such issues. More fundamentally, administrative datasets may not record measures of interest in an optimal manner and may simply not include measures of interest that would be very valuable from a research perspective.

From an ethical perspective, the use of administrative data for research purposes raises issues of privacy, and of how data should be used. Many administrative datasets, such as the National Minimum Dataset (NMDS) dataset held by New Zealand's Ministry of Health, include data for individuals who had little choice about participation due to their ill-health. When data are held for individuals who have not had an opportunity to decline participation in research, it is even more vital that no harm be done to such individuals by that research. The primary harm that can potentially occur through the

use of administrative data for research is violation of privacy through the potential identification of individuals. Researchers, or those who gain access to data held by researchers can potentially use data to harm those individuals either through criminal means (e.g. blackmail), social means (e.g. gossip) or by using the information to make decisions unfavourable to the individual (e.g. an insurance company refusing insurance to an individual having obtained access to administrative data pertaining to that individual's disease risk). A more generalised 'subjective' harm may also occur simply from the knowledge that personal data are potentially available to entities who might use it in some harmful manner, usefully described in legal terms as "the apprehension of information-driven injury" (Calo 2011).

In New Zealand privacy issues have recently been a major concern, with administrative data and private records being accidently provided by government officials via email and loopholes in computer systems making administrative datasets publicly available (Savage 2012, Young 2012). This emphasises the concerns that researchers must keep in mind when designing studies that utilise administrative data.

Organisational approaches demonstrate the importance of maintaining privacy. For example, Statistics New Zealand avoid violations of privacy by greatly restricting access to individual level data collected at the Census (not technically an administrative dataset as its primary function is research), while data aggregated the level of geographical units such as meshblock or Census Area Unit are more widely available, because the process of aggregation dramatically reduces the potential for identification. New Zealand's Ministry of Health maintains a number of health-related administrative datasets linked by National Health Index (NHI) numbers for each individual (discussed later in this Chapter). Many steps are taken to preserve anonymity: names and addresses will not be provided to researchers who obtain data extracts, and NHI numbers are uniquely encrypted for each data extract, meaning that data from different extracts cannot be linked.

It is also important to consider the many benefits of research incorporating administrative data. Administrative data are low cost and may allow researchers to address research question that are of great policy significance which cannot be realistically explored using randomised controlled trials given the great cost that this would entail. The ethical issues that the use of administrative data entails can partially be addressed using strict anonymisation. In addition some study designs may allow individuals to opt out of large-scale research projects.

3.1.2. Quasi-experimental design

The WUNZ:HS evaluation study design is best described as quasi-experimental, with pre-test and post-test data collected for treatment group and matched control group individuals/ households (Shadish, Cook et al. 2002). Quasi-experimental study designs

produce results that are less useable for inferring causality than those produced by a randomised controlled trial, the gold standard for study design (Victora, Habicht et al. 2004) . Randomised controlled trials, by virtue of the randomisation process, are able to more reliably control for both known and unknown confounders, producing results with greater plausibility. The goal of quasi-experimental designs is to best approximate the characteristics of a randomised controlled trial(Shadish, Cook et al. 2002) .

A quasi-experimental study with a control group but only post-test data, is likely to produce data of limited value due to the potential of selection bias effects to distort results (Shadish, Cook et al. 2002). Selection bias, essentially the ways in which the characteristics of the treatment and control group differ that are correlated with the outcome of interest, may partially be controlled for by collecting data on potential confounders, but uncertainty will always remain regarding the impact of unmeasured or imperfectly measured confounders on results. The WUNZ:HS evaluation study design limits the potential impact of selection bias by collecting pre-test data. By incorporating pre-test information, for example pre-retrofit hospitalisation costs, the pre-existing differences between the treatment group and the control group that are not the result of treatment can be partially accounted for.

This design choice entailed obtaining administrative data for a time period that included a significant amount of pre-test and post-test exposure. January 2008 was identified as a suitable start date, meaning that as the WUNZ:HS programme began in July 2009 there would be at least one and a half years "pre-test" data for most individuals (less or even no pre-test data might be obtained for individuals born after Jan 2008). Unfortunately due to the limitations imposed by the evaluation timeframe (completion by August 2011) less "post-test" data would thereby generally be collected, with useable post-test data only available up until October 2010 for the initial evaluation. The extension addressed this issue, obtaining data for each individual up until December 2011. However, the quasi-experimental design of the WUNZ:HS evaluation has some limitations that it is difficult for such studies of this type to avoid.

3.1.3. Limitations of the WUNZ:HS evaluation design

In addition to the inherent limitations of administrative data, perhaps the greatest, although most difficult to quantify, limitation of the study design is potential selection bias that cannot be fully controlled for by using pre-test data. For example it is possible that the health status of treatment group participants may have influenced their decision to participate in WUNZ:HS, either of their own volition or by placing them in contact with organisations encouraging and/or funding participation in WUNZ:HS for vulnerable individuals. If so, the well-known regression to the mean effect (Bland and Altman 1994, Shephard 2003) has the potential to introduce a difficult to eliminate distortion to analysis, for example individuals who have recently been hospitalised will have high health costs during the period in which that hospitalisation occurred

and simply due to the period of relative good health that may be expected to follow a hospitalisation will incur relatively low health costs during this subsequent period.

Thus, if the experience of a period of poor health is a motivating factor for participation in WUNZ:HS then there is some danger that the timing of the treatment might produce the appearance of a beneficial effect of treatment as a period of lowered health costs will naturally coincide with the post-treatment period. The availability of pre-treatment data for each individual would not allow us to fully account for this effect as the same pattern of events and costs would not apply equally to control group individuals. This problem is partially addressed by the fact that there is a large period of pre-treatment and post-treatment data for each individual, limiting the impact of short term regression to the mean effects somewhat.

Potential alternatives to the evaluation design are briefly explored in the discussion chapter, Chapter Nine.

3.2. Dataset creation

EECA initially provided anonymised data for each of the retrofits it had funded during the period 1st July 2009 (the start date for WUNZ:HS) – May 30th 2010, including addresses and the details of each retrofit, and other information such as whether the home was owner-occupied or renter-occupied and whether the retrofit was for a Community Services Card holder or not; 46,655 unique addresses were identified.

These addresses were supplied to Quotable Value who matched them to the addresses in their database, 37,163 of the addresses were able to be matched to a QV listing, a match rate of 79.7%. The match rate likely reflects problems such as inaccurately entered addresses in either dataset, particularly for problematic addresses such as recently subdivided properties, units or rural addresses (Telfar Barnard, Preval et al. 2011). The 79.7% match rate is a concern as it both reduced the power of the study and also raised the possibility that there was something unusual about the unmatched addresses and their occupants that meant that excluding them might bias the study in some way. For example, rural/small-town addresses might be harder to match due to idiosyncrasies of address. If rural homes were, in general, of poorer quality, or occupants less healthy then we might expect a greater relative benefit from receiving treatment. Excluding such addresses and their matched controls would thus bias any finding towards the null. The opposite could equally apply; however, because we do not have QV data for the unmatched addresses it was not possible to form any conclusions at all about the characteristics of unmatched homes or their occupants.

The 37,163 addresses then needed to be matched to suitable control addresses. The Group identified the physical characteristics of homes that would most influence their health impact and which approximated housing quality data held by QV, this process was informed by evidence on the impact of New Zealand home characteristics on health (Telfar-Barnard 2010) . A scoring system was devised to rank potential control homes, with points assigned to each category depending on the quality of the match:

Table 3 QV Variables

QV variable	**Possible values/units**	Maximum Points	Notes
Census area unit	Any land based CAU in NZ	10	Mandatory match
Category	LI: Lifesytle RA: Residential appartment RC:Residential covered flat RD: Residential dwelling RF: Residential OYO RH: Residential home and income RN: Residential n (not known) RR: Residential purpose-built flat	10	Mandatory match
House Type	Apartment, bach, bungalow (post-war), contemporary, cottage, pre-war bungalow, quality bungalow, quality old, Spanish bungalow, state rental, terrace appartments, townhouse, unit, villa.	10	Mandatory match
Levels (single/multi-story)	Single Multi	10	Mandatory match
Decade	Any decade prior to 2000	10	
Floor Area	m^2	10	
Bedrooms	Number of bedrooms	5	
Main Roof Garages	Number of garages under main-roof	5	
Building	Matched on condition(average, fair, good, mixed, poor, missing) and construction: (Aluminium, brick, concrete, fibre cement, glass, malthoid/fab, mixed material, missing, plastic roughcast, steel/iron, stone, tile profile, weatherboard)	10	Match score 10 if both condition and type match, 7 if only one matches, 0 if neither
Roof	Matched on condition(average, fair, good, mixed, poor, missing) and type: (Aluminium, brick, concrete, fibre cement, glass, malthoid/fab, mixed material, missing, plastic roughcast, steel/iron, stone, tile profile, weatherboard)	10	Match score 10 if both condition and type match, 7 if only one matches, 0 if neither
Modernised	Y/N	10	

Adapted from (Telfar Barnard, Preval et al. 2011)

Matches on Census Area Unit (CAU), property category, house type and levels were mandatory. Dwelling construction decade could vary by up to three decades, with the

score dropping by 2 points for each decade distance from the target decade. Floor area could vary by up to 50%, with the score being reduced by 1 point for each 10% difference from the target. Number of bedrooms and main roof garages were scored with a maximum of score 5, subtracting 1 for each variance in the number of bedrooms and garages between the treatment home and the potential control home. The categories "wall materials and condition" and "roof materials and condition" were given a score of 10 when both the Quotable Value dataset records for materials and condition matched, 7 where only building materials matched and 0 if materials didn't match. A score of 10 points was assigned if the "modernised" indicator matched and 0 if it did not match.

Quotable Value applied the dwelling match protocol described above to identify up to 10 control addresses for each dwelling, selecting in order of matching score and randomising when there were "ties" that would lead to more than 10 homes being selected. 31,423 treatment dwellings (67.4% of all treated dwellings) were initially able to be matched to at least one control address; 269,110 control dwellings were selected and the distribution of matches was as follows:

Table 4. Match of treatment and control dwellings, by count and percentage

Number of controls	Count (treatment)	% of treatment cohort	% of total treatment addresses (includes those not matched with QV data)
10	22520	71.7%	48.3%
9	962	3.1%	2.1%
8	948	3.0%	2.0%
7	1003	3.2%	2.1%
6	964	3.1%	2.1%
5	973	3.1%	2.1%
4	958	3.0%	2.1%
3	1043	3.3%	2.2%
2	985	3.1%	2.1%
1	1067	3.4%	2.3%
0	9492	0.0%	20.35%
Total	40915	100.0%	88%

(Telfar Barnard, Preval et al. 2011)

Variation in the number of matches is a concern because in so far as there were relationships between home characteristics (including location) and both the number of matches found and health outcomes following treatment, biases could be introduced to analysis because more data would come from control group individuals with certain home characteristics which would distort results. As discussed later in Chapter Four it was possible to partially address this issue by selecting and including only four control

homes and their occupants when more than four matches were found for a given treatment home.

Following the initial identification of the study cohort, 7,141 dwellings which EECA identified as having received treatment under the WUNZ:HS programme between May 2010 (the cut-off point for inclusion in our treatment group) and Dec 2010 were later dropped from the control group. Dropping these 7,141 dwellings resulted in dropping a further 22 treatment group dwellings, who no longer had at least one matched control.

Quotable Value provided the Group with detailed home characteristics for each cohort address based on the categories set out in Table 3. Descriptions of the key variables are set out in Section 3.2. Due to the anonymisation necessary to meet ethics approval addresses were not supplied, limiting the potential to link household level records with additional geographically referenced datasets to the Census Area Unit level, which was provided for each household.

3.2.1. Ministry of Health Data

Having identified the initial cohort addresses, it was then necessary to identify the occupants of these households. In order to ensure the anonymity of study participants the Ministry of Health provided addresses for every individual (see Table 5 below) in the National Health Index (NHI) dataset (as at October 2010), with a unique newly created identifier, to QV. QV matched these addresses to its list of those records still in the study cohort (31,401 treatment dwellings and 261,969 control dwellings).

We were then able to identify cohort dwellings that could not be matched with at least one NHI record. This resulted in the exclusion of 1,492 treatment group dwellings and 26,724 control group dwellings. Removing these dwellings resulted in the further removal of 164 treatment group dwellings, that no longer had at least one matched control group dwelling, and 9,318 control group dwellings that no longer had a matched treatment dwelling.

Once the study cohort of 29,745 treatment and 225,927 control addresses were identified QV returned this list to the Ministry of Health, who then provided the Group with NHI-based demographic data, NMDS data, and prescription dispensation data for the period Jan 2008 – Dec 2010 for all individuals identified as occupying a cohort dwelling as at October 2010. The 29,745 treatment homes were linked with 110,360 individual occupants and the 225,927 control homes were linked with 863,350 individual occupants.

Hospitalisation and prescription use data were gathered from Jan 2008 despite the fact that WUNZ:HS began on the 1st of July 2009 in order to provide pre-test data for participants, as per the quasi-experimental study design described at the start of this

chapter. Table 5 presents a brief description of the data provided. Each dataset is described in much greater depth in subsequent subsections of this Chapter.

Table 5. Administrative health data

	Description
NHI data	The NHI database contains demographic information about every person in New Zealand who has received health or disability services. Each person is assigned a unique encrypted NHI number. Demographic data provided: date of birth, date of death (if applicable), ethnicity variables, sex, and residency status.
NMDS data	The NMDS database includes details of every public hospital admission in New Zealand. Patients are identified by their NHI number. NMDS data provided: details of hospitalisation events for each cohort member including ICD codes and cost information.
Pharms data	The Pharms database compiles all of the claims made by pharmacists for reimbursement following the dispensation of a prescription subsidised by the New Zealand government agency PHARMAC. Individuals are identified by their NHI number. Pharms data provided: pharmaceutical product codes and cost details.

3.2.2. The MED extension

The MED funded an additional analysis of the health impacts of WUNZ:HS. This analysis required the obtaining of an additional years' worth of NHI, NMDS and Pharms data for cohort members (so data were now held for the period 1st of Jan 2008 – 31st Dec 2011). It was also necessary to obtain EECA records for the period 1st of Dec 2010- 31st of Dec 2011 in order to identify control addresses that were no longer suitable due to having participated in the programme and to identify any further retrofits carried out for treatment addresses.

EECA provided 53,433 distinct addresses that had received retrofits. These addresses were matched by QV with their own dataset; 32,703 could be suitably matched, a 61% match rate. After various data cleaning processes 32,483 unique usable home records with accurate QV identifiers were identified.

I matched these 32,483 QV identified records with the cohort of 255,672 homes. I found that 8,193 control group homes had received treatment thus making them ineligible for continued inclusion in the study. This in turn meant that the 29,513 individuals occupying these homes were no longer eligible and were removed from the study. Removal of these 8,193 control homes also left 35 treatment homes without at least one

matched control home. Dropping these 35 treatment homes mean removing a further 138 treatment individuals from the cohort.

The new cohort contained 247,444 homes and 944,059 occupants. Finally a test was run matching newly acquired NHI demographic data (July 2012) with the original NHI data. I compared demographic data for the 944,059 individuals remaining in the study. Initial observation of the data revealed a slight discrepancy in variables such as date of birth, sex and ethnicity. Contact with the Ministry of Health (MoH) confirmed that very minor discrepancies in the data are a natural artefact of the processes whereby the NHI dataset is updated (discussed further in Subsection 3.2.2.6.). Individual contact with hospitals and regular electronic "sweeps" of primary health databases are opportunities for newer data to replace previously held data and for data entry errors and mistakes to enter the system (and for inaccuracies to be corrected). Although it would not be possible to assess whether the original demographic data, or the more recently supplied data, were more accurate, it was necessary to formulate rules for the degree of discrepancy that seemed acceptable before an individual had to be removed from the cohort (i.e. seemed likely to be from a different individual).

Sex seemed too important to accept a mismatch. Therefore I removed 288 individuals who were mismatched on sex (having confirmed that the "sex" field in NHI records would not be likely to change following gender reassignment surgery). Date of birth matching was more difficult, it being easy to imagine a single typing error which could lead to a mismatch. After consultation with MoH staff, I established a protocol for date of birth matching. I removed individuals, if date of birth in the two records differed by more than five years. I also removed individuals, if more than one field (day, month, or year) did not match the original; 382 further individuals were dropped using this method (9 having already been dropped due to sex mismatch) (Preval 2012).

Ethnicity is a more complex variable (evidenced by a higher percentage of mismatches in the NHI data presented in Table 6) with numerous coding options and more potential variability in the way that individuals may self-identify over time and thus potential for the variable to differ (see Section 3.2 for further discussion). After consultation with the MoH I concluded that it would be too difficult to formulate a meaningful matching rule for this variable and thus did not remove any cohort members due to discrepancies (Preval 2012). This left the cohort at 943,389 individuals from 247,413 homes. Table 6 sets out the matching process.

Table 6. Matching based on MoH demographic data

	Date of birth	Birth protocol	Sex	Ethnicity
Match	941,939 (99.78%)	943,668 (99.96%)	943,771 (99.97%)	917,047 (97.14%)
Don't match	2,120 (0.22%)	391 (0.04%)	288 (0.03%)	27,012 (2.86%)
Total	944,059	944,059	944,059	944,059

(Preval 2012)

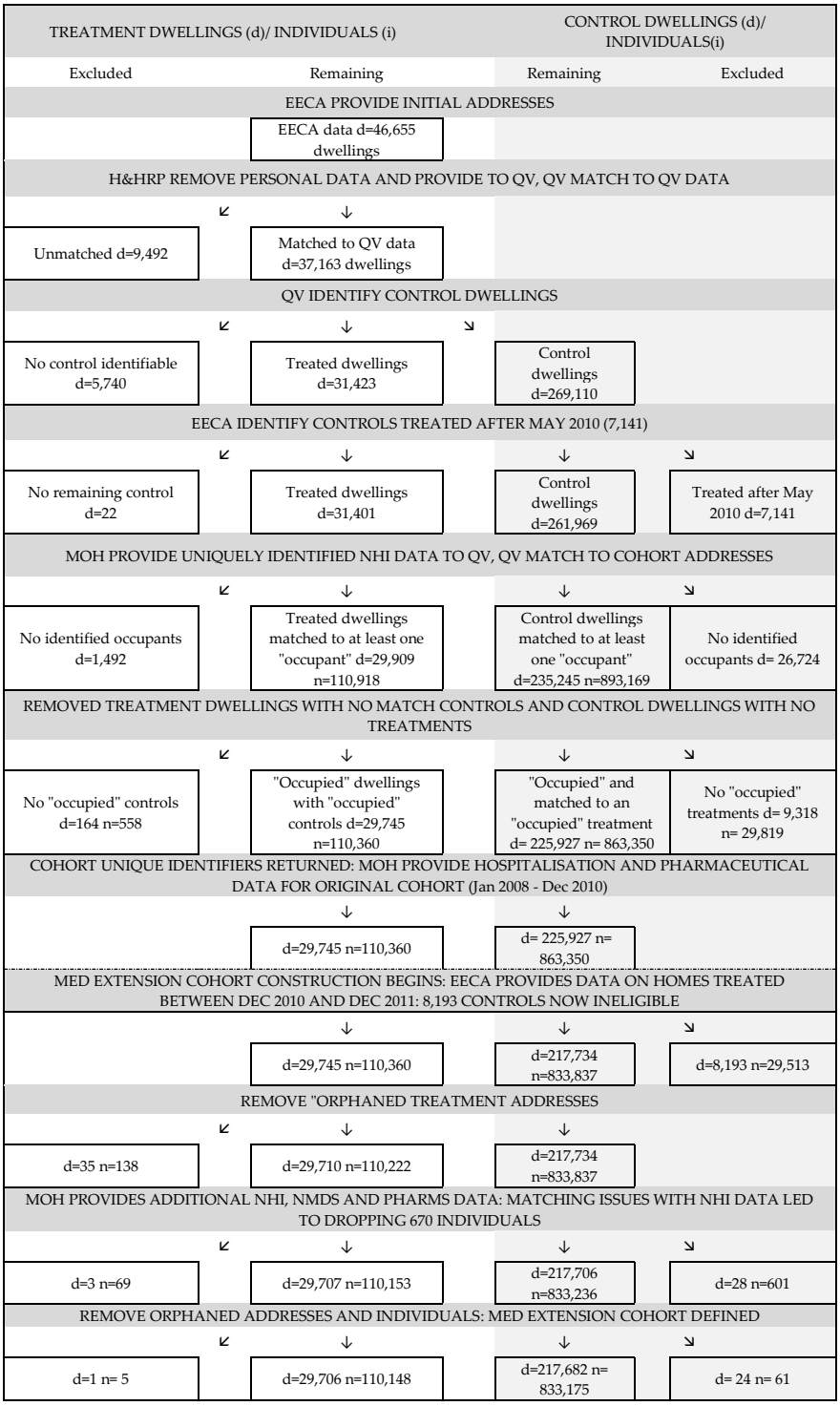
Finally, I tested whether these removals left any homes "empty" and thus any treatment homes without an occupied matched control home or vice versa. Twenty four control homes with 61 occupants were dropped from the study due to their matched treatment home now having no occupants and 1 treatment home with 5 occupants had to be dropped from the study, because it had no matched occupied controls (Preval 2012).

The adjusted cohort was 943,323 individuals from 247,388 homes, of whom 110,048 individuals were in 29,706 treated homes and the remainder were in control homes. This cohort was the basis for the analysis presented in Preval (2012) and is the starting point for this thesis.

3.2.3. Data summary

The data gathering process described in Subsections 3.1.1-3.1.2 is summarised in Figure 1 below. Section 3.2 includes descriptions of each data source and key variables. Appendix 2 contains a full summary of every variable collected including source and a brief description.

Figure 1. Dataset creation



3.3. Dataset description

This section describes the datasets and key variables that were used for the evaluation of the health effects of WUNZ:HS presented in Telfar Barnard, Preval et al. (2011) and Preval (2012)which form the basis of this thesis. Following the various matching and data gathering processes described in Section 3.2 the following data were held for each individual/household –NHI data (demographics), NMDS data, Pharms data, QV data (home characteristics, location etc.), and EECA data (work done, Community Services Card status etc. (treatment group only)).

Where possible 2006 Census data are provided as a comparator in order to address issues of generalisability. One concern about using Census data to provide such policy context is that the treatment group (and possibly the control group given matching process is likely to select households with similar tenure given the many physical similarities matched on such as location) has a higher proportion of home-owning households than the Census population, which is problematic because individuals from home-owning households are likely to be significantly different than those from non-home-owning households in terms of socioeconomic status, due to affordability issues (Law and Meehan 2013).

Census tenure data are available at a household level. Household level Census tenure data shows that in 2006 approximately 62% of households occupied dwellings that they either owned or partially owned or were held by their family's family trust/s. By contrast 88.72% of treatment group households occupied homes that were designated "non-rental" according to EECA's data on treatment group homes. It would be informative to compare the makeup of the treatment and control groups for values such as ethnicity, sex and age with Census data stratified by tenure, however it is not possible to link household level Census tenure data with individual level data using publicly available datasets.

3.3.1. NHI

The National Health Index number (NHI number) is a unique combination of letters and numbers that is assigned to each person using health and disability support services, typically at birth (records for new-borns are submitted by hospitals and midwives). The National Health Index (NHI) contains the name, address, date of birth, sex, NZ resident status, ethnicity and date of death (if applicable) and various medical warning flags and donor information for each person with an NHI number (MoH 2012b).

The NHI has many functions: "…individuals can be positively and uniquely identified for the purposes of treatment and care, and for maintaining medical records. This means that healthcare providers can be sure that they are talking about the same person, thereby reducing the chance of making a clinical decision based on wrong

information"(MoH 2012b). Patient identification is particularly important in a mobile society such as New Zealand (Morrison and Nissen 2010), where individuals may often change primary or secondary healthcare providers. NHI numbers link several datasets including the NMDS and Pharms datasets used in this thesis.

NHI data for individuals is regularly updated as a result of individual contact with health services. Public hospital admissions and primary healthcare records (e.g. GP records) are opportunities for patient information to be recorded; these data will then be used to update the NHI when hospitals provide details of admissions to the MoH or the MoH carries out one of its regular electronic "sweeps" of primary healthcare databases. The majority of fields such as sex and birthdate would not be expected to change; however the address field is likely to change often, reflecting the highly mobile nature of New Zealand society (Morrison and Nissen 2010).

As discussed in Section 3.1 each update has the potential to introduce inaccuracies (or to correct previous errors). In addition, internal MoH algorithms regularly resolve issues with multiple NHIs assigned to a single individual by assigning primary status to one of the NHI numbers, while multiple individuals may attempt to use the same name in order to access healthcare (for example non-residents using the identity of a friend or family member) which may lead to data discrepancies.

The WUNZ:HS evaluation was designed based on the potential to link NHI records with QV data using the address field, however, as dictated by the anonymity preserving study design, the Group was not provided with NHI addresses. QV data, discussed later in this Section, includes mesh-block level address data for each cohort home, which provides sufficient detail to model any location-related factors such as external temperature or socio-economic status of neighbourhood while maintaining anonymity.

The key NHI fields available for analysis are date or birth, ethnicity, sex and date of death. We can also derive a field called "number_of_occupants" based on the number of records supplied for each treatment and control address.

3.3.1.1. Age

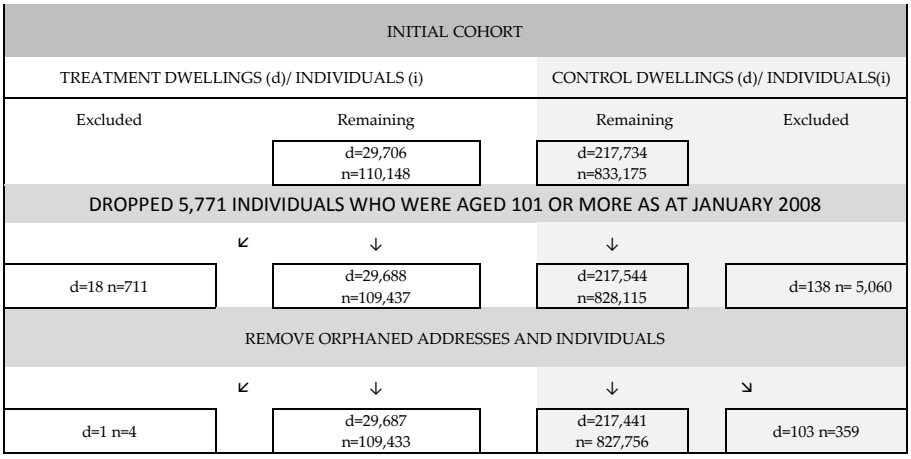
Age is defined in the dataset simply by making a calculation based on date of birth, a variable supplied in the NHI. Each individual contributes 48 rows of data for the 48 months of the study, Jan 2008 – Dec 2011. Each of these rows contains an age variable calculated based on date on birth. A person will age on the month that contains their birthday i.e. a person who turns 30 on the 29th of June 2009 will be recorded as being aged 30 in their June 2009 row. Prior to birth a person's age is recorded as "." and between birth and the month in which their first birthday takes place their age is recorded as "0".

It is useful to compare the age structures of the treatment and control groups and that of the New Zealand population (as per the 2006 Census). Comparison of the treatment and control groups naturally informs modelling as differences in age structure can be expected to be a potential confounder given the vulnerability of very young and elderly people to respiratory ailments and of the elderly to cardiovascular ailments. Comparison with the New Zealand population is important for purposes of generalisation.

Before these comparisons were made an initial visual analysis of cohort age data indicated that there were a number of implausible ages in the cohort data, such as 125. This suggested that an initial "cull" would be useful, and also raised concerns about the reliability of the birthdate field. Although age data available from the 2006 Census are not easily available divided into an over 100 category, according to Statistics New Zealand "[a]s of 2011, there are likely to be 400–500 centenarians currently living in New Zealand. Of these, fewer than 40 would be aged 105 years or over"(Statistics New Zealand 2011). By contrast the cohort contained 6,185 individuals aged 100 or more (0.72% of the treatment group and 0.67% of the control group) and 3,753 individuals over 105 as at January 2008, the first month of the Study: this clearly suggests that the birthdate records of many of these individuals are not reliable. I chose to remove cohort members who were aged 101 or older as of January 2008: while some of these records would be valid it was not clear that such a removal would introduce bias, and indeed there was little impact on the key result of the original reports and of this thesis, the changes in hazard ratios for mortality presented in Chapter Eight.

Removal of these individuals also led to the removal of 363 individuals from 104 homes that no longer had an occupied matched treatment or control home. Figure 2 details the removals:

Figure 2: Removal of people aged 101 or more as at Jan 2008

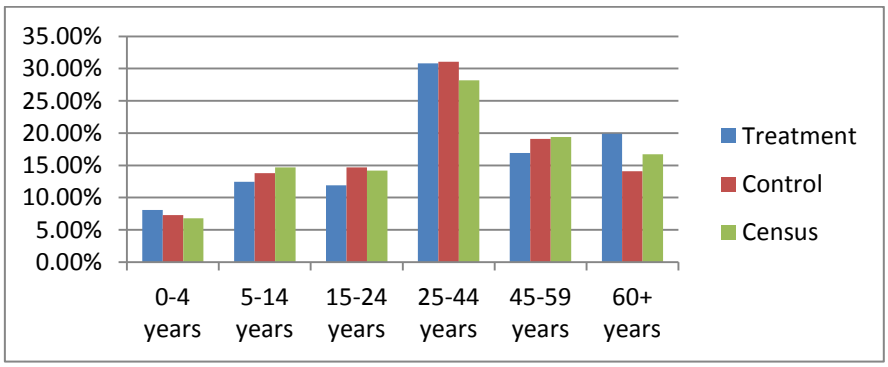


Having removed these individuals it is now possible to compare the age structure of the treatment group and control group on January 2008 (looking at those who were alive at this time) and the New Zealand population (as at the 2006 Census). Table 7 sets out the comparison:

Table 7: Age structure of treatment group and control group as at Jan 2008 and 2006 Census Population

	Treatment		Control		Census 2006	
Age group	N	%	n	%	n	%
0-4 years	8,493	8.07%	58,241	7.27%	275,034	6.80%
5-14 years	13,095	12.45%	110,549	13.80%	592,365	14.70%
15-24 years	12,517	11.90%	117,550	14.67%	570,960	14.20%
25-44 years	32,409	30.81%	248,900	31.07%	1,133,739	28.20%
45-59 years	17,769	16.89%	152,996	19.10%	779,226	19.40%
60+ years	20,914	19.88%	112,941	14.10%	671,718	16.70%
Total	**105,197**	**100.00%**	**801,177**	**100.00%**	**4,023,042**	**100.00%**

Figure 3: Age structure of treatment and control group as at 2008 and 2006 Census population



The key point of interest in the above table and graph is that the treatment group has a larger proportion of people aged 60 years and older than both the control group and

the Census population. Application of the two-sample Wilcoxon rank-sum (Mann-Whitney) test for trend suggests that there is a difference in the age distribution of the treatment and control groups that is highly statistically significant ($Z=-28.956$, $p>|Z|=0.000$). It is unsurprising that a statistically significant result was found given the very large numbers involved. Older people may be disproportionately represented because they prioritise warmth and were more likely to have access to capital under the policy settings which underpinned WUNZ:HS and thus were more likely than the average person to participate in WUNZ:HS, they may also be disproportionately represented because they were targeted by service providers or community groups as ideal candidates for WUNZ:HS, finally they may have been more likely to participate given that they typically have Community Service Cards which result in higher discounts.

The greater proportion of older people in the treatment group is a potential threat to the internal validity of the study, even if difference-in-difference-based models can account for differences in baseline characteristics. One example of how it could introduce bias is if temperatures were lower in 2008 (the "before" period for all cohort members) than in 2009, 2010 and 2011[6]. Because hospitalisation for elderly people is correlated with external temperature (Telfar-Barnard 2010) a model that adjusted for hospitalisation rates or costs prior to treatment could still show a positive change in hospitalisation rates or costs even if treatment did not have any impact because the treatment group would be more affected by the improvement in winter temperatures on average. Introducing an age variable into models, stratifying by age or standardising by age would be ways of addressing this concern.

3.3.1.2. Number of occupants

The number of occupants of a given study home is calculated by summing the NHI records linked to that address as at Oct 2010 (when the NHI data was provided to QV). In this subsection I look at the number of occupants per home on the first of January 2008. In interpreting these figures it is important to note that of 293,370 cohort address identified by QV 28,216 had to be dropped because they did not match any NHI address provided by the MoH. This is like to reflect a combination of address-matching issues and also the fact that some addresses were not linked to any suitable individual with an NHI number on Oct 2010. In January 2008 the average number of occupants was 3.67, with 99% of homes having 12 or fewer occupants. If 28,216 records with 0 occupants are added the mean is reduced to 3.29.

Statistics New Zealand report that the average New Zealand household as at Census 2006 was 2.6. It may not be surprising that the mean is higher for the study cohort, this

[6] In fact average national temperatures appear stable during the four years of the Study based on NIWA summary data. (https://www.niwa.co.nz/climate/summaries/annual)

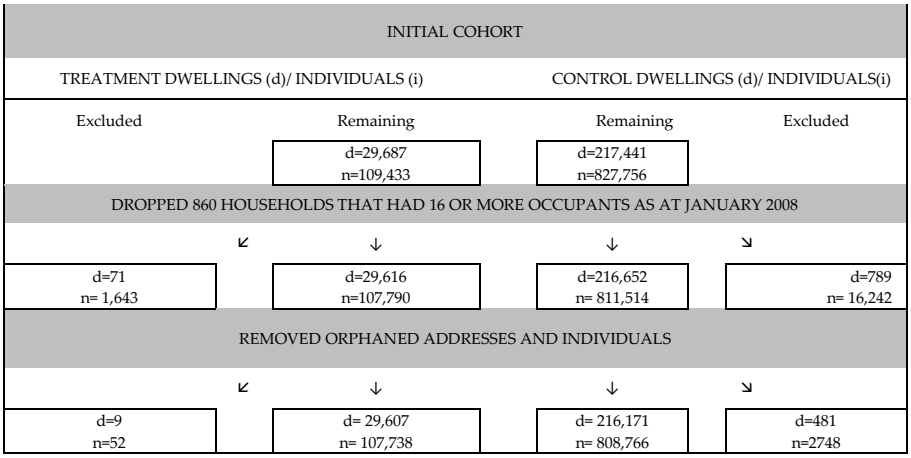
is likely to reflect the types of households that participated in WUNZ:HS in terms of home characteristics, life-stage and household size (the control group was selected by matching on housing characteristics and many of these characteristics such as physical size can be assumed to be associated with household size). In particular, apartments within large apartment complexes could not receive insulation under the programme (although they might obtain heating retrofits or other retrofits), and typical households occupying such apartments are plausibly smaller than the national average.

Turning to extreme values, households with over 12 occupants are plausible; however, as household size increases there is an increasing possibility that the size reflects a distortion due to the timing of interactions with hospitals or primary healthcare providers: e.g. people whose last interaction occurred when they were living at the cohort address but who have since moved (see 3.2.2.6 for a more detailed discussion). Analysis of data from a study of crowding in Housing New Zealand[7] properties (Baker, Zhang et al. 2006) shows that only 0.1% of properties had more than 12 occupants as opposed to 0.9% in the current study. Housing New Zealand tenants are likely to be different in many ways from the study cohort and thus more likely to experience overcrowding given the association between poverty and overcrowding, on the other hand the data from Baker, Zhang et al. were based on official figures and did not include occupants that Housing New Zealand was not aware of.

It is difficult to assess what a plausible maximum number of occupants is, but values such as 205 (the largest number of occupants linked to a cohort address) clearly reflect a number of unsuitable possibilities: the address may be a boarding house, apartment complex, halfway house or a retirement home: if so it seems likely that something problematic took place at the matching stage i.e. conflation of a single room or apartment with the entire building in either the NHI data or the QV data. In order to avoid the introduction of a great deal of statistical noise I excluded 860 addresses that had more than 15 occupants as at January 2008. Removing these addresses had the following consequences, set out in Figure 3:

[7] Housing New Zealand is a Crown agent that provides housing services for vulnerable New Zealanders

Figure 4: Removal of households that had 16 or more occupants as at Jan 2008



3.3.1.3. Ethnicity

There are two potential sources of ethnicity data.

The first is the NHI dataset and the second is the NMDS dataset discussed in Section 3.2.3. The available NHI data includes four ethnicity fields, the first three ethnicity fields (*ethnicg1- ethnicg3*) record coded ethnicity, with *ethnicg1* mandatory, and the second two fields opportunities to record additional ethnic groups. The final field *ethnicgp* represents the result of a prioritised ethnicity process. Under prioritised ethnicity an individual is coded Māori (21) if any ethnicity column is coded Māori; Pacific (30-39) if any column is coded Pacific and none are Māori; Asian (40-49) if any column is coded Asian and none are Māori or Pacific; Other (not 11 and not 21, 30-30 or 40-49) if any column is not NZ European (11) and none are Māori, Pacific or Asian; and NZ European if the only column with data is NZ European.

The second source of ethnicity data is the NMDS. The NMDS dataset also contains the four field ethnicity structure described above.

Prioritised ethnicity is a controversial approach. Utilised by Statistics New Zealand in its analysis of Census data from 1991 to 2001 prioritised ethnicity was favoured, because it enabled the analysis to:

> *(a) be consistent with Recommendation 4 of the report of the Review Committee on Ethnic Statistics;*
>
> *(b) be consistent with the definition of Māori found in the Māori Affairs Amendment Act 1974 and the Electoral Amendment Act 1980 ("a person of the Māori race of New Zealand, and includes any descendant of such a person");*

(c) ensure that important but numerically small groups are not submerged in the dominant majority; and

(d) ensure that groups (such as Māori and Pacific Island groups) about whom policy decisions are commonly made, requiring information to inform those decisions, and which have in the past been shown statistically to be disadvantaged in some way, continue to be identified for monitoring purposes"

Department of Statistics (1993) quoted in Cormack and Robson (2010) p 14.

Considered problematic, prioritised ethnicity was officially replaced in 2005 with "total count" and "single and combined response" measures (Cormack and Robson 2010), although it has remained in use. Statistics NZ moved away from prioritised ethnicity, because it imposed a categorisation on individuals that was not always consistent with self-identified ethnicity (which includes multiple-ethnicities and/or personal "ranking" of ethnicity) and because it resulted in the under-counting of Pacific Islanders who also identified as Māori (Cormack and Robson 2010).

While acknowledging these concerns this thesis takes a prioritised approach because it allows for the creation of an exclusive baseline denominator (unlike the "total ethnicity" approach) and utilises ethnicity categories that are likely to be of interest from a policy perspective (unlike the "single and combined response" approach). Following the approach taken in (Telfar Barnard, Preval et al. 2011) I have used an "ever" approach to utilising ethnicity data. Essentially a variant of the prioritised ethnicity approach described above, this thesis creates a categorical ethnicity variable *ethn* for each individual: Other (0), Māori (1), Pacific Islander (2) and Asian (3). Individuals are coded Māori if they were coded 21 in either dataset, coded Pacific if they were coded 30-39 in either dataset and not coded 21 in either dataset, coded Asian if they were coded 40-49 in either dataset and not 21 or 30-39 in either dataset, and finally Other if not coded Māori, Pacific or Asian. Combining NZ European and non-NZ European Other into a single field is useful because non-NZ European Other is too small a field to be statistically useful and not generally of interest from a policy perspective.

Table 8 sets out the comparison of ethnicity as at Jan 2008 (chosen because this date is closest to the 2006 Census) and naturally excludes cohort individuals who were not born at this time:

Table 8: Ethnic composition of treatment group and control group as at Jan 2008 and 2006 Census Population

	Treatment		Control		Census	
	N	%	N	%	N	%
Other (0)	81,089	78.31%	590,176	75.40%	2,727,717	70.66%
Māori (1)	10,753	10.38%	97,560	12.46%	565,347	14.65%
Pacific (2)	4,495	4.34%	43,653	5.58%	226,275	5.86%
Asian (3)	7,211	6.96%	51,313	6.56%	340,800	8.83%
Total	103,548	100.00%	782,702	100.00%	3,860,139	100.00%

There is a higher proportion of Other in the both the treatment group and control group relative to the Census population. This greater proportion appears to have been made up with a reduction in the proportion of Māori and Asian people and a smaller drop in the proportion of Pacific Islanders. It seems likely that some of the difference between the treatment and control group and the Census population is driven by the high proportion of home-ownership in the treatment group (and thus to some extent in the control group given that home ownership is likely correlated with location, house-type etc.) given that home ownership is correlated with socioeconomic status and socio-economic status is correlated with ethnicity. For example, research has shown that Pacific Island people (with lower socioeconomic status on average) have lower rates of home-ownership (Grimes and Young 2009).

Comparison of the treatment and control group suggest small but meaningful differences which could potentially have an impact on internal validity such as slightly lower proportions of Māori and Pacific Islanders in the treatment group. As with age, this difference could introduce bias if conditions which have an impact on ethnic groups differentially change over the course of the study (for example changes in unemployment rates), even when a difference in difference approach is used.

Including the "ethn" variable in models, stratifying by ethn or standardizing by ethn would be ways of addressing such concerns.

3.3.1.4. Sex

Sex in the NHI refers to biological sex, rather than gender, which is described as a person's social role (masculine or feminine) (MoH 2009b). Sex in the NHI may be categorised as male, female, unknown or indeterminate.

Table 9 sets out the sex variable in the treatment and control groups and in the Census population.

Table 9: Sex composition of treatment group and control group as at Jan 2008 and 2006 Census Population

	Treatment		Control		Census	
	N	%	N	%	N	%
Male	49,258	47.57%	384,776	49.16%	1965621	48.80%
Female	54,234	52.38%	397,435	50.78%	2062338	51.20%
Unknown	56	0.05%	491	0.06%		
Indeterminate	0	0.00%	0	0.00%		
Total	103,548	100.00%	782,702	100.00%	4,027,959	100%

As might be predicted there is little systematic difference between the treatment and control groups and the Census with regard to sex distribution. This is not surprising given that household composition by sex is unlikely to be influenced by factors such as tenure. If is worth noting that the Census data do not include the NHI defined categories Unknown and Indeterminate.

The slight differences in sex distribution between the treatment and control groups could possibly introduce some bias even in difference-in-difference-based models as there is a correlation between sex and risk of certain relevant conditions including cardio-vascular conditions (and thus likelihood of cardiovascular-related hospitalisations, pharmaceutical use and mortality) which could also be correlated with factors that vary over the course of the study such as external winter temperatures. As with age and ethnicity this could be addressed by including "sex" in models, stratifying by sex or standardizing by sex.

3.3.1.5. All causes mortality

Date of death is naturally important for any mortality-rate-related calculations. It is also important for establishing the exposure of cohort members. The NHI data includes a date of death variable. This variable is sourced from the Office of Births, Deaths and Marriages (MoH 2009b). According to the Department of Internal Affairs website "*[t]he funeral director or person responsible for the burial or cremation of the body, or their authorised agent, must notify Births, Deaths and Marriages of the death within three working days of the burial or cremation*" (Department of Internal Affairs 2013). Assuming this information is accurate, given that NHI data was collected in mid-2012 (for the second time) it seems likely that the date of death field should be reliable with regard to deaths during the Study period of Jan 2008 – Dec 2011.

One of the major limitations of NHI mortality data is that there is no inclusion of cause of death. Cause of death is recorded in the Mortality Collection, a separate NHI linked dataset maintained by the Ministry of Health, and coded using standard ICD-10 coding in up to seven fields (MoH 2009a), however when the WUNZ:HS evaluation was first designed it was decided not to attempt to obtain Mortality Collection data due to an

anecdotal "lag" of up to two years before deaths are recorded. Such differences would have made the data of limited use, with the possibility of regional or ethnic differences in reporting lag an additional concern. Further concerns with regard to the Mortality Collection included uncertainty about the accuracy of its cause of death fields, with some evidence suggesting that typically only a single cause of death is recorded which is problematic if the final cause of death differs from the underlying cause, or if there are multiple causes. Although the Mortality Collection data has some limitations one of its benefits would be to exclude accidents: a source of statistical noise that analysis based on the NHI date-of-death field must contend with. Accident rates are unlikely to be influenced by the intervention, particularly traffic accidents, which compromise a large proportion of accident mortality.

During the latter part of this thesis careful consideration was given to the value of attempting to obtain the Mortality Collection data. The difficulties discussed above, when combined with greater ethical hurdles due to the specificity of mortality data, which when linked with Census Area Unit data available for the cohort could risk the anonymity of cohort members, led me to choose not to try and obtain Mortality Collection data.

3.3.1.6. Accuracy of NHI address field

As described in Section 3.1, the cohort was created by linking the NHI address of each NHI individual with the EECA treated addresses and matched control group addresses identified by QV. Linking the NHI addresses of all NHI individuals as at Oct 2010 with the QV dataset was the basis of the study design described in Section 3.1.

There are reasons to be concerned about the accuracy of NHI addresses on a given date, given that the field is updated through contact with health services and many people access health services infrequently. This concern is magnified by the fact that New Zealanders are particularly residentially mobile (Morrison and Nissen 2010). However, given the high proportion of non-renters in the treatment group (88.72%) we can expect them to be less mobile than the general population (Housing New Zealand Corporation 2005, Richards 2007), and the investment that typically accompanies participation in the WUNZ:HS programme suggests some commitment to remaining in a single location. These considerations do not apply to control group members to the same extent although having been selected from the same Census Area Unit as their matched treatment houses may make them more likely to be homeowners because they may live in areas with high proportions of home-owners.

It is very difficult to assess the accuracy of the NHI address field at a given date given the large numbers involved which would preclude survey and the absence of a comparable dataset (personal communication, MoH official, 2010). In the case of this thesis, however, one potentially useful measure of the accuracy of the NHI address

field over the period of the Study does exist. This is the "Domicile Code" field of the NMDS dataset discussed later in Section 3.2.3. Domicile Codes are based on the 1991 and 1996 Census Area Units (CAUs) and are assigned to each hospitalisation record based on the patient's usual residential address provided at the time of hospitalisation. It is possible to compare the Domicile Codes assigned at hospitalisation with the Domicile Code that matches the 2006 CAUs that were provided for each cohort address by Quotable Value. Unfortunately, this comparison may be of limited value given that it will not capture changes within CAUs, however any differences by group may be instructive. Table 10 sets out the quality of match over time, by treatment group, by ethnicity, by sex and by age-group (as at October 2010).

Table 10: Proportion of NMDS records that are mismatched based on comparison of NMDS Domicile Code and QV matched Domicile Code by year, treatment group, ethnicity, sex and age-group

Mismatch	Category					
	Year					
	2008	**2009**	**2010**	**2011**		
N	69.11%	81.67%	89.76%	75.95%		
Y	30.89%	18.33%	10.24%	24.05%		
	100%	100%	100%	1100%		
	Group					
	Treatment	**Control**				
N	82.72%	78.78%				
Y	17.28%	21.22%				
	100%	100%				
	Ethnicity					
	Other	**Māori**	**Pacific**	**Asian**		
N	81.28%	73.47%	77.08%	78.88%		
Y	18.72%	26.53%	22.92%	21.12%		
	100%	100%	100%	100%		
	Age-group					
	0-4	**5-14**	**15-29**	**30-44**	**45-59**	**60+**
N	75.6%	79.32%	64.24%	75.45%	82.68%	88.5%
Y	24.4%	20.68%	35.76%	24.55%	17.32%	11.5%
	100%	100%	100%	100%	100%	100%

Table 10 shows that there is a higher proportion of mismatches for years further from 2010, the year in which NHI address data were provided, which seems entirely consistent with what would be predicted given residential mobility over time. This is concerning because the 2008 time period makes up a reasonable proportion of the *before* period for most cohort members.

The greater proportion of mismatched control group hospitalisations suggests, as hypothesised, that the treatment group is slightly less residentially mobile than the

control group. If the degree of mismatch had been material, this would have been concerning, as residential mobility may be a marker for other unobserved/unobservable differences between the two groups, and is relevant for consideration of the success of the matching process in producing comparable cohorts. There are higher proportions of mismatched addresses for Māori than other ethnicities which is consistent with lower rates of home-ownership (which is related to residential mobility). It is interesting to note that the proportion of mismatched hospitalisation records lowers for those aged 45-59 and 60+, presumably reflecting reduced residential mobility due to home-ownership, settled life-stage etc. This result is reassuring given that a reduction in cardio-vascular hospitalisation and mortality for older people (the group primarily at risk) is one of the plausible outcomes for the WUNZ:HS programme, however it is a concern that there is a high mismatch rate for younger people, particularly those aged 0-4 who are particularly vulnerable to respiratory ailments.

It is not clear how to deal with mismatched hospitalisation address records. While it is important to note that we expect that individuals who move will be replaced by similar individuals in terms of average health characteristics, intuitively it makes sense to remove individuals with such records from the study, or to exclude from analysis the periods in which their address records are shown not to be consistent, because while including such records would not create an obvious systematic bias removing them could reduce statistical noise. However, on reflection this would introduce a problematic distortion to the analysis of cost data, because no such filter has been applied to non-hospitalised individuals: this could reduce the estimated difference in average hospitalisation costs for treatment and control groups post-intervention, a very problematic outcome for a cost-based analysis. Therefore, I have included the individuals who have mismatched hospitalisation address records in the analysis while accepting that this will introduce statistical noise.

3.3.2. NMDS data

The NMDS is a database of hospitalisation events that occur in New Zealand. Records of hospitalisation discharges at public hospitals are provided electronically to the Ministry of Health within 21 days of discharge, as are public-funded events such as birth or geriatric care at private hospitals (MoH 2010). The NMDS also contains some information about privately funded hospitalisations/ hospitalisation events (MoH 2010). NMDS data are available for each cohort member for the period Jan 2008 – Dec 2011 and includes ICD-10 coded diagnostic categories, detailed cost data and other relevant information.

In order to use NMDS data to explore the impact of WUNZ:HS on hospitalisation costs it is necessary to identify hospitalisations that might conceivably be affected by

insulation and/ or improved space heating. This process is described in the following subsection.

3.3.2.1. Hospitalisation categories and exclusions

The NMDS contains a number of hospitalisation types that can safely be excluded from this analysis. Firstly, there is strong reason to exclude hospitalisations at private hospitals that are not publicly funded, because such events typically have a lag in reporting which means that the data will be incomplete if collected soon after the events occur; in addition reporting may vary by hospital. Fortunately, privately funded events only make up a small proportion of events (approximately 6.4% in 2009/2010) (MoH 2012a). This exclusion was carried out as part of the WUNZ:HS evaluation study design; private hospitalisations were not provided to the Group.

Secondly, there are a large range of hospitalisation costs that are associated with birth that can safely be assumed to have very little to do with the WUNZ:HS programme and should be excluded since birth rates will be driven by events nine months earlier. Using the field "diag01" which is the primary diagnosis, we can exclude events that are clearly birth-related. Diagnostic fields in the NMDS are coded using standard ICD-10 codes[8]. Events with a primary diagnosis code beginning with "O" (ICD-10 Chapter Fifteen (Pregnancy, childbirth and the perineum)), "P" (ICD-10 Chapter Sixteen (Certain conditions originating in the perinatal period)) or "Q" (ICD-10 Chapter Seventeen (Congenital malformations, deformations and chromosomal abnormalities)) can be eliminated. In addition there is a separate event type field which codes birth events as "BT", these events can also be eliminated (Telfar Barnard, Preval et al. 2011).

Thirdly, waiting list admissions, coded in the admission type field as "WN" are excluded. Waiting list admissions are typically related to chronic conditions that are not subject to change over time (Telfar Barnard, Preval et al. 2011). As such, including such admissions would be at the very least a source of statistical noise. However, more concerning is the possibility that chronic conditions that lead to waiting list hospitalisations might also make a person particularly likely to be identified as a suitable candidate for WUNZ:HS. If so, regression to the mean effects might systematically bias analysis, making WUNZ:HS appear to reduce hospitalisation costs depending on the relative timing of hospitalisation and WUNZ:HS participation (Telfar Barnard, Preval et al. 2011).

Other exclusions include hospitalisations coded "N" in the New Zealand resident indicator field because non-New Zealand residents can be considered too different from the rest of the study cohort and less likely to have occupied the study address

[8] ICD-10 is a worldwide standard for classifying and coding diseases and health problems (WHO 2014).

during the entire period of the study; removing these hospitalisations can be expected to reduce statistical noise.

Accidents are an interesting category: an *a priori* plausible link could be made between changes in some categories of accidents and indoor temperatures: perhaps vulnerable individuals are more physically coordinated at higher temperatures and thus less likely to fall? In this thesis, this small possibility had to be weighed against the lack of data available on the location of events and the large proportion of accidents that are unlikely to have any link with indoor conditions: as such, admissions that had the purchaser code "A0" indicating that the Accident Compensation Corporation of New Zealand was the principal health service purchaser (MoH 2009b) have been excluded from analysis.

Finally, it is useful to exclude the majority of admissions with a primary diagnosis code beginning with "Z" (ICD-10 Chapter Twenty-One (Factors influencing health status and contact with health services)). This Chapter primarily consists of hospitalisation events linked to conditions that do not change or that are unlikely to be linked to housing (Telfar Barnard, Preval et al. 2011), however there are a number of ICD-10 Chapter Twenty-One 21 categories that are relevant from a cost perspective such as cardiac rehabilitation (Z50.0), post-surgery convalescence (Z54.0), and other and unspecified convalescences (Z54.8 and Z54.9). Hospitalisation events with these primary diagnosis codes are included in the analysis. Other types of convalescence such as post chemotherapy convalescence are excluded.

Possible exclusions were transfers (between or within hospitals) coded as "T" in the admission source field and readmissions (defined as another admission with the same primary diagnosis code within a defined period such as 30 days): analysis that focused on rates of hospitalisation would typically exclude such events as not being new events, however a cost-based analysis does not need to make this distinction: transfers and readmissions contribute to costs and should not be excluded. Table 11 sets out the key exclusions:

Table 11: Key Hospitalisation Exclusions

Category	Coding	Reason for exclusion
Birth events	Event type="BT"	Not adverse health events
Non-New Zealand residents	New Zealand resident indicator="N"	Not relevant to study
ICD-10 Chapter Fifteen (Pregnancy, childbirth and the perineum)	Clinical code starting with "O"	Majority not adverse health events; distribution of hospitalisations driven by birth rate and events nine-months prior.
ICD-10 Chapter Sixteen (Certain conditions originating in the perinatal period)	Clinical code starting with "P"	Majority non-relevant health events; distribution of hospitalisations driven by birth rate and events nine-months prior.
ICD-10 Chapter Seventeen (Congenital malformations, deformations and chromosomal abnormalities)	Clinical code starting with "Q"	Distribution of hospitalisations driven by birth rate and events nine-months prior.
ICD-10 Chapter Twenty-One (Factors influencing health status and contact with health services)	Clinical code starting with "Z" unless "Z50.0","Z54.0", "Z54.8","Z54.9"	Non-relevant health events; most do not represent conditions subject to change, others are not related to conditions likely to be impacted by WUNZ:HS i.e. post chemotherapy convalescence
Accidents	Purchaser_code "A0"	Majority of accidents unlikely to be associated with indoor temperature or dampness.

Having identified exclusions it is now useful to discuss key hospitalisation categories. A natural starting point is total hospitalisations. This category has the benefit of capturing all of the potential changes in hospitalisation costs resulting from participation in WUNZ:HS in a single useful measure. It is not ideal however, because it will naturally include noise from categories that cannot be expected to be influenced by the WUNZ:HS programme such as cancer-related hospitalisations. Two useful categories used in the WUNZ:HS evaluation were respiratory hospitalisations (primary diagnosis code beginning with "J", ICD 10 Chapter Ten (Diseases of the Respiratory System)) and circulatory hospitalisations (primary diagnosis code beginning with "I", ICD 10 Chapter Nine Diseases of the Circulatory System). These categories capture the types of hospitalisation most likely to be impacted by WUNZ:HS-related improvements in indoor temperature and reduced dampness and associated reduction in allergens discussed in Subsection 2.1.4. Other plausible categories include infectious disease and mental health-related hospitalisations, however there is much less evidence for these categories, and inclusion of too many additional categories of interest may risk charges of data dredging. Table 12 sets out the Hospitalisation Categories used in this thesis.

Table 12: Hospitalisation Categories.

Outcome	Description
Total hospitalisations	All hospitalisations other than exclusions (see above)
Circulatory illness	ICD-10 Chapter Nine
Respiratory illness	ICD-10 Chapter Ten

Adapted from (Telfar Barnard, Preval et al. 2011)

3.2.2.2. Cost variables in the NMDS

Hospitalisation costs are dictated by a number of factors such as the types of procedures required, complications and length of stay. Each NMDS hospitalisation event record includes comprehensive information relating to the cost of that event. A cost weight figure is given for each hospitalisation event: cost weights are calculated by factoring in DRG code (a categorisation system for severity), length of stay and in some cases other procedures and technical codes (MoH 2010). The cost weight can then be multiplied by a standard cost figure for that particular financial year in order to derive the cost of that hospitalisation.

Although not a cost variable, the length of stay for each hospitalisation, which can be calculated by the difference in "event start date" and "event end date" variables has the potential to be used as a starting point for valuing time lost to hospitalisation. This interesting possibility was not explored in the original WUNZ:HS health evaluations but is discussed in Chapter Five.

3.3.2.3. Possible issues for NMDS data

As with any dataset there are potential issues around data entry errors. There is no obvious reason to think that such errors could introduce systematic bias into this analysis. There are also possible issues around variation in how events are coded both over time (including code creep[9]) and by District Health Board (DHB). It is possible that such changes might introduce a systematic bias, but the matched cohort and difference-in-difference design should account for at least some of these potential distortions, as should the short time frame of four years.

3.3.3. Pharms data

Pharms is a "data warehouse" currently owned and maintained by the Ministry of Health which records claims by pharmacies for reimbursement for dispensed pharmaceuticals subsidised by PHARMAC[10]. When prescriptions are presented at pharmacies for subsidised items individuals pay a standard co-payment (unless they meet certain criteria) and details of the prescription are entered into pharmacy electronic records including NHI details if these have been recorded. When the patient co-payment does not cover the cost of the prescription for a subsidised item a claim for reimbursement is sent to the Ministry of Health. Reimbursement claims are recorded in the Pharms dataset and can be linked via NHI with other Ministry held datasets.

The ability to link Pharms data with changes in home insulation/ heating has a great deal of potential. A brief literature review suggests that only one housing intervention study has investigated the impact of improved housing on total pharmaceutical costs, the Watcombe Study, and this study was limited both by timeframe and also small study numbers which meant that results might "…. be influenced by one or two heavy service users whose conditions may have no connection with their housing situation" (Barton, Basham et al. 2007). Other research including my own (Preval, Chapman et al. 2010) has assessed some pharmaceutical costs, but only a limited subset (asthma medication). The opportunity to assess the impact of WUNZ:HS on usage of a comprehensive range of pharmaceuticals for a large cohort is thus unique.

[9] Code creep is a term without a single formal definition, but a useful definition, cited by Seiber (2006), is "changes in hospital record keeping practices to increase case mix and reimbursement" (Steinwald and Dummit 1989).

[10] PHARMAC is a Crown agency that decides which medicines and related products are subsidised for use in the community and public hospitals.

3.3.3.1. Pharmaceutical cost data

The Pharms data provided included considerable cost-related data, however several additional steps were required to fully cost each prescription. As one of my roles within the Group was to interpret and analyse Pharms data, I will reproduce almost verbatim the process that I carried out in Telfar Barnard, Preval et al. (2011). The details of the costing procedure were established using various government sources (New Zealand Health Information Service 2006 , PHARMAC 2007) and personal communications (personal communications with Chris Peck (PHARMAC) 2011; Chris Lewis (MoH) 2011). I later co-authored a technical report with a research pharmacist based on the processes described below (Foster, Preval et al. 2011). It is important to note that all costs/prices exclude GST (goods and services tax); from a cost-benefit perspective taxes represent transfer payments and should be excluded.

The following paragraphs in italics are directly quoted from Telfar Barnard, Preval et al. (2011) and are my own work. The key fields utilised from the Pharms data provided were:

- *HPAC cost ex supplier excluding GST (A): Baseline cost of pharmaceutical dispensed (for a given dispensation) based on subsidised price listed in Pharmaceutical Schedule and number of units dispensed.*
- *Retail subsidy excluding GST (B): HPAC cost ex-supplier plus a 4% or 5% retail mark-up.*
- *Dispensing fee value excluding GST(C): Each dispensation incurs a dispensation fee which community pharmacies charge the government. More complex or problematic dispensations (for example requiring special preparation) incur a higher fee.*
- *Patient contribution excluding GST (D): A small fee that patients may pay at time of dispensation towards the cost of a prescription in some circumstances. The fee does not apply to some groups including children younger than 6, or to repeat prescriptions.*
- *Reimbursement cost excluding GST (E): This is the total that the government pays to a community pharmacy for a given dispensation. It is calculated in the following way:*

$$E = B + C - D$$

In order to calculate the total cost of a given dispensation it is necessary to include both patient costs and government costs. In addition to the patient contribution described above, patients must pay for any portion of a given prescription which is not subsidised, and will typically also be charged a mark-up of 86% by the dispensing pharmacy including GST, this is a recommended mark-up set out in the Pharmaceutical Schedule but pharmacies may mark-up as they see fit in line with commercial imperatives: data are not collected by PHARMAC on such mark-ups (PHARMAC 2007). The majority

of products subsidised by PHARMAC are fully subsidised and do not require a patient contribution of this type.

To estimate the non-subsidised portion of a given dispensation I utilised the following available fields:

- *Formulation id: A six digit code that serves to identify the active chemical ingredient in a product, the amount present, and the product type (tablet, solution, injection).*
- *Price(F): This field contains a generic pack price associated with the relevant formulation id (e.g. the price of 100 tablets if the product is wholesaled in 100 tablet packs).*
- *Subsidy (G): This field indicates the subsidy that would be paid for a generic price pack associated with a formulation id (e.g. the subsidised price that the government would pay for 100 tablets if the product is wholesaled in 100 tablet packs). The majority of dispensations in our data set involved products with a 100% subsidy.*

I assumed that ((F-G)/F) represented the unsubsidized proportion associated with a given formulation at the time of dispensing and used the following formula to estimate the pre-mark-up unsubsidised cost of a given dispensing:

- *Unsubsidised cost excluding GST (H) = ((F-G)/F)*A where A is the HPAC cost ex supplier ex GST (see definition above).*

I then calculated:

- *Unsubsidised cost including PHARMAC suggested mark-up excluding GST (J) = H*1.86/1.125 or H*1.86/1.15*
  (GST was 12.5% Jan 2008-Sept 2010, and 15% Oct 2010 onwards)
- *The total cost to the nation of a given dispensing excluding GST (K) = E (government) + D (patient) + J (patient)*

A final complication is that PHARMAC negotiates additional confidential rebates from pharmaceutical companies. These rebates vary from year-to-year, do not apply equally to all products, and for reasons of commercial sensitivity the details of these rebates are not made public (and cannot be ascertained from Pharms data). I made an estimate of the average rebate using the PHARMAC Annual reports which report total rebates for community pharmaceuticals. In the 2007/2008 financial year gross expenditure was $751.71 million reduced by estimated supplier rebates of $114.89 (a 15.2% reduction). The estimated rebate reduction for the 2008/2009 financial year was 14.3% and in 2009/2010 8%. The estimated reduction for 2010/2011 was 14.15% (115.8/818.6).The

estimated reduction for 2011/2012 was 17.46% (161.5/924.8). I calculated cost figures assuming that these rebates apply in the following way:

- *The total cost to the nation of a given dispensing excluding GST (L) = E (government) + D (patient) + J (patient) – applicable rebate *(E+D)*

(Telfar Barnard, Preval et al. 2011, pp 21-23).

3.3.3.2. Pharmaceutical categories

As with hospitalisation events, it is important to establish categories of pharmaceuticals that seem most likely to be impacted by changes in insulation and heating. Following the discussion of hospitalisation categories these are respiratory-related pharmaceuticals and circulatory illness-related pharmaceuticals. It is also useful to establish a total pharmaceutical category that captures all changes in pharmaceutical use, while acknowledging that this category would likely include a great deal of statistical noise from pharmaceutical categories that are unlikely to be affected by the intervention. It was not practical to exclude pharmaceuticals associated with birth as most pharmaceuticals have many uses.

In order to categorise the Pharms data it was necessary to link each formulation id with the SiMPle database, an online database made available by Pharmac which contains extensive details for every product that has been subsidised by Pharmac during its operation. The key information acquired from the SiMPle database was a three level ATC (Anatomic Therapeutic Chemical) code classification for each formulation id. The Anatomical Therapeutic Chemical (ATC) Classification System is used to classify drugs and is controlled by the WHO Collaborating Centre for Drug Statistics Methodology(WHO 2012).

Expert advice was sought, utilising a comprehensive list of ATC codes from the SiMPle dataset in order to identify pharmaceuticals whose usage rates might theoretically be altered by a change in insulation or heating (personal communications with Professor Julian Crane, University of Otago, Wellington, 2011 and Dr Scott Metcalfe, Pharmac, 2011). Table 13 sets out the categories used.

Table 13: Pharmaceutical Categories

Outcome	Description
Total dispensations	All dispensations
Circulatory illness-related dispensations	ATC Code Level 1: Cardiovascular System OR ATC Code Level 1: Blood and Blood Forming Organs, Chemical name "Aspirin" OR ATC Code Level 3: "HMG CoA Reductase Inhibitors (Statins)"
Respiratory illness-related dispensations	Chemical name "Prednisone" OR ATC Code Level 3: "Inhaled Corticosteroids", "Inhaled Corticosteroids with Long-Acting Beta-Adrenoceptor Agonists", "Beta-Adrenoceptor Agonists", "Inhaled Beta-Adrenoceptor Agonists", "Inhaled Anticholinergic agents", "Inhaled Beta-Adrenoceptor Agonists with Anticholinergic Agents", "Methylxanthines", "Other Bronchodilators" and "Cough Preparations"

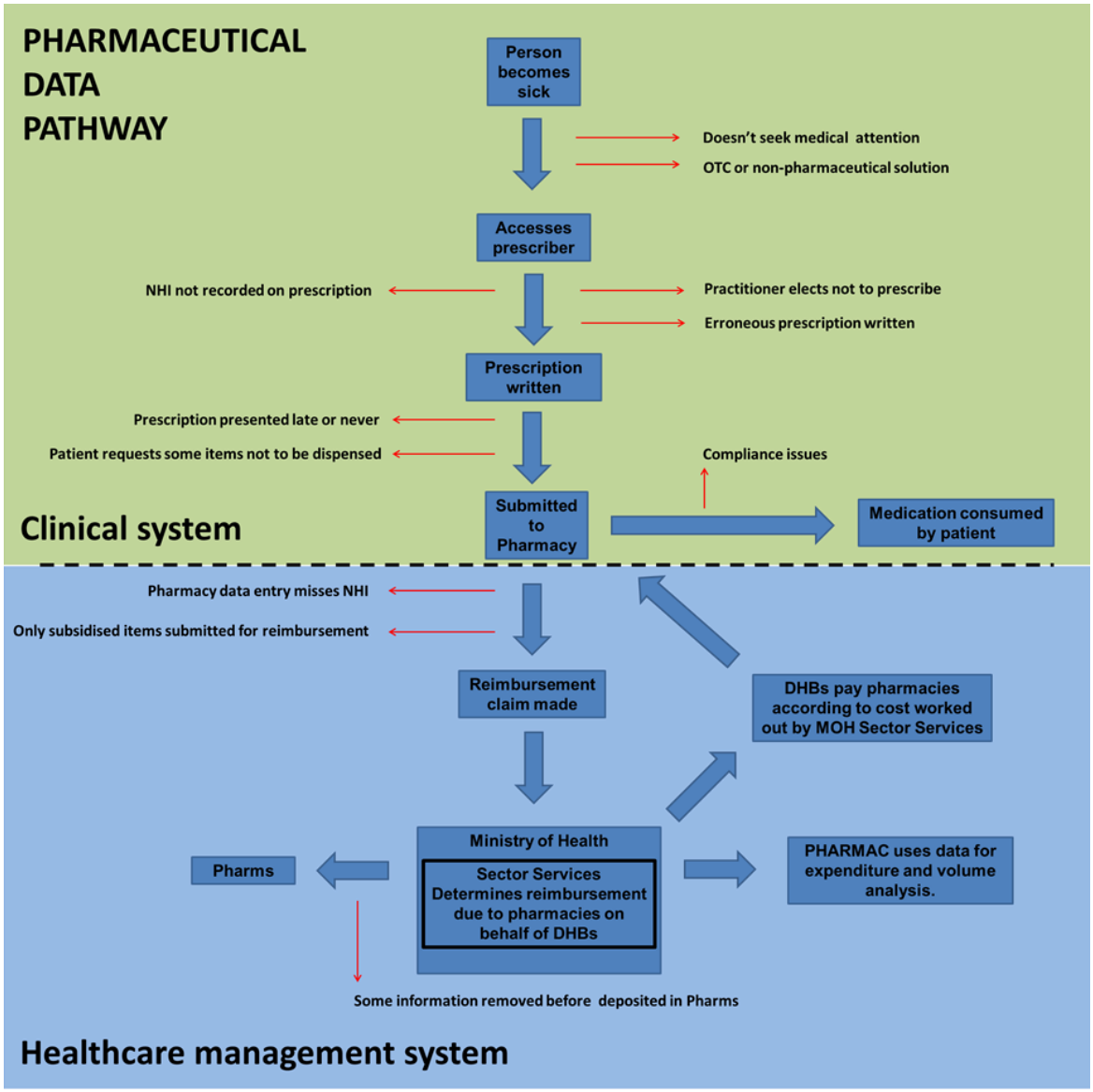
(Telfar Barnard, Preval et al. 2011)

3.3.3.3. Possible issues for Pharms data

As part of my work in the Department of Public Health I co-supervised a former pharmacist and current medical student, Daniel Wellington on a "summer student" project aimed at understanding the value and limitation of Pharms data, with the goal of publishing a paper on the ways in which New Zealand's pharmaceutical data could be improved from both a research and a monitoring perspective. Our work was very useful for both the WUNZ:HS evaluation and this thesis in identifying the limitations of the data.

The following figure, produced as part of an intended technical report summarises the process by which Pharms data are created, highlighting the many opportunities for data to go unrecorded or be miss-recorded.

Figure 5: Pharmaceutical Data Pathway



(Daniel Wellington, used with permission)

As Figure 3 demonstrates, there are many reasons to be wary about the comprehensiveness of the data contained within Pharms. Many of the "leakages" can be assumed to apply equally over time and to both the treatment and control groups such as erroneous prescriptions, unfilled prescriptions and missed NHIs on prescriptions. More of a concern is the leakage labelled "only subsidised items submitted for reimbursement". This leakage covers prescriptions that are fully paid for by the patient's co-payment, which are never claimed for and thus not included in Pharms. There is an untestable possibility that there may be some impact of WUNZ:HS on medicines that typically are fully funded by patient co-payments; analysis of Pharms data will not capture such changes. It is also possible that changes in co-payment rates might introduce systematic bias if the two groups differ with regard to prescription rates for products that this would affect, however, fortunately the first

major change in co-payment rates in 20 years was announced in the 2012 Budget, outside the 2008-2011 period of analysis.

There are other issues with Pharms data. There is likely to be an issue of lag, with some repeat prescriptions covering a period of up to three months from the initial prescribing time, thus creating situations where prescriptions dispensed in the *after* period for a given individual were actually written several months before in the *before* period. Such issues are likely to affect prescriptions for individuals from the treatment and control group equally, but could bias results towards the null by effectively diluting the *after* period.

3.3.4. QV data

Following the matching process carried out by QV the following fields were available for each treatment and control address:

Table 14. QV data available for each study home

Category	Description
QPID	Unique QV identifier
Territorial Authority	District or City Council boundary that home falls within.
Suburb	Suburb as per address
Census Area Unit	Stats NZ defined areas – there are approx. 1860, of varying population sizes, covering the whole of NZ
Category	Industrial/residential/commercial
House-type	Apartment/ Bach/ Bungalow (Post-war) / Contemporary/ Cottage/ Pre-war Bungalow/ Quality Bungalow/ Quality Old/ Spanish Bungalow/ State Rental/ Terrace Apartments/ Townhouse/ Unit/ Villa
Floor area (m^2)	Floor area
Number of levels	Number of levels
Number of bedrooms	Number of bedrooms
Decade built	Decade built
Roof material	Roof construction material – vast majority steel or tile
Roof condition	Average/fair/good/mixed/poor/null
Building material (external walls)	Wall construction material
Building condition (external walls)	Average/fair/good/mixed/poor/null
Number of single garages under main roof	How many single garage equivalents under main roof: a double garage would = 2 units
Modernised	Indicates if home is modernised – Y/N

As described in Section 3.1, the matching process means that there is a high degree of similarity between treatment homes and their matched control homes with respect to these variables. While some matches were not mandatory, I would not expect variation in the quality of matches to be a major source of bias or statistical noise, because those fields considered most relevant such as Census Area Unit or house-type were mandatory matches.

However, this thesis is designed to explore whether changes in insulation or heating under the WUNZ:HS study result in changes in health-related costs. The primary causal mechanisms proposed are increased internal warmth and reduced mould and damp, and as such variation in building qualities between the treatment and control group homes could influence the modelled impact of treatment by introducing bias or statistical noise. It is also important to compare cohort building quality with that of the general population because this is relevant to the question of generalisability.

Lucy Telfar-Barnard developed a categorical variable that captures much of the physical risk profile of cohort homes, based on previous research in her doctoral thesis (Telfar-Barnard 2010). Quoting directly from Appendix 2 (Telfar Barnard, Preval et al. 2011, pp 55-56):

> *The authors developed a six-category dwelling health risk typology for this study. This classification was based upon previous research and unpublished analysis of health outcomes by dwelling type, and is based upon QV data for dwelling construction decade; dwelling type; roof and wall condition; and the overall dwelling condition indicator included in the "category" field.*
>
> *The dwelling health risk typology has two components: a binary construction type indicator (new, old), and a ternary condition indicator (high, average, low).*
>
> *Dwellings were assigned to the "old" construction category if they were built before 1950; or belonged to QV house type categories "Pre-war Bungalow", "Spanish Bungalow", or "Quality Old". Otherwise, they were assigned to the "new" construction category, as were pre-1950 dwellings with the QV house type "Contemporary".*
>
> *For the condition indicator, dwellings were initially assigned to the "average" category. Second, any dwellings with a QV overall condition category of "Superior" were reassigned to the "low risk" category. Next, dwellings were reassigned to the "high risk" category if:*
>
> - *the QV dwelling quality indicator was "Poor";*
> - *the Wall condition was "Poor";*
> - *the Roof condition was "Poor";*
> - *the dwelling was built in the 1930s or 1940s; the dwelling type was not "Quality Old", "Quality Bungalow" or "Contemporary"; and the overall condition was not "Superior"; or*
> - *the dwelling was a "Post-war Bungalow" built in the 2000s and its overall condition was not "Superior".*

Finally, dwellings were reassigned to the "low risk" category if the QV dwelling type was "Quality Old", "Quality Bungalow" or "Contemporary" and the dwelling had not previously been assigned to the "high risk" category. The construction type and condition categories were then combined to create the six categories "Old high risk", "Old average risk", "Old low risk", "New high risk", "New average risk", and "New low risk".

It is useful to compare the distribution of the "dwelling health risk" variable for the treatment group and the control group. Naturally, it is not possible to compare this data with Census data, because the "dwelling health risk" is derived from QV data and not recorded in Census data.

Table 15: Distribution of "dwelling health risk" variable for individuals from treatment and control group households

Dwelling health risk category	Treatment group households	Control group households	All cohort households
Old high risk	14,413	94,971	109,384
	13.38%	11.74%	11.93%
Old average risk	16,981	118,203	135,184
	15.76%	14.62%	14.75%
Old low risk	1,344	12,699	14,043
	1.25%	1.57%	1.53%
New high risk	3,160	25,469	28,629
	2.93%	3.15%	3.12%
New medium risk	63,611	492,482	556,093
	59.04%	60.89%	60.68%
New low risk	8,229	64,942	73,171
	7.64%	8.03%	7.98%
Total	107,738	808,766	916,504
	100.00%	100.00%	100.00%

Visual comparison of treatment group and control group individuals in Table 15 suggests that they are similar with regard to the distribution of the "dwelling health risk" variable, which is unsurprising given that the matching process was driven by components of the "dwelling health risk" variable such as decade built which was mandatory. However a chi-square goodness of fit test comparing the observed dwelling health risk category proportions for the treatment group with proportions found in the control group is highly significant (chi-square (5) = 445.64, p = 0.000). This statistical significance of this test result likely reflects the large sample size and is not a cause for concern.

Returning to Table 14 above, one of the key variables from the perspective of my thesis is the CAU field which enables analysis by region or Climate Zone (see subsection 3.2.8). We were also able to use QPIDs (unique QV identifiers) to link QV data with data previously supplied by QV, which was used by Lucy Telfar-Barnard as part of her

thesis in order to identify the meshblock [11] code for each treatment and control home (with a few exceptions where a valid meshblock was not available). Meshblock level data allowed us to identify the NZ2006 Index of Deprivation[12] for the meshblock that contains a given cohort home. Where meshblock level data were not available we used CAU level NZ2006 Index of Deprivation measures as the closest available proxy.

The basis of the evaluation presented in (Grimes, Young et al. 2011) was the potential to obtain metered energy use data for cohort addresses. The evaluations presented in Telfar Barnard, Preval et al. (2011) and Preval (2012) did not include energy use in their models; the limitations of that data (particularly reduced statistical power due to matching and data availability issues) were considered too great at that time. The available data are further described in Section 3.2.7.

Addresses also allow for the possible inclusion in models of temperature either by linking directly with local temperature data. The National Institute of Water and Atmospheric Research (NIWA) collects comprehensive weather data for New Zealand. The potential to link to this data was utilised in the WUNZ:HS evaluation assessment of energy use (Grimes, Young et al. 2011), which linked NIWA data to address data by CAU and then region (described further in Section 3.3.8).

Table 16 sets out the potentially useful datasets that can be linked using address data:

Table 16: Datasets linkable using address data

Dataset	Description	Unit(s) of linkage
NIWA data	Average temperature or average minimum temperature by month or season	Region (via linkage with CAU using Statistics NZ data) and time-period
Metered Energy Use data	Metered electricity use or metered energy use (electricity and reticulated natural gas) per household per month	Home address and time-period
Statistics New Zealand	Area-based measures such as 2006NZDep, region or Climate Zone	Meshblock (via linkage to additional QV dataset using QPID) and CAU

[11] Meshblocks are geographical units containing a median of 87 people. Statistics NZ divided Census 2006 data into approximately 41,393 meshblocks.

[12] 2006NZDep is a measure of deprivation derived from 9 socioeconomic measures recorded in the 2006 Census calculated at the meshblock level. 2006NZDep is calculated using a principal component score, which has been scaled to have a 1000 index point mean and a standard deviation of 100 index points. 2006NZDep is typically reported as an ordinal scale divided into 10ths, with 1 representing the least deprived 10% (the 10% with the lowest principle component scores) and 10 the 10% that are most deprived. 2006NZDep is also available at a CAU level, calculated by averaging the scores of the meshblocks within the CAU(MoH 2007).

3.3.4.1. Possible issues for QV data

It was not possible to run any tests or find any corroboration for the accuracy of QV's data or matching process. As one of the primary data sources for housing information in New Zealand, including ratings calculations made by local authorities, QV data are legally bound to be accurate (at least with regard to data relevant for ratings purposes); however, as a state-owned commercial entity the details of QV's data gathering and verification processes are a commercial secret, meaning that a great deal of trust has to be placed in the organisation. The issue of accuracy applies equally to data for treatment and control group homes and as such it seems likely that the impact of any inaccuracies in data or matching process on the models presented in thesis will be in the form of statistical noise, although bias cannot be ruled out given the opacity of QV's dataset and processes.

3.3.5. EECA data

EECA data were the basis of the dataset creation process. As EECA data only apply to treatment homes (and thus to treatment group individuals) they do have some limitations. It is important to note that the matching process did not directly account for the insulation status or heating status of control group households, which are not known, although the matching process selected homes that were physically similar thus and more likely to have similar insulation status. In addition to uncertain baseline insulation and heating status for control group households, both treatment and control group households may have obtained retrofits outside of the WUNZ:HS programme during the course study period. This concern is further explored in Subsection 9.1.4.

The key data from EECA include details of the types of installation carried out for each retrofit, which allow the treated home and matched control homes (and occupants) to be categorised with regard to work done. The key analytic categories are "insulation" and "heating"; although other categories could be explored such as "draught-proofing".. The detail of these definitions is further explored in Subsection 4.1.4.

EECA data also identifies the status of the home prior to installation with regard to insulation and whether the household participated in the WUNZ:HS programme at the higher level of subsidy due to including an occupant that was a Community Services Card (CSC) holder[13] . CSC status is of significant interest from a policy perspective; with previous programmes exclusively targeting this group it is interesting to explore any difference in health effects by CSC status. The literature would suggest health effects would be smaller in non-CSC homes due to assumed higher pre-treatment indoor temperatures. Ideally the study design would have included a measure of CSC

[13] CSCs are available for permanent legal residents of New Zealand (and refugees) who meet age criteria (usually 18 or over) and meet income criteria (low-middle income or receipt of income support). Income criteria are dependent on relationship status and family size. CSCs qualify holders for a health service cost reductions. Source: www.workandincome.govt.nz

status in the matching process, however such information is not held by QV; location in the same CAU and occupying housing with very similar physical characteristics is a weak proxy for CSC status. Table 17 sets out the key EECA variables from an analytic perspective.

Table 17: Key EECA variables

EECA field	Description	Use of field
Measure	Type of work done: e.g. "ceiling insulation, none existing", multiple rows in the dataset for each household, one for each separate installation procedure.	Categorise treated homes and their matched controls based on work done: key categories: "insulation" and "heating", potentially "insulation and heating". Other categorisations possible for relatively cheap procedures such as draught-proofing, can also identify homes with insulation prior to treatment.
Low income	Defines whether the household participated in the WUNZ:HS programme as a CSC holder.	Categorise treated homes and their matched controls by CSC status – potential to include as a variable in models or to stratify using this variable.
Yrmonth	Variable identifying the year and month a given retrofit took place in –majority of retrofits for a given home reported in same time period suggesting all work done in short timeframe.	Relevant study time for a treatment household (and matched controls) before this month defined as *before*, this month defined as "." and excluded from analysis, study time after this month defined as *after*.

3.3.6. Metered Energy Use Data

Grimes and colleagues (Grimes, Young et al. 2011) collected metered energy use data for cohort addresses from all of New Zealand's energy retailers except one energy company whose data was not in a form that enabled analysis and who declined to participate (Contact Energy, 24.7% New Zealand market-share). Metered energy use data are measured at the ICP (installation control point) i.e. the meter. Submissions on usage volume are made monthly by energy retailers for each electricity and gas ICP to separate central authorities. Submission volumes are expressed in kWh and are based on a combination of actual meter readings, and also modelled and estimated usage (Grimes, Young et al. 2011).

Approximately half (49%) of cohort addresses could be matched with metered energy use data covering the period Jan 2008 – November 2010. Various additional steps were taken to clean the data such as eliminating obviously false records and addresses that switched energy provider during the study (Grimes, Young et al. 2011).

From the perspective of this thesis, there are two primary concerns that make exploring the inclusion of metered energy use data in hospitalisation cost, pharmaceutical cost or mortality models non-viable. Firstly, the initial 49% match rate would mean losing a great deal of statistical power, less of a concern when modelling regular events such as monthly energy use readings, but much more of a concern when modelling rare events such as death or hospitalisation. Secondly, metered energy use data were only collected for the period Jan 2008 – November 2010, while health data are available for the period Jan 2008 – Dec 2011, thus including metered energy use data would also limit statistical power by reducing the length of the study, unless further energy data were collected which would be restrictively difficult without the backing of a government funded evaluation.

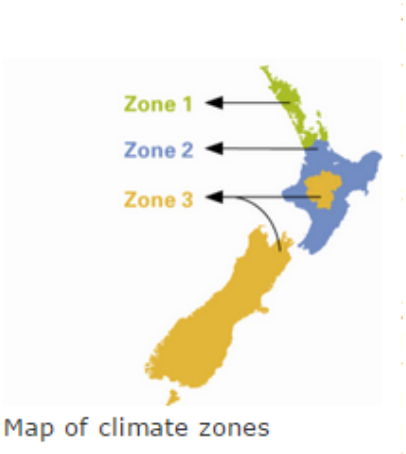
3.3.7. Climate data

There are several useful approaches that can potentially be used to account for the impact of climate on the health of study participants.

Firstly, it is possible to include regional dummy variables in models. This means the introduction of 15 variables for the 16 regions that are represented in the dataset. This approach will not capture variation over time and may also capture variation in measures other than climate such as employment and population health, which may be considered a limitation if the primary goal is capture the impact of climate. There are many reasons to consider the inclusion of a region variable and these are discussed further in Subsection 4.1.4.

Secondly, it is possible to categorise each study home by Climate Zone. Climate Zones are set out in Appendix B of New Zealand Building Standard NZS 4218, which divides the country into three Climate Zones which have different requirements for, for example, minimum ceiling and floor insulation values. Zone One includes the top of the North Island until just south of Auckland, Zone Two includes the remainder of the North Island with the exception of the Central Plateau. Zone Three includes the Central Plateau of the North Island and the entirety of the South Island. Climate Zones are set out in Figure 6.

Figure 6: Climate Zones



(Source: www.dbh.govt.nz/building-code-info-sheets)

The benefits of addressing variations in climate using Climate Zones are a less complex model and also the policy significance of Climate Zones. However, this trade-off will inevitably involve the loss of some information, and the same issues that limit region as a measure of climate variation also apply to Climate Zones.

Thirdly, it is possible to categorise study time, for example to divide study time into winter and non-winter periods. Time-based variables are not appropriate for some analyses such as the Cox Proportional Hazard-based models presented in Chapter Four.

Fourthly, it is possible to include temperature data collected by NIWA in order to capture the impact of variation in temperature over time. NIWA is a Crown Research Institute, a government-owned research organisation. Using the CliFlo data portal (http://cliflo.niwa.co.nz) NIWA make data from 6,500 climate stations set around New Zealand publicly available, with some observations dating back to 1850. There are currently 600 climate stations operating in New Zealand, which provide a vast range of data on temperature, wind, rainfall, humidity and other weather phenomena. Climate station data are available from CliFLo in both raw and statistically aggregated forms.

Given the CAU of each cohort address it is possible, as noted in Subsection 3.2.5, to link study data with CliFlo data; one approach to this linkage was taken by (Grimes, Young et al. 2011). In that paper the CAU of each cohort address was linked to CliFLo data at a regional level. The CAU (or nearest CAU) for each weather station that provided consistent monthly average temperature data over the period studied (Jan 2008 – Nov 2010) was identified. If there was a single suitable weather station in a region then the monthly average temperatures recorded by that station were taken as proxies for all cohort addresses in that region. When there were multiple suitable weather stations within a region the weather station located in the CAU with the highest population density (calculated by dividing the usually resident population established in the 2006 Census by the area of the CAU) was chosen as the proxy for that region (Grimes, Young et al. 2011). Monthly measures of mean temperature were collected for the identified weather station for each of the 16 regions that contained study homes. It is important to note that the use of region-based climate data cannot account for differences in micro-climate, and thus introduces measurement error into the analysis and the potential for the ecological fallacy. Reassuringly, any distortion introduced by measurement error in this context is likely to bias results towards the null.

Other sources of external temperature data are potentially available. For example, Telfar-Barnard (2010) obtained annual mean minimum temperature data at CAU level from Landcare Research, another Crown Research Institute. Greater accuracy would clearly be a benefit of obtaining CAU specific data. However, due to data and time limitations, in this thesis I adopted the approach taken in Grimes, Young et al. (2011), using the same 16 weather stations as proxies for the 16 regions.

More detailed data such as daily measures of minimum and maximum temperature were initially considered, but ultimately rejected. It is also worth noting that high external concentrations of atmospheric pollutants such as ozone and particulate matter have been correlated with rises in mortality and morbidity and may confound analyses

that do not account for them. However, CliFlo does not include access to measures of pollutants at the Climate Station level.

3.3.8. Imputed benefits: GP data, days off work data and days off school data

One of the key limitations of the WUNZ:HS evaluation study design was that it did not enable researchers to directly assess the impact of improved insulation or heating on visits to general practitioners (GPs). Previous research by our research team (Howden-Chapman, Matheson et al. 2007) found an unexpected small but statistically significant *increase* in visits to GPs, based on GP records, following improved insulation for a study population, in which each household contained a member with a respiratory ailment. A subsequent study (Howden-Chapman, Pierse et al. 2008), with a study population in which each household included an asthmatic child (aged 0-12) found a statistically significant decrease in GP visits for asthmatic children. Because the Ministry of Health does not have a database that records individual GP visits linked to patients' NHI numbers it was not possible to assess any change in the rate of GP visits using administrative data in the WUNZ:HS evaluation.

Similarly, it was not possible to assess any changes in days off work or school. Howden-Chapman, Matheson et al. (2007) and Howden-Chapman, Pierse et al.(2008) both report a reduction in days off school as a result of improved insulation and improved heating respectively and Howden-Chapman, Matheson et al. (2007) report a reduction in self-reported days off work for adults following improved insulation. It would not be realistically possible to independently assess this potential benefit of the WUNZ:HS programme for the complete cohort given the thousands of schools and work-places that would be involved and the absence of any NHI linkable administrative data that records absences.

Although it was not possible to directly assess the impact of the programme on GP visits, days off work or days off school, from a cost-benefit analysis (CBA) perspective it was not reasonable to exclude these categories given that previous research had suggested that statistically significant savings in these areas might result from improved insulation and heating. Using work carried out by the *He Kainga Oranga* programme (Chapman 2009), including my own previous work (Preval, Chapman et al. 2010), I attempted to apply results derived from previous randomised controlled trials to the treatment group of the WUNZ:HS evaluation, in order to produce a rough estimate of potential changes in benefits (Telfar Barnard, Preval et al. 2011). Although not a major focus of this thesis, this analysis does contribute to the cost-benefit figures discussed in Chapter Seven and Nine and is relevant to the general topic of this thesis. As such, I have presented the work that I produced for the WUNZ:HS evaluation verbatim. Some of the assumptions differ slightly from those used in this thesis, for example inclusion of people aged 101-104; it does not seem useful to adjust these

figures given that they are presented to illustrate an approach rather than as a central piece of analysis.

It was not possible for the present study to address a number of key potential benefits of improved insulation, heating and health that have been found in previous research carried out by the Housing and Health Research programme (Howden-Chapman, Matheson et al. 2007) . These include reduced frequency of GP visits, reduced days off work and reduced days off school. Statistically significant benefits found in our previous studies are set out in Table 18. The benefits were calculated for the winter period (June –August) for the Housing, Insulation and Health Study (Howden-Chapman, Matheson et al. 2007, Howden-Chapman, Pierse et al. 2008), and (June – Sept) for the Housing, Heating and Health Study (Howden-Chapman, Pierse et al. 2008) but then extrapolated to capture an entire year.

The Housing, Insulation and Health Study was not powered to find a statistically significant reduction in respiratory hospitalisation admissions, but it did find a non-statistically significant reduction in respiratory hospitalisation admissions (p = 0.16 adjusted). The Housing, Heating and Health Study was similarly not powered to find statistically significant results for respiratory hospitalisation outcomes. Comparison of the respiratory hospitalisation admission results for the Housing Insulation and Health Study with the non-statistically significant reduction in asthma hospitalisations and RSV observed in the present Study suggests some basis for extrapolation in the case of insulation.

Table 18: ***Statistically significant health-related savings documented in previous H&HRP analyses***

Feature	***Housing, Insulation and Health Study*** ***Cost-benefit analysis results reported in Chapman, Howden-Chapman et al. (2009)***	***Housing, Heating and Health Study*** ***Cost-benefit analysis results reported in (Preval, Chapman et al. 2010)***
Study Description	*A cluster randomised controlled trial comparing health and energy use outcomes for treatment and control group individuals/ households following a standardised insulation retrofit.*	*A cluster randomised controlled trial comparing health and energy use outcomes for treatment and control group individuals/ households following a standardised heating retrofit (replacement of an inefficient heater with either a heatpump, flued gas heater or pellet burner). Treatment and control group homes were insulated before baseline measures were taken.*
Participant characteristics	*Each participant household included at least one occupant who had symptoms of respiratory disease; households were in predominantly low-income communities.*	*Each participant household included at least one asthmatic child aged 6-12; households were located in the South Island or Lower North Island.*
Reductions in GP visits	*An increase(negative result) of 48 GP visits per 1000 occupants per winter (0.05 visits per occupant) based on GP records.* *Valued at -$3.60 ($2001) per occupant per year when extrapolated to include the entire year.* *[Note: a statistically significant reduction in self-reported GP visits was observed but preference was given to GP records]*	*A reduction of 0.37 self-reported visits to the GP per winter for each asthmatic child, valued at $22.73 per child ($2006) when extrapolated to include the entire year.* *No other statistically significant reductions in GP self-reported GP visits.*
Reductions in days off work	*Adults aged 19-64: a reduction of 102 days off work per 1000 adults per winter (0.10 days per adult)* *Savings valued at $16.84 ($2001)per adult per year based on 80% (to adjust for co-workers "picking up the slack") of the 2001 average wage and extrapolated to include the entire year.*	*None found.*
Reduction in days off school	*Children aged 6-11: a reduction of 512 days off school per 1000 children over winter (0.51 days per child).* *Saving valued at $11.51 ($2001) per child per year when extrapolated to include the entire year, valued based on 1/3 of the youth minimum wage and 7 hours lost.* *Teenagers 12-18: a reduction of 1316 days off school/work per 1000 teenagers per winter (1.32 per teen)* *Savings valued at $59.21($2001) per teenager per year when extrapolated to include the entire year, valued based on*	*A reduction of 1.8 days off over winter terms per asthmatic child (aged 6-12), valued at $41.50 ($2006) when extrapolated to include the entire year, valued based on 1/3 of the gross daily minimum wage.* *Further benefit of reduced caregiver time were valued at $124.50 per asthmatic child per year ($2006)* *No statistically significant changes for children without asthma.*

	2/3 of the youth minimum wage and 7 hours lost.	

Table 19 reports the demographic profile of the treatment households which we have occupant information for as at July 2009 (the start of the WUNZ:HS programme), excluding occupants with no age or an implausible age (>105) or who have died. We also report the proportion of child occupants aged 6-11 that we predict will have asthma (15.1%) (Masoli, Fabian et al. 2004) as this is relevant for the estimation of heating-related benefits.

Table 19: ***Demographic profile of treatment households as at July 2009***

Demographic measure	***n***
Number of dwellings	*29,704*
Number of occupants	*107,421*
Average number of occupants	*3.61*
Children aged 0-5	*9,115 (8.49%)*
Children aged 6-11 estimated asthmatic	*1,168 (1.09%)*
Children aged 6-11 estimated non-asthmatic	*6,567 (6.11%)*
Teenagers aged 12 -18	*9,347 (8.7%)*
Adults age 19-64	*63,121 (58.76%)*
Older people (65+)	*18,103 (16.85%)*

It is possible to apply the data produced by previous studies to the demographic profile set out above. In order to do so we need to update the figures used in our previous cost-benefit analyses to 2011 prices and resolve any inconsistencies between the two approaches. The key inconsistency is the valuing of childcare for sick children, which was costed by the Housing, Heating and Health Study but not the Housing, Insulation and Health Study. We present our results with childcare costed separately, but in general we favour the inclusion of reduced childcare costs and consider our valuing of such costs at 6 hours at minimum wage ($13) to be reasonably conservative. We assume that lost time from school should be valued at 2/3 of 6 hours at the minimum hourly wage ($13) for teenagers and 1/3 for children aged 6-11, the youth minimum wage no longer applying. Adult time off work is valued at 80% of an 8 hour day at an average hourly wage of $25.96 ($2011).

Table 20 presents a summary of the additional benefits imputed per year per 1000 dwellings receiving either an insulation retrofit (floor and or ceiling) or a heating retrofit based on our previous analyses and assuming the demographic distribution set out in Table 20 (there was an average of 3.61 people per dwelling, so when looking at 1000 households we estimated benefits for 3,610 occupants).

Note that for reasons of convenience we conservatively assume that the reduction in days off school reported in the Heating, Housing and Health Study for asthmatic children only apply to ages 6-11 rather than 6-12.

Table 20: ***Additional imputed yearly benefits per 1000 households assuming 3.61 occupants per household and age structure of treatment group as at July 2009***

Age group	*Number of occupants per 1000 households*	*Potential benefit*	*Value per occurrence ($2011)*	*Change in number of occurrences per winter*	*Change in number of occurrences per year (adjusting for cold days in non-winter period)**	*Insulation: Predicted benefit per 1000 households per year**	*Heating: Predicted benefit per 1000 households per year*
0-5	*306.489*	*Reduced medical visits*	*$54.34*	*-0.05*	*-0.0835*	*-$1,390.70*	*$0.00*
6-11 (asthma)	*39.349*	*Reduced medical visits*	*$54.34*	*-0.05 (insulation) & 0.369 (heating)*	*-0.0835 (insulation) + 0.461 (heating)*	*-$178.55*	*$986.29*
		Reduced days off school	*$26.00*	*0.51 (insulation) & 1.8 (heating)*	*0.765 (insulation) + 2.025 (heating)*	*$782.65*	*$2,071.72*
		Associated reductions in caregiver costs	*$78.00*	*0.51 (insulation) & 1.8 (heating)*	*0.765 (insulation) + 2.025 (heating)*	*$2,347.95*	*$6,215.17*
6-11 (without asthma)	*220.571*	*Reduced medical visits*	*$54.34*	*-0.05*	*-0.0835*	*-$1,000.85*	*$0.00*
		Reduced days off school	*$26.00*	*0.51*	*0.765*	*$4,387.16*	*$0.00*
		Associated reductions in caregiver costs	*$78.00*	*0.51*	*0.765*	*$13,161.47*	*$0.00*
12-18	*314.07*	*Reductions in medical visits*	*$54.34*	*-0.05*	*-0.0835*	*-$1,425.10*	*$0.00*
		Reduced days off school/work	*$52.00*	*1.32*	*1.98*	*$32,336.65*	*$0.00*
19-64	*2121.236*	*Reductions in medical visits*	*$54.34*	*-0.05*	*-0.0835*	*-$9,625.16*	*$0.00*
		Reduced days off work	*$166.14*	*0.1*	*0.167*	*$58,855.92*	*$0.00*
65+	*608.285*	*Reductions in medical visits*	*$54.34*	*-0.05*	*-0.0835*	*-$2,760.11*	*$0.00*
Net benefit of insulation retrofit per 1000 households per year (Sum of Predicted Benefits for each age group)						***$95,491.33***	
Net benefit of insulation retrofit per 1000 households excluding childcare savings per year (Sum of Predicted Benefits for each age group)						***$79,981.90***	
Net benefit of heating retrofit per 1000 households per year (Sum of Predicted Benefits for each age group)							***$9,273.19***
Net benefit of heating retrofit per 1000 households excluding childcare savings per year (Sum of Predicted Benefits for each age group)							***$3,058.01***

**Note: negative values indicate an increase in frequency or cost*

The figures presented in Table 20 suggest that, including childcare costs for sick children, we can impute additional health-related benefits of $95,491.33 ÷ 1000 =$95.49 per household per year for a household receiving insulation (floor and or ceiling) and $9,273.19 ÷ 1000 = $9.27 per household per year for a household receiving a heating retrofit given the demographic structure presented in Table 19. It is important to be cautious about these results given the characteristics of the households included in the Housing, Insulation and Health Study (at least one occupant with respiratory illness and located in predominantly low-income communities) and the Housing, Heating and Health Study (South Island and Lower North Island locations only and an asthmatic child aged 6-12). We suggest that given these complications it might be reasonable to place more confidence in these predictions for those households identified as low income households in the present study. For this reason, when estimating the imputed benefits of insulation for all treatment households we predict $95.49/2 = $47.75 annual benefit, reflecting approximately 50% CSC households. We include the full $95.49 for our sub-analysis of CSC households but do not include any imputed benefits for non-CSC households. Likewise we predict an annual benefit of $9.27/2 = $4.64 imputed benefit for improved heating for all treatment households, $9.27 for CSC households and no benefit for non-CSC households.

(Telfar Barnard, Preval et al. 2011, pp 46-52)

The approach presented above was a limited attempt to address health-related costs and benefits which were not directly assessable in the WUNZ:HS evaluation and are not assessable in this thesis.

CHAPTER 4. MODELLING AND VALUING CHANGES IN MORTALITY

Mortality is clearly a significant health outcome, which dominates the quantification of health burden for many important exposures (Lopez, Mathers et al. 2006). Changes found in mortality reported in the WUNZ:HS evaluation (Telfar Barnard, Preval et al. 2011) made up the primary benefit of the intervention reported in Grimes, Denne et al. (2011); it is thus fitting that the first of the three methodology chapters that make up the core of this thesis should address mortality.

Because this is the first chapter that directly addresses modelling it is useful to describe the basic approach taken in this study to model specification, inclusion of potential confounders, potential transformation of continuous variables, and interactions. This approach is set out in Section 4.1 below and applies to Chapters Four, Five and Six.

Having set out the approach to modelling in Section 4.1 this approach is applied to analyse the available mortality data in Section 4.2. In Section 4.3 the question of how to value changes in mortality rates is addressed; the literature on the value of a human life is explored and a range of justifiable values established for cost-benefit purposes. These values are combined with the results of Section 4.2 to calculate annual household level benefits of statistically significant reduced mortality. These annual household level benefits are later combined with statistically significant changes in hospitalisation costs and pharmaceutical use costs in Chapter Seven in order to calculate net health benefits per household. The figures calculated are used to explore the impact of different assumptions on cost-benefit outcomes.

4.1. Approach to modelling in this thesis

Rothman and Greenland (1998) presents a useful overview of the problems faced by researchers who wish to model the relationship between a dependent variable and exposure(s) of interest. A daunting number of models are possible, given several plausible model specifications and many combinations of confounders, continuous variable forms and interactions terms. Greenland (1989) notes that it is clearly not practical to consider every possible model and states that many researchers may respond to the abundance of model options problematically by making fairly mechanical choices of model form, covariates, and continuous variable forms, which means that better models (however better is defined) may never even be considered.

In the following subsections, informed by the literature, I set out my approach to model specification in this thesis, selection of confounders, transformation of continuous variables, selection of interaction terms and study specific issues around matching and clustering. I also characterise the exposures of interest in this thesis.

4.1.1. Initial modelling decisions

Once research questions have been established, the modelling process begins with the choice of the dependent variable and the establishment of criteria for the inclusion of individuals or households in the analysis. In this thesis dependent variables are chosen in order to best quantify changes in health-related *costs* as a result of participation in the WUNZ:HS programme, with the ultimate goal of estimating total household level changes in costs. For example, in the analysis of hospitalisation costs presented in Chapter Five the primary models are based on hospitalisation cost variables, which capture the direct cost of hospitalisation to the health system (see Subsection 3.2.3.2). An alternative approach would be to use the rate of hospitalisations as the dependent variable: however, this approach creates an additional layer of complications when costs then need to be assigned to changes in the rate of events after regression analysis.

The inclusion criteria for individuals or households are driven by the literature. An obvious starting point is to include all cohort members in the analyses, because the goal of the analysis is to capture the average impact of participation in WUNZ:HS. This approach can be slightly modified because there are a small number of treatment households (48, with 156 occupants) that did not receive either insulation or heating under the programme, but only some other treatment such as a moisture barrier. It is reasonable to remove these houses from the analysis (and their matched control houses) as the primary retrofits of interest, as stated elsewhere, are insulation and heating. This approach is further described in Subsection 4.1.3 below.

In the case of the mortality analyses presented in Sections 4.2 and 4.3, the decision was made to further limit analysis to people aged 65 and over with either a *before* period circulatory-illness-related hospitalisation or a *before* period respiratory illness-related hospitalisation, because there is stronger evidence for a relationship between mortality rates and indoor temperatures for this group, thus increasing the sensitivity of the model.

4.1.2. Model specification

Model specification (choice of model form) in this thesis is informed by approaches recommended in the literature and also by decisions made during the evaluation of WUNZ:HS. Where multiple viable model specifications are used in the literature, a preferred specification is chosen and the choice justified. Checks later are carried out to ensure that models constructed are consistent with underlying assumptions.

4.1.3. Exposures of interest

A number of products and services were available under WUNZ:HS including pipe lagging and draught-proofing. However, floor and ceiling insulation and heating made up the majority of the cost of WUNZ:HS and are the key retrofits of interest both theoretically and from a policy perspective, as noted in Subsection 3.2.6. Randomised controlled trials carried out by *He Kainga Oranga* (Howden-Chapman, Matheson et al. 2007, Howden-Chapman, Pierse et al. 2008) typically group floor and ceiling insulation as a single insulation dummy variable, as do all health-related analyses carried out as part of the WUNZ:HS evaluation (Telfar Barnard, Preval et al. 2011, Preval 2012). There are many legitimate reasons to use this grouping including the fact that successive government programmes have provided both types of insulation as a package; this thesis takes the same approach for reasons of simplicity and comparability with previous studies.

Table 21 sets out the proportion of treatment homes with suitable health data that received these retrofits.

Table 21: Summary of insulation and heating retrofits for treatment homes with suitable health data

Retrofit	Number	Percentage
Insulation (floor and/or ceiling) only	22,811	77.05%
Heating only	2,120	7.16%
Heating and insulation (floor and/or ceiling)	4,628	15.63%
Neither heating nor insulation	48	0.16%
Total	**29,607**	**100.00%**

It is unsurprising that such a high proportion of heating retrofits were also insulation retrofits; this reflects the fact that it was necessary to meet minimum standards of floor and ceiling insulation in order to receive a heating subsidy.

4.1.3.1. Timing of exposures of interest

The analyses we carried out in Telfar Barnard, Preval et al. (2011) and I carried out in Preval (2012) made the simplifying assumption that all retrofits were completed during the month in which the first treatment occurred. Study time for each treatment individual/household and for their matched controls was divided into *before* and *after* and the month in which treatment was carried out was excluded from analysis.

This may have been an unjustifiable assumption as a review of the data shows that in a non-trivial proportion of cases there was a multi-month lag from an initial insulation retrofit to receiving a heating retrofit, suggesting separate interactions with the WUNZ:HS programme over time: ignoring this this can be assumed to have reduced the possibility that any benefit of improved heating would be captured by analysis. To avoid this conservative simplification, this thesis treats the timing of insulation retrofits and heating retrofits as entirely separate in the month-based extended model analyses of hospitalisation costs and pharmaceutical costs. However, the Cox Proportional Hazards model used to analyse mortality data requires that covariates have an unchanging baseline value so the simplifying assumption that all retrofits occurred in the same month is retained.

Each month contributed to the analysis by a given individual/household following initial treatment is categorised based on treatment status as at that month, using a four part treatment dummy variable. This variable is "0" for control group individuals/households, "1" for treatment group individuals/households that have currently received insulation only, "2" for treatment group individuals/households that have currently received heating only and "3" for treatment group individuals/households that have received both insulation and heating. An individual/household that receives heating some months after insulation will contribute some time to analysis as "1" and some time to analysis as "3".

The month in which a given treatment occurred is excluded from analysis if that treatment is the first treatment received: i.e. if for a given home the first interaction with the WUNZ:HS programme was to receive an insulation retrofitted in July 2009, this month would be excluded from all analyses, January 2008 – June 2009 would be the *before* period and August 2009 – December 2011 would be the *after* period. When retrofits within an analytical category (insulation or heating) took place over several months the month in

which the first retrofit occurred is still taken to indicate the start of treatment; this is a slightly conservative simplifying assumption.

4.1.4. Selection of potential confounders and model construction

The process of adjusting for potential confounding takes place at multiple stages of a study. Initial decisions about possible sources of confounding are made at the design stage via data collection choices, which are generally based on previous research findings. Large-scale studies such as the WUNZ:HS evaluations and this thesis are also limited by the administrative data available.

Having collected data, and decided on an appropriate model specification, it is then necessary to select model covariates from those available. Walter and Tiemeier (2009) carried out a thorough survey of current epidemiological practice based on the 2008 publications of four major epidemiological journals and found that, of 300 suitable articles, 28% selected variables based on prior knowledge (ideally justified by explicit reference to the literature), 20% based on stepwise selection and 15% on a pre-specified change-in-estimate criterion. More recently developed methods[14] were not used (Walter and Tiemeier 2009).

Despite statistically significant differences in the distribution of well-known confounders such as age, gender and ethnicity between the treatment and control group, the *individual* level analyses we presented in the WUNZ:HS health evaluations (Telfar Barnard, Preval et al. 2011, Preval 2012) made minimal adjustments for confounding (age in the modelling of hospitalisation events, log average annual hospitalisation costs prior to treatment in the modelling of mortality). Household level models did not incorporate confounding at all. We used the following justifications:

1. Matching controlled for many relevant variables such as location, home quality etc. and to a lesser extent socioeconomic status.
2. Due to the large number of study participants, it was inevitable that statistically significant differences would be found in the distribution of potential confounders in the treatment and control group, but these might not reflect important differences.

[14] These methods include shrinkage and penalised regression, which "shrink" coefficient estimates towards values expected from background information or estimated from the data. Coefficients with greater variance are "shrunk" more than stable coefficients (Greenland 2008).

3. The models used in the analysis of hospitalisation costs and pharmaceutical use costs explicitly controlled for non-time varying confounding by including a fixed effect term.
4. The quasi-experimental design mimicked aspects of a randomised controlled trial, producing a similar distribution of measured potential confounders in the treatment and control groups. Including these covariates (which were partially controlled for by quasi-randomisation) while excluding unknown possible confounders would be to risk over-controlling.
5. Due to the large size of the dataset, and consistently low measures of goodness-of-fit for models, even implausible confounders included in models might have statistically significant coefficients if they had even a small correlation with an unmeasured or partially measured covariate that had an association with the dependent variable.
6. Potential biases that could have been created by matching different numbers of control homes for each treatment home in individual level analyses were addressed by randomly selecting only four matched control homes per treatment home in the analysis of individual level data (if more were available in the initial dataset).

This thesis takes a somewhat different approach to the issue of confounding from the approach described above, opting to initially construct a multiple regression model that includes all plausible confounders. There are a number of reasons why:

1. The original analyses did not explicitly account for the matched design of the WUNZ:HS by including the matching variables in analysis as potential confounders. Rothman and Greenland (1998, p 420) note that "*all matching factors should be treated as potential confounders, in that one should assess whether failure to control for them would affect interval estimates*".
2. There is an a priori epidemiological assumption that potential confounding by age and sex should be explored.
3. Using difference-in-difference-based models with no inclusion of potential confounders is problematic because it ignores the possibility that there may be changes over the course of the study (other than treatment) that may differentially impact the treatment and control groups.
4. The large size of the dataset means that, for hospitalisation and pharmaceutical use cost models, it is not necessary to consider trade-offs between precision (width of the confidence intervals of the coefficients of the exposures of interest) and the risk of excluding real confounders.

Using this inclusive approach, I review potential confounders that are available for each analysis, identifying plausible confounders using literature and common-sense. Initial analyses and comparison of the distribution of the variable or variables in question provide initial insight into the likely impact of model inclusion for each potential covariate thus identified, although no plausible confounders are excluded at this point in analysis. Initially I considered adopting arbitrary thresholds for inclusion in multiple regression models such as that analysis of each potential confounder should demonstrate a statistically significant coefficient in the case of simple regression analysis ($p<0.2$) or a Wald test result that is statistically significant ($p<0.2$) in the case of multiple dummy variables, informed by the recommendations of authors such as Hosmer, Lemeshow et al. (2008). However, some model covariates may only have a confounding effect in combination with other covariates in a multiple regression model (Greenland 2008), or when interacted with other confounders, thus exclusion has the potential to produce unreliable estimates of the coefficients and confidence intervals of the exposures of interest in a given model.

4.1.4.1. Transformation of covariates

In carrying out initial analyses and constructing the initial multiple regression model I do not dichotomise or otherwise transform continuous variables into categorical variables. While dichotomisation and categorisation produce model results that are easily presented and interpreted this involves a trade-off in statistical information and power (Royston, Altman et al. 2006). Royston, Altman et al. (2006) also note that dichotomisation of continuous variables causes additional problems when multiple dichotomised continuous variables are included in a multiple regression model which may leads to "spurious" correlation of one or more of the dichotomised variables and the dependent variable as well as spurious interactions between the dichotomised variables. In some cases this choice may have consequences for interpretability, for example, in the case of the NZ2006 Index of Deprivation variable, there are some issues given that it is common to present results using the ordinal scale derived from the continuous variable, or use other categorical variables such as NZ2006Dep quintiles, which have the advantage of ease of interpretation.

Having established the initial multiple regression model, the next stage is to assess the correct form of any continuous variables to be included in the model. A fractional polynomial[15] approach is taken, using the *fracpoly* command under STATA, and the

[15] Fractional polynomials are generated by comparing the goodness-of-fit of models containing various transformations of a given continuous variable (Royston, Ambler et al. 1999, Royston, Altman et al. 2006) . The transformation or transformations that produce the best fitting model are

original variable is optimally transformed. If there are multiple continuous variables in the model then the continuous variable with the strongest association with the dependent variable is assessed first, while temporarily using a dichotomous or categorical form for the remaining continuous variables in the model. This process is repeated for each continuous variable in ascending order of statistical significance measured by |Z| score until all variables are considered.

Transformations that do not result in a change the bounds of the confidence intervals of any exposure of interest by 1% are rejected. This approach is informed by discussion in Greenland (2008) who cites Rothman and Greenland (1998) who suggest a similar but less stringent criterion. Interaction terms are also considered for the model using this criterion.

This approach is adopted because model parsimony is a relevant consideration. Model parsimony is primarily a concern for the Cox Proportional Hazards model presented in Section 4.2. There are two reasons for this. Firstly, despite the large dataset, the effective degrees of freedom of such models are dictated by the number of events (e.g. deaths) rather than the number of individuals in the model. Peduzzi, Concato et al. (1996), proposed that at least 10 events should be available for each model parameter to avoid issues with model fitting algorithms. This suggests that a more cautious approach to the inclusion of additional variables may be prudent. Secondly, when the Cox Proportional Hazards model's assumptions are tested it is necessary to ensure that there are not any statistically significant interactions between functions of study time and model covariates (discussed further in Subsection 4.2.6.1). If the model contains such interactions proportionality is violated and the standard methods of assessing the validity of the Cox Proportional Hazards model no longer apply. However, if time-based interaction terms do not impact the point estimates and confidence intervals of the exposures of interest then it is pragmatic to exclude them from the model in order to retain the simpler model.

4.1.4.2. Interaction terms

Having produced a multiple regression model with transformed continuous covariates it is then important to test for interactions. Interaction terms should be selected by referring to the literature, or possibly by consultation with experts (Hosmer, Lemeshow et al. 2008). Interaction terms are introduced to the multiple regression model separately, and interaction terms that are statistically significant at the ($p<0.1$) level are identified. After constructing a multiple regression model that includes all interaction terms that meet this criterion removal of interaction terms is considered if such removal has no meaningful

---

identified. The approach used in this thesis considers a range of transformations of the continuous variable from ^-3 to ^3. Single term transformations are considered and so are transformations of a second term. 44 combinations are considered in total (Royston and Sauerbrei 2008, StataCorp. 2009).

impact on the confidence intervals of the coefficient(s) of the exposure(s) of interest, as described above. Removal takes place in order from least statistically significant interaction term to most statistically significant interaction term.

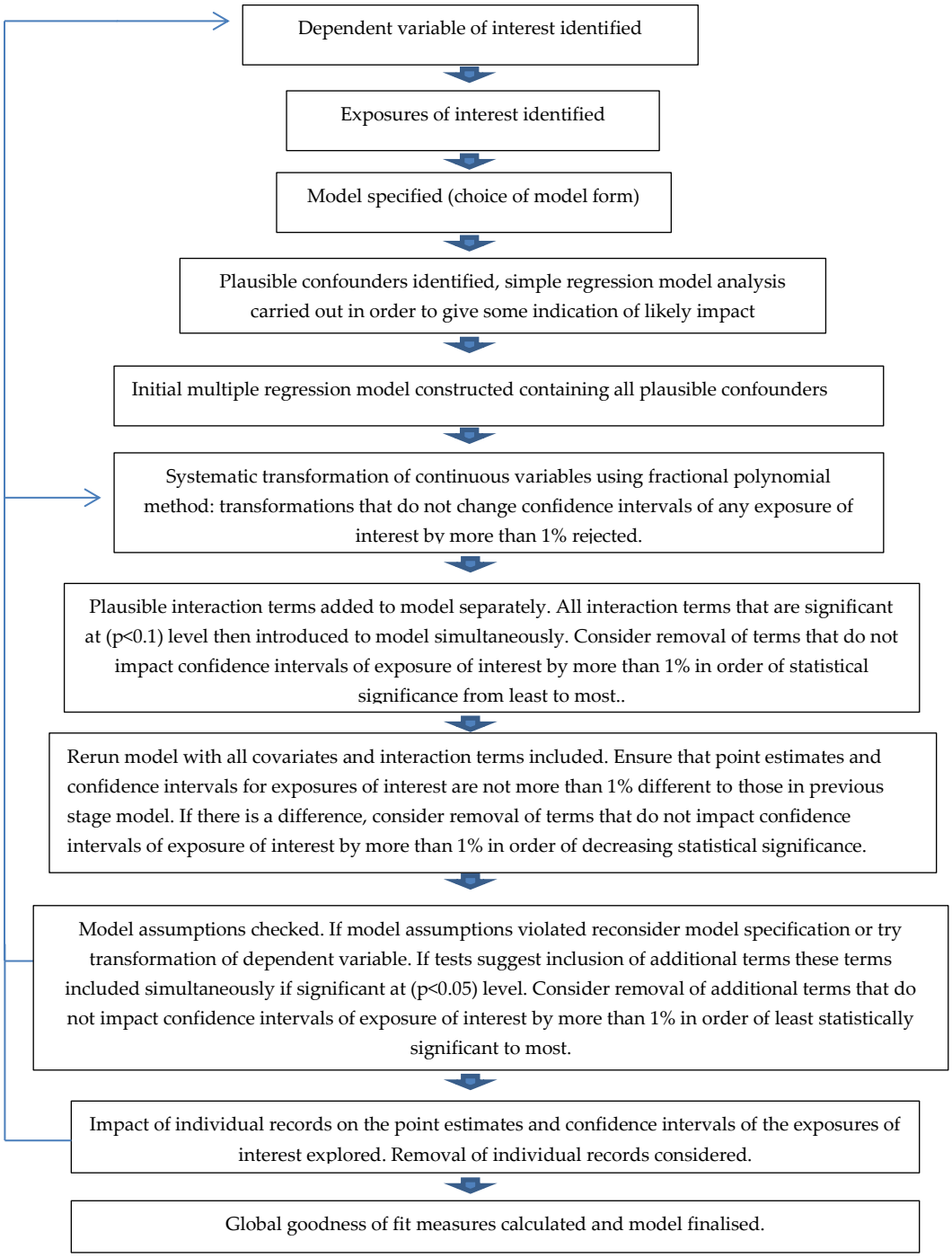
The resulting model can be described as the preliminary final model. At this point all removed variables are added to the model simultaneously to check that that some combination of removed variables not previously considered does not lead to significant confounding.

4.1.4.3. Model assumptions, leverage and goodness-of-fit

This model is then subjected to further tests to ensure that model assumptions are being met. If additional terms are added to the model as a result of this process (for example terms interacting covariates with a function of study time in the case of Cox Proportional Hazard models) removal of such terms is considered if removal would not move the confidence intervals of any exposure of interest more than 1%.

The impact of individual records on the point estimates and confidence intervals of the exposures of interest is considered using appropriate methods, and goodness of fit is assessed. Assuming that these final steps do not lead to further questions or modifications of the model this will be the final model.

Figure 7: Model building process



4.1.5. Accounting for clustering and matching

The WUNZ:HS evaluation study design complicates analysis in two ways. Firstly, clustering is introduced at two levels:

a) At the household level for all analyses, because it is likely that a given household's occupants are more similar to each other than to other study participants with regard to relevant characteristics such as income, genetic disease predisposition, smoking status and second-hand smoke exposure etc.
b) At the level of household and matched control pairings in panel-based models (see Section 5.2 for more detail).

This means that clustering must explicitly be accounted for in analyses or the confidence intervals produced by models for the coefficient(s) of the exposure(s) of interest may be inaccurate. Kirkwood and Sterne (2003) present a number of estimation options for situations where clustering is present:

1. Calculate summary data at the level of the cluster and analyse using standard methods.
2. Use robust standard errors to account for clustering.
3. Use random effects models.
4. Use generalized estimating equations.

In addition, jackknife estimation can be used as a method for accounting for clustering when calculating standard errors. In this thesis I take the second approach listed above, accounting for clustering by using robust standard errors at the cluster level. Where multiple levels of clustering may occur clustering is accounted for using the appropriate higher level cluster.

The matched design of the WUNZ:HS evaluation also introduces complications. Following Rothman and Greenland (1998), it is possible to attempt to control for matching characteristics explicitly by including them as explanatory variables in individual level models (household level models account for matching by including a fixed effect). The matching process described in Section 3.1 includes far too many categorical variables (including 1,574 CAUs which included at least one treatment group dwelling and its matched control group dwellings) and 14 house types (see Subsection 3.3.4). The issue of matching is partially addressed by including covariates in models that have been chosen because they capture key elements of the matching process.

Firstly, "dwelling health risk", a six category composite measure of key physical housing characteristics described in Subsection 3.2.5 is available in order to address the physical

similarities in treatment and matched control group homes. Secondly, a number of plausible variables are considered in order to best capture the impact of geographical location, which can be assumed to include both climate effects and other effects (differences in average health level or employment rates for example), and climate alone.

We adopted an additional method to limit the impact of differences in the number of matches per treatment household on the *individual* level analyses presented in Telfar Barnard, Preval et al. (2011) and Preval (2012). This approach was to limit matching to 4 randomly selected control dwellings (and their occupants) when between 4 and 10 were available and was primarily intended to limit potential bias introduced by differences in the number of matches made. This thesis adopts the same approach and of 216,171 control group dwellings 104,986 were thus randomly excluded from analysis, excluding 396,459 individuals.

4.1.6. Approach to statistical significance

The use of ($p<0.05$) as a binary criterion for meaningful study results, has been heavily criticised (Goodman 1999). This criticism applies equally to the use of 95% confidence intervals if their only function is to examine whether they contain the null hypothesis. Goodman (1999) notes that effect size, sample size, width of confidence intervals, results from previous studies, biological and scientific plausibility and the actual p-values of results should all be considered when interpreting results. This approach is taken in the discussion of individual model formation in this study.

However, while these considerations are very important, there are some situations where an arbitrary p-value-based threshold may have some use. One of the purposes of the analyses presented in this thesis is to produce an estimate of annual household health benefit as a result of receiving a retrofit under WUNZ:HS (presented in Chapter Seven). This involves combining mortality-related benefits, hospitalisation cost savings and pharmaceutical use cost savings. A threshold for inclusion is useful, because to include all factors regardless of statistical significance would produce poor estimates of total benefit. It could be argued that a cost-benefit analysis should only incorporate results that meet the rigorous standards of evidence demanded by modern medical statistics set out in the previous paragraph. However, from a policy evaluation perspective, the results of the WUNZ:HS evaluation are the best evidence available and may not be able to be fully corroborated either by comparison with other studies (given the uniqueness of the evaluation) or thorough examination of biological plausibility. In such a case arbitrary threshold(s) are useful.

In my original analysis in Telfar Barnard, Preval et al. (2011) I summed results significant at the ($p<0.05$) level in order to estimate annual health benefits per household for the purposes of the cost-benefit analysis presented in Grimes, Denne et al. (2011). In this thesis, I sum results that are statistically significant at the ($p<0.05$) and also the ($p<0.1$) levels in order to demonstrate the impact of applying different criteria. Results that are significant only at the ($p<0.1$) level may be regarded as primarily useful for hypothesis generation, but may also have value as "second-tier" evidence which is how I use them here.

4.1.7. Policy relevant stratification

The key variable suitable for stratification from a policy perspective is the Community Services Card status of treatment group individuals/ households. As described in Chapter Two the progress from Energy Wise Home Grants to WUNZ:HS involved the extension of government subsidies from insulation for low income households to insulation and heating for all households occupying a dwelling built prior to 2000. The successor to WUNZ:HS, Warm Up New Zealand: Healthy Homes reversed this policy direction, limiting subsidies to insulation for low income households.

Stratification is addressed by introducing dummy variables into models which capture Community Services Card status.

4.1.8. Adjustment of costs and benefits

All costs are adjusted to Dec 2011 levels (the final month of data). It is necessary to adjust costs to account for inflation. Costs that did not occur during the 2011/12 financial year (1st July 2011- 30th June 2012) are thus adjusted based on the appropriate financial-year average of the Community Health (Inputs) element of the Producer Price Index, an index that measures changes in the prices received by New Zealand producers for their output. This approach to adjustment was taken in previous cost-benefit analyses of *He Kainga Oranga* data (Chapman, Howden-Chapman et al. 2009, Preval, Chapman et al. 2010) and in the analyses of WUNZ:HS health data (Telfar Barnard, Preval et al. 2011, Preval 2012).

Benefits such as mortality avoided are valued using previous research (described in Subsection 4.3). Where figures used in calculation of benefits need to be adjusted to 2011 prices, the Consumer Price Index is used (All Groups). In practice, the choice of Producer Price Index or Consumer Price Index makes little difference as the two indices display extremely similar trends during the Study period.

4.2. Modelling changes in mortality

New Zealand research has shown a statistically significant relationship between winter and all cause excess mortality for elderly people (Davie, Baker et al. 2007, Telfar-Barnard 2010). This relationship was strongest for people who are in the oldest age brackets of the two datasets analysed (Davie, Baker et al. 2007, Telfar-Barnard 2010), although children aged 0-4 were also found to experience statistically significant excess winter mortality to a lesser degree. As discussed in Subsection 2.1.4 it is plausible that this relationship reflects the impact of lower indoor temperatures via a number of mechanisms, with Telfar-Barnard (2010) finding statistically significant excess winter mortality and morbidity ratios for the circulatory (ICD-10 Chapter Nine) and respiratory (ICD-10 Chapter Ten) categories that are most plausibly linked with housing.

The analyses of WUNZ:HS mortality data presented in (Telfar Barnard, Preval et al. 2011, Preval 2012) were limited to people aged 65 and over as at the first month in which the household they occupied received treatment. This decision was based on the findings described above, with 65 being the age at which New Zealand Superannuation becomes universally available for individuals. The cohort of interest was further refined by limiting analyses to either individuals aged 65 and over who had had a circulatory hospitalisation during the *before* period or individuals aged 65 and over who had had a respiratory hospitalisation during the *before* period, where *before* is defined as the period from the start of the study Jan 2008 to the month prior to the first treatment being received (or being received by matched treatment household in the case of control group households/individuals. The decision to limit analyses to two particularly vulnerable cohorts was driven by two important characteristics of the mortality data.

Firstly, as discussed in Subsection 3.3.1.5., the available mortality data are all-cause. This means that cohort choice and analysis must attempt to control for as many factors influencing the rates of non-relevant types of mortality as possible, in order to estimate the change in all-cause mortality that can reasonably be attributed to receiving either insulation or heating under WUNZ:HS.

Secondly, unlike the analysis of hospitalisation costs and pharmaceutical use costs it is not possible to account for *before* period differences in mortality rates between treatment groups and the control group. An approach that used the full range of available data would not be suitable because individuals who died during the *before* period were not able to participate in the WUNZ:HS programme, and we can thus assume their households would also be less likely to participate, therefore treatment group individuals would necessarily have lower *before* period mortality rates, all other things being equal. It was argued in (Telfar Barnard, Preval et al. 2011, p 26) that:

We could not use a basic "difference in difference" approach as described above for mortality data because it would necessarily have included both treatment bias and systematic bias: People could not have sought insulation for their property if they were dead, meaning that the mortality rate in the treatment group would be lower than in the control group before treatment date; and because the treatment group were found to be on average less healthy than the control group, their mortality rate after treatment would be expected to be higher. The difference in difference between before and after and treatment and control groups would therefore be bound to appear adverse, as follows:

$$\left(\frac{Treatment\ mortality\ rate\ after}{Treatment\ mortality\ rate\ before} \div \frac{Control\ mortality\ rate\ after}{Control\ mortality\ rate\ before}\right) = \left(\frac{High}{Low} \div \frac{Average}{Average}\right) > 1$$

The additional argument that treatment mortality rates would be higher *after* due to worse health in the treatment group, while defensible, is not necessary to make the above case: even if mortality rates in the treatment group were only average after treatment the bias described would still apply.

Following the reasoning presented above, this thesis limits primary analyses of mortality to *after* period data for people aged 65 and over with either evidence of a pre-existing circulatory or respiratory condition. The cohorts are referred to as "the circulatory cohort" and "the respiratory cohort" for reasons of clarity and concision.

4.2.1. Model specification

The standard model specification used for the analysis of survival is the Cox Proportional Hazards model. The Cox Proportional Hazard model is widely used in the epidemiological literature for modelling a variety of relationships (time to death, time to relapse etc.). The model can be expressed in the following way:

$$h(t) = h_0(t) * e^{(\beta_1 x_1 + \beta_2 x_{2+} \ldots\ldots + \beta_p x_p)}$$

In this equation h(t) is the hazard of the event of interest i.e. death occurring at study time t, $\beta_1 \ldots \beta_p$ are the coefficients of covariates $x_{1\ldots.}x_p$ respectively, and $h_0(t)$ = the underlying hazard function for individuals for whom $x_{1\ldots.}x_p$ all =0.

Model coefficients estimated using the Cox Proportional Hazard model can be converted to hazard ratios by exponentiation. Hazard ratios can be interpreted in the same way as relative risk ratios, which is useful from an analytic perspective. The hazard ratio (HR) for a one unit increase in a given baseline characteristic (x_1) can be understood by considering the risk of mortality, for two values of x_1, a and b, where $x_{1a} = x_{1b}+1$, holding all other things constant. In this case the HR is literally the ratio of the hazard (h(t)) for x_1=a and the hazard h(t) for x_1=b.

$$HR(t, x = a\ vs\ x = b, \beta_1 \ldots \beta_p) = \frac{h_0(t) * e^{(\beta_1 x_{1a} + \beta_2 x_{2+} \ldots\ldots + \beta_p x_p)}}{h_0(t) * e^{(\beta_1 x_{1b} + \beta_2 x_{2+} \ldots\ldots + \beta_p x_p)}} = \frac{e^{(\beta_1 x_{1a})}}{e^{(\beta_1 x_{1b})}} = e^{\beta_1}$$

The above equation demonstrates the key assumption underlying the Cox Proportional Hazard model, which is that HR(t, x=a vs x=b, $\beta_{1\ldots}\beta_p$) is a constant e^{β_1}, i.e. is does not change over study time (and can thus be described as proportional). $h_0(t)$ is the baseline hazard function (Hosmer, Lemeshow et al. 2008); the relationship between hazard and time when all values of x are zero. The standard Cox Proportional Hazard model assumes that the values of variables for each individual included in models do not change during the period studied; there are fixed baseline characteristics, for example sex, age at start of study, ethnicity or various health characteristics such as baseline BMI (Hosmer, Lemeshow et al. 2008).

In this analysis, study time *(t)* for each cohort individual is defined from the beginning of the "after" period, for their household (either the start of the month when their household was first treated under WUNZ:HS or the start of the month when their matched treatment household was first treated in the case on control group individuals). Each individuals total study time is the time from this initial treatment to either death or censoring on the final date of the study (Dec 31st 2011).

In this thesis I initially explored extensions of the Cox Proportional Hazard model which allowed the value of certain covariates to vary over time (particularly treatment status, ongoing non-relevant hospitalisation costs, temperature related variables and winter/non-winter status) but the resulting models were unduly complex, leading to issues around assessment of model assumptions and fit. It is, however, worth noting that the extended models did not produce markedly different results, which also addresses concerns regarding the appropriateness of the use of non-calendar based study time in this context. The simpler time-invariant models have therefore been retained.

Finally, it is necessary to state my approach to "ties"(when two individuals have the same survival time). Hosmer & Lemeshow (2008) describe three approaches: the exact partial likelihood method and the Breslow and Efron approximations. The authors note that the three approaches produce similar outcomes in terms of model coefficients and standard errors, but state that their preference is to use the Efron approximation if it is available. As Stata provides this approximation, I follow the advice of the authors and take this approach.

4.2.2.Confounding and mortality

It is particularly important to consider confounding factors that might influence relative mortality rates in the post treatment period in the absence of either true randomisation, or

a difference in difference mechanism that can control for both measured and unmeasured confounding. Although the matched study design chosen was successful in producing an similar distribution of potential confounders among both the treatment and control groups (see Section 3.3), when smaller cohorts are created (e.g. 65 and over with a *before* period hospitalisation) it is possible that this will result in a less similar distribution of potential confounders.

The available and plausible candidates for confounding are age, sex, ethnicity, socioeconomic status, baseline health status, housing quality, and location/climate. Each is considered below separately.

4.2.2.1. Age and mortality

One of the most obvious risk factors for mortality is greater age. The initial description of age distribution in Subsection 3.2.2.1 shows that the full treatment group has a higher proportion of people aged over 60 than the control group. Confining analysis to people aged 65 and over limits a major potential source of confounding, but does not necessarily eliminate it; it is important to consider differences in the age structure of the post 65 age group individuals in the treatment and control groups.

Comparison of the age distribution for the four treatment categories (control, insulation only, heating only, insulation and heating) demonstrates there is no statistically significant difference in the distribution of age groups at the ($p<0.1$) level. Table 72 (Appendix 2) presents the full results.

Simple regression analysis of the mortality hazard ratios for the continuous variable "baseline_age" measured in years is the next step in the analysis. Baseline age is defined as the age a person was in the month that the first treatment that the household they occupied was received (or that their matched treatment household received in the case of control group individuals). Simple regression analyses for the two cohorts shows that there is a highly statistically significant relationship between age and mortality hazard. Table 73 (Appendix 2) presents the full results. The hazard ratio for a one year increase in baseline age for the circulatory cohort is 1.08 (95% CI 1.06 – 1.09) and for the respiratory cohort 1.04 (95% CI 1.03 – 1.06). This confirms that age is an important model covariate, although it is not yet clear whether controlling for age will materially alter the estimates of the exposures of interest, given that there are no obvious systematic differences in the age-composition of the different treatment groups and the control group.

4.2.2.2.Sex and mortality

Because women live longer than men on average in New Zealand (for example a female born in 2005-2007 can expect to live to 82.2 and a male 78 (Statistics New Zealand 2013a), any difference in the sex distribution of the cohort could potentially introduce confounding. Table 74 (Appendix 2) sets out the distribution of sex for the two mortality cohorts.

Table 74 shows that there is a difference in the distribution of sex for 'insulation only' individuals vs control group individuals in the circulatory cohort, that is statistically significant at the ($p<0.10$) level. There are no other statistically significant differences in sex distribution. Having carried out this comparison the next step is to carry out simple regression analysis of the dichotomous variable "female". Table 75 (Appendix 2) shows that there is not a strong relationship between sex and mortality hazard in the circulatory cohort, but that there is a very strong relationship between sex and mortality in the respiratory cohort.

Given that there is only a strong relationship between sex and mortality in the respiratory cohort (where sex is evenly distributed between treatment groups and the control group), sex is unlikely to prove a strong confounder of the relationship between treatment group status and mortality in either the analysis of the circulatory or respiratory cohorts.

4.2.2.3. Ethnicity and mortality

There is a well document connection between ethnicity, age and mortality. Statistics New Zealand state that:

> *As a result of differences in death rates, life expectancy at birth for non-Māori exceeded that of Māori by 8.6 years for males and by 7.9 years for females in 2005–07. For males, three-quarters of this difference is due to higher Māori death rates at ages 40–79 years. For females, three-quarters of this difference is due to higher Māori death rates at ages 50–84 years.*

(Statistics New Zealand 2013b)

While similar analyses are not possible for Pacific and Asian sub-populations due to low historical population levels the same report states that mortality rates during the 2005-2007 period were 9.9 per 1000 population for Māori, 8.8 for Pacific, 5.4 for European or Other and 3.7 for Asian (Statistics New Zealand 2013b). Jatrana and Blakely (2008) demonstrated ethnic disparities in mortality risk for individuals aged 65 and over, finding that mortality risk was higher for Māori and Pacific individuals relative to European/Other individuals and that Asian individuals had the lowest mortality risk.

It is clear that ethnicity is an important potential confounder. Table 76 (Appendix 2) sets out the ethnicity distributions for the circulatory and respiratory cohorts. Table 76 indicates that the only statistically significant difference in the ethnic composition of treatment groups and their respective control groups at the ($p<0.1$) level is for the comparison of the 'insulation and heating' group with the control group in the circulatory cohort.

The next stage in analysis is multiple regression analysis of the ethnicity dummy variables. Table 77 demonstrates that a Wald test for the inclusion of all ethnicity dummy variables has a statistically significant association with mortality hazard at the ($p<0.05$) level ($p=0.011$) for the circulatory cohort. A Wald for the respiratory cohort does not find strong evidence for the inclusion of ethnicity dummy variables ($p= 0.118$).

These initial findings suggest that ethnicity dummy variables may be a more important confounder of the relationship between treatment status and mortality hazard for the circulatory cohort.

4.2.2.4. Socio-economic status and mortality

A number of international and New Zealand studies have demonstrated that there is a continuous socio-economic gradient of both morbidity and mortality (Fawcett 2005). This gradient applies to the elderly: Jatrana and Blakely (2008) found that between 34% and 48% of excess mortality for Māori and Pacific individuals aged 65 and over could be explained by five socio-economic factors. It is thus important to explore the ways in which socio-economic status can be included in the analysis of mortality.

The primary measure of socio-economic status available for this thesis, NZDep2006, is available at the meshblock level for each home. Because matching occurred at the CAU level and the NZDep2006 scores of meshblocks within the same CAU are more likely to be similar than a two randomly selected meshblocks, NZDep2006 can be expected to be distributed equally between the treatment groups and the control group in the total cohort. However, when comparing treatment groups and the control group for the circulatory and respiratory cohorts this assumption may no longer apply. Table 78

(Appendix 2) sets out the distribution of NZDep2006 quintiles derived from NZDep2006 scores for the circulatory and respiratory cohorts. There is a weak statistically significant difference in the distribution of NZDep2006 quintiles when comparing the insulation only ($p=0.068$) and heating only ($p=0.054$) groups with the control group in the circulatory cohort, but there was no significant difference at the ($p<0.1$) level in the distribution of NZDep2006 scores within the respiratory cohort.

Simple regression analysis of continuous NZDep2006 variable is the next step in analysis. Simple regression analysis is presented in Table 79. Simple regression analysis shows that the mortality hazard ratio for a 10 unit increase in NZDep2006 score is 1.01 for both cohorts, but this association is significant statistically for the circulatory cohort ($p=0.016$) but not for the respiratory cohort (0.168). Simple regression analysis therefore suggests that NZDep2006 is likely to be a confounder of the relationship between treatment status and mortality hazard for the circulatory cohort.

The other available measure, Community Services Card (CSC) status, is an important marker for socio-economic status in the treatment group (see Subsection 3.3.5. for further discussion), however it is not known in the control group (although the matching process may make control households more likely to have similar CSC holder occupant status this cannot be confirmed). This means that while CSC status is potentially a useful means of stratification from a policy perspective (Subsection 4.1.7.) it is likely to be a weak marker of socioeconomic status for inclusion in models, particularly given that NZDep2006 is available.

4.2.2.5. Health status and mortality

One of the most obvious risk factors for mortality is health status. Health status includes history of or genetic predisposition for potentially fatal diseases (or diseases that while not fatal may contribute to mortality risk in combination with other causes), and the presence of risk factors such as smoking status, alcohol use and obesity, which can increase the risk of potentially fatal disease.

The WUNZ:HS evaluations only had limited access to information about health status. The two sources of information (other than NHI data on demographic factors that are correlated with health status such as age, ethnicity and socio-economic status) are hospitalisation data and pharmaceutical use data. Ultimately the WUNZ:HS mortality evaluations used baseline hospitalisation data to account for differences in the health status of treatment and control group individuals. The alternative, pharmaceutical-use data, were ultimately rejected as problematic due to the limitations of pharmaceutical use data described in Subsection 3.2.4.3. A particular concern would be the exclusion of

hospital-dispensed pharmaceuticals from Pharms, which would quite possibly be most strongly correlated with serious illness and thus most highly correlated with mortality.

Limiting assessment of health status to hospitalisation data, a priori, it seems reasonable to suggest that if you have recently been hospitalised you have a higher mortality risk on average than somebody who has not been hospitalised. This relationship will, of course, be much stronger for hospitalisations relating to life threatening conditions, such as stage four lung cancer, than for hospitalisations relating to a broken ankle.

As described in Section 4.2, *before* period hospitalisation status was the basis for the creation of the two cohorts analysed in this Chapter and also in the previous WUNZ:HS evaluations. The WUNZ:HS evaluations' analyses of mortality rates controlled for confounding by severity of illness by including measures of *before* period *relevant* hospitalisations in models (Telfar Barnard, Preval et al. 2011, Preval 2012). For example, in the modelling of mortality for the circulatory cohort I included a variable "annual_circulatory_cost" which was the log of average annual circulatory costs during the *before* period for each individual plus a small arbitrary number (0.000001) (Preval 2012).

This thesis considers the potential to account for confounding using hospitalisation data in a more comprehensive manner. Unlike the analyses in Telfar Barnard, Preval et al. (2011) and Preval (2012), a broader range of *before* period hospitalisation cost summary measures are considered: average annual cohort-relevant hospitalisation costs (either circulatory or respiratory); and average annual non-relevant hospitalisation costs (all remaining costs).

Analysis of differences in baseline average hospitalisation costs are presented in Table 80 (Appendix 2). There are no differences in average baseline costs that are significant at the ($p<0.1$) level for the circulatory cohort. For the respiratory cohort the only differences that are statistically significant are the difference in average baseline non-relevant costs between the insulation and heating group and the control group, this difference is highly statistically significant ($p=0.007$).

Turning to simple regression analyses of the hospitalisation-based variables, Table 81 (Appendix 2) shows that for the circulatory cohort, an increase in costs of $1000 is associated with an increased mortality hazard for both hospitalisation cost variables (baseline average circulatory costs, baseline average non-relevant costs) and that this relationship is highly statistically significant ($p<0.001$).This is also true for the respiratory cohort.

These initial analyses suggest that while most of the differences in average hospitalisation cost between treatment groups and the control group are not statistically significant, any differences may be significant from a modelling perspective given the strong association between hospitalisation costs and mortality hazard. This suggests differences in health status may be a major potential confounder of the relationship between treatment status and mortality hazard, particularly for the comparison of the "insulation and heating" individuals and control group individuals.

4.2.2.6. Housing characteristics

Accounting for differences in housing characteristics in the treatment and control group is important given the potential for confounding and the importance of accounting for variables used in the initial matching process that underlies the WUNZ:HS evaluation (see Subsection 4.1.4).

Comparison of the distribution of 'dwelling health risk' dummy variables is presented in Table 82 (Appendix 2). The six 'dwelling health risk' dummy variables are derived from compound variable with two components: a binary construction type indicator (new, old), and a ternary condition indicator (high, average, low) which is fully described in Subsection 3.3.4. Table 82 demonstrates that there is no statistically significant difference in the distribution of the dwelling health risk variable when comparing the treatment and control groups, at the ($p<0.1$) level, for either cohort.

Using the 'dwelling health risk' dummy variables multiple regression analysis is carried out and presented in Table 83 (Appendix 2). This Wald test-based analysis suggests that the 'dwelling health risk' dummy variables collectively have a stronger association with mortality hazard in the respiratory cohort ($p=0.006$) than in the circulatory cohort ($p=0.191$).

It appears likely that differences in the distribution of dwelling health risk may be a more important confounder for the relationship between treatment group status and mortality hazard in the respiratory cohort. It is not clear what direction any such confounding will be in.

4.2.2.7. Climate, location and mortality

The relationship between location/climate and mortality rates is complex. Temperate nations such as New Zealand experience excess winter mortality (Davie, Baker et al. 2007, Telfar-Barnard 2010) , but there is also a large body of evidence regarding the short-term correlation between both high (Basu and Samet 2002, Basu 2009) and low (Guo, Li et al. 2013) temperatures and mortality rates internationally. There is also evidence of short term correlation between high temperatures and cardio-vascular, cerebrovascular and respiratory mortality in New Zealand (Hales, Salmond et al. 2000).

There is a range of data available that can be used to capture confounding by climate and/ or location in this thesis. These measures are described in some depth in Subsection 3.3.8.

The location-based variable chosen is Climate Zone. Climate Zones is defined in Subsection 3.3.7. Climate Zone will naturally capture other geographic variations such as variation in socio-economic conditions which may also impact mortality; however, capturing any potential source of confounding is useful. The Climate Zone dummy variables were chosen because the best alternative measure, Regional Council, would involve the inclusion of 15 dummy variables in the initial multiple regression models. Given that the Cox Proportional Hazard model typically requires a ratio of 10 events: 1 model covariate (Peduzzi, Concato et al. 1996), the three Climate Zone dummy variables are preferable given that there are 511 deaths in the circulatory cohort and 379 in respiratory cohort. Time varying measures such as winter, monthly temperature and even daily measures of temperature at a regional level were considered but additional complexity and problems from a modelling perspective meant that after some initial explorations I chose to exclude all time-based variables, as discussed previously.

Table 84 (Appendix 2) demonstrates that there are statistically significant differences at the ($p<0.05$) level in the distribution of individuals in the circulatory cohort by Climate Zone relative to the control group, for individuals with insulation only and heating only at the ($p<0.05$) level. In the respiratory cohort, there is a statistically significant difference in the distribution of individuals by region at the ($p<0.1$) level when comparing the insulation only individuals with control individuals. Regression analyses of the Climate Zone dummy variables presented in Table 85 suggest that it does not have a strong association with mortality hazard.

4.2.3. Initial multiple regression models

The initial Cox Proportional Hazard models include all of the covariates identified in Subsections 4.2.2.1 – 4.2.2.8. The hazard ratios for the exposures of interest for both the circulatory and respiratory cohorts are presented in Table 22. Table 22 does not present the full models, these are presented in Table 86 (Appendix 2).

Table 22: Initial multiple regression mortality models, circulatory and respiratory cohorts

	Variable name	Hazard Ratio (relative to control group)	Robust Standard error	Z	p>|z|	95% CI for HR		Model summary measures
Circulatory cohort	Insulation only	0.717	0.082	-2.92	0.004***	(0.574	0.897)	Log pseudolikelihood = -3909.0558
	Heating only	1.440	0.318	1.65	0.098*	(0.935	2.219)	
	Insulation and Heating	0.645	0.166	-1.7	0.088*	(0.389	1.068)	Wald chi2(18) = 259.91 Prob > chi2 = 0.000
Respiratory cohort	Insulation only	0.827	0.102	-1.53	0.125	(0.650	1.054)	Log pseudolikelihood = -2630.6481
	Heating only	1.197	0.333	0.65	0.517	(0.695	2.064)	
	Insulation and Heating	0.649	0.193	-1.46	0.146	(0.362	1.162)	Wald chi2(18) = 150.93 Prob > chi2 = 0.000

* p<0.1, ** p<0.05, *** p<0.01

Models control for age, sex, ethnicity, two baseline hospitalisation variables, NZDep2006Score, dwelling risk, Climate Zone.

Table 22 shows that receiving insulation leads to reduction in mortality hazard relative to the control group in the circulatory cohort, i.e. the hazard ratio, 0.717 is lower than 1. This reduction is statistically significant at the (p<0.01) level. Interestingly, there is an *increase* in hazard for receiving heating relative to the control group (p = 0.098) and a decrease in hazard for receiving insulation and heating relative to the control group that is statistically significant at the (p<0.1) level. These results appear to vindicate the decision to split treatment into separate exposures of interest as there are different outcomes by treatment type, with a notable similarity between the insulation and 'insulation and heating' groups, but a markedly different outcome for the heating only group. A Wald test for the equality of the insulation and 'insulation and heating' coefficients produced a chi2 score of 0.15 (p>chi2=0.6975), indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. None of the treatments have a statistically significant hazard ratio relative to the control group in the respiratory cohort at the (p<0.1) level.

4.2.4. Transformation of continuous variables

Transformation of the continuous variables in the initial multiple regression model was the next step. The approach used was to transform the continuous variable with the coefficient with the highest |Z| score using the fractional polynomial method, with other less strongly associated continuous variables included in the model in a temporary categorical form. This approach was repeated for the variable with the next highest |Z|

score and so on, until all continuous variables were considered and potentially transformed. In both models the NZ2006DepScore variable was the last variable to be transformed, but the transformed variables did not modify the bounds of the confidence intervals of the hazard ratios of any of the exposures of interest and thus the variable kept its original form. Table 23 details the impact of this process on measures of model fit and the hazard ratios of the exposures of interest. Details of the transformations are not included for reasons of brevity but, for example, the variable age in the respiratory cohort was transformed into two variables, $1/(age)^2$ and $(age)^3$.

Table 23: Fractional polynomial transformation of initial circulatory and respiratory models

	Model	Exposure of interest	Hazard ratio relative to control group	Robust Standard error	Z	p>|z|	95% Confidence Intervals		Measure of fit: Log pseudo likelihood
Circulatory cohort	Baseline non-relevant average annual hospitalisation costs transformed	Insulation only	0.672	0.077	-3.49	0.000	(0.538	0.840)	Wald chi2(19) = 328.25 Log pseudolikelihood = -3897.5803 Prob > chi2 = 0.0000
		Heating only	1.235	0.285	0.92	0.359	(0.786	1.941)	
		Insulation and heating	0.585	0.157	-2	0.046	(0.345	0.990)	
	Baseline age transformed	Insulation only	0.671	0.078	-3.41	0.001	(0.534	0.844)	Wald chi2(20) = 396.31 Log pseudolikelihood = -3866.4891 Prob > chi2 = 0.0000
		Heating only	1.319	0.285	1.28	0.201	(0.863	2.016)	
		Insulation and heating	0.593	0.154	-2.02	0.044	(0.357	0.986)	
	Baseline average annual circulatory hospitalisation costs transformed	Insulation only	0.674	0.079	-3.36	0.001	(0.536	0.848)	Wald chi2(21) = 410.76 Log pseudolikelihood = -3855.9811 Prob > chi2 = 0.0000
		Heating only	1.362	0.294	1.43	0.153	(0.892	2.080)	
		Insulation and heating	0.579	0.149	-2.12	0.034	(0.349	0.959)	
Respiratory cohort	Baseline age transformed	Insulation only	0.835	0.102	-1.48	0.139	(0.658	1.060)	Wald chi2(19) = 176.94 Log pseudolikelihood = -2614.6004 Prob > chi2 = 0.0000
		Heating only	1.124	0.302	0.43	0.664	(0.664	1.902)	
		Insulation and heating	0.636	0.201	-1.43	0.152	(0.342	1.181)	
	Baseline non-relevant average annual hosp costs transformed	Insulation only	0.834	0.102	-1.48	0.139	(0.656	1.061)	Wald chi2(20) = 192.80 Log pseudolikelihood = -2608.1417 Prob > chi2 = 0.0000
		Heating only	1.154	0.312	0.53	0.597	(0.679	1.959)	
		Insulation and heating	0.669	0.211	-1.28	0.202	(0.361	1.240)	
	Baseline average annual respiratory hospitalisation costs transformed	Insulation only	0.839	0.102	-1.44	0.151	(0.661	1.066)	Wald chi2(21) = 220.40 Log pseudolikelihood = -2598.4825 Prob > chi2 = 0.0000
		Heating only	1.154	0.317	0.52	0.603	(0.673	1.977)	
		Insulation and heating	0.655	0.213	-1.3	0.193	(0.347	1.238)	

Having transformed the continuous variables in the two models the next step is to consider interactions.

4.2.5. Interaction terms

Because a large variety of interaction terms are possible it is important to be parsimonious when considering interactions. The most plausible interactions terms, following Hosmer and Lemeshow et al. (2011) are age and sex. These terms are interacted with other plausible interaction terms in the model. As noted in Subsection 4.1.4, the approach taken follows that described in Hosmer and Lemeshow et al. (2011) with interaction terms tested separately and then added into the model simultaneously if they are significant at the ($p<0.1$) level. Removal is then considered if removal does not alter any of the confidence intervals of the exposures of interest by 1% or more.

Table 24 sets out summary figures for interaction terms considered. Based on the results presented in Table 24, I simultaneously included the following variables in the multiple regression circulatory and respiratory models. Circulatory: interaction terms for the interaction of the fractional polynomial transformations of age with transformed baseline average annual circulatory hospitalisation costs. Respiratory: the interaction terms for the interaction of the fractional polynomial transformation of age with ethnicity.

The modified models were then tested to check whether removal of these interaction terms would have any impact on the confidence intervals of the hazard ratios of the exposures of interest. Removal of the interaction terms added to the multiple regression model for the circulatory cohort had very little impact (less than 1% change for all comparisons) on the confidence intervals of the hazard ratios of the exposures of interest. These interaction terms were thus dropped from the model.

The inclusion of interaction terms in the multiple regression model for the respiratory cohort, however, had a much greater impact on the confidence intervals for the hazard ratio of one of the exposures of interest (insulation and heating vs. control). Two interaction terms were retained: the interactions of the second baseline age fractional polynomial (age^3) and the Maori and Pacific Island ethnicity dummy variables.

Table 24: Testing interaction terms for the circulatory and respiratory models

	Interaction category	Variable(s)	df	Wald test, all interaction coefficients=0 (all hazard ratios =1)		
Circulatory cohort	Sex	Baseline average annual circulatory hosp costs	2	chi2(2) = 0.98	Prob > chi2 =	0.6128
		Baseline average annual non-relevant hosp costs	2	chi2(2) = 0.51	Prob > chi2 =	0.7734
		Age	2	chi2(2) = 0.41	Prob > chi2 =	0.8140
		Ethnicity	3	chi2(3) = 7.39	Prob > chi2 =	0.0604*
	Age	Baseline average annual circulatory hosp costs	2	chi2(2) = 11.61	Prob > chi2 =	0.003***
		Baseline average annual non-relevant hosp costs	4	chi2(4) = 5.80	Prob > chi2 =	0.2144
		Ethnicity	6	chi2(6) = 3.32	Prob > chi2 =	0.7679
Respiratory cohort	Sex	Baseline average annual respiratory hosp costs	2	chi2(2) = 0.02	Prob > chi2 =	0.9877
		Baseline average annual non-relevant hosp costs	2	chi2(2) = 3.98	Prob > chi2 =	0.1365
		Age	2	chi2(2) = 0.00	Prob > chi2 =	0.9985
		Ethnicity	3	chi2(3) = 1.59	Prob > chi2 =	0.6607
	Age	Baseline average annual resp hosp costs	4	chi2(4) = 1.20	Prob > chi2 =	0.8776
		Baseline average annual non-relevant hosp costs	4	chi2(4) = 1.20	Prob > chi2 =	0.8776
		Ethnicity	6	chi2(6) = 32.49	Prob > chi2 =	0.0000***

* p<0.1, ** p<0.05, *** p<0.01

Table 25 sets out the hazard ratios and other summary measures for the exposures of interest after interactions have been considered.

Table 25: Preliminary final models, circulatory and respiratory cohorts

	Variable name	HR (vs. control)	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Circulatory cohort[1]	Insulation only	0.673	0.079	-3.37	0.001***	(0.535	0.847)
	Heating only	1.348	0.293	1.38	0.169	(0.881	2.064)
	Insulation and Heating	0.581	0.150	-2.1	0.035**	(0.350	0.964)
Respiratory cohort[2]	Insulation only	0.830	0.100	-1.55	0.122	(0.655	1.051)
	Heating only	1.129	0.313	0.44	0.662	(0.656	1.944)
	Insulation and Heating	0.624	0.208	-1.41	0.158	(0.324	1.200)

* p<0.1, ** p<0.05, *** p<0.01

[1] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone.

[2] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone plus interaction terms for age fractional polynomial (baseline age ^3) and the Pacific and Asian ethnicity dummy variables.

At this point the models were rerun including all previously removed interaction terms simultaneously. The results of the models confirmed that the selection and rejection of interaction terms described above did not contribute materially to the estimates of any of the confidence interval of the exposures of interest in either of the two models, suggesting that the selection and removal processes did not miss complex multi-variable confounding.

The difference between the insulation only and insulation and heating coefficients was again examined using a Wald test for both models. A Wald test for the equality of the insulation and 'insulation and heating' coefficients for the circulatory cohort produced a chi2 score of 0.29 (p>chi2=0.6975), indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. A Wald test for the equality of the insulation and 'insulation and heating' coefficients for the respiratory cohort produced a chi2 score of 0.69 (p>chi2=0.4053), again indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. Finally, combining insulation and 'insulation and heating' into a single treatment dummy variable, addressing the power issues of the smaller respiratory cohort, resulted in a HR of 0.8 (95% CI 0.635 – 1.008), (p=0.058). Given the Wald test results above it is reasonable to interpret this as being driven entirely by the effect of insulation.

The final stage of the modelling process is to examine model assumptions, assess whether any individual records have undue influence on the model and to assess goodness of fit.

4.2.6. *Finalising model*

There are several important stages that must be completed before the model can be finalised.

4.2.6.1. Testing proportional hazards assumption

Underlying the Cox Proportional Hazards model is the assumption that the hazard ratios for model covariates do not change over time in any systematic way (proportionality). A number of methods are available to test this assumption (Hosmer, Lemeshow et al. 2008). I have chosen to examine the scaled Schoenfeld residuals for the models.

Schoenfeld residuals can be calculated for each model covariate for each individual in the model who was not censored e.g. who died. Schoenfeld residuals are the difference between the predicted value for a given variable at the time of death (given the individuals remaining in the risk set at the point of calculation) and the actual value of that variable. When Schoenfeld residuals are graphed against a function of study time (often time or the log of time) any statistically significant non-zero slope shows that the proportional hazards assumption has been violated. The model can be adjusted to account for this violation by including an additional interaction term between the original covariate and an appropriate function of study time, however, when exposures of interest demonstrate non-proportionality this can create problems for interpretability.

I tested the proportionality assumption for all covariates in the model using the "estat phtest" function in Stata, testing for any relationship between the Schoenfeld residuals for each covariate and both study time and then the log of study time. As suggested by Hosmer, Lemeshow et al. (2008) I added interaction terms to the model for all covariates that had a statistically significant relationship with a function of study time at the ($p<0.05$) level using the "tvc texp" options in Stata.

For the circulatory cohort model the following variables demonstrated non-proportionality when their Shoenfeld residuals were graphed against the log of time (t): the Maori ethnicity dummy variable, the Pacific Island ethnicity dummy variable, and two "dwelling health risk" dummy variables ("old and average" and "new and average"). When interaction terms interacting these variables with the log of time were added simultaneously to the model, all terms had statistically significant hazard ratios at the ($p<0.05$) level. However, this inclusion had very little impact on the confidence intervals of any of the three exposures of interest, as demonstrated in Table 26. Consistent with the model building approach described in Subsection 4.1.4, these time-based interaction terms are not retained in the model. Beyond parsimony, an additional benefit of this exclusion is that many of the goodness-of-fit calculations carried out in Subsection 4.2.6.3 are not compatible with time-varying covariates.

For the respiratory cohort model two time-based interaction terms were added to the multiple regression model: the interaction of study time with the Pacific Island ethnicity

dummy variable and with one of the baseline respiratory hospitalisation cost fractional polynomials (^3). As with the circulatory model, these interaction terms were both significant at the ($p<0.05$) level However, they also did not materially alter the estimates of the coefficients of the hazard ratios of the exposures of interest and for the same reasons described in the preceding paragraph were thus excluded from the model going forward.

Table 26: Preliminary final models, circulatory and respiratory cohorts, addition of time-varying covariates

	Variable name	HR (vs. control)	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Circulatory cohort[1]	Insulation only	0.673	0.079	-3.37	0.001***	(0.535	0.847)
	Heating only	1.341	0.292	1.35	0.177	(0.876	2.055)
	Insulation and Heating	0.583	0.150	-2.09	0.036**	(0.352	0.967)
Respiratory cohort[2]	Insulation only	0.823	0.100	-1.61	0.108	(0.649	1.044)
	Heating only	1.138	0.316	0.47	0.642	(0.660	1.960)
	Insulation and Heating	0.633	0.209	-1.38	0.167	(0.331	1.211)

* p<0.1, ** p<0.05, *** p<0.01

[1] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone, and terms interacting the log of study time with Maori and Pacific Island ethnicity dummy variables and "dwelling health risk" dummy variables "old and average" and "new and average".

[2] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone plus interaction terms for the interaction of age and ethnicity and for the interaction of study time and the Pacific Island ethnicity dummy variables and study time and a fractional polynomial transformation of baseline respiratory hospitalisation cost (^3)

4.2.6.2. Assessing leverage, influence, outliers

A number of tools are available for the assessment of influential observations, outliers and high leverage individuals in Cox Proportional Hazards models. When such individuals are identified it is possible to assess what impact removing them from the model would have. Whether such removal can be justified is a more complex question, but it is certainly important to be aware of any individuals that are strongly influencing model output.

Because the goal of this thesis is to estimate the impact of treatment under WUNZ:HS for the study participants (explanatory modelling) the primary importance of model covariates other than the exposures of interest is any confounding effect they may have on the exposures of interest. As such, it is less important to consider the impact of outliers, high leverage individuals and influential observations on the estimates of the hazard ratios and confidence intervals of other model covariates.

Given this focus, I have chosen to concentrate on the calculation of "dfbetas" for each individual, for the coefficients of the three exposures of interest (in each of the two models). Dfbetas measure the impact of including the individual in the model (vs. a model that excludes the individual) in terms of how many standard errors that inclusion moves the estimate of the coefficient in question. A rule of thumb widely stated is to identify

individuals that have dfbeta score for a given covariate that in absolute terms is greater than $|\frac{2}{\sqrt{n}}|$ where n is the number of individuals or observations in the study (StataCorp. 2009).

Table 27: Dfbetas for coefficients of exposures of interest, circulatory and respiratory cohorts

	Variable name	No. of individuals with dfbetas for this covariate that exceed $|\frac{2}{\sqrt{n}}|$	Impact of complete removal
Circulatory cohort[1]	Insulation only	0	N/A
	Heating only	14	Statistically significant protective effect of heating found (HR 0.37, 95% CI (0.16 – 0.85))
	Insulation and Heating	14	Statistically significant protective effect of heating and insulation increased (HR 0.08, 95% CI (0.01 – 0.45))
Respiratory cohort[2]	Insulation only	0	N/A
	Heating only	0	N/A
	Insulation and Heating	0	N/A

[1] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone.

[2] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone plus interaction terms for the interaction of age and ethnicity and for the interaction of study time and the Pacific Island ethnicity dummy variable and study time and a fractional polynomial transformation of baseline respiratory hospitalisation cost $^{(\wedge 3)}$

Table 27 suggests that while the estimates of the coefficients (and thus hazard ratios) for the respiratory cohort are not being driven by outliers, the estimates of the coefficients for the heating vs. control and heating and insulation vs. control dummy variables in the circulatory cohort are being strongly affected by the presence of a small number of study members. Exclusion of these individuals was considered and in both cases complete removal had a large impact on estimated coefficients/hazard ratios.

It was not clear if any removal could be justified in the present study. In general, the topic of removal appears to be a fraught one; some statisticians argue against any removal, while the majority appear to agree that influential records that are easily explainable anomalies (e.g. a multi-choice test result where the participant just chose 'A' for every answer) should be removed. Examination of the outliers in the circulatory cohort did not reveal any obvious data entry errors, and given the limited nature of the administrative data available it did not seem reasonable to attempt to identify combinations of characteristics that seemed biologically implausible.

At this point a pragmatic decision was made to concentrate on the key result of the analyses thus far, the statistically significant ($p<0.05$) protective effect of insulation vs. control in the circulatory cohort, which is not driven by any given study participant (based on dfbetas). The results for heating vs. control and 'heating and insulation' vs. control suggest that some unknown phenomenon is driving estimation of coefficients and hazard ratios, leading to extremely unstable results (with a small cohort) that in one case had the appearance of statistical significance. A conservative approach is not to use such results to assess the benefits of the WUNZ:HS intervention given how susceptible they are to removal. However given the Wald test results reported in Subsection 4.2.3 it may be reasonable to assume that the benefits of insulation accrue to those individuals who occupy houses that received both insulation and heating.

4.2.6.3. Assessment of global goodness-of-fit

A number of measures of global goodness of fit are available for the Cox Proportional Hazard model. It is possible to calculate various measures of explained variation that are analogous to R^2. Choodari-Oskooei, Royston et al. (2012) describe eight measures of explained variation, stating that two in particular, R^2_D and R^2_{PM} produce the best results in their simulation and are most resistant to distortion resulting from excessive numbers of censored study participants. The *Royston and Sauerbrei* R^2_D measure is easily calculated in Stata using the "str2d" command, and is thus adopted as one of the measures of global goodness-of-fit used in this subsection. The finalised circulatory model has an R^2_D value of 0.32. The finalised respiratory model has an R^2_D value of 0.27. These values suggest a reasonably poor fit, no doubt reflecting the lack of information on a range of plausible confounders including baseline smoking status, BMI and other key biomarkers.

Hosmer, Lemeshow et al. (2008) describe a method for assessing global goodness of fit based on assigning each study individual to a "risk-score quantile" based on that individual's modelled risk. If the difference in the observed number of deaths/events and the predicted number of deaths/events for all risk score quantiles is not statistically significant ($p<0.05$) the authors suggest that the model as a whole has "passed the test for a good fitting model" (Hosmer, Lemeshow et al. 2008). The number of risk score groups is calculated by dividing the number of events/deaths by 40, with an upper limit of 10. In Stata the command "stcoxgof" is available to automatically carry out these calculations, and I use it to further explore global goodness-of-fit for the two models.

The results are presented in Table 28 which shows that both models appear to have passed the test for a good fitting model, although the result for the 9th quantile of risk for the circulatory cohort is close to being statistically significant at the ($p<0.1$) level.

Table 28: Quantiles of risk, circulatory and respiratory cohorts

	Quantile of risk	Deaths	Predicted deaths	Z	p-normal	Observations
Circulatory cohort[1]	1	8	10.551	-0.785	0.432	329
	2	16	16.817	-0.199	0.842	329
	3	18	21.564	-0.767	0.443	329
	4	21	26.858	-1.13	0.258	328
	5	29	33.724	-0.813	0.416	329
	6	40	41.559	-0.242	0.809	329
	7	57	50.984	0.842	0.400	328
	8	67	64.535	0.307	0.759	329
	9	102	87.049	1.602	0.109	329
	10	153	157.359	-0.348	0.728	328
	Total	**511**	**511**			**3287**
Respiratory cohort[2]	1	9	12.955	-1.099	0.272	174
	2	18	21.267	-0.708	0.479	173
	3	23	25.917	-0.573	0.567	174
	4	36	30.051	1.085	0.278	173
	5	40	36.982	0.496	0.620	174
	6	45	41.446	0.552	0.581	173
	7	53	50.001	0.424	0.671	174
	8	64	65.344	-0.166	0.868	173
	9	91	95.037	-0.414	0.679	173
	Total	**379**	**379**			**1561**

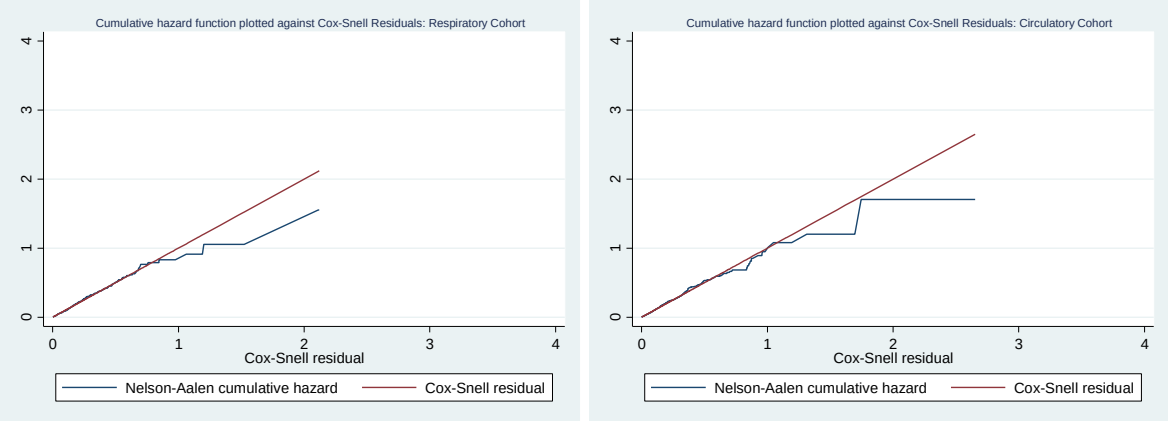
[1] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone, and terms interacting the log of study time with Maori and Pacific Island ethnicity dummy variables and "dwelling health risk" dummy variables "old and average" and "new and average".

[2] Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone plus interaction terms for the interaction of age and ethnicity and for the interaction of study time and the Pacific Island ethnicity dummy variable and study time and a fractional polynomial transformation of baseline respiratory hospitalisation cost $^{(\wedge 3)}$

Finally, it is possible to assess graphically, to what extent the model adequately fits the data by calculating Cox-Snell residuals[16] for each study participant after estimating models, and then rerunning models using each individual's Cox-Snell residual score as a replacement for their study-time (StataCorp. 2009). In the new analysis it is possible to plot the cumulative hazard function (Nelson-Aalen) against the Cox-Snell residuals. The extent to which the cumulative hazard function is consistent with a 45° reference line is a graphical measure of model consistency (StataCorp. 2009). Figure 8 shows that both models appear to demonstrate a reasonable degree of model consistency, particularly as the visible divergences occur towards the right-hand tail of the distribution in both cases. The Stata 11 Base Manual states that the baseline hazard is more variable at the right-hand tail of the distribution "because of the reduced effective sample caused by prior failures and censoring"(StataCorp. 2009).

[16] Cox Snell residuals are defined for an individual i as $rC_i = H_0(t_i) * e^{(\beta_1 x_{1i} + \beta_2 x_{2i} + \ldots\ldots + \beta_p x_{pi})}$, where $H_0(t_i) = \int_0^t h_0(s)ds$ i.e. the cumulative hazard function for that individual at the time of either event or censoring for that individual.

Figure 8: Cumulative hazard function graphed against Cox Snell residuals



4.2.6.4. Stratification by Community Services Card status

The most important stratification from a policy perspective is the CSC status of homes that received a retrofit under the WUNZ:HS programme, as discussed in Subsection 4.1.7. CSC status was accounted for by interacting treatment dummy variables in each model with a binary CSC status variable. Stratifying the models presented in the previous sections produces the following results, presented in Table 29 and Table 30:

Table 29: Results for circulatory model stratified by Community Services Card status

	Variable name	HR (vs. control)	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Community Services Card	Insulation only	0.661	0.086	-3.19	0.001***	(0.512	0.852)
	Heating only	1.325	0.338	1.10	0.271	(0.803	2.190)
	Insulation and Heating	0.590	0.166	-1.87	0.061*	(0.340	1.024)
Non-Community Services Card	Insulation only	0.722	0.172	-1.37	0.171	(0.453	1.150)
	Heating only	1.410	0.563	0.86	0.390	(0.644	3.087)
	Insulation and Heating	0.530	0.334	-1.01	0.313	(0.154	1.821)

* p<0.1, ** p<0.05, *** p<0.01

Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone.

Table 30: Results for respiratory model stratified by Community Services Card status

	Variable name	HR (vs. control)	Robust Standard error	Z	P<|Z|	95% Confidence intervals	
Community Services Card	Insulation only	0.859	0.111	-1.18	0.239	(0.666	1.107)
	Heating only	1.235	0.380	0.69	0.493	(0.675	2.258)
	Insulation and Heating	0.478	0.179	-1.97	0.049**	(0.230	0.995)
Non-Community Services Card	Insulation only	0.703	0.182	-1.36	0.174	(0.422	1.169)
	Heating only	0.735	0.448	-0.51	0.613	(0.223	2.424)
	Insulation and Heating	3.627	2.262	2.07	0.039**	(1.068	12.318)

* p<0.1, ** p<0.05, *** p<0.01

Model controls for age, sex, ethnicity, NZdep2006score, hospitalisation-based variables, dwelling risk score, Climate Zone plus interaction terms for the interaction of age and ethnicity and for the interaction of study time and the Pacific Island ethnicity dummy variable and study time and a fractional polynomial transformation of baseline respiratory hospitalisation cost (^3)

Stratifying treatment dummy variables in the circulatory model allows us to run a Wald test comparing the coefficients of the CSC and non-CSC dummy variables for each treatment type. A Wald test for the equality of the two insulation coefficients produced a chi2 score of 0.11 (p > chi2=0.7373), indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. It is possible that the difference in p-values observed reflects the smaller number of non-CSC individuals. A Wald test for the equality of the two heating only coefficients produced a chi2 score of 0.02 (p>chi2=0.8946), indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. A Wald test for the equality of the two 'insulation and heating' coefficients produced a chi2 score of 0.02 (p>chi2=0.8747), again indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. Finally, Wald tests comparing the coefficients for insulation and 'insulation and heating' for CSC individuals produced a chi2 score of 0.14 (p>chi2=0.7090), and for non-CSC individuals a chi2 score of 0.21 (p>chi2=0.6440), suggesting that there was no evidence to reject the null hypothesis of equality, which was consistent with the non-stratified model.

It is reasonable to conclude that there is no evidence that the CSC status of treatment individuals influences the impact of treatment under WUNZ:HS on mortality for the circulatory cohort. This is not consistent with findings that lower socioeconomic group households take a greater proportion of energy efficiency gains as improved heating (Milne and Boardman 2000), and can be predicted to gain greater health benefits from insulation and heating retrofits; however, this finding may simply reflect an absence of statistical power. The finding may also reflect the fact that matching was not able to account for the CSC status of control group households/individuals, which will bias results towards insignificance.

Turning to the respiratory cohort model, the stratification of the treatment dummy produces two results that are very difficult to reconcile. There is a favourable HR for 'insulation and heating' for those individuals who have a CSC of 0.478:1 that is statistically significant at the ($p<0.05$ level) but an extremely *un*favourable HR for 'insulation and heating' for those who did not have a CSC of 3.627:1, which is significant at the ($p<0.05$) level.

It is highly unlikely that improved heating and insulation could result in much greater mortality risk, and an investigation of the eight individuals who make up this treatment category reveals that 3 have dfbetas that exceed $|\frac{2}{\sqrt{n}}|$,which suggests excessive influence. It seems likely that this result is a statistical anomaly reflecting the extremely small size of the cohort, as it is highly implausible.

Comparison of the model coefficients for the CSC and non-CSC dummies shows that there is no evidence to reject the null hypothesis of equality in the case of insulation (chi2=0.53, prob > chi2=0.4428) and heating (chi2=0.59, prob > chi2=0.4428) but unsurprisingly there is strong evidence against the null hypothesis in the case on insulation and heating (chi2=7.84, prob>chi2=0.0051). Other than this unusual result, which is driven by outliers, the CSC status of individuals does not appear to influence the impact of treatment under WUNZ:HS on mortality in the respiratory model.

4.2.6.5.Summary

Having carried out a robust model construction process, considered the impact of potential violations of the proportional hazards assumption, assessed the impact of influential observations, and considered several measures of global goodness-of-fit it is now possible to assess the meaning of the model results presented in this Section. The key result is a statistically significant reduction in mortality hazard for those individuals in the circulatory cohort who received insulation relative to the control group individuals. The hazard ratio is 0.673 ($p=0.001$). There is also a statistically significant protective effect of receiving both insulation and heating HR 0.58 ($p=0.035$) and a Wald test suggests that there is no evidence to reject the hypothesis that the two HRs are equal. It is interesting to speculate as to why heating alone would not produce a protective effect. It may be that as heating could only be installed in homes that met insulation standards a greater proportion of efficiency gains were taken as energy use savings as pre-treatment temperatures were already relatively high, or the absence of an effect may simply reflect limited statistical power.

Stratification of treatment dummies in the circulatory cohort model by CSC status suggests there is little evidence that CSC status changes the impact of treatment: an unexpected result given that it is expected that lower income households would benefit more from retrofits via greater increases in temperature and thus greater health benefits. The caveat regarding limitations of statistical power noted in the previous paragraph applies equally here.

There are no statistically significant benefits for any treatment category relative to the control group in the respiratory cohort at the ($p<0.1$) level. In general the point estimates of the hazard ratios of the exposures of interest appear roughly consistent with those found in the circulatory cohort, and the lack of statistical significance may reflect the smaller size of the cohort (1561 vs. 3287). In the case of insulation, in particular, the p-value for the HR approaches statistical significance at the ($p<0.1$) level ($p=0.122$), and when insulation and 'insulation and heating' are combined as a single dummy variable the 0.8 HR calculated shows a protective effect of insulation that is statistically significant at the ($p<0.1$) level ($p=0.058$).

Possible causal mechanisms, relating to the effect of insulation in reducing the risk of mortality in the circulatory cohort, include reduced likelihood of ischaemic and cerebrovascular events, reduced severity of such events or a combination of both. However the absence of cause-of-death information means that it is difficult to draw strong conclusions. The causal mechanisms for the protective effect of insulation found in

the respiratory cohort model are likely to be through reduced rates and severity of respiratory illness.

Chapter Seven compiles the results of Chapters Four to Six, which is an opportunity to observe whether there is corroborating evidence in the form of reduced hospitalisation costs or pharmaceutical use costs as a result of receiving insulation under WUNZ:HS, which would strengthen the plausibility of these results. Other evidence in favour of real effects of insulation include the consistency of results for individuals who received insulation and for those who received 'insulation and heating'. However, there remains the possibility that participation in the WUNZ:HS programme is a marker of health status, possibly indicating relatively good health and optimism (and thus lower risk of mortality). This potential selection bias is difficult to address in this thesis, given the timing issues discussed at the start of Section 4.2. I discuss some of the ways that future studies could address this issue in Chapter Nine.

Having established, to the extent possible, that there is a highly statistically significant and clinically significant protective effect of receiving insulation for individuals who are over 65 and have a pre-existing circulatory condition, and a less statistically significant but clinically significant protective effect of insulation for individuals over 65 who have a pre-existing respiratory illness it is now important to consider how these protective effects can be valued and projected over the expected life-time of an insulation retrofit.

4.3. Valuing changes in mortality

It is necessary to consider two factors in order to value the protective effect of insulation observed in the circulatory cohort and respiratory cohorts and thus to include these changes in a cost-benefit analysis. Firstly, how can any change in mortality best be quantified in dollar terms? Secondly, to what extent can the initial observed changes in mortality be extrapolated beyond the study period and over the lifetime of insulation for current cohort members and future home occupants?

4.3.1. The value of a statistical life in New Zealand

One of the most common approaches to quantifying changes in mortality resulting from an intervention is to calculate the Quality Adjusted Life Years (QALYs[17]) or Disability Adjusted Life Years (DALYs[18]) that are gained or lost as the result of the intervention. This approach is often used in cost utility analysis, with change in QALY or DALY per dollar a basis for the comparison of different interventions. This approach is not compatible with cost-benefit analysis, which requires all benefits and costs to be measured in dollar terms, which raises the fundamental question of how we can quantify the value of a human life, often referred to in the literature as the Value of a Statistical Life (VSL). Although there are human capital-based approaches that calculate VSL based on lost productivity these measures fail to account for the full value that individuals place on their lives and thus most current measures of VSL are based on studies that infer how much individuals value their own lives or the lives of others. A number of different approaches can be taken to estimate this value.

Revealed preference studies aim to demonstrate the value that individuals place on their own lives by using models based on pre-existing data. Many such models compare wages for job categories with different levels of mortality risk in order to identify the marginal increase in wage that workers require to take on additional risk: this premium can be extrapolated to estimate the workers' revealed VSL (Viscusi and Aldy 2003). Similarly, the consumer market behaviour method attempts to calculate VSL by identifying the relationship between the price that people pay are willing to pay for safety features in products and the reduction in risk that this produces (Brannon 2004).

[17] A QALY is a measure of life years gained that adjusts for the quality of those years. 10 years in perfect health are equivalent to 10 QALYs, 20 years at 50% health are also equivalent to 10 QALYs.
[18] DALY is a measure that combines years of life lost (relative to an average lifespan) with years of life spent at various levels of disability. DALYs are typically used in population level analyses: for example to quantify and compare the impacts of different diseases.

Stated preference methods Willingness to Pay (WTP) and Willingness to Accept (WTA) attempt to establish the marginal value of a life in dollar terms using carefully designed questionnaires from representative samples of the population (Guria, Leung et al. 2005). WTP studies look at the minimum amount that participants are willing to pay to reduce or avoid a given level of risk, and WTA studies look at the minimum that people are willing to be paid to accept a given higher level of risk (Guria, Leung et al. 2005). Using these methods a VSL figure is derived. VSL estimates do not incorporate an explicit discount rate. There is often a great deal of discrepancy between WTP and WTA-based VSLs, with WTA-based measures typically much larger, possibly reflecting different psychological responses to potential gains and losses (Guria, Leung et al. 2005); this suggests that a conservative option from a cost-benefit perspective is to favour WTP-based values over WTA-based values.

VSL varies greatly from country to country, which reflects differences in GDP as well as variation in preferences. It is thus sensible to focus on VSL figures derived in the New Zealand context. From a New Zealand policy perspective, a transport-related VSL calculated using WTP methods in an influential 1991 paper (Miller and Guria 1991), produced for the then Land Transport Authority, is the official transport sector VSL and has been used for a number of non-transport-related applications (Sanderson, Goodchild et al. 2007, O'Dea and Wren 2010), adjusted for inflation. Other New Zealand candidates include a more recent VSL calculated for the Land Transport Authority using WTA methods (Guria, Jones et al. 2003) and a report for the New Zealand Fire Service (Sanderson, Goodchild et al. 2007) that modifies Miller and Guria's 1991 VSL to account for a lowered risk preference for fire-related VSLs relative to transport-related VSLs based on a survey which showed that participants valued deaths from fire less than transport-related deaths by a ratio of 0.66:1. The findings of Sanderson, Goodchild et al. (2007) are consistent with the widely cited paper by Slovic (1987) which presents risk perceptions within a two-dimensional factor space and shows that, while motor-vehicle accidents and electrical and appliance fires are similarly ranked in terms of the "unknown risk" factor, motor vehicle accidents are ranked more highly in terms of "dread risk".

Given that WTP and WTA methods appear to be favoured internationally over revealed preference methods (Sanderson, Goodchild et al. 2007), and that the three studies described above are the main sources of stated preference-based VSL figures for New Zealanders, it is reasonable to use these three VSL measures as the basis for costing mortality in this thesis, while acknowledging that illness-related VSL may well differ from transport or fire-related VSLs. O'Dea and Wren (2010) provide VSL values for the three studies adjusted to 2008 prices; the current study adjusts these figures to fourth quarter 2011 prices using the Consumer Price Index.

Having established a range of available New Zealand VSLs, it is now necessary to consider how VSL can be converted to a measure called the Value of a Statistical Life Year (VSLY). O'Dea and Wren (2010) provide valuable guidance on how to calculate a VSLY for a given VSL. We must first make an assumption about the age of the person that stated preference study participants had in mind when surveyed (often themselves); given a current life expectancy in New Zealand of approximately 80 years, O'Dea and Wren (2010) assume that the person considered by participants is 40 years old, in good health, and will on average have another 40 years to live. The following formula is then used to calculate VSLY where d is the discount rate applied and n is years of life gained:

$$\sum_{n=1}^{n=40} (VSLY/(1+d)^n) = VSL$$

$$\Rightarrow VSLY = VSL/(\sum_{n=1}^{n=40} \frac{1}{(1+d)^n})$$

Adapted from (O'Dea and Wren 2010, p 24))

It is clear that the higher the discount rate d the higher the VSLY calculated will be for a given VSL. The choice of 40 years of life remaining is not highly sensitive as O'Dea and Wren (2010) note. Table 31 sets out different VSL values calculated using various combinations of the three 2011 Consumer Price Index adjusted VSLs, three different discount rates and three different assumptions about the age of the average person carrying out the original WTP calculations. Discount rates were chosen based on the discount rates used in the original WUNZ:HS evaluation CBA report (Grimes, Denne et al. 2011), 2.5%, 4% and 8%.

Table 31: VSLY values for various combinations of VSL, discount rate and time period

Life years remaining when WTP calculation made	Discount rate	Transport 1991-based VSL (**$ 3,586,946**)	Transport 1998-based VSL ($ **6,073,809**)	Fire 2007-based VSL (**$2,367,384**)
40 (assumes individuals 40 years old)	2.5%	$142,890	$241,958	$94,308
	4.0%	$181,225	$306,870	$119,609
	8.0%	$300,802	$509,351	$198,529
30 (assumes individuals 50 years old)	2.5%	$171,376	$290,192	$113,108
	4.0%	$207,433	$351,249	$136,906
	8.0%	$318,619	$539,521	$210,289
50 (assumes individuals 30 years old)	2.5%	$126,469	$214,151	$83,469
	4.0%	$166,973	$282,737	$110,202
	8.0%	$293,207	$496,491	$193,517

Adapted from Telfar Barnard, Preval et al. (2011, p 37)

Table 31 demonstrates that, given different combinations of assumptions, a large range of VSLY figures can be calculated, from $83,469 to $539,521. VSLY calculated for a given VSL are not strongly influenced by assumptions about the age of the person that WTP/WTA study participants envisaged when carrying out the original studies, but are highly sensitive to choice of discount rate.

The range of discount rates were chosen to be consistent with previous WUNZ:HS evaluation work, following the original CBA in using 4% as the central choice of discount rate, and also reporting values calculated using 2.5% and 8% discount rates as a form of sensitivity analysis. Internationally, a 3% discount rate has been favoured (Aldy and Viscusi 2008, O'Dea and Wren 2010) and Pharmac recommends a 3.5 % discount rate currently (Telfar Barnard, Preval et al. 2011). It is important to note that, counterintuitively, the more conservative discount rate (8 %) typically recommended by institutions such as the New Zealand Treasury results in the calculation of a much greater VSLY. Given a 4% discount rate the three VSLYs used in cost calculations are $181,225, $306,870 and $119,609, with the $181,225 figure derived from the influential work of Miller and Guria (1991) the primary figure used in this thesis, reflecting the fact that the VSL produced in that paper is the official transport sector VSL.

Sensitivity analyses are also carried out to reflect the range of VSLYs calculated above.

4.3.2. Value of a statistical life year: age and health

Before these VSLY values can be applied to the results described in Section 4.2 it is important to consider an additional complication, the age and health status of the WUNZ:HS participants studied. Because the mortality model presented in Section 4.2 focus on individuals aged 65 and over who have pre-existing health conditions it may seem reasonable to question applying unmodified VSLY values to any change in mortality resulting from WUNZ:HS, given that the VSLs that they are based on are assumed to be based on a healthy 40 year old with 40 more years to live.

Viscusi and Aldi (2003) note that VSL measures calculated using *revealed preference* methods demonstrate gradients by both age (older people have lower VSL), by income bracket (lower income people have lower VSL), and by country (VSL is lower in developing nations than in the United States). The authors further refine the relationship between age and revealed preference VSL in a paper which demonstrates an inverse U shaped relationship between VSL and age (Aldy and Viscusi 2008). The two key New Zealand stated preference studies did not find the same inverse U shaped relationship (Miller and Guria 1991, Guria, Jones et al. 2003). Other WTP-based research does suggest that VSL decreases with age, but only for people aged 70 and over, and interestingly does not diminish for people with chronic health problems (Alberini, Cropper et al. 2004).

Methods of valuing life-years gained or lost such as QALY and DALY-based calculations explicitly value time lived with illness or disability less than time lived at full health. If enough information about study participant health was available it might be possible to weight life-years gained as a result of receiving insulation under WUNZ:HS to account for quality of life in a similar way. However, while it might seem reasonable to modify the quantification of any change in mortality resulting from WUNZ:HS in such a way if weightings were available, given WTP-based evidence that people with chronic health problems may not have lower VSL it could be argued that this is inconsistent with using WTP-based values. As O'Dea and Wren (2010, p 28), referring to the inverse U shaped relationship reported by Aldy and Viscusi (2008), note:

> *A possible counter-argument to these empirical results is a non-empirical one. This is that governments, when having to make investment decisions for the community as a whole, are also commonly trying to distribute the benefits equitably, and that this equitable distribution is best done on the basis of valuing life years equally.*

For a mixture of philosophical and pragmatic reasons, in this thesis I have chosen to ignore the possibility that VSLY values could further be modified to take account of the age or health of the relevant cohort. As O'Dea and Wren (2010) state, there is clearly much

potential for New Zealand research to explore the relationship between age and VSL/VSLY which would clarify some of these issues.

4.3.3. Combining the value of a statistical life year with change in mortality hazard

Having established the unmodified VSLY figures that will be the basis of costing any changes in mortality, the costing procedure is fairly straight forward. The following is a modified summary of the approach I took in the original WUNZ:HS evaluation (Telfar Barnard, Preval et al. 2011). In the original report, treatment was not divided into treatment types and thus modifications address the fact that benefits can only be attributed to individuals who received either insulation or insulation and heating (and not heating only). The following process is carried out based on the result for the circulatory cohort. The same process is repeated for the results in the respiratory cohort but only summary tables are presented for reasons of brevity.

4.3.3.1. Valuing change in mortality in the circulatory cohort

It is first necessary to estimate how many individuals per average household would avoid mortality in the first year post-treatment.

(A) = No. of vulnerable individuals in treatment group (65 and over at time of retrofit and circulatory hospitalisation during *before* period) who received either insulation or insulation and heating
= 919

(B) = Risk of mortality per year for a given circulatory cohort control group individual
= Ann. mortality rate per 1000 circulatory cohort control group individuals[19] /1000
= 91.8906/1000
= 0.0918906

(C) = Hazard ratio for mortality for treatment group individuals who received insulation vs. control group
= 0.673

(D) = No. of individuals who avoided mortality in first year post treatment as a result of receiving insulation or insulation and heating
= A*B*(1-C)
= 27.61431988

[19] Annual mortality per 1000 circulatory cohort control group individuals derived from after period mortality data for all circulatory cohort control group individuals.

(E) = No. of individuals *per household* who avoided mortality per household in first year post treatment
= (Chance of an individual avoiding death)* (number of individuals per household)
= (D/ (*All* treatment group individuals))* (Average treatment group household size)
= (D/(100200))* 3.6389367
= 0.001002862

Having established how many individuals per household would avoid mortality in the first year post treatment it was then necessary to make assumptions about how much additional life each individual would gain.

Average remaining expected life years for treatment group individuals who avoided mortality were derived from life tables based on the average age of the vulnerable cohort (Telfar Barnard, Preval et al. 2011) which was 76.67 years. These individuals can expect to live for an additional 11 years on average. Average life gained figure was calculated assuming a 50:50 sex split (which was conservative as there were slightly more females than males, and females typically live longer than males) and thus an average life expectancy for the average treatment group member was calculated. This figure was then halved in order to account for the possibility that these individuals were relatively unhealthy, based on pre-existing health issues. It is important to acknowledge the crudeness of this estimate, which renders the *precision* of the results presented in this section somewhat inherently flawed. When I carried out the original analysis in Telfar Barnard, Preval et al. (2011) it was necessary to make quick and crude estimates. In the intervening years I have not found a better approach to accounting for recent hospitalisation as a predictor of life-years remaining for an elderly New Zealand cohort. The impacts of assuming a 25% reduction in life expectancy or a 75% reduction in life expectancy are also explored, which mitigates the impact of this difficult to justify assumption somewhat.

(F) = Average life years gained per treatment group individual who avoided mortality (mid)
= 11*0.5
= 5.5

(G) = Average life years gained per treatment group individual who avoided mortality (low)
= 11*0.25
= 2.75

(H) = Average life years gained per treatment group individual who avoided mortality (high)
= 11*0.75
= 8.25

Having established a range of conservative figures for (F) – (H) it is then reasonably straight forward to calculate the value of those life years gained. It is assumed that the deaths avoided occurred in the middle of the first year post treatment. Those six months were undiscounted in the original report, an error addressed in this thesis. Subsequent years gained are discounted using an appropriate discount rate.

(J) = Value of change in mortality for a treatment group individual who avoided mortality

$$= 0.5 * \left(\frac{VSLY}{(1+d)^1}\right) + \sum_{n=2}^{n=(F\ or\ G\ or\ H)-1} \left(\frac{VSLY}{(1+d)^n}\right)$$

Table 32: Values of change in mortality per treatment group individual who avoided mortality as a result of receiving insulation (values of J): circulatory cohort

Total years gained per person who avoided mortality	**Discount rate**	Transport 1991-based VSL (**$ 3,586,946**)	Transport 1998-based VSL ($ **6,073,809**)	Fire 2007-based VSL ($**2,367,384**)
(F) 5.5	2.5%	$720,882	$1,220,682	$475,785
	4.0%	$869,988	$1,473,160	$574,195
	8.0%	$1,274,484	$2,158,097	$841,158
(G) 2.75	2.5%	$374,286	$633,785	$247,031
	4.0%	$461,626	$781,676	$304,674
	8.0%	$714,383	$1,209,672	$471,492
(H) 8.25	2.5%	$1,170,473	$1,981,982	$772,517
	4.0%	$1,384,572	$2,344,509	$913,822
	8.0%	$1,930,091	$3,268,242	$1,273,858

Having established a range of values for J with $869,988 the preferred value it is then easy to calculate the total value of life gained at the level of a *treatment group household* as a result of mortality avoided in the first year post treatment by multiplying J by E

(K) = J*E

Table 33: Total value of change in mortality per treatment group household due to mortality avoided during first year post-treatment (values of K): circulatory cohort

Total years gained per person who avoided mortality	Discount rate	Transport 1991-based VSL (**$ 3,586,946**)	Transport 1998-based VSL ($ **6,073,809**)	Fire 2007-based VSL ($**2,367,384**)
(F) 5.5	2.5%	$722.95	$1,224.18	$477.15
	4.0%	$872.48	$1,477.38	$575.84
	8.0%	$1,278.13	$2,164.27	$843.57
(G) 2.75	2.5%	$375.36	$635.60	$247.74
	4.0%	$462.95	$783.91	$305.55
	8.0%	$716.43	$1,213.13	$472.84
(H) 8.25	2.5%	$1,173.82	$1,987.65	$774.73
	4.0%	$1,388.53	$2,351.22	$916.44
	8.0%	$1,935.61	$3,277.60	$1,277.50

At this point a philosophically interesting juncture is reached. A figure K has been produced valuing the total benefit of the mortality avoided per treatment group household in the first year post treatment. However, from a cost-benefit perspective what is required is an estimate of the total benefit of the treatment over the lifetime of the product, which requires an annual benefit estimate.

It might seem reasonable to simply assume that the total benefit K would accrue each year for different individuals occupying the treatment group households; this would require an assumption that there would always be the same proportion of individuals in each treated household that met the criteria (65 and over with recent circulatory hospitalisation).

However, in a real world scenario, given that people continue to live in the same homes it seems possible, particularly given a reasonable assumption that a person who has recently had a hospitalisation of a given type is more likely to have further hospitalisations of that type than an individual who has not recently had a hospitalisation of that type, that a person might avoid mortality as a result of insulation multiple times over a period of years (assuming there is any such benefit). The reason that this was philosophically interesting is that it raised the question of how to value multiple instances of mortality avoidance for an individual. The VSL discussion above suggests that the life-years remaining for a given individual have a fixed value; this implies that if that individual avoids death for a second time they will not gain additional benefit measured in VSLYs beyond that initially gained. To be more explicit, if the remainder of my life is worth two million dollars to me, avoiding a fatal car crash will gain me additional life valued at two million dollars, but avoiding a second car crash on the following day will not gain me additional life worth a

further two million dollars i.e. my remaining life will not now be worth four million dollars to me.

In order to address this concern, assumptions need to be made about the transitioning occupancy of the average insulated home. The most conservative assumption, which seems reasonable given that elderly people tend to have lower levels of residential mobility, is to assume that individuals in the vulnerable cohort remain in the vulnerable cohort until death (i.e. continue to occupy the same treatment home and continue to experience circulatory illness).

The probability that any vulnerable individual avoids mortality in a given year as a result of receiving insulation was calculated for a vulnerable *individual* as:

(L) = Probability a vulnerable treatment individual avoids mortality in a given year due to insulation
= B* (1-C)
= 0.030048226

This figure was the basis for calculating M, a proportional reduction that can be applied to K in order to produce an annual benefit figure that can be applied to multi-year cost-benefit calculations. M identifies the chance that a vulnerable person who avoided mortality might also avoid mortality due to retrofitted insulation *at least once more* during the rest of their lifespan spent in their retrofitted home. The VSLY based benefits gained by such individuals are dropped from analysis in order to avoid double counting, as the benefit of mortality avoided will be captured for that individual by calculation of (K) at the household level in subsequent years in a multi-year cost-benefit analysis.

(M) = Probability that a vulnerable person who avoided death due to insulation retrofit might avoid death due to insulation retrofit at least once more during the remainder of their life.

$= 1- (1-L)^{n}$

(where n is average additional life years predicted for a vulnerable individual)

In the original report (Telfar Barnard, Preval et al. 2011) a figure of 11.5% was calculated for M which was then conservatively rounded to 15%. In this thesis three values for (M) are calculated, assuming that individuals live for an additional 2.75 years (M)= 8.75%, 5.5 years (M)= 14.11% or 8.75 years (M)=21.66%.

M was applied to K in order to produce an on-going estimate of benefit N.

N = K*(1-M)

Table 34: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): circulatory cohort

Total years gained per person who avoided mortality	Discount rate	Transport 1991-based VSL (**$ 3,586,946**)	Transport 1998 - based VSL ($ **6,073,809**)	Fire 2007-based VSL ($**2,367,384**)
(F) 5.5	2.5%	$620.72	$1,051.08	$409.68
	4.0%	$749.11	$1,268.47	$494.41
	8.0%	$1,097.40	$1,858.24	$724.29
(G) 2.75	2.5%	$342.51	$579.98	$226.06
	4.0%	$422.44	$715.32	$278.81
	8.0%	$653.74	$1,106.98	$431.47
(H) 8.25	2.5%	$919.57	$1,557.13	$606.92
	4.0%	$1,087.78	$1,841.95	$717.94
	8.0%	$1,516.36	$2,567.67	$1,000.80

The figures presented in Table 34 are annual mortality benefits per household as a result of improved insulation. These figures can be summed over the predicted life time of an insulation retrofit with appropriate discounting. A complication which it is important to acknowledge is that the effectiveness of insulation may reduce during the course its life-time. How much such a reduction might limit the benefits found above is difficult to estimate. This important research question is further discussed in Subsection 9.1.6.

The primary annual benefit figure is $749.11 for an average household. This figure is calculated using the current transport sector VSL, a 4% discount rate consistent with the original WUNZ:HS cost-benefit evaluation and the central assumption of 5.25 life-years gained.

Stratification of this analysis by CSC status was carried out, driven by the observation that although CSC status does not appear to have an impact on the protective effect of insulation that a higher proportion of vulnerable individuals occupied households that participated in the WUNZ:HS programme at the CSC subsidy level. This means, following steps A –N that greater annual benefits will be calculated for CSC households, due to the higher average numbers of individuals per household who avoided mortality in year one. The converse is true for non-CSC households. The results of these calculations are presented in Table 35.

Table 35: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): circulatory cohort: results by CSC status of household

CSC status of household	Total years gained per person who avoided mortality	Discount rate	Transport 1991 -based VSL (**$ 3,586,946**)	Transport 1998 -based VSL ($ **6,073,809**)	Fire 2007 -based VSL ($**2,367,384**)
CSC	(F) 5.5	2.5%	$806.17	$1,365.11	$532.08
		4.0%	$972.92	$1,647.46	$642.13
		8.0%	$1,425.27	$2,413.43	$940.68
	(G) 2.75	2.5%	$444.85	$753.27	$293.60
		4.0%	$548.65	$929.04	$362.11
		8.0%	$849.06	$1,437.72	$560.38
	(H) 8.25	2.5%	$1,194.31	$2,022.35	$788.25
		4.0%	$1,412.77	$2,392.26	$932.43
		8.0%	$1,969.40	$3,334.81	$1,299.80
Non-CSC	(F) 5.5	2.5%	$339.75	$575.30	$224.24
		4.0%	$410.02	$694.30	$270.62
		8.0%	$600.66	$1,017.10	$396.44
	(G) 2.75	2.5%	$187.47	$317.45	$123.73
		4.0%	$231.22	$391.53	$152.61
		8.0%	$357.82	$605.91	$236.16
	(H) 8.25	2.5%	$503.33	$852.29	$332.20
		4.0%	$595.39	$1,008.18	$392.96
		8.0%	$829.98	$1,405.41	$547.78

4.3.3.2. Valuing changes in mortality in the respiratory cohort

Valuing changes in mortality for the respiratory cohort was based on the 0.8 HR relative to the control group calculated by combining insulation and 'insulation and heating' into a single dummy variable. Replicating the process described in the previous subsection was straightforward: average age was very similar to the circulatory cohort, and it seemed reasonable to predict the same impact on remaining life-years of pre-existing respiratory illness as pre-existing circulatory illness. Key differences included a much lower proportion of vulnerable individuals per household, leading to lower estimates of benefit per household, but also a slightly higher annual mortality rate in the control group (0.146) leading to higher estimates of benefit per household. The primary annual benefit figure is $371.53 for an average household.

Table 36: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): respiratory cohort

Total years gained per person who avoided mortality	**Discount rate**	Transport 1991 -based VSL (**$ 3,586,946**)	Transport 1998 -based VSL ($ **6,073,809**)	Fire 2007 -based VSL ($**2,367,384**)
(F) 5.5	2.5%	$307.83	$521.26	$203.17
	4.0%	$371.51	$629.08	$245.20
	8.0%	$544.24	$921.56	$359.20
(G) 2.75	2.5%	$169.86	$287.63	$112.11
	4.0%	$209.50	$354.75	$138.27
	8.0%	$324.21	$548.99	$213.98
(H) 8.25	2.5%	$456.04	$772.23	$300.99
	4.0%	$539.46	$913.48	$356.05
	8.0%	$752.01	$1,273.39	$496.33

In order to account for differences in the average number of vulnerable occupants per household by the CSC status of that household we can repeat the stratification process described in the previous subsection.

Table 37: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): respiratory cohort: results by CSC status of household

CSC status of household	Total years gained per person who avoided mortality	Discount rate	Transport 1991 -based VSL (**$ 3,586,946**)	Transport 1998 -based VSL ($ **6,073,809**)	Fire 2007 -based VSL ($**2,367,384**)
CSC	(F) 5.5	2.5%	$416.22	$704.80	$274.71
		4.0%	$502.31	$850.57	$331.53
		8.0%	$735.86	$1,246.04	$485.67
	(G) 2.75	2.5%	$229.67	$388.91	$151.58
		4.0%	$283.27	$479.66	$186.96
		8.0%	$438.36	$742.29	$289.32
	(H) 8.25	2.5%	$616.62	$1,044.13	$406.97
		4.0%	$729.41	$1,235.11	$481.41
		8.0%	$1,016.79	$1,721.75	$671.08
Non-CSC	(F) 5.5	2.5%	$143.62	$243.19	$94.79
		4.0%	$173.33	$293.49	$114.40
		8.0%	$253.91	$429.95	$167.58
	(G) 2.75	2.5%	$79.25	$134.19	$52.30
		4.0%	$97.74	$165.51	$64.51
		8.0%	$151.26	$256.13	$99.83
	(H) 8.25	2.5%	$212.77	$360.28	$140.43
		4.0%	$251.68	$426.18	$166.11
		8.0%	$416.22	$704.80	$274.71

4.3.3.3. Accounting for individuals in both cohorts

A final corrective, necessary from a cost-benefit analysis perspective, is to consider how to account for the 96 individuals aged 65 and over who had both circulatory and respiratory hospitalisations prior to receiving insulation or insulation and heating under WUNZ:HS, i.e. were in both cohorts. A simple and pragmatic approach is to avoid double counting any benefits these individuals might receive by adjusting the results presented in Table 36. There were 469 individuals in the respiratory cohort, we can account for double counting by reducing household level benefits for the respiratory cohort by 96/469=20.47%.

Table 38: Ongoing annual benefit of reduced mortality per household as a result of receiving insulation (values of N): respiratory cohort: adjusted to avoid double counting

Total years gained per person who avoided mortality	Discount rate	Transport 1991 -based VSL (**$ 3,586,946**)	Transport 1998 -based VSL ($ **6,073,809**)	Fire 2007 -based VSL ($**2,367,384**)
(F) 5.5	2.5%	$260.19	$440.59	$171.73
	4.0%	$314.01	$531.72	$207.25
	8.0%	$460.01	$778.94	$303.61
(G) 2.75	2.5%	$127.11	$215.24	$83.90
	4.0%	$156.78	$265.47	$103.47
	8.0%	$242.62	$410.83	$160.13
(H) 8.25	2.5%	$362.70	$614.16	$239.38
	4.0%	$429.04	$726.50	$283.17
	8.0%	$598.08	$1,012.74	$394.73

A similar corrective can be applied to CSC stratified results although it is not presented here for reasons of concision.

4.3.4. Summary

Section 4.3 presents a method for valuing changes in mortality rates. A range of values for reduced mortality rates are presented. The assumptions underlying these calculations are laid out and justified. These values are applied to the reductions in mortality found in Section 4.2 in order to produce annual estimates of the financial benefit of changes in mortality resulting from receiving insulation under WUNZ:HS at the household level. The resulting figures can be combined with any estimated changes in household hospitalisation costs or pharmaceutical use costs resulting from receiving a WUNZ:HS retrofit in order to produce summary figures suitable for inclusion in a cost-benefit analysis. Chapter Five addresses hospitalisation costs and Chapter Six addresses pharmaceutical use costs.

Chapter Seven presents the full range of benefits calculated, sums them at the household level and applies them to the cost-benefit model presented in Grimes, Denne et al. (2011).

CHAPTER 5. MODELLING CHANGES IN HOSPITALISATION COSTS

The goal of the analyses presented in Chapters Four, Five and Six of this thesis is to assess whether receiving treatment under WUNZ:HS leads to changes in health-related benefits and costs. While the analysis of mortality data presented in Chapter Four required additional steps in order to produce estimates of changes in costs, NMDS hospitalisation data already includes cost, eliminating the need to estimate costs based on different theoretical foundations.

Cost data are inherently more skewed to high values than hospitalisation count data and in fact there may appear to be a case for analysing changes in the rate of hospitalisation rather than costs. However, we can imagine situations where a treatment reduced the rate of hospitalisation, but those hospitalisations that did occur were more costly on average, meaning that treatment had no benefit in cost terms: analysis based purely on count data would miss such subtleties.

The original goal of this chapter was to both extend the household level analyses of hospitalisation cost data that I carried out in the two WUNZ:HS health evaluations (Telfar Barnard, Preval et al. 2011, Preval 2012) and to explore the potential for individual level analyses of hospitalisation cost data, which were not included in the original reports (Telfar Barnard, Preval et al. 2011).

There are two hospitalisation categories that may plausibly be influenced by treatment under WUNZ:HS; circulatory illness-related hospitalisation and respiratory illness-based hospitalisation. Other potential outcome measures, such as psychiatric hospitalisation, have some slight plausibility, but are not supported by much evidence and thus are not explored here. It is also possible to analyse all-cause hospitalisations, attempting to capture all changes in hospitalisation cost in a single analysis, assuming that differences between treatment groups in other hospitalisation cost categories are random. Thus any difference in total hospitalisation costs may be interpreted as resulting from receiving treatment under WUNZ:HS. In the two WUNZ:HS evaluations results for all-cause hospitalisations, circulatory hospitalisations and respiratory hospitalisations were presented but CBA results were calculated based on all-cause hospitalisation results (Grimes, Denne et al. 2011), which was a conservative decision in context.

Section 5.1 revisits the potential of the dataset for individual analysis. Typical characteristics of hospitalisation cost data are discussed and initial descriptive statistics for the available individual level hospitalisation data are presented to inform this discussion. Section 5.1 presents some initial analysis of a number of potentially suitable model

specifications and presents the results of model diagnostics. These considerations confirmed that the data available are not well suited to individual level modelling.

Section 5.2 describes an innovative approach to modelling hospitalisation costs, panel-based analysis of cost difference data aggregated at the treatment and matched control group household level, and carries out the model building process. The model utilised in the two WUNZ:HS health evaluation papers (Telfar Barnard, Preval et al. 2011, Preval 2012) is extended to explore the impact of climate variation and the separation of the exposure of interest from treatment vs. control to insulation, heating, and insulation and heating. Particular attention is given to the discussion of extreme values.

5.1. Individual level analysis

Health cost and health services utilisation modelling is typically complicated by the distribution of the dependent variable, such as skewness (a heavy right tail) or a high proportion of zero values (Venturini, Dominici et al. 2008). Model specifications such as ordinary least squares regression are often not appropriate for untransformed data of this type because their assumptions are violated, for example the distribution of error terms in such models is unlikely to be approximately normal and the homoscedasticity assumption is also likely to be violated (Dodd, Bassi et al. 2006). Excess zeroes may also be a concern for ordinary least squares models in some cases (Mihaylova, Briggs et al. 2011).

The hospitalisation data from the WUNZ:HS evaluation demonstrates the typical characteristics of health cost data. Table 39 presents summary measures of average annual hospitalisation cost post-treatment for all individuals who had at least one year of pre-treatment data (important as baseline hospitalisation costs are an important potential confounder). Individuals who had no post-treatment data, i.e. who died prior to receiving treatment are excluded from analysis as are occupants of treatment homes that did not receive insulation or heating and their matched control households.

Table 39: Summary measures, all study individuals' post-treatment average annual hospitalisation cost data

Variable name	Mean	Standard deviation	Skewness	Kurtosis	Proportion of zero values
Average annual hospitalisation cost (all cause)	$284.91	2774.85	59.74	7059.98	91%
Average annual circulatory hospitalisation cost	$47.68	854.26	50.52	5133.21	98.7%
Average annual respiratory cost	$27.19	584.47	94.59	13713.31	98.8%

Table 39 shows that the hospitalisation data are highly skewed: they exhibit extreme kurtosis and likely exhibit excess zeroes. Such data are unlikely to be suited for analysis using standard OLS regression, although in some cases a large enough dataset may overcome such problems (Mihaylova, Briggs et al. 2011). A number of approaches to modelling cost data are described in the literature. Common approaches cited include OLS (Briggs and Gray 1998, O'Hagan and Stevens 2003, Briggs, Nixon et al. 2005, Mihaylova, Briggs et al. 2011), OLS using log transformed costs (Briggs and Gray 1998, Manning and Mullahy 2001) and Generalised Linear Models (GLMs) of the Gamma family (where the mean is a function of the square of the variance) with the log link (where the dependent variable is the log of the expected value of cost) (Manning and Mullahy 2001, Dodd, Bassi

et al. 2006, Mihaylova, Briggs et al. 2011) . In addition to this GLM specification many other combinations of family and link function may be used (Manning, Basu et al. 2005, Mihaylova, Briggs et al. 2011), and automated methods have been developed for selecting the best combination of family and link for a given dataset. The problem of excess-zeroes has been addressed by using two-step models, in which the probability of incurring any cost at all is modelled using a logit or Probit model and then a suitable GLM or OLS is used to model costs amongst those who have incurred costs(Duan 1983, Mihaylova, Briggs et al. 2011) .

Choice of model in the literature appears to be driven partly by theory and partly by the characteristics of the data available. If multiple model specifications were appropriate for the data available in this study it would be necessary to consider the criteria to be used for model selection more carefully. However, an initial review of models fitted for the three categories of hospitalisation cost suggest that analysis at an individual level may not be tenable at all, which was the conclusion reached in the original WUNZ:HS health evaluation (Telfar Barnard, Preval et al. 2011). Post-estimation model diagnostics for initial models containing the following plausible confounders: sex, age, ethnicity, pre-treatment hospitalisation costs of the relevant type, NZDep2006score, dwelling risk score and location, showed that none of the standard approaches appeared viable.

I ran initial models using all-cause hospitalisation costs as the dependent variable. I assessed OLS, log-OLS, GLM (Gamma family and log link) and a two-step model using a logit model to determine the probability of incurring a hospitalisation and a GLM (Gamma family and log link) to model costs amongst those who incurred a cost. The models produced had extremely low R^2 and adjusted R^2 scores, and in the case of the two-stage model the model would not converge. I analysed those models that were successfully implemented using the linktest function. The linktest function, available in Stata, is based on work by Pregibon (1979) and involves rerunning a given regression $y= f(X\beta)$with both $X\beta$ and $X\beta^{\wedge 2}$ as covariates (StataCorp. 2009). If the $X\beta^{\wedge 2}$ term is statistically significant this indicates that the independent variables in the model have not been correctly specified, or alternatively that the dependent variable has not been correctly specified (StataCorp. 2009). In this case all rerun models had highly significant $X\beta^{\wedge 2}$ terms, suggesting that there were serious concerns about the specification of either the dependent variable or the independent variables.

I briefly considered some initial inclusion criteria such as age (65 and over) or evidence of pre-existing health issues (a baseline hospitalisation event of the relevant type). The results of these more targeted analyses also indicated that individual analysis would be highly problematic. At this point I concluded that the analysis of individual level hospitalisation

cost data was not viable for this dataset. The rest of this chapter focuses on a more promising analytical approach: household level modelling of hospitalisation costs.

5.2. Household level analysis: Original model

During the evaluation of WUNZ:HS it was concluded that individual level cost analysis would not prove viable. This conclusion led Group members to consider an alternative approach: household level analysis of hospitalisation and pharmaceutical costs. This approach had already proven extremely effective in the analysis of electricity and energy use data produced by Grimes, Young et al. (2011). I modified the approach to suit the characteristics of the cost data available. The benefits of the household level models included a reduction in the number of zero values for the dependent variable, and the centring of the dependent variable around zero.

The resulting household level models produced statistically significant and plausible results, including an estimated reduction in monthly all-cause hospitalisation costs per-household of approximately $5.37 (Telfar Barnard, Preval et al. 2011). This result was not replicated in the MED-funded extension, however, which found no statistically significant change in hospitalisation costs as a result of receiving insulation or heating under the WUNZ:HS programme (Preval 2012).

In this Section I will reproduce the model I used in the two WUNZ:HS evaluation heath reports, with some minor modifications such as separate insulation, heating and insulation and heating dummy variables. In my original analysis the possibility of including additional model covariates in order to address possible confounding was rejected (see Subsection 4.1.4). Here plausible household level confounders are considered.

5.2.1. Model specification

The original model is a difference-in-difference model. The model compares the difference between each treatment group household's average monthly hospitalisation costs and the mean of its matched control group households' monthly hospitalisation costs before and after the intervention. The model is a fixed effects OLS estimator with robust standard errors clustered by treatment/control household pairing.

$$Hosp\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 x_{4it} \ldots + \beta_p x_{pit} + \varepsilon_{it}$$

$Hosp\ Cost\ Diff_{it}$ represents the difference between the average monthly hospitalisation costs of treatment house$_i$ and its control group houses over time t (either *before* or *after*); α_i represents the unobserved individual house fixed effect of house$_i$ relative to its control houses; $insulation_{it}$ is a dummy variable, that is 1 if house$_i$ had retrofitted insulation *only* for the entirety of the after period, $heating_{it}$ is a dummy variable that is 1 if house$_i$ had retrofitted heating *only* for the entirety of the after period; $heating_and_insulation_{it}$ is a dummy variable that is 1 if house$_i$ had both retrofitted insulation and heating for the after

period. The coefficients on the treatment dummies, if statistically significant, indicate the size and direction of any change in the difference in average monthly hospitalisation costs as a result of treatment. As with Cox Proportional Hazard model in Chapter Four it is not possible to fully characterise changes in treatment group status using this model, particularly where there is some gap between the initial insulation retrofit and the additional heating retrofit: the extended model described in Section 5.3 addresses this concern. $x_{4it}...x_{pit}$ are confounders (discussed in Subsection 5.2.3. below); ε_{it} is the residual term, which is assumed to be correlated between periods within houses, but independent between houses.

5.2.2. Model construction and approach to extreme values

Model construction required the summing of the relevant hospitalisation costs (either all-cause, respiratory or circulatory) by household for each month of data. An average hospitalisation cost figure was calculated for all of the matched control households of a given treatment household. Finally the difference in monthly costs between the treatment household, and the average of its matched controls for that month was calculated. Months for a given treatment/control pairing where there were either no individuals occupying the treatment household, or no individuals occupying *any* control households, were dropped (these records reflect deaths during the study period). The month in which the first treatment occurred is also dropped. Finally the monthly difference in cost figures were averaged to produce a before period average monthly cost difference and an after period average monthly cost difference for each treatment/control pairing.

At this point, for each model, treatment/control pairings that had before or after period cost difference records that were below the 1st or above the 99th percentiles were dropped, consistent with my approach in the original WUNZ:HS health evaluations. The results of this removal on measures of skew and kurtosis are set out in Table 40.

Table 40: Summary statistics, dependent cost variables, with and without exclusions

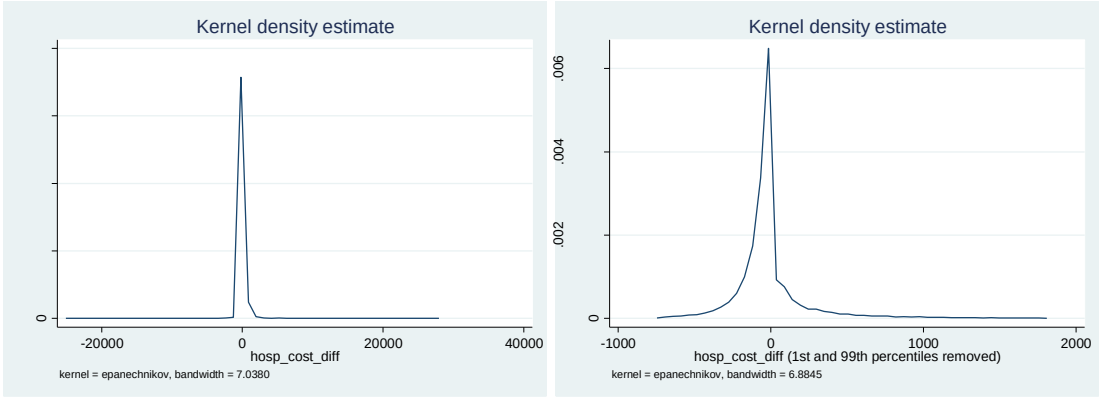
Dependent variable	Inclusion rule	Mean	Std. deviation	Skewness	Kurtosis
Hosp_cost_diff	All pairings	13.95	534.85	10.22	449.01
	Exclude below 1st and above 99th percentiles	-7.76	246.01	2.64	15.98
Circ_cost_diff	All pairings	6.26	219.69	27.99	1634.6
	Exclude below 1st and above 99th percentiles	-3.31	62.14	3.44	29.26
Resp_cost_diff	All pairings	3.15	122.59	27.7	1570.19
	Exclude below 1st and above 99th percentiles	-1.68	32.63	3.59	29.34

Table 40 suggests that the exclusions greatly reduce the standard deviation, skew and kurtosis of all three dependent variables. This is to be expected, and suggests that data quality has been improved by the exclusions. The consistent reduction in mean values suggests that the extreme values above the 99th percentile were further from the original mean than the extreme values below the 1st percentile. This is not necessarily a concern, and is likely to reflect the averaging process by which the dependent variables were constructed, where costs for control households are averaged across up to nine other control households[20], reducing the extremeness of their averaged values.

The impact of removal is also demonstrated by a kernel density plot of hosp_cost_diff for data with and without removal in Figure 9. The kernel density plot of the complete dataset does not visually indicate any obvious cut-off point for extreme data.

[20] It was not necessary to randomly remove control households where there were more than four matches for a given treatment household, as was the case for the individual level datasets used in previous chapter, because the averaging process employed in this chapter means that the number of control households would not impact analysis.

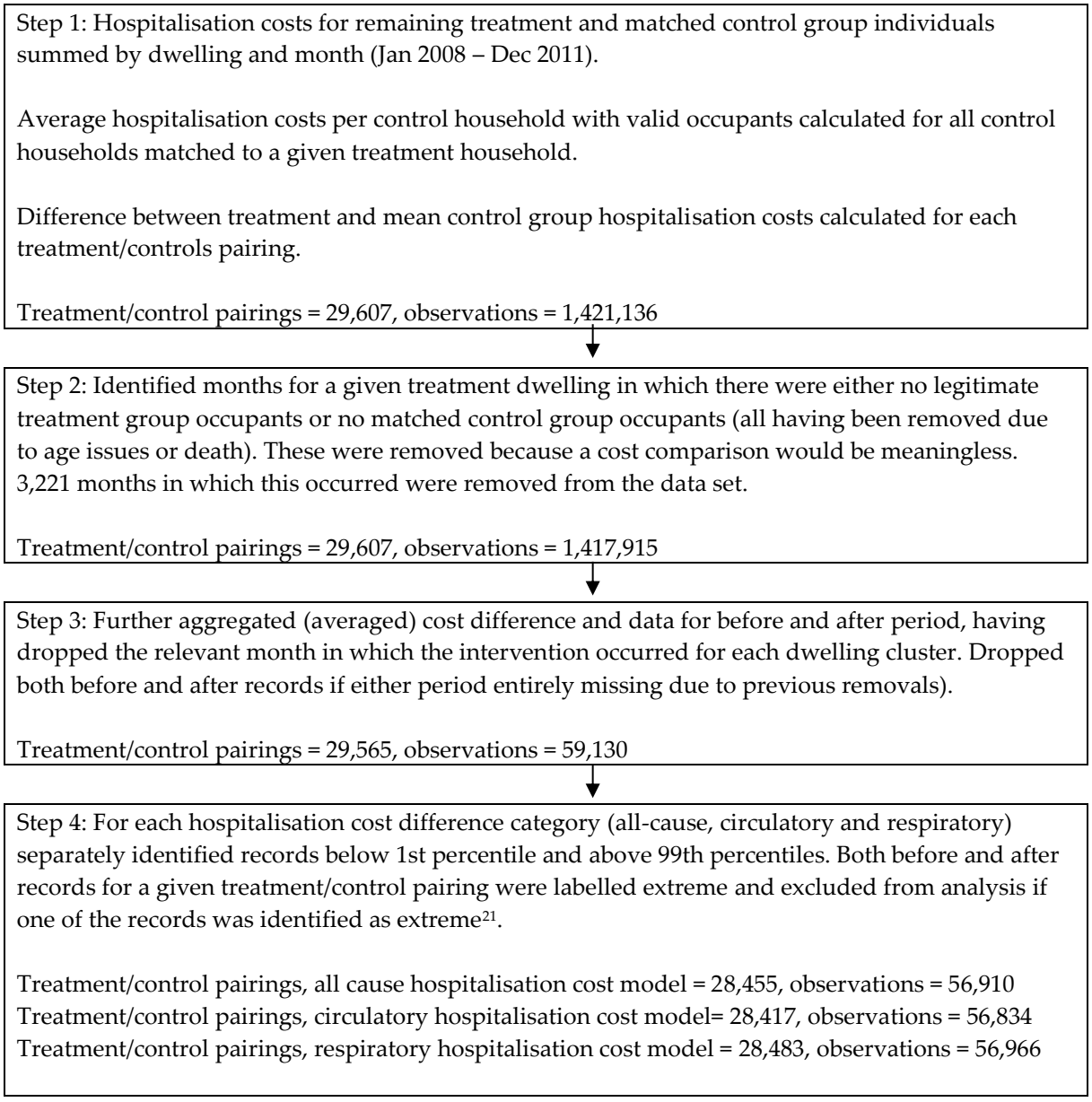
Figure 9: Kernel density plots, hosp_cost_diff, with and without removal of extreme values



Removal of extreme values greatly reduces skewness and kurtosis, but may be considered controversial given that the very nature of hospitalisation cost data is inherently skewed and there are no grounds for removal of extreme values due to data quality concerns. However, the possibility of a few extremely costly hospitalisations greatly distorting analysis is ultimately too important to ignore. Later, in Subsection 5.2.5 the impact of removal is explored in depth with models being rerun both without exclusions and with removal of treatment/control pairings that had before or after period cost difference records that were below the 5th or above the 95th percentiles.

The dataset construction and cleaning process is set out in Figure 10.

Figure 10: Dataset construction, original hospitalisation costs models



[21] Slight variation in the numbers of pairings and observations reflect differences in the numbers of treatment pairing that had an extreme value in both the *before* and *after* periods.

5.2.3. Confounding

Household level confounding is an interesting topic. In the original report potential confounders were excluded entirely from analysis, with confounding by time-invariant potential covariates such as region or differences in dwelling health risk between treatment households and matched control households captured by the fixed effects term in the model (as are unmeasured time-invariant potential confounders). This approach is reasonable for confounders that do not change. However, there are two time-varying covariates that were not considered in the original report which are worth considering here.

5.2.3.1. Occupant characteristics

One of the reasons that occupant characteristics may be an important confounder in a model of this type is that they change over the course of the study. If there were no births or deaths during the study then the difference between the number of occupants in a treatment household and its matched controls would be a constant and thus captured by $\alpha_{i.}$, however this is clearly not the case. There are both known changes in occupant characteristics (birth, death) and also many unknown changes such as the real residence of individuals in the dataset, which will not always be in a study household (see Subsection 3.1.1.6 for more on this concern).

It is not possible to account for unknown changes in occupant characteristics, although the quasi-random design of the WUNZ:HS evaluation limits any confounding effect. However, it is possible to address confounding created by deaths and births in the study cohort. To address these changes a model covariate named $occupant_diff_{it}$ can be calculated for the before and after periods for each treatment/control pairing based on the averaged difference in monthly occupancy of treatment households and their matched control households. It is easy to imagine that changes in $occupant_diff_{it}$ could confound analysis, given that there is very likely to be a strong relationship between the difference in the number of occupants and the difference in costs. Similarly, a figure named age_diff_{it} can be used to address confounding by changes in the average age of treatment and matched controls, which would primarily result from deaths and births. Age_diff_{it} is defined as the difference between the average age of all treatment household occupants and the average age of all matched control household occupants. The distribution of these variables for the complete dataset in the "after" period is set out in below.

Figure 11: Histograms showing distribution of occ_diff and age_diff, "after" period

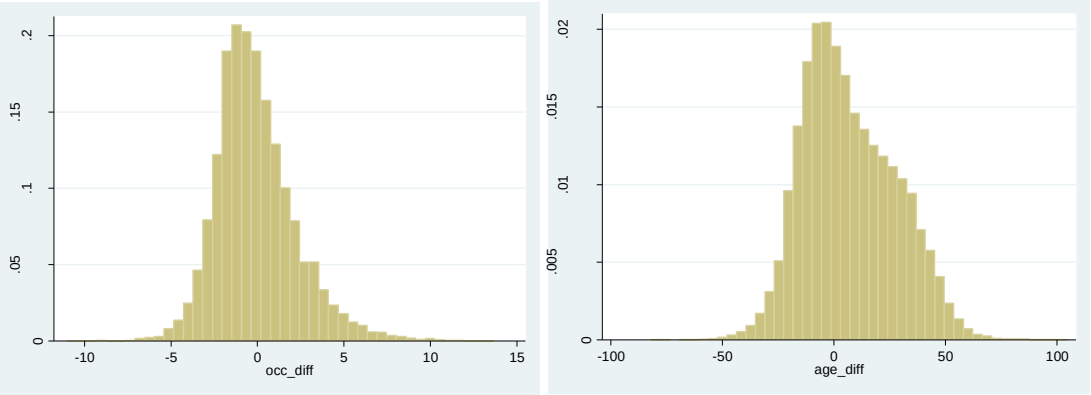


Table 41 presents separate simple regression analyses for these confounders. The models have either all–cause hospitalisation cost difference, circulatory illness-related hospitalisation cost difference or respiratory illness-related hospitalisation cost difference as the dependent variable. The equations for these simple regressions are

1) $Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 occupant_diff_{it} + \varepsilon_{it}$

2) $Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 age_diff_{it} + \varepsilon_{it}$

Table 41: Simple regression analyses, occupant-diff and age_diff variables, treatment/control pairings with values below 1st and above 99th percentiles excluded

Covariate	Dependent variable	Coefficient	Robust standard error	Z	P>|Z|	Confidence intervals	
Occupant_diff	All-cause hosp diff	56.09	7.292	7.69	<0.001***	41.80	70.39
	Circulatory hosp diff	5.51	1.657	3.33	0.001***	2.27	8.76
	Respiratory hosp diff	5.21	1.101	4.73	<0.001***	3.05	7.37
Age_diff	All-cause hosp diff	7.90	0.796	9.93	<0.001***	6.34	9.46
	Circulatory hosp diff	1.33	0.214	6.22	<0.001***	0.91	1.75
	Respiratory hosp diff	0.03	0.118	0.21	0.832	-0.21	0.26

* p<0.1, ** p<0.05, *** p<0.01

In the above table, the first row shows that when *occupant_diff* is regressed against all-cause hospitalisation cost difference using the original fixed effects model it has a coefficient of 56.09, which is interpreted as showing that a one unit increase in the value of

occupant_diff, e.g. a child being born in the treatment group, is associated with a $56.09 increase in average monthly all-cause hospitalisation cost difference.

Any confounding by these variables will depend on the way in which they change over the course of the study. For example, the average values of *occupant_diff* and *age_diff* were higher in the after period than in the before period for insulation only treatment/control pairings. This suggests that the estimate of the coefficient of the insulation variable in an all-cause hospitalisation cost difference model that excluded these covariates would be biased towards showing harm as a result of treatment.

Other potential confounders are difficult to assess. For example, there is a great deal of evidence that ethnicity is linked to risk of circulatory illness in New Zealand. It may be that changes in the composition of treatment and control households (via birth and death) which result in changes in the proportions of different ethnicities within a treatment/control pairing are a source of potential confounding. However, it is not clear how to define a factor variable that effectively captures the ethnic composition of a given treatment/control pairing. In this context, it is important to note that the fixed effects term in each model will capture the impact of many unknown and complex characteristics of treatment/control pairings that are time-invariant.

5.2.4. Initial multiple regression model

The initial multiple regression model specification is:

$$Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 occupant_diff_{it} + \beta_5 age_diff_{it} + \varepsilon_{it}$$

The results of the multiple regression models are presented in Table 42 for each hospitalisation type. F-statistics and R^2 model fit measures are presented.

Table 42: Initial multiple regression models, hospitalisation cost difference models, treatment/control pairings with values below 1st and above 99th percentiles excluded

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	Z	P>|Z|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	-3.33	2.064	-1.61	0.107	(-7.374	0.718)	R^2within= 0.0230	F (5: 28,455) =49.50
	Heating only	3.36	6.897	0.49	0.626	(-10.159	16.878)	R^2between= 0.0373	
	Insulation and heating	-2.17	4.719	-0.46	0.645	(-11.422	7.078)	R^2overall= 0.0248	
Difference in circulatory hospitalisation costs	Insulation only	-1.40	0.563	-2.49	0.013**	(-2.504	-0.297)	R^2within= 0.0071	F(5: 28416) =15.07
	Heating only	-2.33	1.797	-1.3	0.194	(-5.854	1.189)	R^2between= 0.0194	
	Insulation and heating	-3.38	1.335	-2.53	0.011**	(-5.993	-0.758)	R^2overall= 0.0116	
Difference in respiratory hospitalisation costs	Insulation only	-0.50	0.286	-1.73	0.083*	(-1.058	0.065)	R^2within= 0.0022	F=(5:28483) =5.96
	Heating only	-0.12	0.958	-0.13	0.897	(-2.001	1.754)	R^2between= 0.0076	
	Insulation and heating	-0.47	0.624	-0.76	0.447	(-1.697	0.748)	R^2overall= 0.0047	

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

The covariates *age_diff* and *occupant_diff* both had highly statistically significant associations with the dependent variable in each cost difference model, confirming the value of the decision to consider them.

Initial multiple regression models show that there is no statistically significant effect of any treatment type on all-cause hospitalisation cost difference, There are small, but statistically significant reductions in circulatory hospitalisation cost differences of $1.40 per month as a result of receiving insulation (p=0.013) and $3.38 for both insulation and heating (p=0.011). There is a small reduction in monthly respiratory hospitalisation cost differences of $0.50 as a result of receiving insulation which is significant at the (p<0.1) level.

The models show no statistically significant change in cost differences as a result of receiving heating only, which is consistent with results from the Cox Proportional Hazard model presented in the previous chapter.

The models generally show no statistically significant impact of result of receiving both heating and insulation, except in the case of the circulatory hospitalisation model, where "heating and insulation" has a larger coefficient than insulation alone, which would be consistent with an additive effect of heating. However a Wald test for equality of the insulation and 'insulation and heating' coefficients shows that there is little evidence to reject the null hypothesis that the two coefficients are equal (F (1,28,416)=1.87, prob>F =

0.1709). This pattern is replicated in the all-cause and respiratory illness-related hospitalisation cost difference models where Wald test results also show that there is no evidence to reject the null hypothesis that the coefficients of insulation and 'insulation and heating' are identical, with scores of (F(1, 28,454)=0.05, prob > F = 0.8213) and (F(1, 28,482)=0.00, prob >F = 0.9740) respectively.

Given the similarity between the coefficients for insulation and 'insulation and heating' in the all-cause and respiratory illness-related hospitalisation cost difference models it appears likely that the absence of a statistically significant effect reflects limited statistical power, given that a small proportion of households received both insulation and heating.

Comparing the coefficients and statistical significance of the insulation variable in the three models, it is plausible that monthly savings in the all-cause model of $3.33 include the effects of insulation on both circulatory hospitalisation cost savings $1.40 and respiratory hospitalisation cost savings of $0.50, as well as perhaps some additional hospitalisation cost impacts. The coefficient of the insulation dummy is not statistically significant in the all-cause cost difference model (p=0.107), which may reflect greater randomness introduced by the inclusion of a high proportion of hospitalisation cost types not relevant to the treatment. However, when insulation and 'insulation and heating' are combined into a single dummy variable and the all-cause hospitalisation cost model rerun, the variable has a coefficient of -3.15 (p=0.098) which is statistically significant at the (p<0.1) level.

It is important to note the very low R^2 values reported for all multiple regressions in Table 42. There is a concern that the inclusion of additional model covariates in this context will explain a great deal of variation and thus appear statistically significant. This concern needs to be balanced against the danger of excluding real confounders from the analysis.

5.2.5. Transformation of continuous variables and interaction terms

The fractional polynomial approach described in Subsection 4.1.4 was initially applied to the transformation of the continuous variables in the three models. However, results suggested that a fractional polynomial approach was not appropriate, producing models with much poorer fit. It appears likely that applying the fractional polynomial technique to models that had very low measures of goodness-of-fit, led to over-fitting. In addition, it was difficult to assess the biological plausibility of transformed variables given the complexity of the relationship between the dependent variables and the *age_diff* and *occupant_diff* variables. Ultimately it seemed difficult to justify transforming the continuous variables in the model due to these concerns.

The difficulties experienced when transforming continuous covariates gives some credence to the approach taken in the original WUNZ:HS analysis where the inclusion of covariates in cost models was avoided due to the potential for over-fitting. Inclusion of untransformed covariates represents a compromise between the risk of over-fitting and the legitimate reasons to consider confounding described in Subsection 4.1.4.

Finally, there were no plausible interaction terms to consider for any model.

5.2.6. Assessing the impact of extreme value removal

Removal of extreme values, as described in Subsection 5.2.2, will produce better fitting models but, as noted previously, may be controversial. Ultimately extreme value removal is a pragmatic response to extremely skewed data but it is important to explore the impact of removal. In Table 43 both models are run both without any exclusions and then with the removal of all treatment/control pairings with observations with cost values that are below the 5th or above the 95th percentiles.

Table 43: Hospitalisation cost diff models, no exclusions and treatment/control pairings with values below the 5th or above the 95th percentiles excluded

Dependent variable	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	-2.217	4.623	-0.48	0.632	(-11.278	6.844)	R^2within= 0.0139	F (5:29,564) =32.78
	Heating only	-13.048	13.892	-0.94	0.348	(-40.277	14.180)	R^2between= 0.0193	Prob>F =0.0000
	Insulation and heating	19.050	10.832	1.76	0.079*	(-2.182	40.282)	R^2overall= 0.0125	
Difference in all-cause hospitalisation costs (values below 5th or above 95th percentiles removed)	Insulation only	-4.835	1.150	-4.21	0.000**	-(7.088	-2.581)	R^2within = 0.0082	F(5,24374) = 20.27
	Heating only	-5.252	3.727	-1.41	0.159	-(12.558	2.054)	R^2between = 0.0250	Prob > F = 0.0000
	Insulation and heating	-6.447	2.645	-2.44	0.015*	-(11.631	-1.263)	R^2overall = 0.0154	
Difference in circulatory hospitalisation costs	Insulation only	-2.444	2.209	-1.11	0.269	(-6.774	1.886)	R^2within= 0.0050	F(5: 29564) =8.65
	Heating only	-11.236	4.702	-2.39	0.017**	(-20.452	-2.020)	R^2between= 0.0101	Prob > F = 0.0000
	Insulation and heating	0.315	3.475	0.09	0.928	(-6.496	7.126)	R^2overall= 0.0059	
Difference in circulatory hospitalisation costs (values below 5th or above 95th percentiles removed)	Insulation only	-0.572	0.159	-3.59	0.000***	-(0.885	-0.260)	R^2within = 0.0019	F(5,24281) = 9.38
	Heating only	-0.801	0.551	-1.45	0.147	-(1.881	0.280)	R^2between = 0.0006	Prob > F = 0.0000
	Insulation and heating	-1.131	0.364	-3.1	0.002***	-(1.844	-0.417)	R^2overall = 0.0005	
Difference in respiratory hospitalisation costs	Insulation only	1.225	1.070	1.14	0.252	(-0.873	3.323)	R^2within= 0.0028	F(5: 29564) =6.7
	Heating only	0.479	3.280	0.15	0.884	(-5.951	6.908)	R^2between= 0.0036	Prob > F = 0.0000
	Insulation and heating	1.103	1.618	0.68	0.495	(-2.068	4.275)	R^2overall= 0.0024	
Difference in hospitalisation costs (values below 5th or above 95th percentiles removed)	Insulation only	0.011	0.082	0.14	0.892	-(0.150	0.173)	R^2within = 0.0010	F(5,24474) = 4.95
	Heating only	-0.198	0.270	-0.73	0.464	-(0.728	0.332)	R^2between = 0.0016	Prob > F = 0.0002
	Insulation and heating	-0.044	0.187	-0.24	0.813	-(0.411	0.322)	R^2overall = 0.0010	

* p<0.1, ** p<0.05, *** p<0.01

All models control for occupant_diff and age_diff

Comparison of coefficients for particular exposures of interest across models with different exclusions shows that the approach to exclusion of extreme values has a major impact on estimates of some of the coefficients of interest. For example, the coefficient of the heating and insulation variable is 19.05 (p=0.079) without exclusions, -2.17 (p=0.645) with exclusion of values below the 1st or above the 99th percentiles and -6.44 (p=0.015) with exclusion of values below the 5th or above the 95th percentiles. It can be concluded from this example that model results are sensitive to exclusion strategy and that it is important

to be transparent about the impact of different approaches. An a priori decision to exclude values below the 1st or above the 99th percentiles appears to be reasonable and generally provides estimates bounded by the other two approaches to exclusion.

5.2.7. Model specification and global goodness-of-fit measures

In order to address the possibility of model misspecification, leading to inaccurate calculation of p-values and confidence intervals, I calculated bootstrapped estimates of p-values and confidence intervals, using 500 repetitions for each calculation. Bootstrapped estimates were consistent with p-values and confidence intervals calculated using robust standard errors clustered at the treatment/control pairing level, suggesting that model misspecification is not a major concern. Table 44 presents a comparison of p-values and confidence intervals for the difference in all-cause hospitalisation cost model, demonstrating this consistency.

Table 44 All-cause hospitalisation cost difference model with robust clustered standard errors and bootstrapped standard errors, treatment/control pairings with values below 1st and above 99th percentiles excluded

Model	Variable name	Coefficient	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Difference in all-cause hospitalisation costs model with robust clustered standard errors	Insulation only	-3.33	2.064	-1.61	0.107	(-7.374	0.718)
	Heating only	3.36	6.897	0.49	0.626	(-10.159	16.878)
	Insulation and heating	-2.17	4.719	-0.46	0.645	(-11.422	7.078)
Difference in all-cause hospitalisation costs with bootstrapped standard errors	Insulation only	-3.32	2.046	-1.63	0.104	(-7.340	0.683)
	Heating only	3.36	6.973	0.48	0.63	(-10.309	17.028)
	Insulation and heating	-2.17	4.638	-0.47	0.64	(-11.263	6.919)

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

Turning to measures of global goodness-of-fit, all of the models presented in the previous subsections have extremely low measures of goodness-of-fit in terms of R^2 overall, ranging from the 0.024 to 0.0005. This suggests that the models are only explaining a tiny proportion of the variation in hospitalisation costs for the treatment/control pairings. This is perhaps not surprising given the limitations of the data: health relevant time-varying covariates such as relative proportions of smokers per treatment/control pairing or relative levels of obesity are missing from the model as are a great number of other potential covariates. Furthermore, it was argued by (Grimes, Young et al. 2011), who found similar R^2 results in their modelling of energy efficiency, that such model outcomes reflected the success of the matching process, noting that in the absence of a treatment effect any variation should be random if houses are successfully matched, leading to an R^2 of 0, i.e. a

model with no explanatory power. The authors note that they would expect the "signal" of treatment to be dominated by the "white noise" of random variation, thus leading to very low R^2 scores(Grimes, Young et al. 2011).

5.2.8. Stratification by Community Services Card status

Stratification by Community Services Card status of the treated household in each treatment/control pairing is straight forward: treatment status is interacted with CSC status, where the CSC status of the treatment household in a given treatment/control pairing is captured by the variable CSC_i which either equals 1 or 0.

$$Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} * CSC_i + \beta_2 insulation_{it} * (1-CSC_i) + \beta_3 heating_{it} * CSC_i + \beta_4 heating_{it} * (1-CSC_i) + \beta_5 heating_and_insulation_{it} * CSC_i + \beta_6 heating_and_insulation_{it} * (1-CSC_i) + \beta_7 occupant_diff_{it} + \beta_8 age_diff_{it} + \varepsilon_{it}$$

Table 45 presents the results of this stratification of treatment dummy variables.

Table 45: Hospitalisation cost models with treatment stratified by Community Services Card status, treatment/control pairings with values below 1st and above 99th percentiles excluded

Dependent variable modelled	CSC status of treatment household	Variable name	Coefficient	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Difference in all-cause hospitalisation costs	CSC	Insulation	-4.408	3.044	-1.45	0.148	(-10.375	1.559)
		Heating	8.370	10.856	0.77	0.441	(12.908	29.649)
		Insulation and heating	1.432	6.580	0.22	0.828	(-11.465	14.329)
	Non-CSC	Insulation	-2.010	2.618	-0.77	0.443	(-7.141	3.121)
		Heating	-1.689	8.455	-0.2	0.842	(-18.261	14.884)
		Insulation and heating	-7.429	6.466	-1.15	0.251	(-20.103	5.245)
Difference in circulatory hospitalisation costs	CSC	Insulation	-1.564	0.844	-1.85	0.064*	(-3.217	0.090)
		Heating	1.663	2.995	0.56	0.579	(-4.207	7.533)
		Insulation and heating	-4.178	1.811	-2.31	0.021**	(-7.728	-0.628)
	Non-CSC	Insulation	-1.199	0.686	-1.75	0.081*	(-2.545	0.146)
		Heating	-6.313	1.983	-3.18	0.001***	(-10.199	-2.427)
		Insulation and heating	-2.203	1.945	-1.13	0.257	(-6.016	1.610)
Difference in respiratory hospitalisation costs	CSC	Insulation	-0.547	0.443	-1.23	0.217	(-1.416	0.322)
		Heating	-0.175	1.545	-0.11	0.91	(-3.203	2.853)
		Insulation and heating	-0.803	0.880	-0.91	0.361	(-2.527	0.921)
	Non-CSC	Insulation	-0.436	0.330	-1.32	0.186	(-1.083	0.210)
		Heating	-0.073	1.130	-0.06	0.948	(-2.287	2.141)
		Insulation and heating	-0.002	0.839	0	0.998	(-1.646	1.642)

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

Wald tests for the equality of CSC and non-CSC treatment coefficients within each of the three models consistently found no evidence to reject the null hypothesis that the coefficients were equal, and thus no evidence that CSC status has a mediating effect on the

impact of insulation. It is possible that this result reflects limited statistical power to reject the null hypothesis.

The one exception was the comparison of the coefficients for heating in the circulatory hospitalisation cost model. In this case, unexpectedly, heating is associated with a highly statistically significant reduction in costs of $6.31 for non-CSC households (p=0.001), but a non-statistically significant *increase* in costs for CSC households of $1.66. A Wald test for the equality of the coefficients found that the probability of such a difference resulting from chance was (p =0.0263). This result is inconsistent with the other findings in this table and in the previous chapter showing no evidence of an effect of CSC status. The result is also inconsistent with previous research indicating that colder households will take a greater proportion of energy efficiency gains as higher temperatures (Milne and Boardman 2000); as lower-income households are more likely to experience cold indoor temperatures prior to treatment we would predict that they would experience greater temperature improvements and thus gain greater health benefits from a heating retrofit. It is unclear what to make of this apparent anomaly. Speculatively it may reflect the possibility that higher income households are more able to change their heating behaviour in response to receiving their heating retrofit and thus gain great health benefits than CSC households, however this pattern is not found in the comparison of heating coefficients by CSC status in the all-cause and respiratory cost difference models. It is also likely that inability to match on CSC status when selecting matched control households biased results towards a null finding as discussed in Subsection 4.2.6.4.

Comparison of insulation and 'insulation and heating' variable coefficients by treatment type and CSC status found no evidence the reject the null hypothesis of equality, suggesting that it would be legitimate to aggregate these two groups.

5.2.9. Summary

Analysis of household level differences in hospitalisation costs using the original difference-in-difference model suggests that receiving insulation, heating or insulation and heating under WUNZ:HS does not have a statistically significant impact on all-cause hospitalisation costs. However when insulation and heating are combined into a single dummy variable, the variable has a coefficient of -$3.15 ($p=0.098$) which is statistically significant at the ($p<0.1$) level. There are small, but statistically significant reductions in circulatory hospitalisation costs of $1.40 per month as a result of receiving insulation ($p=0.013$) and $3.38 for both insulation and heating ($p=0.011$). There is a small reduction in monthly respiratory hospitalisation costs of $0.50 as a result of receiving insulation which is significant at the ($p<0.1$) level. It is plausible that the causal mechanisms underlying these cost reductions included increased indoor temperatures and reduced dampness and mould leading to fewer or less severe (costly) respiratory or circulatory hospitalisations. These connections between insulation, heating and health were summarised in Subsection 2.1.4.

Model results consistently showed a beneficial impact of insulation and typically no impact of heating. When the coefficients of insulation and 'insulation and heating' variables were compared no evidence was found to reject the hypothesis that coefficients were equal, suggesting that there is a protective effect of insulation, which is gained by households that received insulation or both insulation and heating. It is unclear why heating alone would not show a protective effect given that presumed primary causal mechanisms by which insulation effects health: increased indoor temperatures and reduced dampness, are also expected outcomes of improved heating. In Subsection 4.2.6.5 I speculated that because households that received heating had to have pre-existing insulation that met WUNZ:HS standards, a greater proportion of energy efficiency gains from improved heating might be taken as energy savings rather than increased warmth due to higher baseline temperatures. Similarly the health gains from placing insulation in an uninsulated home are likely to be greater than those for heating because an uninsulated home will on average have lower pre-treatment indoor temperatures than a home that received heating only (and thus had adequate pre-existing insulation), and an increase in temperature from a lower baseline will have greater health benefit.

Stratification of models did not produce much evidence to reject the hypothesis that CSC status does not influence the impact of receiving a WUNZ:HS retrofit, although in one case (heating in the circulatory hospitalisation cost difference model) an impact of CSC status was found. While this general finding is not consistent with predictions based on previous research it may in part simply reflect limited statistical power.

5.3. Household level analysis: Extended model

One of the limitations of the original model analysed above is that it does not allow us to assess the impact of changes in average temperature over the course of the study. It is likely that any effect of treatment will be greater during colder months given the evidence linking colder indoor temperatures with negative health outcomes described in Subsection 2.1.4. It is thus plausible that the inclusion of temperature or time based variables may result in better fitting models of hospitalisation cost difference. In addition, the extended model allows us to more accurately characterise treatment status by month; where previous analyses have necessarily simplified the timing of treatments, where in reality there were sometimes multiple month gaps between the initial insulation retrofit and a subsequent heating retrofit.

The extended model uses the same dataset as the original model, but does not aggregate monthly cost difference records by *before* and *after* period. Initial assessment of the available data showed that circulatory and respiratory hospitalisation cost data were too sparse when analysed at a monthly level for the removal of extreme values used in the original model to be possible (the value of both the 99th and 95th percentiles of the circulatory and respiratory monthly records was 0). For reasons of comparability with the original model the extended model thus focuses on total hospitalisation costs only.

5.3.1. Model specification

The extended model is a difference-in-difference model. The model compares the difference between each treatment group household's average monthly hospitalisation costs and the mean of its matched control group households' monthly hospitalisation costs. The model is a fixed effects OLS estimator with robust standard errors clustered by treatment/control household pairing.

$$Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it+} \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 x_{4it...} + \beta_p x_{pit} + \varepsilon_{it}$$

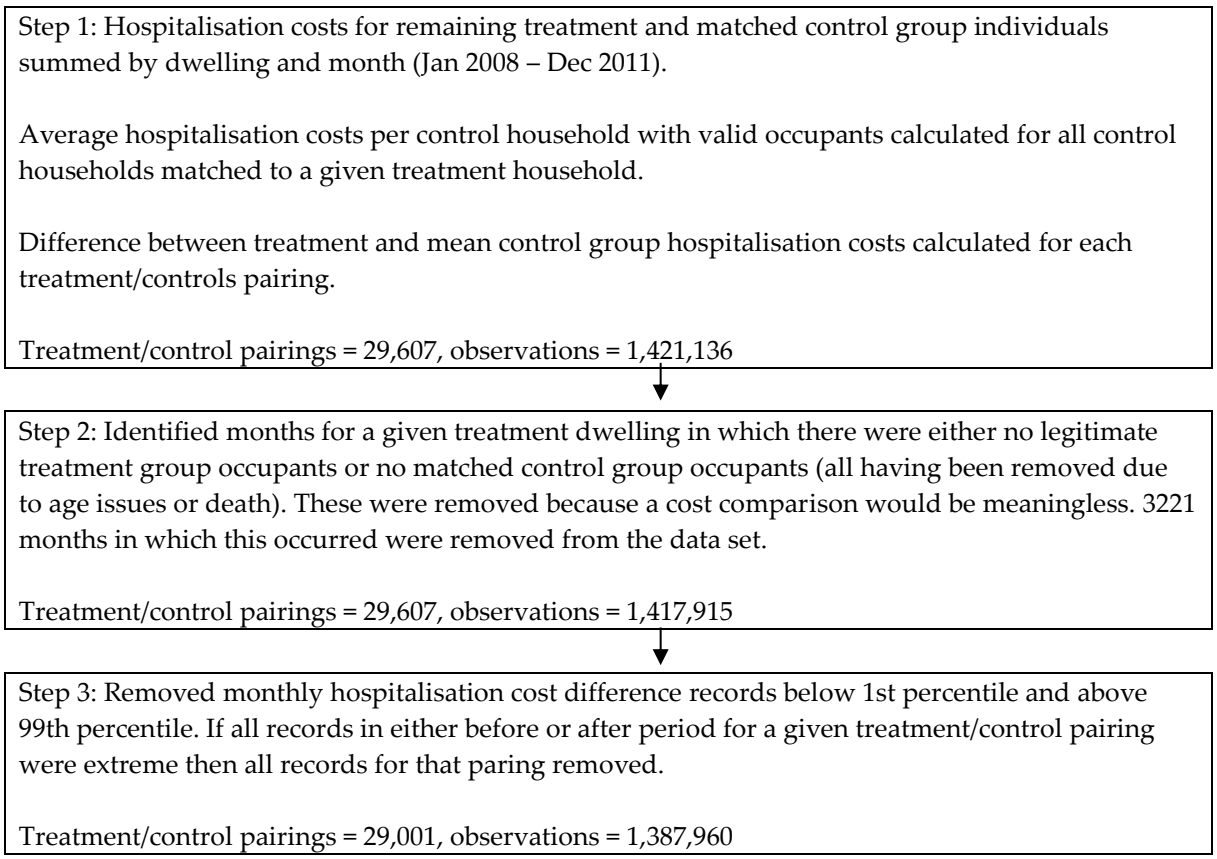
$Hospitalisation\ Cost\ Diff_{it}$ represents the difference between the average monthly hospitalisation costs of treatment house$_i$ and its control group houses for a given year-month t (there are 48 sequential monthly records covering the period Jan 2008 – Dec 2011). The other model covariates are defined as per the original model (see Subsection 5.2.1).

.

5.3.2. Model construction

Model construction follows the process described in Subsection 5.2.2, excluding the final stage which is the aggregation of monthly difference-in-cost data into before and after periods. Rather than exclude all records for treatment/control pairings that have a record that is below the 1st or above the 99th percentiles (necessary when there are only two records for each pairing: before and after), individual months that were below the 1st or above the 99th percentile were excluded, following the approach taken in Grimes, Young et al. (2011). Where there were pairings which had values below the 1st or above the 99th percentiles for either all of the before period or all of the after period these pairings were excluded entirely.

Figure 12: Dataset construction, extended hospitalisation costs model



5.3.3. Confounding

The extended model is an opportunity to consider both the time-variant occupant characteristics analysed in Subsection 5.2.3.1 and climate -based covariates.

5.3.3.1. Occupant characteristics

The covariates described in Subsection 5.2.3.1, *occupant_diff* and *age_diff* are again plausible candidates for inclusion in the extended model. Table 46 presents simple regression analyses for these confounders. The equations for these simple regressions are

1) $Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 occupant_diff_{it} + \varepsilon_{it}$

2) $Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 age_diff_{it} + \varepsilon_{it}$

Table 46: Simple regression analyses, occupant-diff and age_diff variables, extended model, months with values below 1st and above 99th percentiles excluded

Covariate	Coefficient	Robust standard error	t	P >|t|	Confidence intervals	
Occupant_diff	33.255	1.720	19.34	0.000**	(29.883	36.626)
Age_diff	-0.381	.0947	-4.03	0.000**	(-0.567	-0.196)

As with the original model these covariates are highly statistically significant.

5.3.3.2. Season and location

The primary measures of season and location available for inclusion in the extended hospitalisation costs difference model are monthly average temperature by region (measured in ° C) and a winter vs. non-winter dichotomous variable. The logic behind the consideration of these covariates is that there may be a systematic difference in the relative hospitalisation costs of treatment and matched control households over time that is driven by changes in external temperature, irrespective of treatment status in a given month. If such a relationship existed and was not accounted for by the inclusion of suitable covariates then trends in temperature over the course of the study, or the proportion of study time that took place during winter prior to treatment and post-treatment could confound the estimation of effects associated with the exposures of interest. Table 47 presents simple regression analyses for the two most plausible confounders: monthly temperature (measured at the region level and described in Subsection 3.3.7) and a winter vs. non-winter dichotomous variable.

Table 47: Simple regression analyses: winter and monthly average temperature, extended model, months with values below 1st and above 99th percentiles excluded

Covariate	Coefficient	Robust standard error	t	P >|t|	Confidence intervals	
Monthly average temperature	0.389	0.076	5.13	0.000**	(0.240	0.538)
Winter	-2.449	.545	-4.49	0.000**	(-3.519	-1.379)

The monthly average temperature and winter covariates are both highly statistically significant although the interpretation of the coefficients is difficult. It appears that as temperature increases by one degree there is a greater difference between treatment household costs and matched control household costs. Consistent with this, the difference in costs is lower in winter than in the non-winter period. These results may reflect some systematic difference in the characteristics of treatment individuals and their households relative to matched controls or they may reflect the effect of treatment in the after period of the Study (if there are differences in the proportion of winter and non-winter time in the before and after periods), in which case these effects may be negligible in the multiple regression model.

In addition to the inclusion of winter and monthly mean temperature variables it was also be useful to consider interactions between treatment dummies and winter, as it is predicted that the benefit of such variables will be gained primarily in winter. These interaction terms can only be considered in the context of the multiple regression model.

5.3.4. Multiple regression model

The initial multiple regression model can be considered the final model, given the concerns about variable transformation expressed in the Subsection 5.2.5, which apply equally here, and the absence of any plausible interaction terms to consider. Table 48 presents the results of the model, where $monthly_temp_{it}$ represent the mean monthly temperature for the region that household *i* occupies during month *t*:

$$Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 occupant_diff_{it} + \beta_5 age_diff_{it} + \beta_6 winter_t + \beta_7 mean_monthly_temp_{it} \varepsilon_{it}$$

Table 48: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	-3.186	0.694	-4.59	0.000***	(-4.545	-1.826)	R-sq: within = 0.0006	F(7,29000) = 60.61
	Heating only	-0.186	2.236	-0.08	0.934	(-4.569	4.197)	R-sq: between = 0.0165	Prob > F = 0.0000
	Insulation and heating	-2.963	1.511	-1.96	0.050*	(-5.926	-0.001)	R-sq: overall = 0.0012	

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter and monthly average temperature

The multiple regression model suggest that there is a highly statistically significant reduction in average all-cause hospitalisation costs as a result of having a WUNZ:HS insulation retrofit in a given month of $3.19 per month. There is no statistically significant impact of receiving heating alone under the programme but there is a statistically significant (p=0.050) reduction in all-cause hospitalisation costs of $2.96 as a result of having both insulation and heating in a given month. Here it seems natural to consider a Wald test for the equality of the insulation and 'insulation and heating' coefficients. The test produced an F score of 0.02 (F(1,29000)), Prob>F= 0.8269 indicating that there is no evidence to reject the null hypothesis that the two coefficients are equal. This result suggests that there is an impact of insulation retrofits, but no evidence of an impact of heating on hospitalisation costs, whether installed separately or combined with an insulation retrofit. The greater statistical significance of the coefficient of the insulation only treatment dummy variable likely reflects the greater number of treatment/control household pairings where the treatment household received insulation only.

Consideration of interaction terms is the next task. Here the following equation is evaluated:

$$\begin{aligned} \textit{Hospitalisation Cost Diff}_{it} = {} & \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 \\ & occupant_diff_{it} + \beta_5 age_diff_{it} + \beta_6 winter_t + \beta_7 mean_monthly_temp_{it} + \\ & + \beta_8 insulation_{it} * winter_t + \beta_9 heating_{it} * winter_t + \beta_{10} heating_{it} * winter_t \\ & \beta \varepsilon_{it} \end{aligned}$$

Table 49: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded, winter interaction terms added

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	-3.502	0.791	-4.43	0.000***	(-5.053	-1.953)	R-sq: within = 0.0006 R-sq: between = 0.0165 R-sq: overall = 0.0012	F(7,29000) = 42.52 Prob > F = 0.0000
	Heating only	-0.018	2.461	-0.01	0.994	(-4.841	4.805)		
	Insulation and heating	-3.739	1.651	-2.26	0.024**	(-6.974	-0.503)		
	Insulation only*winter	0.943	1.186	0.79	0.427	(-1.381	3.263)		
	Heating only*winter	-0.473	3.095	-0.15	0.879	(-6.540	5.594)		
	Insulation and heating*winter	2.275	2.067	1.10	0.271	(-1.776	6.327)		

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter,monthly average temperature, interaction terms for treatment and winter

Table 49 suggests that there is not a strong case for the inclusion of interaction terms, with none of the reported p-values for the interaction terms close to statistical significance, and a lower F-statistic for the model reported (40.52 vs. 60.56). Alternate specifications, including models that excluded the mean average temperature variable similarly did not reveal any statistically significant interaction between treatment status and winter. Given these concerns, my preference here is to retain the model presented in Table 48, which has the additional advantage of ease of interpretation.

More generally, it is concerning that the interaction terms reported here do not appear to indicate a season based variation in the effect of the insulation and insulation and heating variables. This is problematic from a modelling perspective and is an important caveat to note when assessing the success of the modelling process, perhaps pointing towards selection bias. Concerns about selection bias are explored further in Section 9.1.4.

The full regression output of the models presented in Table 48 and Table 49 are presented in Appendix Two in Table 87 and Table 88 respectively.

5.3.5. Assessing the impact of extreme value removal

Here I follow the process carried out in Subsection 5.2.5, looking at the impact of the approach to extreme values. Table 50 presents the results of running the multiple regression model described above with no removal of extreme values. It was not possible to run the model with values below the 5th and above the 95th percentiles excluded as both the 95th and 90th percentile values of all-cause hospitalisation cost difference were 0, meaning it was not possible to differentiate records in the 90th and 95th percentiles.

Table 50: All-cause hospitalisation cost difference, extended model, no exclusions

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	2.601	4.349	0.60	0.550	(-5.923	11.126)	R-sq: within = 0.0008	F(7,29606) = 60.55
	Heating only	-6.277	13.561	-0.46	0.643	(-32.858	20.304)	R-sq: between = 0.0032	Prob > F = 0.0000
	Insulation and heating	24.503	10.179	2.41	0.016 **	(4.550	44.456)	R-sq: overall = 0.0002	

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter and monthly average temperature

As with the original model, Table 50 demonstrates the importance of exclusion strategy in the extended model. Insulation and heating do not have a statistically significant impact in the no-exclusion model while 'insulation and heating' has a highly statistically significant adverse impact on cost difference (the model results suggest that receiving an insulation and heating retrofit under WUNZ:HS increases monthly all-cause hospitalisation costs at the household level!). These results are clearly inconsistent with the model with exclusions, and seem unlikely to reflect a real effect of treatment.

5.3.6. Model specification and global goodness-of-fit measures

As with the original model presented in Section 5.2 it is useful to consider an alternative approach to the calculation of confidence intervals using bootstrapped standard errors to provide more robust inference in the face of possible model misspecification. Table 51 presents the original multiple regression model with robust standard errors clustered at the treatment/control pairing level and the model with bootstrapped standard errors.

Table 51: All-cause hospitalisation cost difference extended model with robust clustered standard errors and bootstrapped standard errors, months with values below 1st and above 99th percentiles excluded

Model	Variable name	Coefficient	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Model with robust clustered standard errors	Insulation only	-3.186	0.694	-4.59	0.000***	(-4.545	-1.826)
	Heating only	-0.186	2.236	-0.08	0.934	(-4.569	4.197)
	Insulation and heating	-2.964	1.511	-1.96	0.050*	(-5.926	-0.001)
Model with bootstrapped standard errors	Insulation only	-3.186	0.717	-4.44	0.000***	(-4.592	-1.780)
	Heating only	-0.186	2.244	-0.08	0.934	(-4.584	4.213)
	Insulation and heating	-2.964	1.498	-1.98	0.048**	(-5.900	-0.027)

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff, age_diff, winter and monthly average temperature

The high degree of similarity between the standard errors and confidence intervals in Table 51 gives additional reassurance that misspecification of the model is not producing unreliable estimates of model -based standard errors in this case.

Turning to measures of global goodness-of-fit, the extended model has a low overall R^2 score of 0.0012, as with the original model. This suggests that the extended model can only explain a small proportion of the variation in the dependent variable, which, following the argument in Subsection 5.2.7 may reflect the success of the matching process, or may simply reflect the weakness of the effect. Finally, it is of interest to stratify results by the Community Services Card status of the treatment households.

5.3.7. Stratification by Community Services Card Status

The equation for the stratified model is:

$$\begin{aligned} Hospitalisation\ Cost\ Diff_{it} = {} & \alpha_i + \beta_1 insulation_{it} * CSC_i + \beta_2 insulation_{it} * (1\text{-}CSC_i) + \beta_3 heating_{it} * CSC_i \\ & + \beta_4 heating_{it} * (1\text{-}CSC_i) + \beta_5 heating_and_insulation_{it} * CSC_i + \\ & \beta_6 heating_and_insulation_{it} * (1\text{-}CSC_i) + \beta_7 occupant_diff_{it} + \\ & \beta_8 age_diff_{it} + \\ & \beta_9 average_monthly_temperature_{it} + \beta_{10} winter_t + \varepsilon_{it} \end{aligned}$$

The results of stratification of the extended model by the Community Services Card status of the treatment household in each treatment/control pairing are presented in Table 52.

Table 52: All-cause hospitalisation cost difference, extended model, treatment stratified by Community Services Card status, months with values below 1st and above 99th percentiles excluded

Dependent variable modelled	CSC status of treatment household	Variable name	Coefficient	Robust Standard error	Z	P>|Z|	95% Confidence intervals	
Difference in all-cause hospitalisation costs	CSC	Insulation	-3.747	1.015	-3.69	0.000***	-(5.737	-1.758)
		Heating	-2.250	3.365	-0.67	0.504	-(8.846	4.346)
		Insulation and heating	-2.224	2.102	-1.06	0.29	-(6.344	1.896)
	Non-CSC	Insulation	-2.503	0.912	-2.75	0.006***	-(4.290	-0.716)
		Heating	1.881	2.936	0.64	0.522	-(3.874	7.636)
		Insulation and heating	-4.090	2.072	-1.97	0.048**	-(8.152	-0.029)

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter and monthly average temperature

Wald test results showed that there was no evidence to reject the hypotheses that the coefficients of CSC and equivalent non-CSC variables were identical: insulation (F(1,29,000)=0.83, prob > F = 0.3609), heating (F(1,29,000)=0.86, prob > F=0.3551) and 'insulation and heating' (F(1,29,000)=0.40, prob > F = 5267).

There was also no evidence to reject the hypotheses that the coefficients of 'insulation' and 'insulation and heating' were equal for either CSC (F(1,29,000)= 0.43, prob > F = 0.5116) or non-CSC (F(1,29,000)= 0.50, prob > F = 0.4810) treatment/control pairings. The limitations of the CSC variable have been discussed previously, and limit the strength of this conclusion.

5.3.8. Summary

The extend model of all-cause hospitalisation cost differences found that there were statistically significant savings as a result of receiving insulation and 'insulation and heating'. Wald test results showed there was no evidence to reject the null hypothesis that the two coefficients were equal. There was no evidence of any impact of the CSC status of treatment group households on this outcome.

Comparing the original all-cause hospitalisation model and the extended model, results are very consistent, with a monthly saving as a result of insulation in the original model of $3.33 (p=0.107) and of $3.19 in the extended model (p<0.001). Given the greater statistical power of the extended model the difference in p-values is easily explained. The extended model can be interpreted as confirming that the result found in the original model was not the result of unmeasured confounding by temperature or season and in general

strengthens the conclusion that there is a real effect of insulation on all-cause hospitalisation costs.

5.4. Chapter summary

Analysis of hospitalisation costs in this chapter has been limited by a fundamental problem, the problematic nature of hospitalisation cost data, which frustrated attempts to model costs at an individual level. Modelling at the treatment/control household pairing level was more suitable in this context.

Taking all-cause hospitalisation costs to be the primary dependent variable of interest (the assumption made in the original WUNZ:HS evaluations), the original model found no statistically significant effect of treatment ($p<0.1$) for either insulation, heating or insulation and heating. When insulation was combined with 'insulation and heating' into a single dummy variable the result was a saving of $3.15 per month ($p=0.098$). The extended model, which accounted for changes in temperature and season found a highly statistically significant ($p<0.001$) reduction in all-cause hospitalisation costs of $3.18 per household per month, a similar effect size but highly statistically significant, reflecting the greater power of the extended model.

Modelling of differences in circulatory hospitalisation costs found that there was a small statistically significant reduction in circulatory hospitalisation costs of $1.40 per month per treated household as a result of receiving insulation, and $3.38 as a result of receiving both insulation and heating. Modelling of the difference in respiratory hospitalisation costs found a reduction in costs of $0.50 per insulated household per month that was statistically significant at the ($p<0.1$) level. These results can be interpreted as suggesting that reduction in all-cause hospitalisation cost differences are driven by changes in these two plausible cost difference categories.

All household level cost difference models were sensitive to the strategy used to exclude extreme values, meaning that it was important to make the impact of these choices explicit. The exclusion of records below the 1st or above the 99th percentiles appeared reasonable a priori. Having assessed hospitalisation costs it is now time to turn the final health cost category analysed in this thesis, pharmaceutical costs.

CHAPTER 6. MODELLING CHANGES IN PHARMACEUTICAL USE COSTS

As with hospitalisation cost data, analysis of individual level pharmaceutical cost data using the WUNZ:HS evaluation dataset is problematic given issues of extreme kurtosis, skew etc. and so meaningful individual level analysis is considered impractical. Instead, Chapter Six follows the approach presented in Chapter Five, first presenting a slightly modified variation of the original household level pharmaceutical use cost model used in Telfar Barnard, Preval et al. (2011) and then an extended version of the model, which can better categorise the exposures of interest and account for changing season and temperatures.

It is important to note that the pharmaceutical cost data available from the Pharms dataset do not include hospital-dispensed pharmaceuticals, as these are not claimed for using the process that generates the Pharms dataset. Hospital-dispensed pharmaceutical costs are captured by the cost weight code for a hospitalisation event described in Subsection 3.2.2.2.

Section 6.1 presents the modified original model. The model is described and the model building process is carried out. Section 6.2 presents the extended version of the model. In each case the model is described, the model building process is carried out and results presented.

6.1. Original model

The original model used to evaluate pharmaceutical use costs at the household level is essentially identical to the original hospitalisation costs model described in Section 5.2. One difference of interest between the two datasets is that there is only pharmaceutical use data available for the period Jan 2008 – Dec 2010 as an additional year of pharmaceutical data was not sought for the MED funded extension (Preval 2012). This decision was made because the results reported in the original WUNZ:HS health evaluation (Telfar Barnard, Preval et al. 2011) were judged to be of limited policy significance. The pharmaceutical use data is further limited due to concerns noted in the original report about NHI data quality for the period October 2010- Dec 2010, which demonstrated an implausible reduction in hospitalisations for both treatment and control group individuals and was thought to reflect a reporting lag due to the recent nature of the data, thus these months are excluded from analysis (Telfar Barnard, Preval et al. 2011).

6.1.1. Model specification

The original model is a difference-in-difference model essentially identical to the hospitalisation cost difference model. The model compares the difference between each treatment group household's average monthly pharmaceutical use costs and the mean of its matched control group households' monthly pharmaceutical use costs before and after the intervention. The model is a fixed effects OLS estimator with robust standard errors clustered by treatment/control household pairing.

$$Pharms\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 x_{4it\ldots} + \beta_p x_{pit} + \varepsilon_{it}$$

Pharms Cost Diff$_{it}$ represents the difference between the average monthly pharmaceutical use costs of treatment house$_i$ and its control group houses over time t (t is either before or after). The three pharmaceutical use cost categories considered are all cause pharmaceuticals, circulatory illness-related pharmaceuticals and respiratory illness-related pharmaceuticals. Further information about the products that make up these categories is available in Subsection 3.3.3.2.

The covariates α_i, *insulation*$_{it}$, heating$_{it}$, *heating_and_insulation*$_{it}$, $\beta_4 x_{4it\ldots} + \beta_p x_{pit}$, and the residual term ε_{it} are defined identically to those in the original hospitalisation cost difference model (see Subsection 5.2.1).

6.1.2. Model construction and approach to extreme values

The model construction process for the pharmaceutical cost difference models was largely similar to the model construction process for the hospitalisation cost difference models and it is not necessary to present a detailed description of the process here. Having constructed the dataset the approach to the removal of extreme values was the same as in the hospitalisation cost difference model: for each model, all records for treatment/control pairings that had before or after period cost difference records that were below the 1st or above the 99th percentiles were dropped. The results of this removal on measures of skew and kurtosis are set out in Table 53.

Table 53: Summary statistics, dependent pharmaceutical use cost difference variables, with and without exclusions

Dependent variable	Inclusion rule	Mean	Std. deviation	Skewness	Kurtosis
All cause pharmaceutical use cost difference	All pairings	10.589	167.594	10.630	255.868
	Exclude pairings with records below1st or above 99th percentiles	3.871	79.711	1.687	8.701
Circulatory illness-related pharmaceutical use cost difference	All pairings	2.950	29.809	6.590	234.472
	Exclude pairings with records below1st or above 99th percentiles	1.826	20.968	1.944	8.028
Respiratory illness-related pharmaceutical use cost difference	All pairings	1.080	18.523	4.181	39.550
	Exclude pairings with records below1st or above 99th percentiles	0.131	11.913	3.209	17.779

Table 53 suggests that, as with the hospitalisation cost difference dependent variables, exclusions reduce the standard deviation, skew and kurtosis of all three dependent variables, although the degree of reduction is much greater for the all-cause pharmaceuticals than for the other two variables. The pattern of reduced mean values is also repeated here, likely reflecting the averaging process whereby the dependent variables were created.

In Subsection 6.2.5 the impact of removal is explored in depth with models being rerun both without exclusions and with removal of treatment/control pairings that had before or after period cost difference records that were below the 5th or above the 95th percentiles.

The dataset construction and cleaning process is set out in Figure 13.

Figure 13: Dataset construction, original pharmaceutical cost difference models

Step 1: Pharmaceutical use costs for remaining treatment and matched control group individuals summed by dwelling and month (Jan 2008 – Sept 2010).

Average pharmaceutical use costs per control household with valid occupants calculated for all control households matched to a given treatment household.

Difference between treatment and mean control group pharmaceutical use costs calculated for each treatment/controls pairing.

Treatment/control pairings = 29,607, observations = 977,031

↓

Step 2: Identified months for a given treatment dwelling in which there were either no legitimate treatment group occupants or no matched control group occupants (all having been removed due to age issues or death). These were removed because a cost comparison would be meaningless. 1265 months in which this occurred were removed from the data set.

Treatment/control pairings = 29,606, observations = 975,766

↓

Step 3: Further aggregated (averaged) cost difference and data for before and after period, having dropped the relevant month in which the intervention occurred for each dwelling cluster. Dropped both before and after records if either period entirely missing due to previous removals).

Treatment/control pairings = 29,517, observations = 59,304

↓

Step 4: For each hospitalisation cost difference category (all-cause, circulatory and respiratory) separately identified records below the 1st percentile or above the 99th percentile. Both before and after records for a given treatment/control pairing were labelled extreme and excluded from analysis if one of the records was identified as extreme.

Treatment/control pairings, all cause pharmaceutical use cost model
= 28,630, observations = 57,260
Treatment/control pairings, circulatory illness-related pharmaceutical use cost model
= 28,699, observations = 57,398
Treatment/control pairings, respiratory illness-related pharmaceutical use model
= 28,697, observations = 57,394

6.1.3. Confounding

Here the approach to modelling pharmaceutical cost differences follows the approach used in the hospitalisation cost difference models. As with those models two variables are considered occupant_diff and age_diff.

6.1.3.1. Occupant characteristics

The justification for considering the age-diff and occupant_diff covariates is the same as in the hospitalisation cost difference model: these variables can change over the course of the study as people are born and die and could explain a great deal of variation in pharmaceutical cost differences for a given treatment/control pairing over time.

Table 54 presents separate simple regression analyses for these confounders.

Table 54: Simple regression analyses, occupant_diff and age_diff variables, treatment/control pairings with values below 1st and above 99th percentiles excluded

Covariate	Dependent variable	Coefficient	Robust standard error	t	P>|t|	Confidence intervals	
Occupant_diff	All-cause pharmaceutical cost diff	26.111	1.553	16.75	0.000***	(23.057	29.167)
	Circulatory pharmaceutical cost diff	3.717	0.315	11.80	0.000***	(3.099	4.334)
	Respiratory pharmaceutical cost diff	1.073	0.184	5.83	0.000***	(0.712	1.435)
Age_diff	All-cause pharmaceutical cost diff	1.171	0.194	6.05	0.000***	(0.791	1.550)
	Circulatory pharmaceutical cost diff	0.313	0.036	8.68	0.000***	0.242	0.384
	Respiratory pharmaceutical cost diff	0.123	0.026	4.80	0.000***	0.073	0.173

* p<0.1, ** p<0.05, *** p<0.01

6.1.4. Initial multiple regression model

The initial multiple regression model specification is:

$$Pharmaceutical\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 occupant_diff_{it} + \beta_5 age_diff_{it} + \varepsilon_{it}$$

The results of the multiple regression models are presented in Table 55 for each pharmaceutical use type. F-statistics and R^2 model fit measures are presented.

Table 55: Initial multiple regression models, original pharmaceutical cost difference models, , treatment/control pairings with values below 1st and above 99th percentiles excluded

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P<|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all pharmaceutical use costs	Insulation only	-1.879	0.349	-5.38	0.000***	(-2.564	-1.195)	R-sq: within = 0.0427 R-sq: between = 0.0589 R-sq: overall = 0.0546	F(5,28629) = 75.59 Prob > F = 0.0000
	Heating only	-4.156	1.054	-3.94	0.000***	(-6.222	-2.090)		
	Insulation and heating	-0.878	0.768	-1.14	0.253	(-2.385	0.627)		
Difference in circulatory pharmaceutical use costs	Insulation only	-0.502	0.075	-6.67	0.000***	(-0.650	-0.355)	R-sq: within = 0.0292 R-sq: between = 0.0907 R-sq: overall = 0.0858	F(5,28698) = 45.42 Prob > F = 0.0000
	Heating only	-0.736	0.227	-3.24	0.001***	(-1.183	-0.291)		
	Insulation and heating	-0.594	0.159	-3.72	0.000***	(-0.908	-0.281)		
Difference in respiratory pharmaceutical use costs	Insulation only	0.031	.0485087	0.64	0.520	(-0.064	0.126)	R-sq: within = 0.0072 R-sq: between = 0.0103 R-sq: overall = 0.0096	F(5,28696) = 13.93 Prob > F = 0.0000
	Heating only	0.039	.1405254	0.28	0.778	(-0.236	0.315)		
	Insulation and heating	0.056	.1027661	0.55	0.584	(-0.145	0.258)		

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

Table 55 shows that there are highly statistically significant reductions in all pharmaceutical use costs per month as a result of receiving insulation ($1.88, p<0.000) and heating only ($4.16, p<0.000). These results are plausible, although it is surprising that the effect of receiving both insulation and heating appears negligible, which would not have been predicted given the other results. Wald tests for the equality of the insulation and the 'insulation and heating' coefficients suggested that there was not enough evidence to reject the null hypothesis of equality (F(1,28629)=1.45, prob > F= 0.2280) but that there is evidence to reject the null hypothesis when comparing the coefficients for heating and 'heating and insulation' (F(1,28,629)=6.41, prob > F= 0.0114). It is difficult to reconcile these results, unless we conclude that there are differences in the treatment/control pairings that make up the three groups which mediate the impact of treatment. These results are not consistent with the majority of previous results, which found a consistent effect of insulation for the insulation and 'insulation and heating' pairings and no detectable effect of heating. It is possible that these results may reflect different causal pathways by which heating affects pharmaceutical use costs.

Interestingly, the inclusion of the occupant_diff and age_diff terms appears to have greatly modified results presented in Telfar Barnard, Preval et al. (2011), suggesting that there was confounding present in the model used in that report. In Telfar Barnard, Preval et al. (2011) insulation was found to reduce all-cause pharmaceutical use costs by $0.92 while heating did not have a statistically significant impact.

There are highly statistically significant ($p<0.01$) but small reductions in monthly circulatory illness-related pharmaceutical use as a result of all three treatments. The cost reductions were $0.50 (insulation), $0.74 (heating only) and $0.59(insulation and heating) per month. Wald tests do not provide evidence to reject the null hypothesis of equality when comparing any of these coefficients.

There are no statistically significant changes in monthly respiratory illness-related pharmaceutical use cost differences as a result of receiving any treatment under WUNZ:HS.

It is interesting to speculate about what might be driving the much larger savings observed in the all-cause pharmaceutical use model given that the savings are greater than the sum of circulatory-illness and respiratory illness-related pharmaceutical use cost savings. It is possible that the circulatory and respiratory illness-related variables are misspecified and do not contain some important pharmaceutical category on which treatment has an impact, or that some other non-circulatory and non-respiratory health benefit is occurring. This possibility was not further explored at this time.

6.1.5. Transformation of continuous variables and interaction terms

Informed by the experience of exploring the transformation of continuous variables using factional polynomials in the hospitalisation cost difference models, transformation was not considered here. There were also no interaction terms of interest to consider.

6.1.6. Assessing the impact of extreme value removal

In this subsection the impact of potential outliers is considered. In Table 56 I run models both without any exclusions and then with the removal of all values that are below the 5th or above the 95th percentiles.

Table 56: Cost difference models, no exclusions and , treatment/control pairings with values below the 5th or above the 95th percentiles excluded

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all pharmaceutical use costs	Insulation only	-1.121	0.962	-1.17	0.244	-(3.007	0.764)	R-sq: within = 0.0091 R-sq: between = 0.0182 R-sq: overall = 0.0157	F(5,29516) = 24.72 Prob > F = 0.0000
	Heating only	-6.899	2.166	-3.18	0.001***	-(11.145	-2.653)		
	Insulation and heating	-0.437	1.790	-0.24	0.807	-(3.946	3.072)		
Difference in all pharmaceutical use costs (records below 5th or above 95th percentiles removed)	Insulation only	-1.646	0.230	-7.14	0.000***	-(2.097	-1.194)	R-sq: within = 0.0445 R-sq: between = 0.0688 R-sq: overall = 0.0634	F(5,25429) = 97.43 Prob > F = 0.0000
	Heating only	-1.564	0.717	-2.18	0.029**	-(2.969	-0.159)		
	Insulation and heating	-1.566	0.519	-3.02	0.003***	-(2.583	-0.549)		
Difference in circulatory illness-related pharmaceutical use costs	Insulation only	-0.369	0.128	-2.88	0.004***	-(0.619	-0.118)	R-sq: within = 0.0166 R-sq: between = 0.0732 R-sq: overall = 0.0675	F(5,29516) = 35.46 Prob > F = 0.0000
	Heating only	-1.060	0.362	-2.93	0.003***	-(1.769	-0.352)		
	Insulation and heating	-0.215	0.211	-1.02	0.309	-(0.628	0.199)		
Difference in circulatory illness-related pharmaceutical use costs (records below 5th or above 95th percentiles removed)	Insulation only	-0.621	0.058	-10.78	0.000***	-(0.734	-0.508)	R-sq: within = 0.0312 R-sq: between = 0.0815 R-sq: overall = 0.0760	F(5,25635) = 58.58 Prob > F = 0.0000
	Heating only	-0.599	0.182	-3.3	0.001***	-(0.955	-0.243)		
	Insulation and heating	-0.702	0.126	-5.57	0.000***	-(0.949	-0.455)		
Difference in respiratory illness-related pharmaceutical use costs	Insulation only	0.185	0.068	2.71	0.007***	(0.051	0.319)	R-sq: within = 0.0090 R-sq: between = 0.0141 R-sq: overall = 0.0134	F(5,29516) = 20.23 Prob > F = 0.0000
	Heating only	0.215	0.197	1.09	0.274	-(0.171	0.601)		
	Insulation and heating	0.165	0.134	1.23	0.218	-(0.098	0.427)		
Difference in respiratory illness-related pharmaceutical use costs (records below 5th or above 95th percentiles removed)	Insulation only	-0.194	0.028	-7.02	0.000***	-(0.249	-0.140)	R-sq: within = 0.0069 R-sq: between = 0.0092 R-sq: overall = 0.0084	F(5,25690) = 26.20 Prob > F = 0.0000
	Heating only	-0.342	0.080	-4.3	0.000***	-(0.498	-0.186)		
	Insulation and heating	-0.247	0.060	-4.11	0.000***	-(0.364	-0.129)		

* p<0.1, ** p<0.05, *** p<0.01

All models control for occupant_diff and age_diff

It is difficult to draw any conclusions about the optimal exclusion strategy based on these results. It is worth noting that all three models show statistically significant cost reductions for all three treatment types when values below the 5th or above the 95th percentiles of the cost difference variables are removed. This contributes to the conclusion that there is a real effect of treatment that is revealed as extreme values distorting analysis are removed. Interestingly, the respiratory cost difference model shows an increase in costs of $0.18 (p=0.007) in the model with no exclusions, in contrast to the $0.19 savings per month when

exclusions are based on the 5th and 95th percentiles. This highlights the impact of removal strategy and the need to make this impact explicit and transparent.

6.1.7. Model specification and global goodness-of-fit measures

The concern that the pharmaceutical cost difference models are misspecified was addressed by recalculating confidence intervals using a bootstrap approach with 500 repetitions. The resulting confidence intervals and p-values were consistent with those calculated using robust standard errors clustered at the treatment/control pairing level, consistent with an appropriately specified model. Table 57 presents a comparison of p-values and confidence intervals for the all-cause pharmaceutical use cost difference model, demonstrating this consistency.

Table 57 All-cause pharmaceutical use cost difference model with robust clustered standard errors and bootstrapped standard errors, treatment/control pairings with values below 1st and above 99th percentiles excluded

Model standard errors	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals	
Robust clustered standard errors	Insulation only	-1.879	0.349	-5.38	0.000***	(-2.564	-1.195)
	Heating only	-4.156	1.054	-3.94	0.000***	(-6.222	-2.090)
	Insulation and heating	-0.878	0.768	-1.14	0.253	(-2.385	0.627)
Bootstrapped standard errors	Insulation only	-1.880	0.352	-5.34	0.000***	-(2.569	-1.190)
	Heating only	-4.157	1.018	-4.08	0.000***	-(6.152	-2.161)
	Insulation and heating	-0.879	0.794	-1.11	0.268	-(2.436	0.678)

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

Turning to global goodness-of-fit, the original pharmaceutical cost difference models have low measures of goodness-of-fit in terms of R^2 overall. As argued previously, this may reflect the success of the matching strategy.

6.1.8. Stratification by Community Services Card status

Stratification by Community Services Card status is useful from a policy analysis perspective. Table 58 presents the results of this stratification.

Table 58: Hospitalisation cost models stratified by Community Services Card status, treatment/control pairings with values below 1st and above 99th percentiles excluded

Dependent variable modelled	CSC status of treatment household	Variable name	Coefficient	Robust Standard error	Z	P<|Z|	95% Confidence intervals	
Difference in all cause pharmaceutical use costs	CSC	Insulation only	-2.132	0.497	-4.29	0.000***	-(3.107	-1.158)
		Heating only	-4.094	1.535	-2.67	0.008***	-(7.102	-1.086)
		Insulation and heating	-1.461	1.023	-1.43	0.153	-(3.465	0.543)
	Non-CSC	Insulation only	-1.569	0.454	-3.45	0.001***	-(2.459	-0.678)
		Heating only	-4.221	1.434	-2.94	0.003***	-(7.032	-1.410)
		Insulation and heating	0.002	1.140	0	0.999	-(2.232	2.236)
Difference in circulatory pharmaceutical use costs	CSC	Insulation only	-0.744	0.112	-6.65	0.000***	-(0.963	-0.524)
		Heating only	-0.808	0.372	-2.17	0.03**	-(1.538	-0.078)
		Insulation and heating	-0.803	0.223	-3.6	0.000	-(1.240	-0.365)
	Non-CSC	Insulation only	-0.210	0.090	-2.33	0.020**	-(0.387	-0.033)
		Heating only	-0.666	0.258	-2.58	0.010**	-(1.172	-0.159)
		Insulation and heating	-0.286	0.216	-1.33	0.185	-(0.708	0.137)
Difference in respiratory pharmaceutical use costs	CSC	Insulation only	0.115	0.070	1.64	0.100	-(0.022	0.252)
		Heating only	0.241	0.216	1.12	0.264	-(0.182	0.664)
		Insulation and heating	0.082	0.143	0.58	0.564	-(0.198	0.362)
	Non-CSC	Insulation only	-0.070	0.062	-1.13	0.259	-(0.192	0.052)
		Heating only	-0.163	0.179	-0.91	0.362	-(0.514	0.188)
		Insulation and heating	0.018	0.139	0.13	0.899	-(0.254	0.290)

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

Wald test comparisons of CSC and non-CSC coefficients within each model generally find no evidence to reject the null hypotheses that the coefficients are equal. A notable exception is the comparison of the coefficients of insulation in the circulatory illness-related pharmaceutical use cost difference model by CSC status ($F(1,28,698)=14.45$, prob > F = 0.0001), where the larger coefficient for the CSC treatment/control pairings appears consistent with the hypothesis discussed in previous chapters that lower income individuals would be predicted to gain greater health benefits from improved insulation or heating.. In general, there is limited evidence of CSC status modifying outcomes, and certainly no clear pattern of effect in the two examples where there is some evidence of an impact. The limitations of the CSC variable have been discussed previously, and limit the strength of any conclusion drawn.

6.1.9. Summary

Analysis of household level differences in pharmaceutical use costs using the original difference-in-difference model suggests that receiving insulation *or* heating result in small but statistically significant cost savings in the all cost difference model. These savings are not greatly affected by the Community Services Card status of treatment households.

There are small but statistically significant cost savings as a result of receiving all treatment types on circulatory illness-related pharmaceutical use cost differences. Here there is evidence CSC status does have an impact on the influence of insulation.

Finally, there is little evidence of a beneficial impact of any treatment type on respiratory illness-related pharmaceutical use cost differences. This result is somewhat surprising given the evidence that improved temperatures lead to reductions in respiratory symptoms (Howden-Chapman, Matheson et al. 2007, Howden-Chapman, Pierse et al. 2008), which might be hypothesised to also reduce pharmaceutical use costs.

6.2. Household level analysis: Extended model

6.2.1. Model specification

The extended model is a difference-in-difference model, identical to the extended model of hospitalisation cost differences presented in Section 5.3. The model compares the difference between each treatment group household's average monthly pharmaceutical use costs and the mean of its matched control group households' monthly pharmaceutical use costs. The model is a fixed effects OLS estimator with robust standard errors clustered by treatment/control household pairing.

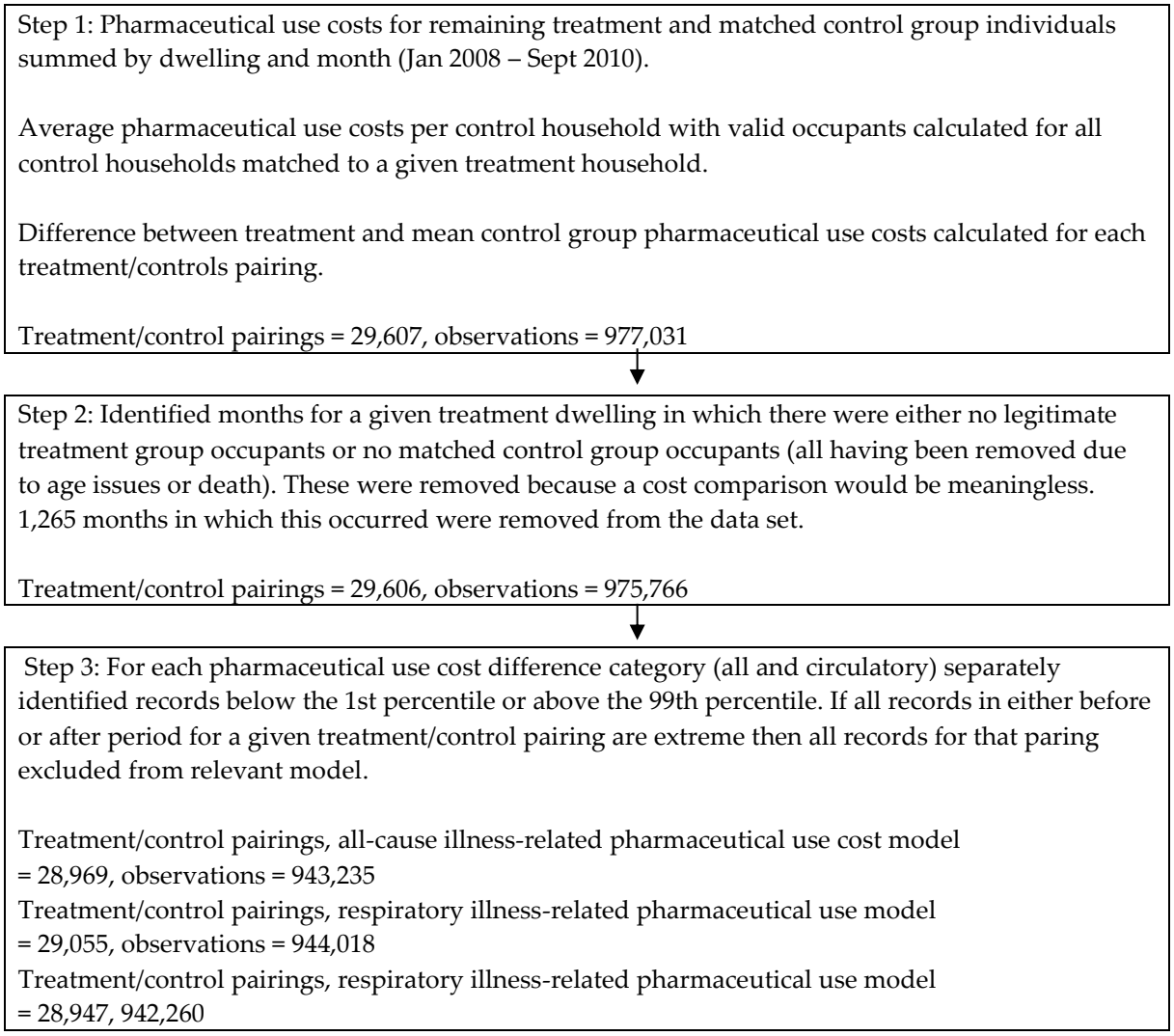
$$Pharm\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 x_{4it} \ldots + \beta_p x_{pit} + \varepsilon_{it}$$

Pharma Cost Diff$_{it}$ represents the difference between the average monthly pharmaceutical use costs (either all or circulatory illness-related only) of treatment house$_i$ and its control group houses for a given month t (there are 33 Study months). The other model covariates are defined identically to those presented in Subsection 5.3.1

6.2.2. Model construction

Model construction follows the process described in Subsection 5.3.2, and is set out in Figure 12:

Figure 9: Dataset construction, extended pharmaceutical use cost difference models



6.2.3. Confounding

The extended model is an opportunity to consider both the time-variant occupant characteristics analysed in Subsection 6.2.3.1 and climate based covariates.

6.2.3.1. Occupant characteristics

The covariates described in Subsection 6.2.3.1, *occupant_diff* and *age_diff* are plausible candidates for inclusion in the extended model. Table 59 presents simple regression analyses for these confounders.

Table 59: Simple regression analyses, occupant-diff and age_diff variables, extended model, months with values below 1st and above 99th percentiles excluded

Dependent variable modelled	Covariate	Coefficient	Robust standard error	t	P >|t|	Confidence intervals	
All pharmaceutical use cost difference	Occupant_diff	23.124	0.789	29.31	0.000***	(21.579	24.671)
	Age_diff	1.225	0.105	11.61	0.000***	(1.0181	1.432)
Circulatory illness-related pharmaceutical use cost difference	Occupant_diff	3.614	0.204	17.69	0.000***	(3.214	4.015)
	Age_diff	0.371	0.029	12.60	0.000***	(0.313	0.429)

As with the original pharmaceutical use cost difference models these covariates are highly statistically significant.

6.2.3.2. Climate and location

The primary measures of climate and location available for inclusion in the extended pharmaceutical use costs difference models are monthly average temperatures by region and a winter vs. non-winter dichotomous variable. Table 60 presents simple regression analyses for the two most plausible confounders: monthly temperature and a winter vs. non-winter dichotomous variable.

Table 60: Simple regression analyses: winter and monthly average temperature, extended mode, months with values below 1st and above 99th percentiles excluded l

Dependent variable modelled	Covariate	Coefficient	Robust standard error	t	P >|t|	Confidence intervals	
All pharmaceutical use cost difference	Monthly average temperature	0.020	0.016	1.29	0.196	(-0.011	0.051)
	Winter	-0.182	0.123	-1.48	0.140	(-0.424	0.060)
Circulatory illness-related pharmaceutical use cost difference	Monthly average temperature	0.007	0.004	1.80	0.072*	(-0.001	0.016)
	Winter	-0.069	0.037	-1.84	0.066*	(-0.142	0.005)
Respiratory illness-related pharmaceutical use cost difference	Monthly average temperature	-0.0005	0.003	0.17	0.867	(-0.005	0.006)
	Winter	-0.018	0.021	-0.85	0.397	(-0.060	0.024)

* p<0.1, ** p<0.05, *** p<0.01

The monthly average temperature and winter covariates are not associated with the pharmaceutical use cost difference dependent variable and it seems unlikely that they will prove to be important confounders in the multiple regression model. Both variables have a statistically significant association with the circulatory illness-related pharmaceutical use cost dependent variable at the ($p<0.1$) level but have very small coefficients, suggesting that any confounding effect that these covariates in the multiple regression model might have may be small in size. Both variables have no statistically significant association with the respiratory illness-related pharmaceutical use cost difference dependent variable.

6.2.4. Multiple regression model

The initial multiple regression model can be considered the final model, given the concerns about variable transformation expressed in Chapter Five and in Subsection 5.2.5,. Table 61 presents the results of the model. Winter based-interaction terms are considered seperately.

Table 61: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
All cause pharmaceutical use cost difference	Insulation only	-2.008	0.249	-8.03	0.000***	(-2.498	-1.518)	R-sq: within = 0.0113	F(7,28968) = 164.55 Prob > F = 0.0000
	Heating only	-3.008	0.807	-3.73	0.000***	(-4.589	-1.426)	R- sq: between = 0.0674	
	Insulation and heating	-1.195	0.568	-2.10	0.036**	(-2.309	-0.081)	R-sq: overall = 0.0403	
Circulatory illness-related pharmaceutical use cost difference	Insulation only	-0.389	0.068	-5.75	0.000***	(-0.523	-0.257)	R-sq: within = 0.0037	F(7,29054) = 67.00 Prob > F = 0.0000
	Heating only	-0.578	0.220	-2.63	0.009***	(-1.009	-0.147)	R-sq: between = 0.1007	
	Insulation and heating	-0.234	0.154	-1.52	0.129	(-0.536	0.068)	R-sq: overall = 0.0472	
Respiratory illness-related pharmaceutical use cost difference	Insulation only	-0.034	0.038	-0.88	0.379	(-0.109	0.041)	R-sq: within = 0.0012	F(7,28946) = 19.52 Prob > F = 0.0000
	Heating only	-0.042	0.124	-0.34	0.732	(-0.286	0.201)	R-sq: between = 0.0124	
	Insulation and heating	-0.016	0.090	-0.18	0.857	(-0.194	0.161)	R-sq: overall = 0.0061	

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff, age_diff, winter and monthly average temperature

The all pharmaceutical use cost difference model shows statistically significant cost savings per month of $2.01 for insulation ($p<0.01$), $3.01 for heating only ($p<0.01$) and $1.20 for 'insulation and heating' ($p<0.05$). These results suggest a protective effect of all three treatment types which results in reduced pharmaceutical usage. Wald test results for the equality of the insulation and 'insulation and heating' coefficients suggest that there is no evidence to reject the null hypothesis of equality, ($F(1, 28,968)=1.76$, prob > F = 0.1847).

Wald test results for the comparison of the 'heating' and 'insulation and heating' coefficients show some evidence to reject the null hypothesis ($F(1, 28{,}968)=3.41$, prob>F = 0.0647). This result is surprising given that the coefficient of heating is larger than the coefficient of 'insulation and heating' (-3.008 vs. 1.195) and given that insulation and heating have real and separate effects we would predict that 'insulation and heating' would have a larger combined effect. We can speculate that this result reflects differences in the characteristics of treatment households that received the different treatment types, or it may simply be a statistical anomaly driven by random variation and the small number of treatment/control pairings.

A similar pattern is repeated in the circulatory illness-related pharmaceutical use cost difference model where there are highly statistically significant coefficients for insulation -$0.39 and -$0.58 for heating and a non-statistically significant cost reduction of -$0.23 for 'insulation and heating that is smaller than the other two coefficients. Here, however, Wald tests for the equality of the 'insulation' and 'insulation and heating' coefficients ($F(1, 29{,}054)$, prob > 0.3482) and the 'heating' and 'heating and insulation' coefficients ($F(1, 29{,}054)$, prob > 0.1977) do not provide evidence to reject the null hypotheses that coefficients are equal.

Finally, as with the original model, the extended model of respiratory illness-related pharmaceutical use cost differences does not demonstrate any statistically significant impact of any treatment type at the ($p<0.1$) level.

Given that the coefficients for each treatment type in the all-cause model are larger than comparable coefficients in the circulatory illness-related model and the respiratory illness-related model combined we can speculate that either the circulatory and respiratory categories have been misspecified or that there is another effect being captured by the all-cause model, perhaps reductions in infectious disease or psychological illness-related pharmaceutical use. Alternatively, the inconsistency may be evidence of model inadequacy.

Finally, it is interesting to note that the climate-related covariates, monthly average temperature and winter vs. non winter, did not have statistically significant coefficients in any of the three multiple regression models, suggesting that there was no confounding that the inclusion of these covariates addressed (see Table 89 and Table 90 Appendix Two for full regression output). However, the extended model may still be of value given its greater statistical power and accurate recording of the treatment status of each treatment household in a given month, which was not possible in the original model given the necessary simplifications required for a *before* and *after* model.

Although it is unlikely that any statistically interaction effects may be found given these initial results, it is still useful to explore any effect of including interaction terms between treatment dummies and winter. The output of evaluate the following regression is presented in Table 62:

$$Pharmaceutical\ Use\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \beta_3 heating_and_insulation_{it} + \beta_4 occupant_diff_{it} + \beta_5 age_diff_{it} + \beta_6 winter_t + \beta_7 mean_monthly_temp_{it} + + \beta_8 insulation_{it} * winter_t + \beta_9 heating_{it} * winter_t + \beta_{10} heating_{it} * winter_t \beta\varepsilon_{it}$$

Table 62: Multiple regression model, extended model, months with values below 1st and above 99th percentiles excluded, winter interaction terms added

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
All cause pharmaceutical use cost difference difference	Insulation only	-1.962	0.278	-7.06	0.000***	(-2.507	-1.418)	R-sq: within = 0.0113 R- sq: between = 0.0674 R-sq: overall = 0.0403	F(10, 28968) = 115.26 Prob > F = 0.0000
	Heating only	-2.855	0.878	-3.25	0.001***	(-4.576	-1.133)		
	Insulation and heating	-0.951	0.639	-1.49	0.137	(-2.203	0.301)		
	Insulation only*winter	-0.113	0.307	-0.37	0.712	(-0.715	0.489)		
	Heating only*winter	-0.356	0.877	-0.41	0.685	(-2.076	1.363)		
	Insulation and heating*winter	-0.511	0.623	-0.82	0.413	(-1.732	0.711)		
Circulatory illness--related pharmaceutical use cost difference	Insulation only	-0.397	0.077	-5.16	0.000***	(-0.547	-0.246)	R-sq: within = 0.0037 R- sq: between = 0.1007 R-sq: overall = 0.0472	F(10, 29054) Prob > F = 0.0000
	Heating only	-0.579	0.233	-2.48	0.013**	(-1.035	-0.122)		
	Insulation and heating	-0.201	0.174	-1.16	0.247	(-0.542	0.139)		
	Insulation only*winter	0.014	0.084	0.17	0.868	(-0.151	0.179)		
	Heating only*winter	0.002	0.221	0.01	0.994	(-0.432	0.435)		
	Insulation and heating*winter	-0.065	0.179	-0.36	0.716	(-0.417	0.286)		
Respiratory illness-related pharmaceutical use cost difference	Insulation only	-0.012	0.045	-0.26	0.795	(-0.099	0.076)	R-sq: within = 0.012 R- sq: between = 0.0124 R-sq: overall = 0.0061	F(10, 28946) =14.35, Prob > F = 0.0000
	Heating only	0.143	0.145	0.98	0.326	(-0.142	0.428)		
	Insulation and heating	-0.017	0.104	-0.16	0.874	(-0.221	0.187)		
	Insulation only*winter	-0.053	0.054	-0.97	0.332	(-0.159	0.054)		
	Heating only*winter	-0.419	0.154	-2.72	0.007***	(-0.721	-0.117)		
	Insulation and heating*winter	-0.006	0.111	-0.06	0.954	(-0.224	0.211)		

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter,monthly average temperature, interaction terms for treatment and winter

Table 62 demonstrates very little evidence of statistically significant interactions between treatment status and winter. Models that include treatment*winter interaction terms also consistently have lower F statistics. A similar finding was reported in Section 5.3.4 and was interpreted as indication some evidence of selection bias or perhaps unmeasured

confounding, given that benefits of treatment under WUNZ:HS are hypothesised to be greater in winter. The single exception to this finding is the interaction term for heating only and winter in the respiratory illness related pharmaceutical use cost difference model. The direction of this interaction suggests an effect of heating in winter that is greater than that observed for heating in summer (which is not statistically significant in this model). This finding is plausible. Given the limited policy significance of this finding, due to the small size of the savings involved, the result is not further explored in this thesis.

6.2.5. Assessing the impact of extreme value removal

Here I follow the process carried out for previous cost difference models in order to assess the impact of the approach to extreme value removal. Table 63 presents the results of running the multiple regression extended regression models described above with no removal of extreme values and with values below the 5th or above the 95th percentiles removed.

Table 63: All-cause pharmaceutical use cost difference, extended model, no exclusions and months above 5th or below 95th percentiles excluded

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all pharmaceutical use costs	Insulation only	-1.320	0.897	-1.47	0.141	(-3.079	0.439)	R-sq: within = 0.0036 R-sq: between = 0.0213 R-sq: overall = 0.0099	F(7,29605) = 46.11 Prob > F = 0.0000
	Heating only	-6.980	2.222	-3.14	0.002***	(-11.334	-2.626)		
	Insulation and heating	-0.726	1.821	-0.4	0.690	(-4.296	2.844)		
Difference in all pharmaceutical use costs (values below 5th or above 95th percentiles removed)	Insulation only	-0.785	0.160	-4.91	0.000***	(-1.098	-0.472)	R-sq: within = 0.0071 R-sq: between = 0.0778 R-sq: overall = 0.0374	F(7,26445) = 178.16 Prob > F = 0.0000
	Heating only	-1.464	0.521	-2.81	0.005***	(-2.485	-0.444)		
	Insulation and heating	-0.714	0.371	-1.93	0.054*	(-1.442	0.013)		
Difference in circulatory illness-related pharmaceutical use costs	Insulation only	-0.383	0.125	-3.08	0.002***	(-0.628	-0.139)	R-sq: within = 0.0027 R-sq: between = 0.0758 R-sq: overall = 0.0347	F(7,29605) = 57.13 Prob > F = 0.0000
	Heating only	-0.959	0.314	-3.05	0.002***	(-1.575	-0.343)		
	Insulation and heating	-0.048	0.219	-0.22	0.825	(-0.477	0.380)		
Difference in circulatory illness-related pharmaceutical use costs (values below 5th or above 95th percentiles removed)	Insulation only	0.232	0.047	4.96	0.000***	(0.140	0.324)	R-sq: within = 0.0031 R-sq: between = 0.0983 R-sq: overall = 0.0406	F(7,28017) = 78.05 Prob > F = 0.0000
	Heating only	0.167	0.157	1.06	0.287	-(0.140	0.474)		
	Insulation and heating	0.263	0.100	2.63	0.009***	(0.067	0.459)		
Difference in respiratory illness-related pharmaceutical use costs	Insulation only	0.169	0.064	2.62	0.009***	(0.043	0.295)	R-sq: within = 0.0017 R-sq: between = 0.0137 R-sq: overall = 0.0083	F(7,29605) = 24.33 Prob > F = 0.0000
	Heating only	0.284	0.199	1.43	0.153	(-0.105	0.673)		
	Insulation and heating	0.110	0.136	0.81	0.416	(-0.156	0.376)		
Difference in respiratory illness-related pharmaceutical use costs (values below 5th or above 95th percentiles removed)	Insulation only	-0.211	0.018	-11.97	0.000***	-(0.246	-0.177)	R-sq: within = 0.0012 R-sq: between = 0.0099 R-sq: overall = 0.0032	F(7,28935) = 76.58 Prob > F = 0.0000
	Heating only	-0.238	0.057	-4.19	0.000***	-(0.350	-0.127)		
	Insulation and heating	-0.277	0.041	-6.69	0.000***	-(0.358	-0.196)		

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff, age_diff, winter and monthly average temperature

Removal strategy has a meaningful impact on the results of the extended all pharmaceutical use cost difference model. The model with no exclusions only shows a statistically significant saving as a result of receiving heating: this saving is larger than that found in the model with exclusion of the 1st and 99th percentiles ($6.98 vs. $3.01). The model with exclusions at the 5th and 95th percentiles shows highly statistically significant ($p<0.01$) savings as a result of receiving insulation ($0.78) and heating ($1.46). There are also savings as a result of receiving 'insulation and heating' ($0.71) that are statistically significant at the ($p<0.1$) level.

The circulatory illness-related pharmaceutical cost difference model finds statistically significant savings as a result of receiving insulation or heating with no exclusions but finds statistically significant ($p<0.01$) *increases* in cost differences as a result of receiving insulation or 'insulation and heating' when records containing dependant variables below the 5th or above the 95th percentiles are excluded.

The circulatory illness-related pharmaceutical cost difference model find a statistically significant ($p<0.01$) *increase* in costs as a result of receiving insulation with no exclusions, but finds statistically significant ($p<0.01$) decreases in cost differences as a result of receiving all three treatment categories when records containing dependant variable below the 5th or above the 95th percentiles are excluded.

There is no obvious pattern in the impact of different exclusion strategies on the size and sign of the coefficients of treatment variables. There is a general trend toward results with greater statistical significance as the exclusion strategy becomes broader. As with the discussion of exclusions strategies elsewhere in this chapter and in the previous chapter, the importance of the discussion here is to make the impact of a priori model building choices apparent.

6.2.6. Model specification and global goodness-of-fit measures

Here I follow the approach to previous models and test for misspecification by recalculation model standard errors. Table 64 presents the original multiple regression model for all pharmaceutical use cost differences with robust standard errors clustered at the treatment/control pairing level and the model with bootstrapped standard errors.

Table 64: Extended model with robust clustered standard errors and bootstrapped standard errors, months with values below 1st and above 99th percentiles excluded

Dependent variable	Standard error method	Variable name	Coefficient	Robust standard error	t	P>|t|	95% Confidence intervals	
All pharmaceutical use cost difference	Robust clustered standard errors	Insulation	-2.008	0.249	-8.03	0.000***	(-2.498	-1.518)
		Heating	-3.008	0.807	-3.73	0.000***	(-4.589	-1.426)
		Insulation and heating	-1.195	0.568	-2.10	0.036**	(-2.309	-0.081)
	Bootstrapped standard errors	Insulation	-2.008	0.239	-8.4	0.000***	(-2.477	-1.540)
		Heating	-3.008	0.844	-3.56	0.000***	(-4.662	-1.354)
		Insulation and heating	-1.195	0.567	-2.11	0.035**	(-2.306	-0.083)

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter and monthly average temperature

The high degree of similarity between the standard errors and confidence intervals in Table 64 gives additional reassurance that misspecification of the model is not producing unreliable estimates of standard errors for these this model. Finally, it is useful to stratify results by the Community Services Card status of the treatment households.

6.2.7. Stratification by Community Services Card Status

The results of stratification of the extended models by the Community Services Card status of the treatment household in each treatment/control pairing are presented in Table 65.

Table 65: All-cause hospitalisation cost difference, extended model, stratified by Community Services Card status, months with values below 1st and above 99th percentiles excluded

Dependent variable modelled	CSC status of treatment household	Variable name	Coefficient	Robust standard error	Z	P>|Z|	95% Confidence intervals	
Difference in all cause pharmaceutical use costs	CSC	Insulation	-2.268	0.353	-6.42	0.000*	(-2.961	-1.575)
		Heating	-3.148	1.218	-2.58	0.01**	(-5.535	-0.760)
		Insulation and heating	-0.510	0.786	-0.65	0.517	(-2.051	1.031)
	Non-CSC	Insulation	-1.662	0.335	-4.97	0.000***	(-2.318	-1.007)
		Heating	-2.867	1.052	-2.72	0.006***	(-4.930	-0.804)
		Insulation and heating	-2.238	0.783	-2.86	0.004***	(-3.773	-0.702)
Difference in circulatory pharmaceutical use costs	CSC	Insulation	-0.475	0.099	-4.8	0.000***	(-0.669	-0.281)
		Heating	-0.515	0.366	-1.4	0.16	(-1.233	0.203)
		Insulation and heating	-0.314	0.219	-1.44	0.151	(-0.742	0.114)
	Non-CSC	Insulation	-0.278	0.084	-3.29	0.001***	(-0.443	-0.112)
		Heating	-0.641	0.241	-2.66	0.008***	(-1.113	-0.169)
		Insulation and heating	-0.114	0.201	-0.57	0.571	(-0.507	0.280)
Difference in respiratory pharmaceutical use costs	CSC	Insulation	0.032	0.055	0.59	0.557	(-0.075	0.140)
		Heating	0.138	0.187	0.74	0.46	(-0.228	0.504)
		Insulation and heating	0.058	0.126	0.46	0.643	(-0.188	0.304)
	Non-CSC	Insulation	-0.120	0.050	-2.43	0.015**	(-0.217	-0.023)
		Heating	-0.225	0.163	-1.38	0.167	(-0.544	0.094)
		Insulation and heating	-0.128	0.124	-1.03	0.302	(-0.371	0.115)

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff and age_diff

Wald test comparisons of CSC and non-CSC coefficients within each model generally find no evidence to reject the null hypotheses that the coefficients are equal. The single exception to this rule is the comparison of insulation coefficients in the respiratory pharmaceutical use cost difference model ($F(1, 28{,}946)=4.35$, prob> F = 0.0370). Comparison of the two coefficients shows that there is a non-statistically significant *increase* in cost difference of $0.03 for CSC households that received insulation and a statistically significant ($p<0.05$) decrease in cost difference for non-CSC households of $0.12. This result seems difficult to reconcile with predictions of greater health benefits for lower-income individuals/households. However, the Pharms dataset record prescriptions which often cost individuals money, even when individuals receive greater levels of subsidy due to income status. This means that the ability to respond to illness by obtaining prescriptions is strongly mediated by income level. We may speculate that statistically significant (although small) savings for the non-CSC treatment group are the result of reductions in non-essential prescription costs. Such prescriptions might be less frequently obtained by

CSC-individuals and households due to income constraints, meaning that improved temperatures would have less impact on these costs. It is import to note, however, the limitations of CSC stratification, discussed previously, which will bias results towards the null due to misclassification.

6.2.8. Summary

Analysis of household level differences in all-cause pharmaceutical use costs using the extended model suggests that receiving insulation, heating and 'insulation and heating' result in small, but statistically significant cost savings in all three cost difference models. These results are not greatly affected by the CSC status of treatment households. There are small but statistically significant cost savings as a result of receiving insulation or heating on circulatory illness-related pharmaceutical use cost differences, which are also not influenced by CSC-status. Finally, there is limited evidence of a beneficial impact of any treatment type on respiratory illness-related pharmaceutical use cost differences.

As with the previous model these results are sensitive to the exclusion strategy used.

6.3. Chapter summary

The exploration of pharmaceuticals use cost differences in this chapter proved surprisingly fruitful given initial expectations that any benefits found would be very small in size based on the results calculated in Telfar Barnard, Preval et al. (2011). Inclusion of the *occupant_diff* and *age_diff* variables produced models that found a greater impact of treatment than the previous models presented, in some cases similar in size to that found in the hospitalisation cost difference models presented in the previous chapter. It is difficult to draw strong conclusions about underlying causal mechanisms given that changes in costs may be driven by both rates of prescriptions, length of prescriptions and types of drugs prescribed.

Comparison of results for the original and extended pharmaceutical use cost difference models show fairly similar results in terms of size direction and statistical significance of coefficients, although results for 'insulation and heating' varied in statistical significance between the original and extended models, which may reflect smaller numbers (relative to the number of households that received insulation only) and thus greater sensitivity to subtle differences in exclusion strategy, or more accurate modelling of treatment timing.

Having completed a thorough examination of mortality-related costs, hospitalisation costs and pharmaceutical use costs it is now possible to compare these results, to aggregate results that meet inclusion criteria ($p<0.05$ or $p<0.1$) at an annual household level, and to examine the impact of these combined results on the outcomes of the original cost-benefit analysis presented in Grimes, Denne et al. (2011). This is the topic of Chapter Seven.

CHAPTER 7. COMBINING HEALTH BENEFITS

This chapter presents the results of Chapters Four, Five and Six together in order to consider to what extent they contradict or corroborate each other. In the second section results are combined in order to produce annual estimates of benefit per household suitable for inclusion in a cost-benefit analysis such as that carried out by Grimes, Denne et al. (2011). Finally, different estimates of annual health benefit per household are substituted into the original cost-benefit analysis in order to explore the impact of different modelling decisions and exclusion strategies on the benefit:cost ratios calculated.

7.1. Comparison of health benefits

Table 66 presents the coefficients of the insulation, heating and 'insulation and heating' for mortality, hospitalisation cost and pharmaceutical use cost models. Where both the extended and original models are available the extended results are presented.

Table 66 Summary table of model coefficients ordered by treatment type

Variable name	Model	Units	HR or Coeffi-cient	SE	P-value	95% CI		Beneficial
Insulation	All-cause hospitalisation costs (extended model)	Dollars	-3.186	0.694	0.000**	(-4.545	-1.826)	Y
	All-cause pharmaceutical use costs (extended model)	Dollars	-2.008	0.249	0.000***	(-2.498	-1.518)	Y
	Circulatory hospitalisation costs	Dollars	-1.4	0.563	0.013**	(-2.504	-0.297)	Y
	Circulatory illness-related mortality	Hazard Ratio	0.673	0.079	0.001**	(0.535	0.847)	Y
	Circulatory illness-related pharmaceutical use costs (extended model)	Dollars	-0.389	0.068	0.000***	(-0.523	-0.257)	Y
	Respiratory hospitalisation costs	Dollars	-0.5	0.286	0.083*	(-1.058	0.065)	Y
	Respiratory illness-related mortality	Hazard Ratio	0.83	0.1	0.122	(0.655	1.051)	Y
	Respiratory illness-related pharmaceutical use costs (extended model)	Dollars	-0.034	0.038	0.379	(-0.109	0.041)	Y
Heating	All-cause hospitalisation costs (extended model)	Dollars	-0.186	2.236	0.934	(-4.569	4.197)	Y
	All-cause pharmaceutical use costs (extended model)	Dollars	-3.008	0.807	0.000***	(-4.589	-1.426)	Y
	Circulatory hospitalisation costs	Dollars	-2.33	1.797	0.194	(-5.854	1.189)	Y
	Circulatory illness-related mortality	Hazard Ratio	1.348	0.293	0.169	(0.881	2.064)	N
	Circulatory illness-related pharmaceutical use costs (extended model)	Dollars	-0.578	0.22	0.009***	(-1.009	-0.147)	Y
	Respiratory hospitalisation costs	Dollars	-0.12	0.958	0.897	(-2.001	1.754)	Y
	Respiratory illness-related mortality	Hazard Ratio	1.129	0.313	0.662	(0.656	1.944)	N
	Respiratory illness-related pharmaceutical use costs (extended model)	Dollars	-0.042	0.124	0.732	(-0.286	0.201)	Y
Insulation and heating	All-cause hospitalisation costs (extended model)	Dollars	-2.963	1.511	0.050*	(-5.926	-0.001)	Y
	All-cause pharmaceutical use costs (extended model)	Dollars	-1.195	0.568	0.036**	(-2.309	-0.081)	Y
	Circulatory hospitalisation costs	Dollars	-3.38	1.335	0.011**	(-5.993	-0.758)	Y
	Circulatory illness-related mortality	Hazard Ratio	0.581	0.15	0.035**	(0.350	0.964)	Y
	Circulatory illness-related pharmaceutical use costs (extended model)	Dollars	-0.234	0.154	0.129	(-0.536	0.068)	Y
	Respiratory hospitalisation costs	Dollars	-0.47	0.624	0.447	(-1.697	0.748)	Y
	Respiratory illness-related mortality	Hazard Ratio	0.624	0.208	0.158	(0.324	1.200)	Y
	Respiratory illness-related pharmaceutical use costs (extended model)	Dollars	-0.016	0.09	0.857	(-0.194	0.161)	Y

* p<0.1, ** p<0.05, *** p<0.01 HR = Hazard Ratio

When Table 66 is considered, a strong pattern suggesting a beneficial effect of treatment emerges. There are only two instances in the above table where model coefficients did not indicate a beneficial effect of receiving treatment under WUNZ:HS: the HRs for the heating treatment variable in the analysis of circulatory illness-related mortality and in the analysis of respiratory illness-related mortality. These HRs were not statistically significant. Of the twenty two covariates/HRs indicating a protective effect of treatment, four were significant at the ($p<0.01$) level, ten were statistically significant at the ($p<0.05$) level and twelve were significant at the ($p<0.1$) level.

These results appear to strengthen the case for a real protective impact of receiving treatment under WUNZ:HS, particularly for insulation, where six of eight results indicated a protective effect significant at the ($p<0.1$) level. Although the coefficients of 'insulation and heating' were generally less statistically significant, Wald tests carried out in the previous chapters generally found that there was no evidence to distinguish the coefficients of 'insulation' and 'insulation and heating' in models, which can be interpreted as suggesting that there is a beneficial effect of 'insulation and heating' driven by insulation, which was not statistically significant due to the lower numbers of households who received this treatment combination. Therefore models had limited power to detect this benefit.

Turning to heating, model coefficients for heating did not indicate statistically significant beneficial effects for the most part, which may reflect limited statistical power due to relatively small numbers of households who received heating only. A number of theories as to why heating model coefficients might not be statistically significant were discussed in previous chapters, including the possibility that receiving a heating retrofit *only* implied that pre-existing insulation was present and thus that baseline temperatures were already relatively high. This would limit the health impact of further improvements. The exception to this general finding was the analysis of pharmaceutical use cost data, where heating was found to have a highly statistically significant beneficial effect. It is unclear why this should be the case, but we can speculate that improved temperatures due to heating may have an impact on all-cause pharmaceutical use costs and circulatory illness-related pharmaceutical use costs via different causal pathways than those for hospitalisation costs and mortality. It may also be the case that the denser pharmaceutical cost difference data had greater power to detect health effects of heating due to the greater number of non-zero cost difference records in the dataset relative to the hospitalisation dataset, reflecting the higher frequency of subsidised prescriptions relative to hospitalisations.

The key result of the previous three chapters, from a cost-benefit and thus policy perspective, was the confirmation of the original finding in the WUNZ:HS evaluations

(Telfar Barnard, Preval et al. 2011) that receiving insulation led to a reduction in the risk of mortality for individuals aged 65 and over who had a baseline period circulatory hospitalisation. While it is difficult to draw strong conclusions about the causal processes underlying this result given the absence of cause-of-death data, the consistent household level reductions in circulatory hospitalisation costs and circulatory illness-related pharmaceutical use costs for households that received insulation only, relative to their matched control households, are consistent with this outcome. The small hospitalisation and pharmaceutical use cost savings involved may reflect the fact that elderly people make up a small proportion of the individuals in each study household.

7.2. Annual household health benefits

Having presented a summary of the key results of the three previous chapters it is now possible to aggregate benefits in order to estimate the annual benefits per household of receiving insulation or heating. For the most part previous results did not find an effect of treatment for 'insulation and heating' that could be distinguished from the effect of insulation alone and so it is not particularly valuable to separately sum benefits for households that received both 'insulation and heating'. Limiting the summing of benefits to insulation and heating only is consistent with the approach in the original health evaluations and has the additional benefit of allowing me to produce results compatible with the original CBA, thus making it straightforward to assess the impact of different assumptions on the original benefit:cost ratio calculated.

Table 67 sums all-cause hospitalisation-related cost reductions, all-cause pharmaceutical use cost reductions, mortality-related benefits and the imputed GP visit and reduced time of work and school benefits described in Subsection 3.3.8 in order to produce estimates of annual benefit per household for insulation and heating.

Annual household benefits are calculated by summing all benefits that are statistically significant at either the ($p<0.1$) or ($p<0.05$) levels in order to assess the impact of different inclusion thresholds. The decision to focus on all-cause hospitalisation and pharmaceutical use cost results and to ignore circulatory and respiratory model results is pragmatic, reflecting the small differences in the size of benefits calculated, relative to mortality-related benefits.

Mortality benefits are presented using the favoured assumptions underlined in Table 34 and Table 38, calculated using the current New Zealand transport sector VSL assuming a 4% discount rate and that individuals who avoided mortality gained, on average, 5.5 additional years of life.

Table 67 Annual health-related benefit per household of WUNZ:HS insulation and heating retrofits for all households

Retrofit type	Benefit type	Annual saving
Insulation	All-cause hospitalisation costs (extended model)	$38.23***
	All-cause pharmaceutical use costs (extended model)	$24.10***
	Imputed benefits	$47.75**
	Annual benefit of reduced mortality per household as a result of receiving insulation - circulatory cohort (assumes 4% discount rate)	$749.51***
	Annual benefit of reduced mortality per household as a result of receiving insulation - respiratory cohort (assumes 4% discount rate)	$314.01*
	Sum annual benefit per household (p<0.05)	$859.19
	Sum annual benefit per household (p<0.1)	$1,173.20
Heating	All-cause hospitalisation costs (extended model)	$2.23
	All-cause pharmaceutical use costs (extended model)	$36.10***
	Imputed benefits	$4.64 **
	Sum annual benefit per household (p<0.05)	$40.74
	Sum annual benefit per household (p<0.1)	$40.74

* p<0.1, ** p<0.05, *** p<0.01

The results of Table 67 show a large annual health-related benefit per household as a result of receiving a WUNZ:HS insulation retrofit, regardless of whether the criterion used is a p-value of ($p<0.05$) or ($p<0.01$). These results are clearly driven by the mortality-related benefits calculated in Subsection 4.3. Heating is shown to have a much smaller annual benefit, driven by a reduction in pharmaceutical use costs.

Table 68 presents annual household benefits stratified by CSC status. The difference in annual benefits is almost entirely driven by differences in the mortality-related benefits calculated for CSC and non-CSC households which reflect differences in the average proportion of vulnerable individuals per household. There are also differences in the imputed benefits by CSC status, based on the calculations reported in Subsection 3.3.8, which limit imputation of benefits to CSC households due to the targeted nature of the RCTs that inform these calculations. There is no statistical justification for stratification based on the results reported in Chapters Four, Five and Six.

Table 68: Annual health-related benefit per household of WUNZ:HS insulation and heating retrofits for CSC households and non-CSC households

CSC households		
Retrofit type	Benefit type	Annual saving
Insulation	All-cause hospitalisation costs (extended model)	$38.23***
	All-cause pharmaceutical use costs (extended model)	$24.10***
	Imputed benefits	$95.49 **
	Annual benefit of reduced mortality per household as a result of receiving insulation - circulatory cohort (assumes 4% discount rate)	$972.92 ***
	Annual benefit of reduced mortality per household as a result of receiving insulation - respiratory cohort (assumes 4% discount rate)	$399.49*
	Sum annual benefit per household (p<0.05)	$1,130.74
	Sum annual benefit per household (p<0.1)	$1,530.23
Heating	All-cause hospitalisation costs (extended model)	$2.23
	All-cause pharmaceutical use costs (extended model)	$36.10***
	Imputed benefits	$9.27 **
	Sum annual benefit per household (p<0.05)	$45.37
	Sum annual benefit per household (p<0.1)	$45.37
Non-CSC households		
Retrofit type	Benefit type	Annual saving
Insulation	All-cause hospitalisation costs (extended model)	$38.23***
	All-cause pharmaceutical use costs (extended model)	$24.10***
	Imputed benefits	$0.00
	Annual benefit of reduced mortality per household as a result of receiving insulation - circulatory cohort (assumes 4% discount rate)	$410.02***
	Annual benefit of reduced mortality per household as a result of receiving insulation - respiratory cohort (assumes 4% discount rate)	$137.85*
	Sum annual benefit per household (p<0.05)	$472.35
	Sum annual benefit per household (p<0.1)	$610.20
Heating	All-cause hospitalisation costs (extended model)	$2.23
	All-cause pharmaceutical use costs (extended model)	$36.10***
	Imputed benefits	$0.00
	Sum annual benefit per household (p<0.05)	$36.10
	Sum annual benefit per household (p<0.1)	$36.10

* p<0.1, ** p<0.05, *** p<0.01

Because mortality benefits are such an important component of annual household health-related benefits it is useful to explore how different modelling assumptions influence the calculation of the annual household level health benefit associated with insulation. A variety of mortality-related benefit model outcomes are combined with the hospitalisation cost savings, pharmaceutical use cost savings and imputed benefits presented in Table 67 in order to demonstrate the impact of mortality model assumptions.

Table 69: Combined annual health benefits of insulation, all households, for a variety of mortality cost modelling assumptions

P-value threshold	Total years gained per person who avoided mortality	Discount rate	Transport 1991 -based VSL	Transport 1998 -based VSL	Fire 2007 based VSL
P<0.05	5.5	2.50%	$730.80	$1,161.16	$519.76
		4.00%	**$859.19**	$1,378.55	$604.49
		8.00%	$1,207.48	$1,968.32	$834.37
	2.75	2.50%	$452.59	$690.06	$336.14
		4.00%	$532.52	$825.40	$388.89
		8.00%	$763.82	$1,217.06	$541.55
	8.25	2.50%	$1,029.65	$1,667.21	$717.00
		4.00%	$1,197.86	$1,952.03	$828.02
		8.00%	$1,626.44	$2,677.75	$1,110.88
P<0.1	5.5	2.50%	$1,119.38	$1,299.78	$1,030.92
		4.00%	$1,173.20	$1,390.91	$1,066.44
		8.00%	$1,319.20	$1,638.13	$1,162.80
	2.75	2.50%	$986.30	$1,074.43	$943.09
		4.00%	$1,015.97	$1,124.66	$962.66
		8.00%	$1,101.81	$1,270.02	$1,019.32
	8.25	2.50%	$1,221.89	$1,473.35	$1,098.57
		4.00%	$1,288.23	$1,585.69	$1,142.36
		8.00%	$1,457.27	$1,871.93	$1,253.92

* p<0.1, ** p<0.05, *** p<0.01

Table 69 demonstrates the large range of annual health benefit estimates that can be calculated using different modelling assumptions, from $2,677.75 to $336.14. Of course the range of values would be greater still if statistical uncertainty was considered, but I have chosen to avoid presentation of confidence intervals here for reasons of concision. I have underlined my preferred value of $859.19, which is calculated using the (p<0.05) threshold and thus conservatively excludes benefits derived from reduced mortality in the respiratory cohort, which were not statistically significant at the (p<0.05) level.

7.2.1. Cost-benefit analysis

The cost-benefit analysis of the WUNZ:HS programme presented in Grimes, Denne et al. (2011) combined health results with estimates of the impact of insulation and heating on energy use (Grimes, Young et al. 2011) and a sophisticated analysis of industry impacts (Denne and Bond-Smith 2011), and programme cost data provided by EECA. Annual benefits were modelled to accrue over the lifetime of insulation and heating retrofits and future benefits were discounted appropriately. The model produced an estimate of the

total benefits and costs of the WUNZ:HS programme which required predictions regarding numbers of retrofits that would occur beyond the period of the evaluation.

It is possible to crudely substitute the results produced in this chapter for the health results used in the original cost-benefit analysis. Table 70 presents the results of such an approach, which is limited by a lack of access to the full calculations used by Grimes, Denne et al. (2011). Total health benefit estimates are derived for 18 different combinations of model assumptions and are used to substitute for one of the total health benefit figures used to calculate the primary benefit cost ratio in the original analysis. The health benefit figure in the original analysis was $1,561 million (Grimes, Denne et al. 2011), which was calculated assuming a 4% discount rate, using all-cause hospitalisation and pharmaceutical use cost model results (which was a conservative approach in the original context) and a central assumption regarding additionality[22]. Substitution is limited to total health benefit figures estimated in this thesis using mortality benefit estimates calculated using a 4% discount rate.

The purpose of this exercise is to illustrate the impact of different modelling choices on the benefit:cost ratios calculated rather than to contrast these results with the original benefit:cost ratio calculated (3.9:1) (Grimes, Denne et al. 2011); these differences, which reflect differences in mortality calculations and pharmaceutical use cost results, are of limited interest here. All benefit:cost ratios indicate strong evidence for the continuation of the WUNZ:HS programme, although in an ideal policy environment these results would need to be contrasted with benefit:cost ratios for alternative courses of action.

Table 70 WUNZ:HS evaluation benefit:cost ratios for a variety of modelling assumptions

P-value threshold	Total years gained per person who avoided circ or respiratory mortality	Transport 1991 -based VSL	Transport 1998 -based VSL	Fire 2007 based VSL
P<0.05	5.5	6.4:1	10.2:1	4.6:1
	2.75	4.1:1	6.2:1	3.0:1
	8.25	8.9:1	14.4:1	6.2:1
P<0.1	5.5	8.6:1	10.1:1	7.9:1
	2.75	7.7:1	8.5:1	7.2:1
	8.25	9.6:1	11.7:1	8.5:1

7.2.2. Summary

[22] Additionality, in this context, refers to the proportion of retrofits carried out under WUNZ:HS that were not predicted to occur in the absence of the programme (Grimes, Denne et al. 2011)

The results presented in the first section of this chapter, derived from different modelling approaches, appear to be largely consistent with each other and thus strengthen the case that that there was a real health benefit of receiving treatment under WUNZ:HS, particularly for insulation. The second section explored different approaches to aggregating benefits derived from mortality, hospitalisation cost difference and pharmaceutical use cost difference models. This demonstrated that the results are sensitive to the approach taken to estimating mortality-related benefits and the p-value threshold chosen for inclusion in this context. This can be inferred from the wide range of benefit cost ratios presented in Table 70. Nevertheless, even the most conservative set of assumptions leads to a benefit: cost ratio of 3.0: 1 in Table 70; i.e. benefits are three times greater than costs.

CHAPTER 8. POLICY COMPARISON OF WUNZ:HS AND AUSTRALIA'S HIP

Chapter Eight presents a policy-based comparison of the WUNZ:HS programme and Australia's ceiling insulation -based Home Insulation Programme (HIP). The two programmes represent a natural experiment; similar policies carried out by similar countries at almost the same time, but with very different outcomes. WUNZ:HS was considered a success story, having achieved its targets earlier than hoped, while the HIP ended early and in disarray with work carried out under the programme requiring costly assessment and remediation (Hawke 2010, Kortt and Dollery 2012).

At first glance, the subject of this chapter represents a change of focus for this thesis, given previous chapters' focus on the evaluation of the health benefits of the WUNZ:HS programme. However, the policy comparison of the two programmes enriches the previous chapters in two key ways.

Firstly, by identifying key differences between the two programmes that appear to have contributed to their relative success, it is possible to more accurately characterise WUNZ:HS, which is important if other nations or future New Zealand governments wish to operate similar programmes. In understanding the ways in which the two programmes differed we can identify key policy design and implementation issues, which are relevant to the success of programmes of this type. This is extremely relevant to the generalisation of the results presented in Chapters Four - Seven: it is clear that a programme similar to WUNZ:HS could fail to achieve any of the health outcomes presented, and indeed the evident energy savings (Grimes, Young et al. 2011), if its design or implementation more closely resembled the HIP.

Secondly, by comparing the two programmes the value and influence of policy evaluations such as the WUNZ:HS evaluation and high quality research more generally can be explored.

The evaluation of health cost data presented in the earlier chapters of this thesis and as part of the WUNZ:HS evaluation (Telfar Barnard, Preval et al. 2011, Preval 2012) can be understood as part of the rational public policy cycle, contributing to the policy evaluation of a programme which can then either be continued, modified or discontinued. The rational public policy cycle, although frequently criticised as being too conceptually linear, can be summarised in the following way:

1. Agenda-setting: how problems come to the attention of policy makers.
2. Policy formation: how problem solving options are formulated by policy makers.

3. Decision-making: the choice and enactment of a policy or policies (potentially including inaction).
4. Policy implementation: how policies are put into effect.
5. Policy evaluation: evaluation of effectiveness of policy by policy makers (may include many types of evaluation including economic evaluation), cycle returns to agenda setting as policy is continued, modified or discontinued.

Adapted from Howlett, Ramesh et al. (1995, p11).

Howlett, Ramesh et al. (1995) acknowledge that presenting the policy cycle as a single iterative loop is misleading, minimising the complex reality of the policy process. However, they argue that dividing the policy process into discrete parts facilitates understanding and allows comparative studies of different stages of the process to be carried out in a well-defined manner. This is particularly relevant to this chapter, given its focus on the comparison of the WUNZ:HS programme and the HIP. Taking the insight of Howlett and Ramesh et al (1995) as a starting point, I first describe the HIP in some detail, placing both it and WUNZ:HS in the international economic context of the time. Key programme characteristics are then summarised for ease of reference. The substance of Chapter Eight, informed by review of key Australian government grey literature, three published papers, New Zealand grey literature and a key informant interview with EECA official Jenny Lackey[23] is a policy-based comparison of the two programmes in the light of relevant theory, highlighting the problems faced by HIP and the ways in which these problems were avoided or negotiated by the WUNZ:HS programme. This process illuminates key differences that contributed to the outcomes of the two programmes.

[23] Jenny Lackey was a Project Manager for the WUNZ:HS programme for the entire period of its operation. She participated in a key informant interview in 2010, provided grey literature and also reviewed this chapter in 2014 in order to check it for accuracy with regard to her comments.

8.1. HIP: Context and overview

Both the WUNZ:HS programme and the HIP were developed in the context of the "Great Recession" of the late 2000s. The recession is widely thought to have originated in the United States, precipitated by a dramatic collapse in value of residential mortgage-backed securities and related collateralised debt obligations, as the U.S. housing bubble burst in 2007. The resulting international recession was characterised by reduced financial liquidity, dramatic rises in unemployment and reduced international trade.

New Zealand and Australian-owned banks were not heavily exposed to residential mortgage–backed securities and collateralised debt obligations (CDOs), however the loss of liquidity that characterised the crisis had a direct effect on the banks of both nations as credit dried up internationally. A number of non-bank financial institutions collapsed in both countries, driven by the shrinking of the two nation's respective property bubbles, loss of liquidity and "runs-on-the bank" caused by loss of investor confidence; both introduced depositor guarantee schemes in order to avoid further catastrophic runs. More generally the global recession had an impact on both nations by reducing demand for exports and increasing unemployment. New Zealand's economy was in recession from December 2007 to March 2009, while Australia was in recession for a single quarter (January 2009 – March 2009). Both countries fared well relative to many other OECD nations, due to relatively sound financial institutions and steady Chinese growth increasing demand for exports. New Zealand also benefited from Australian demand for exports.

Although New Zealand and Australia avoided the worst of the recession, it was not always clear that this would be the case, and faced with rising unemployment, particularly in their respective building/construction sectors, New Zealand and Australia followed the United States and the United Kingdom in adopting Keynesian economic policies intended to stimulate growth and reduce unemployment. Australia's then Prime Minister Kevin Rudd announced the $42 billion dollar Nation Building and Jobs Plan on February the 3rd 2009, a mixture of initiatives intended to immediately stimulate the economy and create jobs while investing in capital that would provide longer term benefits. The Nation Building and Jobs Plan was the largest of several stimulus measures undertaken by the Government between October 2008 and May 2009 with an estimated combined value of approximately $AUS 90 billion dollars (Senate 2010). New Zealand's stimulus initiatives during the recession were smaller in scale but similar in intent, comprising measures valued at approximately $NZ 7 billion dollars (NZPA 2008).

At $NZ 347 million the WUNZ:HS programme was a small, but significant part of the total stimulus initiatives undertaken by successive New Zealand governments. At $AUS 2.7

billion the HIP represented a smaller, but still significant proportion of total Australian stimulus measures. The programmes differed from other stimulus measures such as school and home building in Australia and road building in New Zealand in that they attempted to increase consumer demand for a product which was perceived to have positive externalities (Senate 2010). 900,000 homes were estimated not to meet current minimum standards for ceiling and floor insulation in New Zealand before the start of the WUNZ:HS programme and 2.7 million Australian homes were estimated to lack suitable ceiling insulation before the start of the HIP (Senate 2010).

The HIP programme began on the 3rd of February 2009, following unsuccessful proposals for similar schemes in 2007 and 2008. The basis of the HIP was a $1600 rebate for consumers for ceiling insulation. Insulation was to be installed by private installers. The stated goals of the HIP were to generate an economic stimulus, while creating jobs and supporting small businesses, as well as improving the energy efficiency of homes with the intention of reducing heating and cooling bills by $200 per year per home and reducing Australia's greenhouse gas emissions by around 4.56 million tonnes per year by 2020 (Hawke 2010, Senate 2010).

The Australian goal of improved energy efficiency was likely to be driven largely by reductions in air conditioning costs/improved cooling efficiency, rather than reduced heating costs/ improved heating efficiency. Australia is in general much warmer than New Zealand with the majority of its major cities experiencing 500- 1000 heating degree days[24] or less (base of 18°C) annually (Bureau of Meteorology 2013) while major New Zealand cities have heating degree days ranging between 1150 (Auckland), 2400 (Christchurch) and 2600 (Dunedin) (base of 18°C) (Lloyd 2006). Australia, in turn, experiences a much greater number of cooling degree days in the majority of its population centres (Bureau of Meteorology 2013).

The goal of improving energy efficiency was consistent with Australia's commitments under the Kyoto Protocol during the 2008-2012 period, which were to limit carbon dioxide emissions to 108% of 1990 levels, or to offset any excess emissions. New Zealand likewise had a commitment to reduce its emissions to 1990 levels. Emissions reductions were explicitly part of the rationale behind the HIP (Senate, 2010) but less clearly part of the

[24] Heating degree days are a measure of the heating (all sources including insulation and space heating) required to achieve a given base temperature. At a city level they may simply be calculated as the difference between daily average temperature and a given base temperature (Carbon Trust 2012). Degree days may then be summed over a period of time that is of interest (such as a year). More complex calculations may be made at an individual building level if more detailed data is available.

rationale behind WUNZ:HS, perhaps reflecting some political ambivalence regarding this issue on the part of the ruling National Party-led coalition government.

The goal of improving comfort is listed in the Senate Committee report (Senate 2010), but other health benefits are not explicitly discussed in the grey literature.

The HIP was intended to operate for two and a half years and to retrofit 2.2 million homes with ceiling insulation, while providing 9,800 jobs during the duration of the programme(ANAO 2010). The HIP was administered by the then Department of Energy, Water, Heritage and the Arts (DEWHA), a federal government department, via a Project Control Group which included a Deputy Secretary from DEWHA, who reported directly to the Environment Minister and representatives from the Office of the Coordinator Genera, which is part of the Department of the Prime Minister and Cabinet (Senate 2010).

Phase 1 of the HIP lasted from the 3rd of February 2009, until July 1st 2009 during which time approximately 53,000 retrofits were carried out with numbers increasing exponentially from month to month (Senate 2010). During this phase installers did not have to register with the programme and consumers only needed to obtain two quotes in order to qualify for a rebate. Consumers claimed rebates directly from DEWHA, which processed and handled all claims. Phase 1 operated smoothly, only slightly hampered by DEWHA's limited capacity and experience in processing high volumes of claims (Lewis 2010). Of some concern was the announcement in June 2009 that the Australian Competition and Consumer Commission was investigating claims that some installers were circumventing the requirement that consumers obtain two quotes (Lewis 2010).

During Phase 1 a consultation process was carried out, which shaped the design of the main programme, Phase 2. Key participants in the consultation process included installers, insulation manufacturers, industry bodies, state officials and training organisations (ANAO 2010, p 66). A comprehensive risk assessment was carried out in April 2009 by consultants Minter Ellison which informed programme design and implementation. Key decisions regarding appropriate qualifications for installers, the means of delivering the rebate and products and installation methods suitable for inclusion in the programme were made following the consultation and risk assessment processes. The Australian National Audit Office reported that *"the design of Phase 2 was strongly influenced by the clear riding instruction of the Commonwealth Coordinator General to reduce red tape and commence work on projects as soon as possible"* (ANAO 2010, p 28).

Phase 2 was significantly different from Phase 1. Key differences included: installer registration online with the registration criteria "prior industry experience; a trade qualification in an approved trade; or insulation specific training" (ANAO 2010, p 104);

householders only needing to obtain a single quote; and payment being made directly by DEWHA to installers. Medicare Australia processed installer registrations and payments in conjunction with DEWHA, which addressed concerns about DEWHA's capacity and experience in this area.

Phase 2 continued the exponential increase in the number of monthly installations (from 78,375 in July 2009 – 209,267 in November 2009 at the peak of the programme).As Phase 2 progressed the programme began to face problems, which were difficult to address in a timely manner, because of the high numbers of installations being carried out (Hawke 2010) and because of the overriding stimulus goal (ANAO 2010). A spate of house fires, believed to have resulted from inappropriate installation of insulation above down-lights came to national attention (Senate 2010), as did on-going allegations of installer fraud and poor quality workmanship (ANAO 2010). Four workers died during installation, three of electrocution related to aluminium foil insulation and the fourth from overheating (Senate 2010). Audits confirmed poor quality work and fire/ electrocution hazards. In addition, the pace of demand meant that Australian production of insulation materials could not meet demand and that it was necessary to import insulation, primarily from Asia, somewhat limiting the stimulus effect of the programme (much of the Chinese-made insulation was later found not to meet advertised specifications) (Lewis 2010).

Although DEWHA responded to these problems by altering incentives, products, rules and requirements, the conflicting goal of achieving stimulus objectives appeared to limit and slow its responses (ANAO 2010). Following numerous changes to the programme, on the 19th of February 2010 the HIP was suspended in disarray, having insulated only 1 million of the 2.2 million homes originally intended. The HIP was formally discontinued on the 22nd of April 2010 following a highly critical report by Dr Allan Hawke (Lewis, 2010).

The failures of the programme meant that significant auditing and remediation was required which necessitated the creation of a Foil Insulation Safety Programme, a Home Insulation Safety Programme which involve investigating and correcting safety hazards resulting from the installation of foil and non-foil insulation respectively and finally an Insulation Industry Assistance Programme that attempted the minimise the harm caused by the premature end of the HIP on installers (ANAO 2010, Kortt and Dollery 2012). All foil installations were ultimately required to be physically checked and a proportion of non-foil installations. The Department of Climate Change and Energy Efficiency (DCCEE) took responsibility from DEWHA for the closure of the programme and remediation activities. Environment Minister Peter Garrett eventually lost responsibility for all remaining energy efficiency programs (Hall 2010). In late 2013 the Liberal Government led

by Prime Minister Tony Abbot announced a comprehensive Royal Commission enquiry into the HIP. The terms of reference of the enquiry include a focus on the assessment of risk by politicians and senior officials with regard to health and safety, and on consultation and risk assessment procedures used (Royal Commission into the Home Insulation Programme 2013). The Commission will report back to the government by the 30th of June 2014.

8.2. Comparison of WUNZ:HS and HIP

The many challenges the HIP faced have been thoroughly described in three comprehensive reports: the government commissioned Hawke Report (Hawke 2010), an Australian National Audit Office report (ANAO 2010) and a Senate Committee Report (Senate 2010) as well as three papers (Dollery and Hovey 2010, Lewis 2010, Kortt and Dollery 2012). The three reports provide excellent and detailed evidence about the operation of HIP and the problems that lead to its demise. Lewis (2010) compares the operation of the HIP programme with two standard Australian public policy texts, the *Australian Policy Handbook* (Althaus, Bridgman et al. 2007) and *Beyond the Policy Cycle* (Colebatch 2006). Dollery and Hovey (2010) and Kortt and Dollery (2012) approach the HIP by placing it in the context of government-failure typologies derived from public choice literature. From this body of grey and academic literature, there is little disagreement about the key problems and characteristics that led to the premature end of the HIP, although there is some dispute on factual matters, for example the degree to which HIP retrofits actually increased the rate of residential fires.

Although the challenges faced by the HIP have been described in depth, previous research has not taken advantage of the natural policy experiment made possible by considering how New Zealand's concurrent WUNZ:HS programme negotiated the design and implementation issues that the HIP faced. In order to do this I have selected key failings of the Australian HIP from the available literature and described the extent to which they do or do not apply to WUNZ:HS, using the idealised policy cycle as a framing device. The comparison of the two programmes focuses on Phase Two of the HIP as Phase One of the programme was clearly a "*learning while doing*" process (Hawke, 2010, p 25), intended to inform the design and risk assessment of the main phase of the programme, Phase Two.

It is useful to summarise key features of the two programmes for ease of reference prior to detailed discussion of the differences between the two programmes that may explain their different outcomes. This summary is presented in Table 71 below.

Table 71: WUNZ:HS and HIP key features

	HIP	WUNZ:HS
Start date – intended completion date	3 Feb 2009 – 31 Dec 2011	1 July 2009 – 30 June 2013
Actual dates	Phase 1 (3 Feb 2009 – 30th Jun 2009) Phase 2 (1st Jul 2009 – 19th Feb 2010) Phase 3 (19th Feb 2010 – remediation only)	1 July 2009 – 30 June 2014
Programme funding	$AUS 2.7 billion (reduced to $AUS 2.45, 2nd Nov 2009)	$NZ 347 million
Primary programme implementation	DEWHA	EECA
Intended outcome	Improved ceiling insulation for 2.2 million homes	$NZ 347 million spent towards improving insulation and space heating in homes
Predicted co-benefits	Economic stimulus Reductions in energy use and energy efficiency improvements Reduced CO_2 emissions Improved comfort	Health savings Economic stimulus Reductions in energy use and energy efficiency improvements Reduced CO_2 emissions Improved comfort
Retrofits available under scheme	Ceiling insulation	Ceiling and floor insulation (mandatory to have both if possible), vapour barriers, draught excluders, hot water pipe lagging, cylinder wraps, Clean heating (either a heat pump, wood burner, pellet burner or a reticulated gas heater)
Funding available to participants/consumers	$AUS 1600 per home for ceiling insulation (reduced to $AUS 1200, 2nd Nov 2009), paid by central government.	33% central government funding for ceiling and floor insulation, draught proofing, moisture barriers and remediation (maximum $NZ 1300), (60%* funding for low income households or landlords of low-income households – no maximum) $500 central government funding towards clean heater ($1200* if low-income household) *Additional funding /more complex funding arrangements available from charities, local government or lines companies
Service delivery model	Participants/consumers contracted directly with installer. **Phase 1 (3 Feb 2009 – 30th Jun 2009):** Claim required two quotes and a site inspection (Senate 2010). Money was paid upfront by participants /consumers who then claimed a rebate processed and paid by DEWHA. **Phase 2 (1st Jul – 19th Feb 2010):** Installers were paid rebate amount directly following claim via Medicare Australia. Consumers often did not pay any money at all. **Phase 3 (19th Feb 2010 - 30th Sept 2012)(DCCCE 2012):** Foil Insulation Safety Programme inspected all foil installations and either remove or install safety switch Home Insulation Safety Programme – inspected 15% of non-foil insulations Insulation Industry Assistance Programme – provided employment assistance to "reputable" installers affected by end of HIP (DCCCE 2012)	Contract between consumers and service providers. Consumers paid service providers directly (less rebate), service providers obtain rebate from EECA via claim. Consumers were able to pay via mortgage extension, interest free loan or via rates. Consumers in some cases had access to additional funding from third parties such as local government and charity organisations.
Number of installers/service providers:	10,834 registered, 7451 made one or more claim under programme (ANAO 2010)	60 companies/community trusts (some of which sub-contracted)
Criteria for inclusion of installers/service providers in programme	**Phase 1 (3 Feb – 30th Jun):** Installers did not need to register to be included in the programme **Phase 2 (1st Jul – 19th Feb 2019):** Installers registered online, were required to have either: "prior industry experience; a trade qualification in an approved trade; or insulation specific training." From 1st July 2009 to 17th Dec 2009 this applied to supervisors only, then all installers (ANAO 2010)	Energy Efficiency and Conservation Authority (EECA) held tendering process prior to start of programme with an additional request for proposals in September 2011. Contracts were subject to annual review. Criteria to be considered for inclusion included the capacity to deliver retrofits to agreed quality and safety standards. Industry experience was a key criteria (J. Lackey, personal communication 2010)
Performance auditing	Data analytics Inspect dwellings (roof inspection) Obtain information from installer (desktop audit) (ANAO 2010)	Service providers required to complete detailed audit following completion of retrofit. EECA carried out separate auditing and inspection procedures (J. Lackey, personal communication 2010)
Sanctions	"*Deregistration would apply to those installers who were fraudulent or intentionally breached the program's terms and conditions*" .(ANAO 2010, p 132)	Variety of sanctions available under contract with service providers including withholding of payments and expulsion from the programme (J. Lackey, personal communication 2010)
Related programmes	Energy Efficient Homes Package also included: Low Emission Assistance Plan for Renters (LEAPR) ("folded" into HIP 1st Sept 2009) Solar Hot Water Rebate	WUNZ:HS was predated by the smaller and more narrowly targeted EnergyWise Home Grants Programme (2003-2009). WUNZ:HS replaced by Warm Up New Zealand: Healthy Homes (2013-) Other related programmes such as Clean Heat ended in 2012 (J. Lackey, personal communication 2014).

8.2.1. Agenda-setting, policy formation and decision making

The Global Financial Crisis provided the context for setting the agenda for both the Australian and New Zealand programmes, with economic stimulus a common goal. Improved energy efficiency, reductions in energy use and thus reductions in CO_2 emissions were also goals of both programmes. However, while the HIP programmes had a stated goal of improving "comfort", my review of the grey literature does not suggest that expected health benefits were a major driver of the HIP, while these were an explicit part of the rationale for the New Zealand programme (Brownlee 2009).

Review of the HIP literature suggests that one of the fundamental explanations of the flawed decisions around the design and implementation of the HIP was the prioritising of the economic stimulus goal over other considerations such energy savings, or possible health outcomes as well as the safety of installers. The design and implementation of WUNZ:HS demonstrated a difference in priorities, with high barriers to programme entry and thorough auditing requirements which naturally compromised the speed of the economic stimulus. Following specific examples of differences in programme design, which may explain the different outcomes of the two programmes, I turn to the more general question of what might explain the differences in the programme design choices made.

8.2.1.1. Entry to programmes and installer characteristics

The differences in the ease of entry and resulting characteristics of installer organisations under the two programmes is useful starting point for specific design comparisons.

Registration as an installer organisation under the HIP at the start of Phase Two required a business to have a valid ABN (Australian Business Number) and to agree to the *Terms and Conditions of Registration.* Over 10,000 organisations registered although only 7,841 completed at least one installation (Senate 2010). The high take-up of the HIP limited the ability of Medicare to confirm that registered installers met programme entry requirements in a timely manner. By contrast, there were only about 60 service provider under the WUNZ:HS programme, 13 of whom had operated as service providers under previous programmes(J. Lackey, personal communication 2010) . Being a service provider under the programme required completing an annual review as a condition for contract extension, and acceptance of stringent auditing requirements (EECA 2009b).

The relative ease of entry to the HIP meant that a number of dishonest operators were able to enter the HIP, and operate unchecked, either billing for work not completed, or otherwise deliberately circumventing programme regulations. Hawke (2010, p viii) cites, "*disturbing claims about the high level of fraud perpetrated by unscrupulous operators. Despite*

some safeguards against fraud, no one foresaw the possible extent of potential malfeasance which was simply alarming [.]" Such small operators would not have been likely to have gained entry to the WUNZ:HS programme, given the extensive tendering process, and any service provider that did gain entry to the programme would have more to lose if found to be engaging in fraudulent practices.

A much greater concern for the HIP was the installing of insulation by unqualified and/or inexperienced installers, which was linked to the poor quality workmanship, dangerously installed insulation and the deaths of four installers which ultimately led to the end of the HIP. The HIP *Terms and Conditions of Registration* initially included requirements that all installers have generic occupational health and safety training and receive an installer pocket book with key safety and best practice information (from August 2009 onwards) and either have or be *supervised* by somebody with: "prior industry experience; a trade qualification in an approved trade; or insulation specific training" (ANAO 2010, p 104). The decision to require only supervisors to meet minimum competency standards was made by DEWHA in consultation with a range of stakeholders, and the trade-off between requiring minimum standards for all workers and just supervisors was explicitly considered, however; *" it would not be possible to implement, given the short period of time available and the need for low barriers to entry to the program"* (ANAO 2010, p 107). It was later reported that supervision was often minimal, meaning that unqualified and inexperienced installers were placed in very difficult situations (Lewis 2010). On February 12, 2010, just before the suspension of the programme, training was tardily made mandatory for all installers.

The request for proposals for WUNZ:HS did not explicitly set minimum installer competencies, but stated that when considering proposals 35% weight would be given to "capability" which included "*proven personnel and experience or capability in installing EECA approved products*" and "*evidence of ability to complete work to the Standards outlined in EECA's Audit and Quality Manual*" (EECA 2009b, p 12).

The fact that neither programme required minimum competencies for installers can be linked to the fact that the insulation industries in Australia (Senate 2010) and New Zealand (J. Lackey, personal communication 2010) were not highly regulated, beyond standard workplace and occupational health and safety rules, prior to the start of the two programmes and that home-owners could install their own insulation without any experience or qualification at all . Industry bodies in Australia opposed the imposition of formal training for all installers citing a culture of learning on the job (ANAO 2010). Turning to New Zealand, Lackey stated that historically, in New Zealand, insulation has been low-paid, unpleasant and prone to high turn-over (personal communication, 2010).

Non-mandatory installer qualifications were considered under WUNZ:HS (J. Lackey, personal communication 2010); however, ultimately the WUNZ:HS policy was based on demanding high quality retrofits but allowing service providers to make decisions about how best to achieve these standards. This approach also entailed a vigorous auditing process to ensure quality and also health and safety outcomes were met. Interestingly, poor quality workmanship was a concern early in the WUNZ:HS programme, even for service providers that had experience operating under previous EECA-administered programmes (J. Lackey, personal communication, 2010), meaning that installers working for organisations with a vested interest in continued programme participation were carrying out substandard work. This points to a fundamental issue with the insulation retrofit industry: the information asymmetries that mean that it is difficult for consumers to adequately assess the work carried out. It is difficult to visually assess the thickness of installed insulation or to assess coverage in inaccessible areas of ceilings and floors. This, in turn, encourages opportunistic sub-optimal retrofits. The solution EECA adopted was more rigorous auditing.

Initially service providers carried out a post installation audit and in addition EECA audited 5% of retrofit sites. This figure was lifted to 10% during the period when concerns about quality were being raised both internally within EECA (J. Lackey, personal communication, 2010) and by the media (Gibson 2010). Lackey also described more extensive actions taken when early audits revealed "safety relevant failures"; these included requiring service providers to check all previous work. It is unclear what proportion of installations were physically inspected during the operation of the HIP, although the programme design included site inspections and roof inspections intended to ensure that programme guidelines were being met and that installations met quality standards (ANAO 2010).

The difference in the outcomes of the two programmes cannot be explained by the programme entry requirements alone, but must also be linked to the rebate structures of the two programmes, which drove high uptake for the HIP and created perverse incentives and limited consumer vigilance in Australia

8.2.1.2. Rebate structure

Although the rebate structures of the two programmes changed over time the core difference between the rebate structures of the two programmes appears to be the fact that Phase Two of the HIP did not require a household contribution unless the total cost of an insulation retrofit exceeded a pre-set limit (initially $AUS 1,600, later reduced to $AUS 1,200), while WUNZ:HS only paid for a fixed proportion of insulation retrofit costs (in some cases with a maximum rebate size), and made a relatively small contribution to total heating retrofit costs (up to $NZ 500 or $NZ 1200 depending on Community Services Card status of occupants)[25]. In Australia, Phase Two of the HIP did not require any financial interaction between households and central government if the total cost of the retrofit was within the current limit, while Phase One required households to pay installers and then obtain a rebate (Senate, 2010). There was a massive increase in the rate of retrofits in the first month of Phase Two of the programme when installer companies were paid directly by Medicare (Senate 2010), suggesting that the rebate structure was highly influential on the speed of the uptake of the programme.

The vast majority of HIP retrofits during Phase Two were below the threshold for household contribution, meaning that households had little motivation to make informed decisions about which installation company to choose, which materials were appropriate or even to check to see that the contracted work was being carried out adequately (which is difficult even when householders are motivated, as noted above) (Senate 2010). Suggestively, when the maximum rebate was reduced from $AUS 1,600 to $AUS 1,200, the average total amount charged for each retrofit dropped, suggesting that installer companies were well aware of these incentives and either were willing to use cheaper (and presumably suboptimal) materials, or otherwise lower costs so that their customers did not need to contribute to the costs, had initially been favouring unnecessarily costly materials or installation choices, or most worryingly were simply overcharging for retrofits prior to the change in rebate size and thus able to reduce charges without changing the nature of the services provided.

By contrast the WUNZ:HS rebate structure can be assumed to have encouraged a relatively high degree of consumer vigilance. As with many of the other problematic aspects of the Australian HIP design, there was awareness of these issues amongst senior officials during the design phase, including explicit mention of these issues in the Minter Ellison risk register published in April 2009.

The rebate structure not only contributed to misaligned incentives, but also drove the

[25] A certain proportion of retrofits for properties occupied by CSC holders received additional third party funding meaning the occupants/landlords were not required to contribute financially.

rapid uptake of the HIP as noted above. While both the WUNZ:HS and the HIP programmes experienced higher initial demand levels than predicted, the number of retrofits carried out under the HIP rose greatly in each subsequent month peaking in November 2009, while demand levels in the WUNZ:HS programme began to fluctuate, requiring careful management by EECA (J. Lackey, personal communication, 2010). The rapid uptake of the HIP can be linked to many of the problems faced by the programme.

8.2.1.3. Product inclusion and installation methods

Although the two programmes had different scopes, the common element of both programmes was insulation. There are two insulation-related products that are particularly interesting as a basis of comparison, foil insulation and down-light covers.

Foil Insulation

In Australia, foil ceiling insulation was linked to many of the problems which troubled the HIP. Electrocution killed three installers of foil insulation, who inadvertently pierced live wiring with metal staples during installation (Senate, 2010), and poorly installed foil retrofits meant that many ceilings were potential fire or electrocution hazards later requiring checking, modification or removal under the FISP programme (ANAO, 2010). Prior to Phase Two industry consultation raised awareness of the risks of foil insulation but, as with installer qualifications, some members of the established insulation industry opposed changes to current practice which included the use of metal staples, citing the lack of previous issues with these products (Hawke 2010). The Senate Committee report concluded that it was the combination of inexperienced installers and products requiring skilled installation that led to the foil-related safety issues. Metal staples were removed from the programme in November 2009 by Minister Peter Garrett and foil insulation was entirely removed from the programme in February 2010 shortly before the end of the programme (Senate 2010).

In New Zealand, foil insulation was not available under the WUNZ:HS programme, although foil floor insulation had been available under the previous Energy Wise Home Grants programme until mid-2008 (J. Lackey, personal communication, 2010). In 2007, in New Zealand, several installers died of electrocution while installing foil insulation, (not installers operating under EECA administrated programmes), and Lackey stated that both safety and other issues such as a lack of durability led to the eventual exclusion of foil from EECA-funded programmes including WUNZ:HS. Interestingly, during a stakeholder meeting early in Phase One of the HIP the New Zealand deaths were discussed but this and other evidence did not dissuade officials from including foil in the programme (Senate 2010).

Downlights

Downlights[26] are the second insulation-related product of interest. In Australia, downlight covers were not mandatory under the HIP programme until November 2009, and prior to this a stringent building standard (AS/NZ 3000:2007) was adopted by the programme that mandated that downlight clearance be 200 mm. In mid-September 2009 a warning was issued to householders about the risks of insulation and downlights by the New South Wales Fire Brigades and the Office of Fair Trading (Hawke, 2010). A targeted inspection of 15,000 homes under the HISP revealed that 7.6% were fire-risks (Senate, 2010). Interestingly, the Australian Government contended that this rate of risk was not higher than in the general industry (particularly given the fact that inspections targeted poor quality installers), and the link between fire risk and the HIP has been contested (Senate, 2010). The Senate Committee report also presents several examples of industry warnings about downlight safety issues that failed to impact the initial design of Phase Two (Senate, 2010).

In New Zealand, downlight covers were also not mandatory under the WUNZ:HS programme and the clearance standard under the programme was lower than the HIP at 150mm (EECA 2009c). However, downlights were a concern under the WUNZ:HS programme; early audits revealed safety issues related to clearance (J. Lackey, personal communication, 2010). EECA was able to successfully address these issues with higher audit rates and sanctions for these "safety relevant failures" (J. Lackey, personal communication, 2010).

[26] Downlights are light fixtures that are recessed into ceilings. When insulation is installed above or adjacent to a downlight without a suitable downlight cover this creates a fire-risk via overheating. Not insulating above and adjacent to downlights reduces the effectiveness of ceiling insulation greatly thus downlight covers, which allow insulation to be installed over downlights, are an important consideration.

8.2.1.4. Differences in policy context

Having considered specific differences in the design of the two programmes, one of the key themes is the prioritisation of the stimulus goal under the HIP relative to the WUNZ:HS programme. WUNZ:HS was expected to produce both economic stimulus and a variety of co-benefits including health benefits (Howden-Chapman and Preval 2014). The HIP was primarily designed as a stimulus measure, with predicted energy co-benefits. In order to better understand the differences in priorities, it is useful to consider the wider context of official and public perception of the benefits of insulation and other retrofits in the two countries.

EECA is a good starting point for discussing differences in the policy context. Prior to the WUNZ:HS programme, EECA had been promoting and part-funding insulation retrofits for fourteen years. During this time EECA developed a productive long-term relationship with *He Kainga Oranga,* the Housing and Health Programme of the Department Public Health Department, University of Otago, Wellington and had co-funded the retrofitted insulation package, designed by EECA, used in the Housing, Insulation and Health Study, which began in 2001. This was consistent with EECA's core goal of improving energy efficiency (broadly defined to include any policy which changes energy use and increases *net benefits).* EECA's retrofit programmes and promotion of insulation, as well as research carried out by *He Kainga Oranga* in conjunction with EECA and many other agencies are likely to partially explain why at the time that WUNZ:HS was being considered (2008-2009) the connection between insulation and health was accepted across most political parties in New Zealand and by the general public. This assertion can be supported by two pieces of evidence.

Firstly, as discussed in Section 2.2., WUNZ:HS was implemented following a change in government. The previous Labour-led coalition government had proposed a $1 billion dollar insulation package that was intended to be rolled out over 15 years. While the programmes differed in certain key ways, this does indicate cross-party acceptance of a both a proven link between housing and health and also of the economic rationality of improving housing. Further evidence is provided by the fact that, in 2005, the Green Party's election campaign included insulation retrofits in its health platform rather than in its energy platform, a difference in approach from that taken by other parties at that time (Carroll, Blewden et al. 2008).

Secondly, two attempts have been made to monitor and measure the impact of *He Kainga Oranga's* research on policy makers and public perception and both suggest that it has been highly influential. The results of an unpublished review of New Zealand news articles and parliamentary speeches demonstrate that over the period 1996 -2010 there was

a steady rise, peaking in 2008, in the number of articles and speeches discussing both housing and health, and a rise in the proportion of such articles and speeches that used keywords associated with research (Fougere, Bennett, Howden-Chapman, unpublished). This rise coincides with the informal dissemination and then formal publication of key results from randomised controlled trials carried out by *He Kainga Oranga.* A report titled *The Social Sciences and Policy-Research Use* (Carroll, Blewden et al. 2008*)* produced by the Centre for Social and Health Outcomes Research and Evaluation, Massey University, North Shore also documented the influence of *He Kainga Oranga*'s research on policy makers and the public, stating that as research findings were disseminated, "*[o]ver time this flow brought about conceptual changes, altering the policy focus on household energy efficiency from a 'supply side' to a 'demand side' approach, and shifting the discourse around energy-efficient homes from one primarily of cost savings to an emphasis on health benefits"* (Carroll, Blewden et al. 2008, p 48).

The influence of this research may be explained by the close relationship between *He Kainga Oranga* and central government including its ongoing relationships with organisations such as EECA and the Ministry of Health (Carroll, Blewden et al. 2008). Many efforts were made disseminate research finding to policy makers (including early drafts of a paper which found a favourable ratio of benefits to costs for insulation and was ultimately published as (Chapman, Howden-Chapman et al. 2009)) and indeed research was structured to answer questions with policy relevance (Carroll, Blewden et al. 2008). In terms of the "national conversation", the common-place notion that cold and damp housing caused poor health has also been increasingly replaced by the idea that there is a research proven link between poor housing quality and health (Carroll, Blewden et al. 2008), as demonstrated by the rise in news articles linking research with housing and health.

Having argued that there was strong political and public acceptance of the health benefits of insulation retrofits in New Zealand when the WUNZ:HS programme was developed, it is necessary to ask if there was a different perception in Australia. I did not find research directly addressing this very specialised question but, based on keyword searches of the HIP-related literature that this chapter is based on, I conclude that health was not part of the HIP agenda, beyond some mentions of improved comfort, and that this likely reflects the lack of widespread acceptance by Australian politicians/policy officials and, presumably, the general public, of the connection between insulation and health. A case could certainly have been made for the health benefits of insulation in Australia, as many parts of Australia experience low winter temperatures and Australia does experience excess winter mortality (de Looper 2002), as well as potentially harmful extremely high temperatures which insulation might mitigate.

Ultimately, the reason for highlighting these differences in policy context is to suggest that, when decisions that might compromise quality of work were made when the HIP was designed, there may have been less "push-back" by politicians, officials and even the community who valued the co-benefits of insulation than there might have been in the New Zealand context. Beyond this difference in policy context, there are several key differences in government structure and organisational structure that may also potentially explain differences in the outcomes of the two programmes.

8.2.1.5. Government structure

Australia is a much larger country than New Zealand (22.7 million vs. 4.4 million population) and has three layers of government (federal, state and local) rather than New Zealand's two (central and local). Both the HIP and WUNZ:HS were funded and administered by central government. Review of the two programmes suggests that central government administration was not ideal, in the case of the HIP, but perhaps more appropriate in the case of WUNZ:HS.

Kortt and Dollery (2012) cite literature on fiscal federalism to argue that, because the HIP was intended to have benefits at both the local or regional level (power savings, employment) and at the national level (carbon dioxide emission reductions), there is not an a priori case for which level of government should fund and/or administer the programme. However, they note that central government has an advantage in terms of revenue gathering (via taxation) while state government may have a comparative advantage in the administering of programmes that require local knowledge and local experience. This suggests that the HIP would have functioned better as a centrally funded but state-administered programme, particularly given that many states had previous experience administering similar programmes, such as water conservation and clean heat programmes (Kortt and Dollery 2012), and a highly regulated programme operated by the State of Victoria, which provided a number of different retrofits and achieved relatively high quality outcomes based on audit results (Hawke, 2010). This argument is supported by the many complaints and concerns that State governments voiced during the design and implementation stages of the HIP, including concerns about costs resulting from the HIP such as contractual dispute resolution and fires (allegedly) which were incurred by the States, rather than the Federal Government (Lewis 2010, Senate 2010).

By contrast, New Zealand is effectively the size of a single Australian state, and relatively homogeneous, particularly with regard to factors such as climate. EECA, the implementing agency for WUNZ:HS, had considerable previous national experience administering similar programmes and thus, using the criteria presented in (Kortt and Dollery 2012), New Zealand's centralised approach to both revenue gathering and

programme administration may be considered optimal.

8.2.1.6. Differences in programme formation and organisational structure

At first glance the two programmes appear to have been formed in quite a similar manner, by a central government agency with strong direction from politicians and senior officials. However, the HIP operated with what is consistently described as a top-down model, Lewis (2010, p 86) states,*"[t]he Rudd government's approach to the HIP was very much top-down; DEWHA and other levels of government were expected to fall in line*". The Rudd government gave a great deal of input to the formation and ongoing operations of the HIP. The Federal Government consistently prioritised speed of stimulus over concerns from officials within DEWHA, who favoured a slower roll out over five years and from parties outside DEWHA, such as state governments and various industry bodies (Lewis 2010). DEWHA itself had limited experience in the design or administration of programmes of this type, which may explain some of its lack of success in moderating the stimulus imperative.

By contrast, WUNZ:HS appears to have been formed using a more collaborative approach; (J. Lackey, personal communication 2010) provided some insights in this regard. According to Lackey, the WUNZ:HS programme was developed following a request for proposals from "major government departments" by Prime Minister John Key after the "Job Summit", an event designed to tackle unemployment during the recession (J. Lackey, personal communication 2010). Following an initial proposal by EECA and its monitoring agency, the then Ministry of Economic Development, a more detailed proposal was requested. Strategic high level planning was driven by the Ministry of Economic Development and Cabinet officials, while EECA was able to provide detailed advice on programme specifics, using its previous experience administering similar previous programmes. The WUNZ:HS programme was developed iteratively, with some key decisions driven by the then Minister of Energy and Resources Gerry Brownlee, such as the expansion of the proposed programme to non-low-income households and making floor insulation mandatory where possible (J. Lackey, personal communication 2010). In addition, many of the foundations of WUNZ:HS were laid by the previous government who planned a scheme similar in many respects.

This iterative approach involving greater effective input from experienced policy officials seems to have produced a programme that was better designed and more closely monitored.

8.2.1.7. Consultation, risk assessment and independent policy advice

Consultation occurred prior to Phase Two of Australia's HIP, but it appeared to have little impact on the key elements of programme design described earlier. The comprehensive risk assessment carried out by consultants Minter Ellison in April 2009 similarly seems to have had little impact of programme design; many of the concerns highlighted in the register appear to have been ignored, with Minister Peter Garrett later accused of not having seen the register (Lewis, 2010).

Turning to New Zealand's WUNZ:HS, there was a long history of interaction between EECA and established service providers prior to the creation of WUNZ:HS. Lackey stated that both formal and informal consultation was standard practice (J. Lackey, personal communication 2010). For example, during the design of the WUNZ:HS evaluation the Group avoided research that would require interaction with customers and service providers based on EECA's consultation with service providers who did not wish to be further encumbered with time-consuming paper-work.

In order for consultation, or policy formation more generally, to function optimally it is necessary that independent sources of research and advice are available. In the case of Australia, noting the impact of the lack of an organisation which could provide these services, the Senate Committee suggested that the government "*should establish a dedicated and industry-independent program to research insulation systems and help develop efficient and effective insulation policy*" (Senate, 2010). By contrast, in New Zealand there are at least two organisations that fill this role.

Firstly, New Zealand has *He Kainga Oranga,* a publicly funded university-based research organisation much discussed throughout this thesis which has provided a great deal of independent, policy-relevant research on the value of insulation and heating retrofits.

Secondly, New Zealand has the Building Research Association of New Zealand Incorporated (BRANZ). BRANZ Ltd is an independent organisation funded by a levy on consented building projects with a value of over $NZ 20,000. This funding must be used to "*promote and conduct building-related science and research*" (BRANZ, 2010, p3) and is distributed to a number of research providers which are independent of government although not industry independent. BRANZ has contributed to the work of *He Kainga Oranga* and according to its own 2010 Annual Review, "[t]*he EECA-managed Warm Up New Zealand insulation programme, for example, owes a debt to BRANZ research on household energy use from the Household Energy End-use project (HEEP), as well as to BRANZ testing and assurance for insulation products, and technical knowledge about how to correctly specify and install insulation and the consequences of poor installation*" (BRANZ 2010, p 15).

EECA has ongoing relationships with both organisations, which has ensured that both the research findings and technical output of these organisations is readily accessed by policy officials and also that the research that these organisation produce has policy relevance. However, while it is tempting to agree with the Senate Committee that Australia would benefit from such organisations, it might take many years before the research findings of such an organisation were widely accepted by senior officials and politicians, and the policy landscape was therefore changed. As stated previously, the unwillingness of the Rudd Government and senior officials to consider modifications that might compromise the stimulus effect of the HIP (many of which might also have been suggested by an independent organisation of the type described), led them to give little weight to the results of consultation and external risk assessment. The difference between the outcomes of the two programmes cannot be explained by differences in the levels of consultation, risk assessment or indeed an absence of independent expert advice alone, but also by the degree to which consultation, risk assessment and independent advice were heeded.

8.2.2. Policy implementation

Many of the difficulties faced during the implementation of Australia's HIP resulted from choices made during the design phase of the programme. For example, the implementation of the HIP was made challenging because of the massive increases in monthly retrofits during the period July 2009 – Nov 2009 which resulted from specific design choices such as the rebate structure and low barriers to entry into the programme for installers. Systems such as the Medicare-operated rebate and registration systems did not function optimally, in part due to high levels of demand. This in turn meant that reports of poor workmanship and fraud could not be addressed in a timely manner as backlogs of complaints, and issues to be addressed quickly accumulated, thus making the HIP's monitoring and auditing programme, which included desk-based audits and physical inspections, relatively ineffective (ANAO 2010).

In New Zealand, EECA also experienced many difficulties in the early stages of WUNZ:HS, including capacity issues, but these difficulties did not overwhelm its systems (J. Lackey, personal communication 2010). One of the early challenges faced by the WUNZ:HS programme was poor quality workmanship, as described in Subsection 8.2.1.2. However, EECA was able to improve performance by raising external audit rates and in some cases demanding that offenders check all previous work. Such an approach would be much harder to implement with many small providers and greater uptake rates.

While the two programmes did have different rates of uptake, differences in programme

implementation were also influenced by differences in organisational experience. EECA's long history of administering similar programmes to WUNZ:HS gave it an advantage relative to DEWHA, who, as a policy-development -based agency, had not implemented a programme of this kind. The Senate Committee report notes that DEWHA did not have staff with experience in the insulation industry and that the management structure was not suitable for a programme of the size of the HIP. The Committee report also states that until November 2009 the HIP was managed from within a division with many other diverse areas of responsibility (Senate, 2010). By contrast, EECA has a much more limited remit than DEWHA, which meant that staff were already familiar with issues around insulation and heating retrofits and better equipped to deal with difficulties as they arose.

In Australia, in addition to capacity issues, there appears to have been reluctance by politicians and senior officials to take effective action to address problems during the implementation phase if this might compromise the speed of the roll-out, again reflecting the over-riding stimulus goal (Lewis 2010). Although numerous modifications to the programme did occur, many changes, such as the requirement that all installers have training, which was implemented two days before the suspension of the programme, clearly happened too late.

8.2.3. Policy evaluation

There does not appear to have been a formal evaluation process built into Phase Two of the Australian HIP prior to the start of the programme, although Phase One of the programme did serve as a design and process evaluation tool, as did the various consultation processes, stakeholder meetings and the creation of the Minter Ellison Risk Register, many of which were informed by the operation of Phase One. Ultimately the quickly commissioned independent Hawke Review served as a design and process evaluation for Phase Two, helping to finalise the suspension of the programme, however as noted in the Senate Committee report, *"[t]he government's move to commission an independent review of the HIP (the Hawke Review) was too little, too late and should have been undertaken earlier so that the findings could be used to improve the HIP".* Naturally, ad hoc evaluation occurred throughout Phase Two of the programme leading to changes to registration, product availability and training requirements.

By contrast, the WUNZ:HS evaluation, which was outcome -based, had funding allocated from the start of the programme, and was timed so that results could lead to programme modification or even end the programme.

Unlike design and process evaluations such as the Hawke Report, outcome evaluations such as the WUNZ:HS evaluation, which is the primary focus of this thesis, assess whether

the implementation of the programme achieved its stated goals. In the Australian case, it seems likely that the absence of a formal outcome evaluation process reflects the haste with which the HIP and other stimulus programmes were rolled-out. In addition, it may have seemed less necessary to formally analyse the outcomes of a programme such as HIP because the primary goals of the programme, economic stimulus and job creation for the construction industry during a time of recession might seem to be able to be quantified relatively easily. The Australian National Audit Office report states that, *" [i]n terms of outcomes, it has been estimated that between 6000 to 10 000 jobs have been created. While, clearly, the creation of these jobs was an important outcome in the face of the downturn in the economy, these jobs were shorter-lived than intended due to the early closure of the program…."*. However, the comprehensive report produced by Denne and Bond-Smith (2011), which attempted to capture the industry impact of the WUNZ:HS programme, makes clear that precisely evaluating such impacts is far from straight forward.

8.2.4. Conclusion and recommendations

The comparison of the HIP and the WUNZ:HS programme has the potential to inform future insulation retrofit subsidy -based programmes, both in New Zealand and Australia, and also internationally. This chapter has identified several key differences between the programmes that contributed to their differing outcomes. These differences could be useful levers for policy makers when designing future programmes, as the highly favourable benefit:cost ratios calculated in Grimes et al. (2011) may not have been achieved had the design or implementation of the programme been closer to that of the HIP.

One of the repeated themes of this chapter was the HIP's consistent prioritising of the desired economic stimulus over safety and quality. Political and senior official buy-in to the value of high quality work is vital if health and energy outcomes are to be achieved, because there is so much potential in programmes of this kind for rort, or low quality workmanship if programme design and implementation choices do not account for the nature of the insulation industry or create an unmanageably quick up-take. In the absence of such an extreme focus on stimulus, some of the following recommendations might appear self-evident.

Key recommendations:

- A substantial timeframe for policy design, with strong input by non-government stakeholders and meaningful engagement with international and local research.
- Development of ongoing relationships with major insulation installation organisations.

- Rebate designs which encourage consumer vigilance and engagement: the consumer should contribute directly financially if this is possible. This is also likely to limit uptake speed which is potentially beneficial. This imperative should be balanced against the need to target low-income families who may be unable to contribute financially but who are more likely to gain health benefits from an insulation retrofit.
- Training should be pragmatically addressed: ultimately the programme must be designed to ensure that all installers are trained/experienced or properly monitored by trained/experienced supervisors. This may or may not entail mandatory training.
- Limit programme participation to larger organisations and impose entry requirements based on quality, safety and capacity, thus reducing the likelihood of "cowboys" etc.
- Rigorous auditing procedures and meaningful sanctions imposed in a timely manner.
- Incorporate timely process and outcome evaluations, such as the economic evaluation of the WUNZ:HS programme.
- Need some ability and willingness to control programme rollout speed.
- Account for local/regional knowledge and differences in conditions. This suggests that central government implementation may not be optimal in larger countries.

CHAPTER 9. DISCUSSION AND CONCLUSION

The primary aim of this thesis was to examine methodological issues around the analysis of administrative data in large-scale public health interventions, using data available from the WUNZ:HS evaluation as the basis of a case study and exploring a subset of plausible approaches. The secondary aim of this thesis was the exploration of issues of programme design and evaluation generalisability using a policy-based comparison of the WUNZ:HS and HIP programmes as a case study. Beyond the exploration of specific methods of statistical and econometric analysis, and retrofit programme specific recommendations, it is interesting to consider the broader implications of this work.

9.1. Discussion

The discussion section first summarises the key results of this thesis and what it adds to the body of literature that explores the impact of insulation and heating retrofits on health. Secondly, it places this thesis in the broader context of the literature on large-scale public health interventions. Thirdly it explores the strengths and limitations of the WUNZ:HS evaluation design and dataset and addresses the lessons learnt from the WUNZ:HS HIP comparison. Finally it notes future research directions suggested by this thesis.

9.1.1. Key results in the context of housing, heating and health literature

The most prominent and important result of this thesis was the finding of a highly statistically significant reduction in mortality hazard for individuals aged 65 and over with a baseline circulatory hospitalisation who received insulation, relative to control group individuals (HR of 0.673; 95% CI 0.535 - 0.847; $p<0.001$). A similar effect of insulation, although less statistically significant, was observed for individuals aged 65 and over with baseline respiratory hospitalisations relative to the control group (HR 0.8; 95% CI 0.635 – 1.008; $p=0.058$). A review of two meta-analyses of the relevant literature did not find any research on the impact of insulation or heating retrofits on mortality (Thomson, Thomas et al. 2009, Thomson, Thomas et al. 2013).

These results have some plausibility given New Zealand's pattern of excess winter mortality and morbidity (Davie, Baker et al. 2007, Telfar-Barnard 2010). In the case of individuals with pre-existing cardio-vascular illness plausible mechanisms underlying these findings may include reductions in the number of cardiovascular events occurring due to thrombosis due to haemoconcentration or other cardiovascular reflexes induced by low temperature (Keatinge, Donaldson et al. 1997) in the treatment group, based on the reasonable assumption that insulation retrofits led to increased indoor temperatures. In the case of individuals with pre-existing respiratory illness, plausible causal mechanisms driving the observed difference in mortality include reductions in the frequency of severe respiratory illness in the treatment group driven by warmer indoor temperatures as cold temperatures impact immune resistance and are associated with longer bacterial survival time in droplets (Keatinge, Donaldson et al. 1997). The limitations of these findings are discussed in Section 9.1.4: essentially the absence of cause-of-death data and the limitations of a quasi-randomised study of this type limit the strength of the conclusions that can be drawn.

The other key findings were statistically significant reductions in all-cause monthly hospitalisation costs at the household level of $3.18 per month (95% CI 1.83 - 4.54; $p<0.001$) as a result of receiving an insulation retrofit. A review of the literature suggests that this is

the first example of research involving a control group which has found statistically significant changes in hospitalisation costs as a result of an insulation retrofit based intervention. A previous English intervention study did not find any statistically significant impact of improved housing (including insulation and central heating retrofits) on hospitalisation costs (Barton, Basham et al. 2007), but this finding may be explained by a much smaller sample size. The statistically significant reductions in all-cause hospitalisation costs as a result of receiving an insulation retrofit observed are consistent with New Zealand research which found statistically significant reductions in all-cause acute hospitalisation risk following participation in the Healthy Housing Programme (which included a suite of advice and home modifications including insulation retrofits) for individuals aged 5-34 (Jackson, Thornley et al. 2011), and non-statistically significant but suggestive reductions in hospitalisation rates for individuals who occupied homes that received insulation as part of the Housing, Insulation and Health RCT (Howden-Chapman, Matheson et al. 2007). Although the underlying mechanisms are difficult to ascertain when data is analysed at the household level, the separate analyses of circulatory illness and respiratory illness-related hospitalisation cost differences strengthens the conclusions drawn, as the original model -based analysis found statistically significant reductions in circulatory illness-related costs ($1.40, p=0.013) and respiratory illness-related hospitalisation costs ($0.50, p=0.083). Results also seem consistent with the reductions in respiratory symptoms and GP visits (self-reported) seen in previous RCTs (Howden-Chapman, Matheson et al. 2007, Howden-Chapman, Pierse et al. 2008), and with the health results reported in meta-analyses of housing retrofit studies (Thomson, Thomas et al. 2009, Thomson, Thomas et al. 2013).

The findings of reduced all-cause pharmaceutical use *costs* at the household level for both insulation retrofits ($2.01; 95% CI 1.52 – 2.50; p<0.001) and heating retrofits ($3.01; 95% CI 1.43 – 4.59; p<0.001) also appears to be a novel result. The only directly comparable study revealed by a literature review, the Watcombe Housing Study, did not find any statistically significant change in individual level pharmaceutical use costs as a result of participation in an intervention which involved a suite of housing upgrade including heating and insulation (Barton, Basham et al. 2007). The difference in outcomes might, again, reflect the much greater size of the WUNZ:HS evaluation cohort. The reductions in pharmaceutical use costs seem consistent with previous He Kainga Oranga research finding reductions in respiratory symptoms following an insulation retrofit (Howden-Chapman, Matheson et al. 2007) and a heating retrofit (Howden-Chapman, Pierse et al. 2008). The Housing, Heating and Health RCT also reported statistically significant (at p<0.1 level) reductions in asthma reliever puffs and frequency of night-time use (Howden-Chapman, Pierse et al. 2008) following a clean heating retrofit.

Having placed the key health-related results found in this study in the broader context of the published literature it is now useful to attempt to place the results of this thesis in the broader context of the large-scale public health intervention research.

9.1.2. Thesis in broader context of large-scale public health intervention evaluation research

The body of large-scale public health intervention evaluation literature is vast. A literature review suggests that the specific health-related results reported in this thesis are novel and contribute to the body of knowledge on the impact of housing retrofit interventions. The analysis of health cost data using a household level difference-in-difference fixed effects-based model also appears to be novel, and proved a useful method for addressing some of the problematic characteristics of health cost data. This approach may prove useful for future researchers dealing with similar datasets; however, it is only appropriate when a dataset is created using a matched design and where the intervention is delivered at the level of the household/home.

The natural policy experiment that was the HIP WUNZ: HS comparison was an opportunity to consider the ways in which policy design decisions and implementation challenges can influence the outcomes of large-scale public health interventions. Although it is difficult to state this conclusively given the large body of grey literature in this field, an initial search suggests that the combination of thorough statistical analysis and opportunistic policy-based programme comparison that this thesis represents is unique in the public health intervention evaluation literature, if only because it is rare that such an ideal natural experiment occurs.

It is now interesting to turn to the specific strengths and weaknesses of the WUNZ:HS evaluation (and thus the dataset that the statistical analyses in this thesis are based on).

9.1.3. Strengths of the WUNZ:HS evaluation

A useful starting point for the discussion of the strengths and limitations of the WUNZ:HS evaluation is to briefly review of the circumstances that led to our Group successfully receiving funding to carry out the WUNZ:HS evaluation. The Group won the contract to carry out the evaluation by making a successful proposal which was designed with available funding and specified timeframes in mind. Although funding for an evaluation was built into the WUNZ:HS programme from the outset, the design of the programme did not specifically facilitate evaluation, for example by incorporating randomisation. With this background in mind, it is interesting to consider some of the strengths and weaknesses of the WUNZ:HS evaluation and dataset, and to what extent lessons can be drawn from these strengths and weaknesses.

The greatest strength of the WUNZ:HS evaluation design was the ability to link EECA-provided address data with QV held housing condition data and NHI coded household occupant health data in order to identify suitable matched controls and to construct a study cohort with pre-test and post-test data available for nearly one million individuals. This large dataset was valuable because it enabled the assessment of the impact of WUNZ:HS retrofits on rare events such as mortality and hospitalisation. From a research perspective, as noted above, this dataset was a unique opportunity to extend research on the impact of housing retrofits to evaluate outcomes that previous research did not have the statistical power to assess.

A related strength of the WUNZ:HS design was its ability to obtain anonymised hospitalisation, pharmaceutical use and mortality data for study addresses and matched controls without requiring consent from each anonymised study "participant". Because consent processes are time consuming and resource intensive, had the design required consent to be obtained the evaluation would either have involved a smaller cohort or been much more costly (and thus unlikely to receive funding). Consent processes would have likely slowed the speed of the evaluation, reducing its ability to inform decisions about the ongoing WUNZ:HS programme. Another extremely important benefit of the anonymised design was the elimination of potential biases that might be introduced by differences in the characteristics of individuals who did or did not consent to participate in the study: we can imagine that impoverished or unwell individuals/households might be less likely to participate leading to results with limited generalisability.

9.1.4. Limitations of the WUNZ:HS evaluation

A number of the limitations of the WUNZ:HS design were discussed in Chapter Three. These included loss of data at each stage of the matching process, uncertainty about the validity of NHI derived addresses and the decision not to attempt to obtain cause of death data.

Data quality is a key concern when using administrative data to evaluate an intervention. As discussed in Chapter Three, administrative data will almost inevitably be of poorer quality than the data produced by a well-run RCT, due to the absence of double entry and other data quality protocols. This reflects the different purposes for which these different data are collected. In general, we might expect errors in administrative data to introduce statistical noise to analysis; however, it is also possible that patterns in errors may be differential and thus introduce a bias into analysis. Issues with matching addresses in the initial dataset construction process meant that only 79.7% of treatment addresses provided by EECA were able to be linked with QV data. It is likely that these unlinked addresses were problematic due to idiosyncratic addresses or data entry errors in either the EECA or

QV datasets. It is difficult to form a strong hypothesis about the nature of any bias that these mismatches introduced because no health or housing data was available for unmatched households and their occupants. More generally, it is difficult to assess the nature of any bias resulting from inaccurate or missing administrative data; however the WUNZ:HS evaluation did provide one potential opportunity to assess the impact of inaccurate or inconsistent data: the comparison of the CAUs (Census Area Units) of NHI address records from central NHI records with those provided with hospitalisation records. The NHI address field was the basis for linking EECA provided household addresses and matched control group addresses with the NHI linked health data of occupants. The implications of unmatched NHI data were explored in Subsection 3.3.1.6 by comparing hospitalisation record CAUs with NHI CAUs, and while there were clearly patterns in the rate of mismatches by age and ethnicity, it was reassuring to note a similar proportion of mismatches for the treatment and control groups (17% vs. 21%) and it is difficult to form a strong hypothesis about the direction or strength of any bias that mismatched addresses might produce. Instead, it seems likely that mismatches simply introduced statistical noise into analysis, reducing the power of the evaluation. It was also reassuring, given the large contribution of reduced mortality-related benefits for people aged 65 and over to the calculation of total annual household health benefits, that individuals aged 60 and over had the lowest rate of mismatches (11.5%).

A more serious data-related concern was the decision not to obtain cause-of-death data, due to concerns about an anecdotal lag, limitations of the Mortality Collection data, and then later due to concerns about anonymity (see Subsection 3.3.1.5). The mortality models constructed in Chapter Four attempted to address this issue by limiting analysis to vulnerable individuals, aged 65 and over who had a hospitalisation of a plausible type in the *before* period of the study; however, some uncertainty about the plausibility of the results found, and likely causal mechanisms, will always remain in the absence of cause-of-death data. Despite the possible limitations of the Mortality Collection data, I would recommend that future evaluations of this type carried out in New Zealand should attempt to obtain this data as it would have strengthened mortality-based analyses.

A related limitation, which is, in fact, a limitation of any cost-benefit analysis which relies on VSLY based calculations, and indeed perhaps any cost-benefit analysis at all, is the spurious precision of benefit:cost estimates. In my consideration of the valuation of mortality avoided in this thesis it was necessary to make some informed assumptions about the life-gained by those individuals who avoided mortality as a result of participating in WUNZ:HS. While clearly sign-posted as assumptions, the impact of such assumptions may be difficult to keep track of as part of a complex multi-benefit cost-benefit analysis, and render the *precision* of calculations elsewhere in an analysis somewhat

irrelevant. More generally, it is a concern that while cost:benefit study authors may attempt to provide a range of benefit:cost estimates, reflecting uncertainty, policy makers who interpret such results may lack the capacity to account for the range of plausible results and focus entirely on the central estimate.

Another limitation of the WUNZ:HS design was that, although it was possible to stratify analyses based on the CSC status of treatment households (or the CSC status of "matched treatment household" in the case of control group individuals in the individual level mortality analyses), we were not able to obtain accurate information about the true CSC status of control group households and individuals. This biases the results of stratified analyses towards the null and thus limits the value of the interpretation of stratified analyses from a policy perspective. We did not explore the possibility of attempting to link CSC data held by Work and Income with QV's address data. In retrospect it may have been possible to obtain this data but, because it was not essential to the primary task of the evaluation which was the assessment of the WUNZ:HS programme as a whole, it may have been difficult to justify the expense and time that such matching would likely involve.

Another unavoidable limitation of the WUNZ:HS evaluation design was the absence of individual level health relevant data beyond that available through NHI linked datasets. Potential confounders not captured included smoking behaviour, ongoing health conditions and BMI. Of course, it would not have been possible to obtain such data without a multiyear data collection programme involving consent processes, interviews and costly measurements, which would have been unfeasible given the timeframe and budget available and/or necessitated a much smaller study cohort, which would have lacked the power of the current evaluation.

Similarly, the absence of information about the insulation or heating status of control group households means that these important potential matching variables or confounders cannot be accounted for. The difference-in-difference based study design used to analyse hospitalisation and pharmaceutical-use costs may limit the impact of this issue somewhat because fixed effects terms in these models can account for the unchanging impact of treatment and matched control housing characteristics (which includes the insulation and heating characteristics of the control group households). However, fixed effects terms cannot account for unknown changes in the insulation and heating status of the control group households within each matched pair over the course of the study (or indeed any change in the insulation or heating status of the treatment group households that was not part of WUNZ:HS). In addition, the Cox Proportional Hazards model used in Chapter Four does not account for baseline/pre-treatment

mortality rates and thus there is greater potential for misclassification of control group households to impact on the analysis. In this case, if a proportion of the control group individuals occupied homes that were recently retrofitted with insulation then the true health impact of the study would be attenuated; this would bias results towards the null. Such biases are a serious concern, but more acceptable than biases in the opposite direction.

Considering the absence of data on potential confounders and/or matching variables leads naturally to the largest concern about the WUNZ:HS evaluation design, which is the possibility of selection bias, leading to differences in relevant characteristics of the treatment and control groups. Assessment of the distribution of potential confounders age, sex, and ethnicity, in Chapters Three and Four, did not reveal large differences by treatment group status, suggesting that the quasi-experimental design of the study had achieved a reasonable match of treatment and control groups. It can be inferred that the distribution of unmeasured potential confounders might also be equal.

In addition to quasi-randomisation, selection bias was addressed by the inclusion of potential confounders in multiple regression models. Model choice also addressed selection bias. The difference-in-difference-based fixed effects OLS estimator models presented in Chapters Five and Six eliminated confounding by time invariant characteristics of treatment households and their matched controls. Unfortunately, it was not possible to address the effect of baseline differences in mortality risk in the mortality analyses presented in Chapter Four, because the inclusion of mortality data from the *before* period would unavoidably lead to biased results, given the inability of people who died to participate in the WUNZ:HS programme and the reduced likelihood that the families of such people would participate. In this context, analysis had to rely on post-treatment data, and thus greater information on possible individual level confounders such as smoking status or BMI would have been particularly useful.

A greater concern, from a modelling perspective, is that the quasi-experimental WUNZ:HS evaluation design may be unable to avoid a more subtle selection bias, which could not be addressed using a difference-in-difference approach or even by the incorporation of very detailed data on individual-level potential confounders. An example of such a bias is the possibility that the decision to participate in WUNZ:HS was driven by, or at least indicates, a particular health status; for example, awareness of the value of insulation following a recent illness or targeting by third party organisations who encouraged vulnerable individuals to consider participation in the programme, or even funded their participation. In such cases, regression to the mean effects (Bland and Altman 1994, Shephard 2003) might bias analyses toward the calculation of a beneficial effect.

Addressing this type of selection bias would have required either the incorporation of some randomisation in the WUNZ:HS *programme* design which could then have been taken advantage of in an evaluation, or a full RCT-based evaluation. The former option would have been preferable, although incorporating some form of randomisation to the programme might have been impractical, given the complex structure of the WUNZ:HS programme which involved approximately sixty service providers. A full RCT could be imagined, incentivising household participation in the trial by funding some or all of the costs of programme participation and then randomising the timing of entry into the programme, with some households receiving treatment early (the treatment group) and others later (the control group). This approach could also avoid some of the weaknesses of an entirely administrative and commercial dataset-based evaluation, allowing for the confirmation of occupant identities and the collection of data not linkable to NHI numbers; however, it would either involve a smaller cohort and thus have limited statistical power to assess rare events, or it would be much more costly. A further disadvantage of an RCT is the possibility of limited generalisability (Victora, Habicht et al. 2004), as households who were willing to participate in the trial might not be representative of the average household who participated in the programme.

Ultimately, when the difficulties of alternative approaches are considered, it becomes apparent that many of the weaknesses of the WUNZ:HS programme are inextricably linked to its main strengths: the construction of a very large cohort using a quasi-experimental design in a timely and cost-effective manner. Given the absence of programme-based randomisation, and the constraints of funding and timeliness, the innovative quasi-experimental design of the WUNZ:HS evaluation was likely the best approach available.

Consideration of the strengths and weaknesses of the WUNZ:HS evaluation and dataset leads to a related topic, the extent to which design and process-based evaluations can contribute to the interpretation of the results of an outcome evaluation of this type.

9.1.5. Lessons of from the WUNZ:HS HIP policy comparison

The WUNZ:HS HIP policy comparison was a fascinating exercise, a natural experiment that illuminated salient characteristics of the two programmes and the complex relationship between research, consultation, risk assessment and decision making. It is interesting to consider the ways in which different programme design and implantation choices might have influenced the outcome of the WUNZ:HS programme, perhaps even resulting in widespread dissatisfaction and early programme termination. Identifying key design choices and implementation strengths of the WUNZ:HS programme was made easier by comparison with the HIP, for which programme design and implementation flaws had been well documented. In the absence of this counterfactual it would have been more difficult to identify programme elements that were key to the WUNZ:HS programme's success.

This is relevant when we place ourselves in the position of a policy maker considering evidence from an outcome evaluation carried out in a different time or place, and attempting to use this evidence to inform decisions around a possible intervention. Such a policy maker would be wise to consider any design and/or process evaluations of previous interventions, documentation of decisions around design and process from such interventions, or to consult the original policy makers. This would inform key programme design choices (for example whether middle-income households should contribute financially to a planned retrofit programme). Beyond noting the general principle that similar outcomes are more likely for programmes that make similar key design and implementation choices it is difficult to make general recommendations; the programme design and implementation elements that were key to success (or failure) must be identified in each case. In this context it is interesting to note the value of careful documentation of design and implementation by government organisations, which will then potentially be available to inform future policy makers considering similar programmes. Of course, one of the key lessons of the WUNZ:HS HIP comparison was that previous evidence is only one of the factors weighed by decision makers, and a policy imperative such as economic stimulus may ultimately override other concerns.

Having briefly considered the generalisation of the lessons of the WUNZ:HS HIP comparison the final topic of discussion is the future research directions suggested by this thesis.

9.1.6. Future research directions

The primary research agenda suggested by this thesis is the further exploration of the key result, reductions in mortality for vulnerable individuals as a result of receiving insulation. This could happen by gaining ethics approval to attempt to obtain cause of death data for the cohort (which might not be possible due to anonymity concerns), or by collecting cause-of-death data in a similar future evaluation, trial or cohort study. As large-scale retrofit programmes of this type with funded evaluations are rare, and retrofit-based RCTs may be unlikely to have sufficient statistical power to analyse mortality data given the high cost of such trials, opportunities to explore this area may realistically be limited and ultimately the results presented in this thesis may constitute the best available evidence for a long time to come. Similarly, because cost-based hospitalisation and pharmaceutical data proved difficult to analyse at an individual level in this thesis, it would be interesting to assess whether a future trial or evaluation with a larger cohort and/or more information about participant characteristics would have sufficient statistical power to reliably model changes in costs as a result of receiving an insulation or heating retrofit at an individual level. Individual level analysis might more thoroughly illuminate the causal pathways by which changes in insulation and heating impact hospitalisation and pharmaceutical use costs. Similarly, the application of other regression methods such as quantile regression, which was not explored in this thesis, might better account for the problematic characteristics of the dataset.

Another important area for future research in the New Zealand context would be to attempt to verify the accuracy of NHI-based addresses, leading to greater confidence in the use of this administrative dataset for research purposes. The most likely source of comparison would be New Zealand Census data, however, as both NHI and Census data is anonymised and carefully secured it is difficult to imagine this occurring without co-operation between the organisations that own the two datasets, and given that the primary function of the NHI dataset is administrative, it is unlikely that the impetus to resolve this question would come from the Ministry of Health, who maintain these data.

Future research quantifying the Value of a Statistical Life (VSL) in New Zealand would also be particularly valuable. Section 4.3 and Chapter Seven of this thesis demonstrate the impact that different mortality modelling assumptions can have on benefit:cost calculations. Any research that further standardised these calculations would be very useful. Of particular interest would be the derivation of VSL figures for a range of mortality categories including “death by illness”; comparison of New Zealand transport VSLs (Miller and Guria 1991, Guria, Jones et al. 2003, Guria, Leung et al. 2005) and fire-related VSL (Sanderson, Goodchild et al. 2007) confirms the likelihood, consistent with the international literature, that New Zealanders value death by illness differently to death

from other causes. A unified body of research on VSLs, presenting WTP (and/or WTA) survey-based VSLs for a range of likely mortality categories would be a very useful resource for researchers attempting to assign dollar values to changes in mortality, as would calculation of standardised VSLYs based on explicit assumptions about the average age of survey participants, their implicit discount rates, and predictions about their average future lifespans. Such tools would greatly aid cost-benefit researchers, standardising assumptions so that programmes with mortality co-benefits could be more easily evaluated and compared with other programmes, thus reducing the risk that idiosyncratic mortality-benefit evaluation choices might overly influence results and comparisons. It would also be useful to further explore issues around any variations in VSL by age and health status in New Zealand, although it might ultimately be more prudent and politically neutral for researchers to avoid any such value-based adjustments when calculating benefits and costs.

A further research question noted briefly in Subsection 3.3.5 of this thesis is the potential to evaluate the impact of draught-proofing under WUNZ:HS on health and also energy outcomes. Draught-proofing is very cheap and it if a clear benefit could be shown, however small, it could well be shown to be an effective policy option for future retrofit programmes. However, the effectiveness of draught-proofing might be difficult to distinguish from that of insulation, given the high proportion of households that received both. This is a possible area for future research and one that the WUNZ:HS dataset may facilitate.

Another future research question suggested by this thesis and my own previous work in the field of cost-benefit analyses of heating retrofits is an exploration of whether energy and health gains from retrofits observed immediately after retrofitting continue to accrue over the working life of retrofit products (as cost-benefit analyses assume in the absence of strong evidence to the contrary). This would also require field-work to assess what the real-world working life-time of various retrofits is (typically estimated as 20-30 years for insulation, 10-15 for space heating), and if there is reduced performance over the course of the life-time of the retrofit product. Anecdotal evidence from an expert in the field (personal communication, name withheld) is that heat-pump maintenance by owners is less frequent than that typically recommended by manufacturers (who recommend yearly servicing), and insulation is known to “settle” over-time, reducing its effectiveness. One avenue for such research would be a follow-up evaluation of the WUNZ:HS cohort: such an evaluation might produce persuasive results if assessing ongoing energy use impacts of WUNZ:HS retrofits but might be of limited value in the assessment of ongoing health impacts due to uncertainty about the accuracy of NHI address matching, which was shown in Subsection 3.3.1.6 to decrease in quality over time.

A final avenue for future political or sociological research suggested by this thesis would be a more thorough exploration of the influence of *He Kainga Oranga* on housing policy in New Zealand. The success of *He Kainga Oranga*, culminating in the WUNZ:HS programme, appears to be a paradigmatic example of high quality research appropriately contributing to the rational policy cycle. However, the lower level of funding for the subsequent WUNZ:HS Healthy Homes programme ($100 million vs. $347 million), and the current absence of any budgeted funding beyond the close of that programme appears to represent non-rational decision making, given that only 25% of the homes in New Zealand with inadequate insulation before WUNZ:HS now have adequate insulation (personal communication, source withheld) and such favourable benefit: cost results have been calculated previously. Of course, it is only conjecture to describe this as non-rational policymaking given the absence of information about the decision making process and the costs and benefits of alternative courses of action.

A complete description of the mechanisms of policy influence would need to consider both "rational" and "non-rational" explanations for both the success and then the seeming subsequent loss of influence of *He Kainga Oranga*'s research. Research would need to trace in greater detail the progress of ideas and influence, via both rational and "real world" processes such as media influence, relationships and political strategy. This research would ideally involve unguarded participation from politicians, senior officials and *He Kainga Oranga* researchers. The subject has been documented to a limited extent in an article (Fougere, Bennett, Howden-Chapman, unpublished) and a report (Carroll, Blewden et al. 2008*)*, however, it may be the case that a more detailed account may only be possible many years in the future, when concerns about the political implications for participants in such research might have receded.

9.2. Conclusion

This thesis carried out an exploration of methodological issues associated with the analysis of health cost data using the WUNZ:HS evaluation dataset as the basis of a case study. Methodological issues regarding the use of administrative data were discussed, suitable approaches to assigning value to changes in mortality were evaluated and a subset of suitable models for evaluating problematic hospitalisation and pharmaceutical use cost data were explored.

Specific conclusions included the finding of statistically significant reductions in all-cause mortality for individuals aged 65 and over who had a baseline circulatory hospitalisation and who occupied a household that received a WUNZ:HS insulation retrofit (HR of 0.673; 95% CI 0.535 - 0.847; $p<0.001$). There was also a statistically significant reduction in all-cause mortality for individuals aged 65 and over who had a baseline respiratory hospitalisation (HR 0.8; 95% CI 0.635 – 1.008; $p=0.058$). Households that received insulation had reductions in monthly all-cause hospitalisation costs relative to their matched controls of $3.18 per month (95% CI 1.83 - 4.54; $p<0.001$). Households that received insulation had reductions in monthly all-cause pharmaceutical use costs relative to matched control group households of $2.01 (95% CI 1.52 – 2.50; $p<0.001$). The results indicating a health-protective effect of insulation are consistent with each other, plausible, and consistent with previous research (Howden-Chapman, Matheson et al. 2007). The only statistically significant impact of heating on health detected in the current analysis was that households that received heating demonstrated reductions in monthly all-cause pharmaceutical use costs of $3.01 ($p<0.001$) (95% CI 1.43 – 4.59).

The primary estimate of the annual household-level health benefits of insulation was approximately $860 using a conservative ($p<0.05$) threshold and $1,173 using a ($p<0.1$) threshold. The annual household-level health benefits of heating were calculated to be approximately $41 using either threshold. There was a great degree of variation in estimates of the annual household benefits of insulation and thus considerable variation in estimated benefit cost ratios for the programme as a whole, driven by different choices in the valuation of changes in mortality.

Comparison of the WUNZ:HS programme and Australia's HIP programme suggested that differences in programme outcomes were driven by a combination of differences in programme design and implementation. Some of these differences could be traced to a top-down focus on economic stimulus, which overrode other concerns in the Australian context. The New Zealand policy context was noteworthy for the widespread acceptance of the co-benefits of insulation, which can speculatively be linked to differences in the design and implementation of the WUNZ:HS programme. In addition to informing a

number of housing retrofit specific recommendations presented in Chapter Eight, this comparison also reinforced the importance of correctly characterising previous interventions when using outcome evaluations as a source of evidence for future policies.

REFERENCES

Alberini, A., M. Cropper, A. Krupnick and N. B. Simon (2004). "Does the value of a statistical life vary with age and health status? Evidence from the US and Canada." Journal of Environmental Economics and Management **48**(1): 769-792.

Aldy, J. E. and W. K. Viscusi (2008). "Adjusting the value of a statistical life for age and cohort effects." The Review of Economics and Statistics **90**(3): 573-581.

Althaus, C., P. Bridgman and G. Davis (2007). The Australian Policy Handbook. Crows Nest, NSW, Allen & Unwin.

ANAO (2010). Home Insulation Programme: Audit Report No. 12 2010-11. Canberra, Australian National Audit Office.

Baker, M., M. Keall, E. L. Au and P. Howden-Chapman (2007). "Home is where the heart is—most of the time." Journal of the New Zealand Medical Association **120**(1264).

Baker, M., A. McDonald, J. Zhang and P. Howden-Chapman (2013). Infectious diseases attributable to household crowding in New Zealand: A systematic review and burden of disease estimate. Wellington, He Kainga Oranga/ Housing and Health Research Programme, University of Otago.

Baker, M., J. Zhang, P. Howden-Chapman, T. Blakely, K. Saville-Smith and J. Crane (2006). Housing, crowding and health study: characteristics of cohort members and their hospitalisations: Interim Report. Wellington, He Kainga Oranga/Housing and Health Research Programme.

Barck, C., T. Sandström, J. Lundahl, G. Hallden, M. Svartengren, V. Strand, S. Rak and G. Bylin (2002). "Ambient level of NO2 augments the inflammatory response to inhaled allergen in asthmatics." Respiratory Medicine **96**(11): 907-917.

Barton, A., M. Basham, C. Foy, K. Buckingham and M. Somerville (2007). "The Watcombe Housing Study: the short term effect of improving housing conditions on the health of residents." Journal of Epidemiology and Community Health **61**(9): 771.

Bastings, L. (1958). Handbook on the Insulation and Heating of Buildings with special reference to Dwellings. Wellington, D.S.I.R.

Basu, R. (2009). "High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008." Environ Health **8**(1): 40.

Basu, R. and J. M. Samet (2002). "Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence." Epidemiologic reviews **24**(2): 190-202.

Bland, J. M. and D. G. Altman (1994). "Regression towards the mean." BMJ: British Medical Journal **308**(6942): 1499.

Brannon, I. (2004). "What is a life worth." Regulation **27**: 60.

BRANZ (2010). Annual Review 2010. Wellington, BRANZ Inc.

Briggs, A. and A. Gray (1998). "The distribution of health care costs and their statistical analysis for economic evaluation." Journal of Health Services Research & Policy **3**(4): 233-245.

Briggs, A., R. Nixon, S. Dixon and S. Thompson (2005). "Parametric modelling of cost data: some simulation evidence." Health economics **14**(4): 421-428.

Brownlee, G. (2009). "Boost for warmer, drier, healthier Kiwi homes." Retrieved 05/05/2010, 2010, from http://www.eeca.govt.nz/news/boost-warmer-drier-healthier-kiwi-homes.

Bureau of Meteorology. (2013). "Average annual & monthly heating and cooling degree days." Retrieved 15/10/2013, 2013, from http://www.bom.gov.au/jsp/ncc/climate_averages/degree-days/index.jsp?maptype=1&period=an&product=hdd18#maps.

Calo, M. R. (2011). "The Boundaries Of Privacy Harm." Ind. LJ **86**: 1131-1162.

Carbon Trust. (2012). "Degree Days." Retrieved 05/02/2015, 2015, from http://www.carbontrust.com/resources/guides/energy-efficiency/degree-days.

Carrer, P., M. Maroni, D. Alcini and D. Cavallo (2001). "Allergens in indoor air: environmental assessment and health effects." The Science of the Total Environment **270**(1-3): 33-42.

Carroll, P., M. Blewden and K. Witten (2008). The Social Sciences and Policy-Research Use Building Research Capability in the Social Sciences (BRCSS) Auckland, Centre for Social and Health Outcomes Research and Evaluation, Massey University**:** 81.

Chapman, R. (2009). Estimating the health cost of leaky homes. Do damp and mould matter? Health Impacts of Leaky Homes. P. Howden-Chapman, J. Bennett and R. Siebers. Wellington, Steele Roberts.

Chapman, R., P. Howden-Chapman, H. Viggers, D. O'Dea and M. Kennedy (2009). "Retrofitting houses with insulation: a cost-benefit analysis of a randomised community trial." Journal of Epidemiology and Community Health **63**(4): 271-277.

Chauhan, A. J., H. M. Inskip, C. H. Linaker, S. Smith, J. Schreiber, S. L. Johnston and S. T. Holgate (2003). "Personal exposure to nitrogen dioxide (NO2) and the severity of virus-induced asthma in children." The Lancet **361**(9373): 1939-1944.

Choodari-Oskooei, B., P. Royston and M. K. Parmar (2012). "A simulation study of predictive ability measures in a survival model I: Explained variation measures." Statistics in medicine **31**(23): 2627-2643.

Colebatch, H. (2006). Beyond the policy cycle, Allen & Unwin, Sydney.

Cormack, D. and C. Robson (2010). Classification and output of multiple ethnicities: issues for monitoring Māori health. Wellington, Te Rōpū Rangahau Hauora a Eru Pōmare.

Craig, P., P. Dieppe, S. Macintyre, S. Michie, I. Nazareth and M. Petticrew (2008). "Developing and evaluating complex interventions: the new Medical Research Council guidance." BMJ **337(a1655)**.

Davie, G., M. Baker, S. Hales and J. Carlin (2007). "Trends and determinants of excess winter mortality in New Zealand: 1980 to 2000." BMC Public Health **7**(1): 263.

DCCCE (2012). Annual Report 2011-2012. Canberra, Department of Climate Change and Energy Efficiency.

de Looper, M. (2002). Seasonality of Death. Bulletin No. 3. Bulletin. Canberra, AIHW.

Denne, T. and S. Bond-Smith (2011). Impacts of the NZ Insulation Fund on Industry & Employment.

Department of Internal Affairs. (2013). "Deaths: Frequently Asked Questions." Retrieved 24/05/2013, 2013, from http://www.dia.govt.nz/diawebsite.nsf/wpg_URL/Services-Births-Deaths-and-Marriages-Deaths-Frequently-Asked-Questions?OpenDocument.

Department of Statistics (1993). New Zealand Standard Classification of Ethnicity. Wellington, Department of Statistics.

Dodd, S., A. Bassi, K. Bodger and P. Williamson (2006). "A comparison of multivariable regression models to analyse cost data." Journal of evaluation in clinical practice **12**(1): 76-86.

Dollery, B. and M. Hovey (2010). "Australian federal government failure: The rise and fall of the Home Insulation Program." Economic Papers: A journal of applied economics and policy **29**(3): 342-352.

Drummond, M. F., M. J. Sculpher, G. W. Torrance, B. J. O'brien and G. L. Stoddart (2005). Methods for the economic evaluation of health care programmes, Oxford University Press, USA.

Duan, N. (1983). "Smearing estimate: a nonparametric retransformation method." Journal of the American Statistical Association **78**(383): 605-610.

EECA (2007). Review of EnergyWise Home Grants.

EECA. (2009a). "Funding for Insulation and Clean, Efficient Heating." from http://www.energywise.govt.nz/funding-available/insulation-and-clean-heating.

EECA (2009b). Request for Proposal for service providers to seek funding under the ENERGYWISE™ Warm Up New Zealand: Heat Smart Programme. Wellington, New Zealand, EECA.

EECA (2009c). Warm Up New Zealand: Heat Smart Quality and Audit Manual: Release version 29 September 2009. Wellington, EECA.

EECA. (2012). "Warm Up New Zealand: Heat Smart to focus on insulation " Retrieved 18/06/2013, 2013, from http://www.eeca.govt.nz/news/warm-new-zealand-heat-smart-focus-insulation.

EECA. (2013). "Warm Up New Zealand: Healthy Homes." Retrieved 17/06/2013, 2013, from http://www.energywise.govt.nz/funding-and-programmes/insulation-programme.

Fawcett, J. (2005). Socioeconomic Variation in Mortality in New Zealand 1981-1999: Cross-National And Times Series Comparisons PhD, University of Otago, Dunedin.

Foster, R., N. Preval, T. Blakely, N. Wilson and D. O'Dea (2011). Costing of Pharmaceuticals in New Zealand for Health Economic Studies: Backgrounder and Protocol for Costing: Technical Report No.6. Wellington, Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme, Department of Public Health, University of Otago, Wellington. **20**.

French, L. J., M. J. Camilleri, N. P. Isaacs and A. R. Pollard (2007). "Temperatures and heating energy in New Zealand houses from a nationally representative study—HEEP." Energy and Buildings **39**(7): 770-782.

Gibson, E. (2010, 16/02/2010). "Insulation scheme plagued by poor work." New Zealand Herald Retrieved 16/03/2010, from http://www.nzherald.co.nz.

Gillespie-Bennett, J., N. Pierse, K. Wickens, J. Crane, S. Nicholls, D. Shields, M. Boulic, H. Viggers, M. Baker and A. Woodward (2008). "Sources of nitrogen dioxide (NO2) in New Zealand homes: findings from a community randomized controlled trial of heater substitutions." Indoor Air **18**(6): 521-528.

Gluckman, P. (2011). Towards better use of evidence in policy formation: a discussion paper. Auckland, Office of the Prime Minister's Science Advisory Committee.

Goodman, S. N. (1999). "Toward evidence-based medical statistics. 1: The P value fallacy." Annals of internal medicine **130**(12): 995-1004.

Greenland, S. (2008). "Invited commentary: variable selection versus shrinkage in the control of multiple confounders." American Journal of Epidemiology **167**(5): 523-529.

Grimes, A., T. Denne, P. Howden-Chapman, R. Arnold, L. Telfar-Barnard, N. Preval and C. Young (2011). Cost Benefit Analysis of the Warm Up New Zealand: Heat Smart Programme.

Grimes, A. and C. Young (2009). Pacific Peoples' Homeownership in New Zealand. Motu Note. Wellington, Motu Economic and Public Policy Research. **3**.

Grimes, A., C. Young, R. Arnold, T. Denne, P. Howden-Chapman, N. Preval and L. Telfar-Barnard (2011). "Warming Up New Zealand: Impacts of the New Zealand Insulation Fund on Metered Household Energy Use."

Guo, Y., S. Li, Y. Zhang, B. Armstrong, J. J. Jaakkola, S. Tong and X. Pan (2013). "Extremely cold and hot temperatures increase the risk of ischaemic heart disease mortality: epidemiological evidence from China." Heart **99**(3): 195-203.

Guria, J., W. Jones, M. Jones-Lee, M. Keall, J. Leung and G. Loomes (2003). The New Zealand values of statistical life and of the prevention of injuries. Wellington, Land Transport Safety Authority.

Guria, J., J. Leung, M. Jones-Lee and G. Loomes (2005). "The willingness to accept value of statistical life relative to the willingness to pay value: evidence and policy implications." Environmental and Resource Economics **32**(1): 113-127.

Hales, S., C. Salmond, G. I. Town, T. Kjellstrom and A. Woodward (2000). "Daily mortality in relation to weather and air pollution in Christchurch, New Zealand." Australian and New Zealand Journal of Public Health **24**(1): 89-91.

Hall, L. (2010). "Peter Garett demoted after botched insulation programme." The Age Retrieved 03/03/2011, from http://www.theage.com.au.

Hawke, A. (2010). Review of the Administration of the Home Insulation Program. Canberra, Department of Climate Change and Energy Efficiency.

Hopton, J. L. and S. M. Hunt (1996). "Housing conditions and mental health in a disadvantaged area in Scotland." Journal of Epidemiology and Community Health **50**(1): 56-61.

Hosmer, D. W., S. Lemeshow and S. May (2008). Applied survival analysis: regression modeling of time to event data. New York, John Wiley & Sons.

Housing New Zealand Corporation (2005). Building the Future: the New Zealand housing strategy. Wellington, Housing New Zealand Corporation.

Howden-Chapman, P. (2004). "Housing standards: a glossary of housing and health." Journal of Epidemiology and Community Health **58**(3): 162-168.

Howden-Chapman, P., T. Chandola, M. Stafford and M. Marmot (2011). "The effect of housing on the mental health of older people: the impact of lifetime housing history in Whitehall II." BMC Public Health **11**(1): 682.

Howden-Chapman, P., A. Matheson, J. Crane, H. Viggers, M. Cunningham, T. Blakely, C. Cunningham, A. Woodward, K. Saville-Smith and D. O'Dea (2007). "Effect of insulating existing houses on health inequality: cluster randomised study in the community." BMJ **334**(7591): 460-469.

Howden-Chapman, P., N. Pierse, S. Nicholls, J. Gillespie-Bennett, H. Viggers, M. Cunningham, R. Phipps, M. Boulic, P. Fjallstrom and S. Free (2008). "Effects of improved home heating on asthma in community dwelling children: randomised controlled trial." BMJ **337**(a1411): 848-862.

Howden-Chapman, P. and N. Preval (2014). Cobenefits of Insulating Houses: Research Evidence and Policy Implications. Wellbeing: A Complete Reference Guide, Volume II, Wellbeing and the Environment. London, Wiley-Blackwell. **II:** 607-626.

Howden-Chapman, P., H. Viggers, R. Chapman, D. O'Dea, S. Free and K. O'Sullivan (2009). "Warm homes: drivers of the demand for heating in the residential sector in New Zealand." Energy Policy **37**(9): 3387-3399.

Howlett, M., M. Ramesh and A. Perl (1995). Studying public policy: Policy cycles and policy subsystems. Cambridge, Cambridge University Press.

Hunt, S. (1990). "Emotional distress and bad housing." Health and Hygiene **11**: 72-79.

IEA (2007). Mind the Gap: Quantifying Principal Agent Problems in Energy Efficiency. Paris, International Energy Agency.

Isaacs, N. P. (1993). Thermal Efficiency in New Zealand Buildings: A Historical Overview. Wellington, New Zealand., Centre for Building Performance Research. Victoria University, Wellington, New Zealand.

Jackson, G., S. Thornley, J. Woolston, D. Papa, A. Bernacchi and T. Moore (2011). "Reduced acute hospitalisation with the healthy housing programme." Journal of Epidemiology and Community Health **65**(7): 588-593.

Jatrana, S. and T. Blakely (2008). "Ethnic inequalities in mortality among the elderly in New Zealand." Australian and New Zealand Journal of Public Health **32**(5): 437-443.

Keatinge, W. R., G. C. Donaldson, K. Bucher, G. Jendritsky, E. Cordioli, M. Martinelli, L. Dardanoni, K. Katsouyanni, A. E. Kunst and J. P. Mackenbach (1997). "Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe." The Lancet **349**(9062): 1341-1346.

Kortt, M. A. and B. Dollery (2012). "The Home Insulation Program: an example of Australian government failure." Australian Journal of Public Administration **71**(1): 65-75.

Lah, O. (2009). The Climate for Change: The conditions for effective climate change policies: A case study on residential home insulation policies in New Zealand and Germany. Master of Environmental Studies, Victoria University of Wellington.

Last, J. M. (2001). A dictionary of epidemiology. New York, Oxford University Press.

Law, D. and L. Meehan (2013). Housing Affordability in New Zealand: Evidence from Household Surveys New Zealand Treasury Working Paper 13/14. Wellington, New Zealand Treasury.

Lewis, C. (2010). "The Home Insulation Program Policy Debacle: Haste Makes Waste." Public Policy **5**(2): 83.

Lloyd, B. (2006). "Fuel Poverty in New Zealand." Social Policy Journal of New Zealand(27): 142-155.

Lloyd, B., T. Bishop and M. Callau (2007). Retrofit alternatives for State Houses in Cold Regions of New Zealand. Dunedin, NZ, Energy Management Group, Physics Dept., University of Otago.

Lloyd, C. R., M. F. Callau, T. Bishop and I. J. Smith (2008). "The efficacy of an energy efficient upgrade program in New Zealand." Energy and Buildings **40**(7): 1228-1239.

Lopez, A. D., C. D. Mathers, M. Ezzati, D. T. Jamison and C. J. Murray (2006). "Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data." The Lancet **367**(9524): 1747-1757.

Manning, W. G., A. Basu and J. Mullahy (2005). "Generalized modeling approaches to risk adjustment of skewed outcomes data." Journal of health economics **24**(3): 465-488.

Manning, W. G. and J. Mullahy (2001). "Estimating log models: to transform or not to transform?" Journal of health economics **20**(4): 461-494.

Masoli, M., D. Fabian, S. Holt and R. Beasley (2004). "The global burden of asthma: executive summary of the GINA Dissemination Committee Report." Allergy **59**(5): 469-478.

McChesney, I., I. Cox-Smith and L. Amitrano (2008). Status report on thermal insulation in New Zealand: A Status Report for Beacon Pathway Limited. Auckland, Beacon Pathway Incorporated.

McMahon, L. F. and V. Chopra (2012). "Health Care Cost and Value." JAMA: The Journal of the American Medical Association **307**(7): 671-672.

Mihaylova, B., A. Briggs, A. O'Hagan and S. G. Thompson (2011). "Review of statistical methods for analysing healthcare resources and costs." Health Economics **20**(8): 897-916.

Miller, T. and J. C. Guria (1991). The value of statistical life in New Zealand: market research on road safety. Wellington, Land Transport Division, Ministry of Transport.

Milne, G. and B. Boardman (2000). "Making cold homes warmer: the effect of energy efficiency improvements in low-income homes A report to the Energy Action Grants Agency Charitable Trust." Energy Policy **28**(6-7): 411-424.

Ministry for the Environment (2005). Warm Homes Technical Report: Social Drivers: Phase 1: Interim Progress Report. Wellington, New Zealand Ministry for the Environment.

MoH. (2007). "NZDep2006 Index of Deprivation." Retrieved 12/06/2013, 2013, from http://www.health.govt.nz/publication/nzdep2006-index-deprivation.

MoH (2009a). Mortality Collection Data Dictionary. Wellington, Ministry of Health.

MoH (2009b). National Health Index Data Dictionary Version 5.3. Wellington, Ministry of Health.

MoH (2010). National Minimum Dataset (Hospital Events): Data Dictionary. Wellington, Ministry of Health.

MoH (2012a). Hospital Events 2008/09 and 2009/10. Wellington, Ministry of Health.

MoH. (2012b). "National Health Index overview." Retrieved 01/05/2013, 2013, from http://www.health.govt.nz/our-work/health-identity/national-health-index/national-health-index-overview.

Morrison, P. and K. Nissen (2010). "Moving in and out of areas of deprivation: evidence from the New Zealand census." New Zealand Population Review **36**: 55-80.

Mudarri, D. and W. J. Fisk (2007). "Public health and economic impact of dampness and mold." Indoor Air **17**(3): 226-235.

New Zealand Health Information Service (2006). Pharmaceutical Claims Data Mart (Pharms) Data Dictionary Version 4.0 Wellington, New Zealand Health Information Service.

NZPA. (2008, 26/11/2008). "Stimulus package to inject $7 billion into NZ economy." National Business Review Retrieved 20/10/2010, from http://www.nbr.co.nz.

O'Hagan, A. and J. W. Stevens (2003). "Assessing and comparing costs: how robust are the bootstrap and methods based on asymptotic normality?" Health economics **12**(1): 33-49.

O'Dea, D. and J. Wren (2010). New Zealand Estimates of the Total Social and Economic Cost of "All Injuries" and the Six Priority Areas Respectively, at June 2008 Prices: Technical Report Prepared for NZIPS Evaluation. Wellington, New Zealand, Accident Compensation Corporation.

O'Sullivan, K. C., P. L. Howden-Chapman and G. Fougere (2011). "Making the connection: The relationship between fuel poverty, electricity disconnection, and prepayment metering." Energy Policy **39**(2): 733-741.

OECD. (2011). "Health at a Glance 2011: OECD Indicators." Retrieved 12/12/2012, from http://dx.doi.org/10.1787/health_glance-2011-en.

PCE (2000). Getting more from less: A review of progress on energy efficiency and renewable energy initiatives in New Zealand. Wellington, Parliamentary Comissioner for the Environment.

Peduzzi, P., J. Concato, E. Kemper, T. R. Holford and A. R. Feinstein (1996). "A simulation study of the number of events per variable in logistic regression analysis." Journal of Clinical Epidemiology **49**(12): 1373-1379.

PHARMAC (2007). New Zealand Pharmaceutical Schedule. Wellington, PHARMAC.

Pilotto, L. S., M. Nitschke, B. J. Smith, D. Pisaniello, R. E. Ruffin, H. J. McElroy, J. Martin and J. E. Hiller (2004). "Randomized controlled trial of unflued gas heater replacement on respiratory health of asthmatic schoolchildren." International Journal of Epidemiology **33**(1): 208-211.

Pregibon, D. (1979). Data analytic methods for generalized linear models. PhD Dissertation, University of Toronto.

Preval, N. (2012). An extended analysis of the impacts of retrofitted insulation and new heaters on health services utilisation and costs, and mortality: Evaluation of Warm Up New Zealand: Heat Smart Extension.

Preval, N., R. Chapman, N. Pierse and P. Howden-Chapman (2010). "Evaluating energy, health and carbon co-benefits from improved domestic space heating: A randomised community trial." Energy Policy **38**(8): 3965-3972.

Richards, S. (2007). Occasional Papers in Housing No. 3: Housing needs of the future. Occasional Papers in Housing, Nelson Tasman Housing Trust. **3**.

Rose, G. (2001). "Sick individuals and sick populations." International Journal of Epidemiology **30**(3): 427-432.

Rothman, K. J. and S. Greenland (1998). Modern epidemiology. Philadelphia, Lipppincott-Raven.

Royal Commission into the Home Insulation Programme. (2013). "Terms of Reference." Retrieved 1/15/2014, 2014, from http://www.homeinsulationroyalcommission.gov.au/About/Pages/TermsofReference.aspx.

Royston, P., D. G. Altman and W. Sauerbrei (2006). "Dichotomizing continuous predictors in multiple regression: a bad idea." Statistics in medicine **25**(1): 127-141.

Royston, P., G. Ambler and W. Sauerbrei (1999). "The use of fractional polynomials to model continuous risk variables in epidemiology." International journal of epidemiology **28**(5): 964-974.

Royston, P. and W. Sauerbrei (2008). Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables, John Wiley & Sons.

Rychetnik, L., P. Hawe, E. Waters, A. Barratt and M. Frommer (2004). "A glossary for evidence based public health." Journal of Epidemiology and Community Health **58**(7): 538-545.

Sanderson, K., M. Goodchild and A. Slack (2007). The Value of Statistical Life for Fire Regulatory Impact Statements. Wellington, BERL.

Savage, J. (2012). "Winz staff fired over privacy breaches." New Zealand Herald Retrieved 14/04/2014, 2014, from http://www.nzherald.co.nz.

Science Media Centre. (2008). "$1 billion fund for household insulation." Science Media Centre Retrieved 18/06/2013, 2013, from http://www.sciencemediacentre.co.nz/2008/08/27/1-billion-fund-for-household-insulation/.

Seiber, E. E. (2006). "Physician code creep: evidence in Medicaid and State Employee Health Insurance billing." Health care financing review **28**(4): 83-93.

Senate (2010). Environment, Communication and the Arts Reference Committee. Energy Efficient Homes Package (Ceiling Insulation). Canberra, Senate.

Shadish, W. R., T. D. Cook and D. T. Campbell (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA, Houghton Mifflin.

Shephard, R. J. (2003). "Regression to the Mean." Sports medicine **33**(8): 575-584.

Shortt, N. and J. Rugkåsa (2007). "'The walls were so damp and cold' fuel poverty and ill health in Northern Ireland: Results from a housing intervention." Health and Place **13**(1): 99-110.

Slovic, P. (1987). "Perception of risk." Science **236**(4799): 280-285.

StataCorp. (2009). Stata 11 Base Reference Manual. College Station, TX, Stata Press.

State Services Commission. (2013). "Gateway Review Process." State Services Commission Retrieved 14/06/2013, 2013, from http://www.ssc.govt.nz/gateway.

Statistics New Zealand. (2011). "National Population Estimates: December 2010 Quarter." Statistics New Zealand Retrieved 06/05/2013, 2013, from http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationEstimates_HOTPDec10qtr/Commentary.aspx.

Statistics New Zealand. (2013a). "Demographic Trends: 2012." Statistics New Zealand Retrieved 01/07/2013, 2013, from http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/demographic-trends-2012/deaths%20and%20life%20expectancy.aspx.

Statistics New Zealand. (2013b). "New Zealand Life Tables: 2005-2007." Statistics New Zealand Retrieved 02/07/2013, 2013, from http://www.stats.govt.nz/browse_for_stats/health/life_expectancy/new-zealand-life-tables-2005-07/chapter-2-national-trends-in-longevity-and-mortality.aspx.

Steinwald, B. and L. A. Dummit (1989). "Hospital case-mix change: sicker patients or DRG creep?" Health Affairs **8**(2): 35-47.

Telfar-Barnard, L. (2010). Home truths and cold admissions: New Zealand housing attributes and excess winter hospitalisations. PhD dissertation, University of Otago.

Telfar Barnard, L., N. Preval, P. Howden-Chapman, R. Arnold, C. Young, A. Grimes and T. Denne (2011). The impact of retrofitted insulation and new heaters on health services utilisation and costs, pharmaceutical costs and mortality: Evaluation of Warm Up New Zealand: Heat Smart.

Thomson, H., S. Thomas, E. Sellstrom and M. Petticrew (2009). "The health impacts of housing improvement: a systematic review of intervention studies from 1887 to 2007." American Journal of Public Health **99**(S3): S681.

Thomson, H., S. Thomas, E. Sellstrom and M. Petticrew (2013). Housing improvements for health and associated socio-economic outcomes. The Cochrane Library, The Cochrane Collaboration. **2**.

Tunnicliffe, W. S., P. S. Burge and J. G. Ayres (1994). "Effect of domestic concentrations of nitrogen dioxide on airway responses to inhaled allergen in asthmatic patients." The Lancet **344**(8939-0): 1733-1736.

Venturini, S., F. Dominici and G. Parmigiani (2008). "Gamma shape mixtures for heavy-tailed distributions." The Annals of Applied Statistics **2**(2): 756-776.

Victora, C. G., J.-P. Habicht and J. Bryce (2004). "Evidence-based public health: moving beyond randomized trials." American Journal of Public Health **94**(3): 400-405.

Viscusi, W. K. and J. E. Aldy (2003). "The value of a statistical life: a critical review of market estimates throughout the world." Journal of risk and uncertainty **27**(1): 5-76.

Walter, S. and H. Tiemeier (2009). "Variable selection: current practice in epidemiological studies." European journal of epidemiology **24**(12): 733-736.

Wanless, D. (2004). Securing good health for the whole population. London, HM Treasury and Department of Health, Stationery Office.

WHO (2012). WHO Collaborating Centre for Drug Statistics Methodology, Guidelines for ATC classification and DDD assignment 2013. Oslo, WHO Collaborating Centre for Drug Statistics Methodology.

WHO. (2014). "International Classification of Diseases (ICD)." Retrieved 7/16/2014, 2014, from http://www.who.int/classifications/icd/en/.

Wilkinson, P., K. R. Smith, M. Davies, H. Adair, B. G. Armstrong, M. Barrett, N. Bruce, A. Haines, I. Hamilton, T. Oreszczyn, I. Ridley, C. Tonne and Z. Chalabi (2009). "Public health benefits of strategies to reduce greenhouse-gas emissions: household energy." The Lancet **374**(9705): 1917-1929.

World Health Organization (1987). Health impact of low indoor temperatures: Report on a WHO meeting. Copenhagen, WHO.

Young, A. (2012). "Inquiry into ACC privacy breach." Retrieved 04/04/2014, 2014, from http://www.nzherald.co.nz.

APPENDIX 1. Ethics Committee Approvals

1.1. Ethics approval: WUNZ:HS health evaluation

Health and Disability Ethics Committees

Multi-region Ethics Committee
Ministry of Health
133 Molesworth Street
PO Box 5013
Wellington 6145
Phone (04) 470 0655
(04) 470 0646
Fax (04) 496 2340
Email: multiregion_ethicscommittee@moh.govt.nz

2 March 2010

Dr Philippa Howden-Chapman
Dept Public Health
Wellington School of Medicine and Health Sciences
University of Otago
PO Box 7343. Wellington South

Dear Philippa

Ethics ref: **MEC/10/EXP/002**
Study title: Evaluation of the New Zealand Insulation Fund
Investigators: Dr Philippa Howden-Chapman, Dr Richard Arnold, Ms Lucy Telfar Barnard

The above study has been given ethical approval by the Multi-region Ethics Committee.

Accreditation
The Committee involved in the approval of this study is accredited by the Health Research Council and is constituted and operates in accordance with the Operational Standard for Ethics Committees, April 2006.

Progress Reports
The study is approved until **30th of June 2011**. The Committee will review the approved application annually and notify the Principal Investigator if it withdraws approval. It is the Principal Investigator's responsibility to forward a progress report covering all sites prior to ethical review of the project in **March 2011**. The report form is available at http://www.ethicscommittees.health.govt.nz. Please note that failure to provide a progress report may result in the withdrawal of ethical approval. A final report is also required at the conclusion of the study.

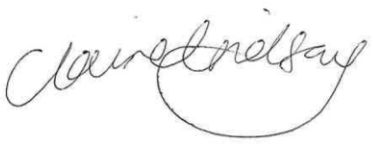
Amendments
It is also a condition of approval that the Committee is advised of any adverse events, if the study does not commence, or the study is altered in any way, including all documentation eg advertisements, letters to prospective participants.

Please quote the above ethics committee reference number in all correspondence.

It should be noted that Ethics Committee approval does not imply any resource commitment or administrative facilitation by any healthcare provider within whose facility the research is to be carried out. Where applicable, authority for this must be obtained separately from the appropriate manager within the organisation.

We wish you well with your study.

Yours sincerely



Claire Lindsay
Administrator
Multi-region
Email: claire_lindsay@moh.govt.nz

1.2. Ethics approval: Jenny Lackey interview

COPY

Reporting Sheet for use ONLY for proposals considered at departmental level

UNIVERSITY OF OTAGO

Form devised May 1995; updated May 1997; June 1998; May 1999, Dec 2000, June 2002

ETHICAL APPROVAL AT DEPARTMENTAL LEVEL OF A PROPOSAL INVOLVING HUMAN PARTICIPANTS (CATEGORY B)

PLEASE read the important notes appended to this form before completing the sections below

NAME OF DEPARTMENT: PUBLIC HEALTH

TITLE OF PROJECT: Econometric, Statistical and Policy Analysis of Large Scale Public Health Interventions

PROJECTED START DATE OF PROJECT: 7th Feb

STAFF MEMBER RESPONSIBLE FOR PROJECT: Dr. Michael Keall (Supervisor)

NAMES OF OTHER INVESTIGATORS OR INSTRUCTORS: Nicholas Preval PhD Candidate, Adjunct Professor Arthur Grimes (supervisor), Professor Philippa Howden-Chapman (supervisor)

BRIEF DESCRIPTION OF THE AIMS: Please give a brief summary (approx. 200 words) of the nature of the proposal:-

The proposal involves an interview with Jenny Lackey, Project Manager, Warm Up New Zealand: Heat Smart of the Energy Efficiency & Conservation Authority (EECA). I hope to gather information about specific aspects of the Warm Up New Zealand: Heat Smart programme which relate to failures of Australia's Home Insulation Programme in order to explore why these failures did not apply in the New Zealand case. It is hoped that this interview will inform an article on the topic as part of Nicholas Preval's PhD thesis.

BRIEF DESCRIPTION OF THE METHOD: Please include a description of who the participants are, how the participants will be recruited, and what they will be asked to do:-

Participant: Jenny Lackey, Project Manager, Warm Up New Zealand: Heat Smart of Energy Efficiency & Conservation Authority (EECA).

COPY

Jenny Lackey will be interviewed. The interview will be semi-structured. Jenny will be given a list of questions prior to the interview.

DETAILS OF ETHICAL ISSUES INVOLVED: Please give details of any ethical issues which were identified during the consideration of the proposal and the way in which these issues were dealt with or resolved:-

The key ethical concern was that the interviewee should not be represented in any published material as a "Government spokesperson". Discussion has been had with her about a suitable way to reference any quotes or information resulting from the interview which did not compromise the interviewee or EECA. For example (Policy official, 2011), or (Lackey, 2011). She will be shown transcripts of the interviews to check accuracy.

ACTION TAKEN

- ☑ Approved by Head of Department
- ☐ Approved by Departmental Committee
- ☐ Referred to University of Otago Human Ethics Committee
- ☐ Referred to another Ethics Committee
  Please specify: .....

DATE OF CONSIDERATION: 19/2/11

Signed (Head of Department): .....

Please attach copies of any Information Sheet and/or Consent Form

1.2.1. Information form

Reporting Sheet for use ONLY for proposals considered at departmental level

(**Note:** Not all of the suggestions or headings on this template will necessarily apply to all projects)

03/02/2011

Warm Up New Zealand: Heat Smart and Australia's Home Insulation Programme: A Comparison
INFORMATION SHEET FOR PARTICIPANTS

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you of any kind and we thank you for considering our request.

What is the Aim of the Project?

The aim of the project is produce an article exploring the failures of Australia's Home Insulation Programme and why they did not apply to New Zealand's Warm Up New Zealand: Heat Smart programme. The information gathered may also contribute to a PhD thesis.

What will Participants be Asked to Do?

Should you agree to take part in this project, you will be asked to participate in a semi-structured interview. The interview should take between half an hour and an hour.

Please be aware that you may decide not to take part in the project without any disadvantage to yourself of any kind.

Can Participants Change their Mind and Withdraw from the Project?

You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?

Data collected will be your responses to questions. The interview will be audio-taped. Although you will be provided with intended questions prior to the interview, this project involves an open-questioning technique where the precise nature of the questions which will be asked have not been determined in advance, but will depend on the way in which the interview develops. Consequently, although the University of Otago Human Ethics Committee is aware of the general areas to be explored in the interview, the Committee has not been able to review the precise questions to be used.

In the event that the line of questioning does develop in such a way that you feel hesitant or uncomfortable you are reminded of your right to decline to answer any particular question(s) and also that you may withdraw from the project at any stage without any disadvantage to yourself of any kind.

The information gathered in the interview will inform an article comparing and contrasting the Warm Up New Zealand: Heat Smart programme with Australia's Home Insulation Programme and a chapter of Nick Preval's PhD thesis.

The interview data will be transcribed by Nicholas Preval and nobody other than Nicholas Preval and his PhD supervisors (Dr Michael Keall, Adjunct Professor Arthur Grimes and Professor Philippa Howden-Chapman) will have access to the interview tape or transcript.

The data collected will be securely stored in such a way that only those mentioned above will be able to gain access to it. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve your anonymity if that is your preference. You are most welcome to request a copy of the results of the project should you wish.

What if Participants have any Questions?

If you have any questions about our project, either now or in the future, please feel free to contact either:-

Nicholas Preval	or	Dr Michael Keall
Department of Public Health		Department of Public Health
University Telephone Number:- 04-918-5246		University Telephone No. :- 04 918 6794

1.2.2. Consent form

Reporting Sheet for use ONLY for proposals considered at departmental level

---

(**Note:** Not all of the suggestions on this template will necessarily apply to all projects; for some projects, additional information may also be required)

Warm Up New Zealand: Heat Smart and Australia's Home Insulation Programme: A Comparison

CONSENT FORM FOR PARTICIPANTS

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:-

1. my participation in the project is entirely voluntary;

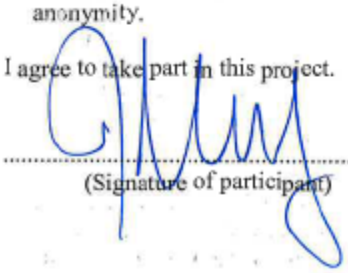
2. I am free to withdraw from the project at any time without any disadvantage;

3. the audio-tapes will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed;

4. this project involves an open-questioning technique where the precise nature of the questions which will be asked have not been determined in advance, but will depend on the way in which the interview develops and that in the event that the line of questioning develops in such a way that I feel hesitant or uncomfortable I may decline to answer any particular question(s) and/or may withdraw from the project without any disadvantage of any kind.

5. The results of the project may be published and available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.

I agree to take part in this project.



..... (Signature of participant)

9/3/11

..... (Date)

APPENDIX 2. Additional Tables

Table 72: Age structure, circulatory and respiratory cohorts

	Circulatory cohort				Respiratory cohort			
Age bracket	**Treatment type**				**Treatment type**			
	Insulation only	Heating only	Insulation and heating	Control	Insulation only	Heating only	Insulation and heating	Control
65-74	334	40	56	1,034	186	21	30	530
	42.39%	51.28%	42.75%	45.15%	46.38%	47.73%	44.12%	50.57%
75-84	358	28	57	950	174	20	27	394
	45.43%	35.90%	43.51%	41.48%	43.39%	45.45%	39.71%	37.60%
85+	96	10	18	306	41	3	11	124
	12.18%	12.82%	13.74%	13.36%	10.22%	6.82%	16.18%	11.83%
Total	788	78	131	2,290	401	44	68	1,048
	100.00%	100.00%	100.00 %	100.00%	100.00%	100.00 %	100.00%	100.00 %
Test statistic (comparison treatment group and control)	Pearson chi2(2) = 3.7837 Pr = 0.151	Pearson chi2(2) = 1.2115 Pr = 0.546	Pearson chi2(2) = 0.2948 Pr = 0.863		Pearson chi2(2) = 4.1735 Pr = 0.124	Pearson chi2(2) = 1.6685 Pr = 0.434	Pearson chi2(2) = 1.6019 Pr = 0.449	

Table 73: Estimated Hazard Ratios for Time to Death for one year increase in age, circulatory and respiratory cohorts

	Variable	Change	Hazard Ratio for Change (2 dp)	Wald Test p-value
Circulatory cohort	Baseline age	1 year	1.08 (95% CI 1.06 – 1.09)	<0.001***
Respiratory cohort	Baseline age	1 year	1.04 (95% CI 1.03 – 1.06)	<0.001***

* p<0.1, ** p<0.05, *** p<0.01

Table 74: Sex structure, circulatory and respiratory cohorts

	Circulatory cohort				Respiratory cohort			
Sex	Treatment type				Treatment type			
	Insulation only	Heating only	Insulation and heating	Control	Insulation only	Heating only	Insulation and heating	Control
Male	433	35	67	1,173	215	25	33	538
	54.95%	44.87%	51.15%	51.22%	53.62%	56.82%	48.53%	51.34%
Female	355	43	64	1,117	186	19	35	510
	45.05%	55.13%	48.85%	48.78%	46.38%	43.18%	51.47%	48.66%
Total	788	78	131	2,290	401	44	68	1,048
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Test statistic (comparison treatment group and control)	Pearson chi2(1) = 3.2628 Pr = 0.071*	Pearson chi2(1) = 1.2175 Pr = 0.270	Pearson chi2(1) = 0.0003 Pr = 0.986		Pearson chi2(1) = 0.6040 Pr = 0.437	Pearson chi2(1) = 0.5082 Pr = 0.476	Pearson chi2(1) = 0.2013 Pr = 0.654	

Table 75: Estimated Hazard Ratios for Time to Death by sex, circulatory and respiratory cohorts

	Variable (deaths, individuals)	Hazard Ratio (female vs. male) (2 dp)	Wald Test p-value
Circulatory cohort	Female (235, 1579) Male (276, 1708)	0.92 (95% CI 0.77- 1.09)	0.32
Respiratory cohort	Female (152, 750) Male (227, 811)	0.69 (95% CI 0.56-0.85)	<0.001***

* p<0.1, ** p<0.05, *** p<0.01

Table 76: Ethnicity structure, circulatory and respiratory cohorts, as at first month *after*

	Circulatory cohort				Respiratory cohort			
Ethnicity	**Treatment type**				**Treatment type**			
	Insulation only	Heating only	Insulation and heating	Control	Insulation only	Heating only	Insulation and heating	Control
Other (0)	701	70	119	2,027	326	35	57	847
	88.96%	89.74%	90.84%	88.52%	81.30%	79.55%	83.82%	80.82%
Māori (1)	49	8	4	156	50	7	6	115
	6.22%	10.26%	3.05%	6.81%	12.47%	15.91%	8.82%	10.97%
Pacific (2)	21	0	1	58	20	0	3	62
	2.66%	0.00%	0.76%	2.53%	4.99%	0.00%	4.41%	5.92%
Asian (3)	17	0	7	49	5	2	2	24
	2.16%	0.00%	5.34%	2.14%	1.25%	4.55%	2.94%	2.29%
Total	788	78	131	2,290	401	44	68	1,048
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Test statistic (comparison treatment group and control)	Pearson chi2(3) = 0.3644 Pr = 0.948	Pearson chi2(3) = 4.9494 Pr = 0.176	Pearson chi2(3) = 9.8149 Pr = 0.020**		Pearson chi2(3) = 2.5966 Pr = 0.458	Pearson chi2(3) = 4.4345 Pr = 0.218	Pearson chi2(3) = 0.7076 Pr = 0.871	

Table 77: Estimated Hazard Ratios for Time to Death for ethnicity, circulatory and respiratory cohorts

	Variable(deaths, individuals)	Hazard Ratio (vs. other) (2 dp)	Wald Test p-value
Circulatory cohort	Other (439, 2917)		0.011**
	Māori (49, 217)	1.58 (95% CI 1.18-2.14)	
	Pacific (15, 80)	1.36 (95% CI 0.78 - 2.27)	
	Asian (8, 73)	0.71 (95% CI 0.36 – 1.41)	
Respiratory cohort	Other (306, 1265)		0.118
	Māori (50, 178)	1.19 (95% CI 0.88 - 1.60)	
	Pacific (21, 85)	1.09 (95% CI 0.68 - 1.74)	
	Asian (2, 33)	0.23 (95% CI 0.06 – 0.91)	

* p<0.1, ** p<0.05, *** p<0.01

Table 78: NZdep2006 quintiles, circulatory and respiratory cohorts

	Circulatory cohort				Respiratory cohort			
NZdep2006 quintile	**Treatment type**				**Treatment type**			
	Insulation only	Heating only	Insulation and heating	Control	Insulation only	Heating only	Insulation and heating	Control
Quintile 1	101	18	20	357	52	8	8	128
	12.82%	23.08%	15.27%	15.59%	12.97%	18.18%	11.76%	12.21%
Quintile 2	137	11	31	464	63	6	12	191
	17.39%	14.10%	23.66%	20.26%	15.71%	13.64%	17.65%	18.23%
Quintile 3	176	14	27	476	74	6	17	220
	22.34%	17.95%	20.61%	20.79%	18.45%	13.64%	25.00%	20.99%
Quintile 4	212	14	35	580	110	10	14	266
	26.90%	17.95%	26.72%	25.33%	27.43%	22.73%	20.59%	25.38%
Quintile 5	162	21	18	413	102	14	17	243
	20.56%	26.92%	13.74%	18.03%	25.44%	31.82%	25.00%	23.19%
Total	788	78	131	2,290	401	44	68	1,048
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Test statistic (comparison treatment group and control)	Pearson chi2(4) = 8.7404 Pr = 0.068*	Pearson chi2(4) = 9.2790 Pr = 0.054*	Pearson chi2(4) = 2.0895 Pr = 0.719		Pearson chi2(4) = 3.1858 Pr = 0.527	Pearson chi2(4) = 4.2590 Pr = 0.372	Pearson chi2(4) = 1.1801 Pr = 0.881	

Table 79: Estimated Hazard Ratios for Time to Death for 10 unit increase in NZDep2006, circulatory and respiratory cohorts

	Variable	Change	Hazard Ratio for Change (2 dp)	Wald Test p-value
Circulatory cohort	NZDep2006	10 units	1.01 (95% CI 1.00-1.02)	0.016**
Respiratory cohort	NZDep2006	10 units	1.01 (95% CI 1.00- 1.02)	0.168

* p<0.1, ** p<0.05, *** p<0.01

Table 80: Baseline hospitalisation costs per year, circulatory and respiratory cohorts

	Treatment type	Variable	Average	Two sample Wilcoxon rank-sum (Mann-Whitney) test vs. control group P > |Z|
Circulatory cohort	Insulation	Baseline average annual circulatory hospitalisation costs	$4847.58	0.6148
		Baseline average annual other hospitalisation costs	$1805.29	0.1747
	Heating only	Baseline average annual circulatory hospitalisation costs	$3933.57	0.5461
		Baseline average annual other hospitalisation costs	$2136.91	0.7551
	Insulation and heating	Baseline average annual circulatory hospitalisation costs	$4309.17	0.2751
		Baseline average annual other hospitalisation costs	$2197.55	0.5158
	Control	Baseline average annual circulatory hospitalisation costs	$4652.40	
		Baseline average annual other hospitalisation costs	$1905.49	
Respiratory cohort	Insulation	Baseline average annual respiratory hospitalisation costs	$3509.69	0.9917
		Baseline average annual other hospitalisation costs	$2755.27	0.8003
	Heating only	Baseline average annual respiratory hospitalisation costs	$4375.95	0.1010
		Baseline average annual other hospitalisation costs	$1854.52	0.4223
	Insulation and heating	Baseline average annual respiratory hospitalisation costs	$3150.29	0.9370
		Baseline average annual other hospitalisation costs	$1317.18	0.0070***
	Control	Baseline average annual respiratory hospitalisation costs	$3319.58	
		Baseline average annual other hospitalisation costs	$3124.37	

* p<0.1, ** p<0.05, *** p<0.01

Table 81: Estimated Hazard Ratios for time to death for hospitalisation measures, circulatory and respiratory cohorts

	Variable name	Variable description	Change	HR for Change (3dp)	Wald Test p - value
Circulatory cohort	Baseline_avg_ circ_cost	Average cost per year of circulatory hospitalisations during *before* period	$1000	1.015 (95% CI 1.009 - 1.021)	<0.001***
	Baseline_avg_non_relevant_cost	Average cost per year of non-circulatory hospitalisations during *before* period	$1000	1.039 (95% CI 1.022 - 1.057)	<0.001***
Respiratory cohort	Baseline_avg_ resp_cost	Average cost per year of circulatory hospitalisations per year during *before* period	$1000	1.023 (95% CI 1.012 – 1.034)	<0.001***
	Baseline_avg_ non_relevant_cost	Average cost per year of non-circulatory non-respiratory hospitalisations during *before* period	$1000	1.028 (95%CI 1.028 – 1.036)	<0.001***

* p<0.1, ** p<0.05, *** p<0.01

Table 82: Housing risk category, circulatory and respiratory cohorts, as at first month *after*

	Circulatory cohort				Respiratory cohort			
Risk category	**Treatment type**				**Treatment type**			
	Insulation only	Heating only	Insulation and heating	Control	Insulation only	Heating only	Insulation and heating	Control
Old and high	71	7	11	204	52	4	7	100
	9.01%	8.97%	8.40%	8.91%	12.97%	9.09%	10.29%	9.54%
Old and average	79	8	16	198	37	3	8	101
	10.03%	10.26%	12.21%	8.65%	9.23%	6.82%	11.76%	9.64%
Old and low	4	0	0	15	0	0	0	6
	0.51%	0.00%	0.00%	0.66%	0.00%	0.00%	0.00%	0.57%
New and high	19	3	2	59	21	2	1	36
	2.41%	3.85%	1.53%	2.58%	5.24%	4.55%	1.47%	3.44%
New and average	547	49	91	1,591	261	34	50	733
	69.42%	62.82%	69.47%	69.48%	65.09%	77.27%	73.53%	69.94%
New and low	68	11	11	223	30	1	2	72
	8.63%	14.10%	8.40%	9.74%	7.48%	2.27%	2.94%	6.87%
Total	788	78	131	2,290	401	44	68	1,048
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Test statistic (comparison treatment group and control)	Pearson chi2(5) = 2.2777 Pr = 0.810	Pearson chi2(5) = 3.1371 Pr = 0.679	Pearson chi2(5) = 3.4511 Pr = 0.631		Pearson chi2(5) = 9.1356 Pr = 0.104	Pearson chi2(5) = 2.4211 Pr = 0.788	Pearson chi2(5) = 3.0699 Pr = 0.689	

Table 83: Estimated Hazard Ratios for time to death for "dwelling health risk" dummy variables, circulatory and respiratory cohorts

	Variable	Description	Change	Hazard Ratio for Change (2 dp)	Wald Test p-value
Circulatory cohort	Dwelling health risk	Dwelling health risk dummy variable	DHR2 (Old average risk) vs. DHR1 (Old high risk) DHR3 (Old low risk) vs. DHR1 (Old high risk) DHR4 (New high risk) vs. DHR1 (Old high risk) DHR5 (New average risk) vs. DHR1 (Old high risk) DHR6 (New low risk) vs. DHR1 (Old high risk)	0.853 (95% CI 0.575 – 1.266) 0.851 (95% CI 0.278 – 2.604) 1.061 (95% CI 0.582 – 1.932) 0.899 (95% CI 0.676 – 1.196) 0.573 (95% CI 0.370 – 0.887)	0.191
Respiratory cohort	Dwelling health risk	Dwelling health risk dummy variable	DHR2 (Old average risk) vs. DHR1 (Old high risk) DHR3 (Old low risk) vs. DHR1 (Old high risk) DHR4 (New high risk) vs. DHR1 (Old high risk) DHR5 (New average risk) vs. DHR1 (Old high risk) DHR6 (New low risk) vs. DHR1 (Old high risk)	0.594 (95% CI 0.382 – 0.922) 1.041 (95% CI 0.267 – 4.051) 0.953 (95% CI 0.567 – 1.603) 0.700 (95% CI 0.521 – 0.940) 0.340 (95% CI 0.185 – 0.626)	0.006**

* p<0.1, ** p<0.05, *** p<0.01

Table 84: Climate Zone, circulatory and respiratory cohorts

	Circulatory cohort				Respiratory cohort			
Regional Council	**Treatment type**				**Treatment type**			
	Insulation only	Heating only	Insulation and heating	Control	Insulation only	Heating only	Insulation and heating	Control
Climate Zone 1	196	13	21	555	101	16	15	255
	24.87%	16.67%	16.03%	24.24%	25.19%	36.36%	22.06%	24.33%
Climate Zone 2	346	29	50	1,032	193	16	29	471
	43.91%	37.18%	38.17%	45.07%	48.13%	36.36%	42.65%	44.94%
Climate Zone 3	246	36	60	703	107	12	24	322
	31.22%	46.15%	45.80%	30.70%	26.68%	27.27%	35.29%	30.73%
Total	788	78	131	2,290	401	44	68	1,048
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Test statistic (comparison treatment group and control)	Pearson chi2(2) = 0.3242 Pr = 0.850	Pearson chi2(2) = 8.6220 Pr = 0.013	Pearson chi2(2) = 13.7936 Pr = 0.001		Pearson chi2(2) = 0.3242 Pr = 0.850	Pearson chi2(2) = 8.6220 Pr = 0.013	Pearson chi2(2) = 13.7936 Pr = 0.001	

Table 85: Estimated Hazard Ratios for time to death for "Climate Zone" dummy variables, circulatory and respiratory cohorts

	Variable	Description	Change	Hazard Ratio for Change (2 dp)	Wald Test p-value
Circulatory cohort	Climate Zone	Climate Zone dummy variable	Climate Zone 2 vs. Climate Zone 1 Climate Zone 3 vs. Climate Zone 1	1.06 0.94	0.628 0.631
Respiratory cohort	Climate Zone	Climate Zone dummy variables	Climate Zone 2 vs. Climate Zone 1 Climate Zone 3 vs. Climate Zone 1	0.95 1.01	0.700 0.942

* p<0.1, ** p<0.05, *** p<0.01

Table 86: Initial multiple regression models, circulatory and respiratory cohorts

	Variable name	HR	Standard error	z	P>|z|	95% Confidence intervals	
Circulatory cohort	treat_dummy1	0.717	0.082	-2.92	0.004***	(0.574	0.897)
	treat_dummy2	1.440	0.318	1.65	0.098*	(0.935	2.219)
	treat_dummy3	0.645	0.166	-1.7	0.088*	(0.389	1.068)
	base_age	1.087	0.007	12.59	0.000***	(1.073	1.101)
	female	0.750	0.068	-3.16	0.002***	(0.628	0.896)
	ethn1	2.039	0.345	4.21	0.000***	(1.463	2.843)
	ethn2	1.933	0.555	2.3	0.022**	(1.101	3.392)
	ethn3	0.855	0.283	-0.47	0.635	(0.446	1.637)
	avg_pre_treat_chap9_cost_1000	1.014	0.003	4.98	0.000***	(1.009	1.020)
	avg_pre_treat_non_relevant_1000	1.045	0.009	5.24	0.000***	(1.028	1.063)
	final_nzdep_score_10	1.006	0.005	1.21	0.228	(0.996	1.017)
	risk_cat2	0.973	0.209	-0.13	0.900	(0.638	1.484)
	risk_cat3	1.407	0.689	0.7	0.486	(0.539	3.676)
	risk_cat4	1.169	0.364	0.5	0.615	(0.636	2.151)
	risk_cat5	1.049	0.163	0.31	0.759	(0.774	1.422)
	risk_cat6	0.715	0.167	-1.43	0.152	(0.453	1.131)
	climate_zone_dummy_2	1.055	0.124	0.46	0.646	(0.838	1.329)
	climate_zone_dummy_3	0.994	0.127	-0.05	0.961	(0.773	1.277)
Respiratory cohort	treat_dummy1	0.827	0.102	-1.53	0.125	(0.650	1.054)
	treat_dummy2	1.197	0.333	0.65	0.517	(0.695	2.064)
	treat_dummy3	0.649	0.193	-1.46	0.146	(0.362	1.162)
	base_age	1.047	0.007	6.46	0.000***	(1.033	1.062)
	female	0.706	0.075	-3.27	0.001***	(0.573	0.870)
	ethn1	1.399	0.243	1.93	0.053*	(0.995	1.967)
	ethn2	1.283	0.326	0.98	0.326	(0.780	2.112)
	ethn3	0.235	0.162	-2.1	0.036**	(0.061	0.910)
	avg_pre_treat_chap10_cost_1000	1.030	0.005	5.79	0.000***	(1.020	1.041)
	avg_pre_treat_non_relevan~1000	1.022	0.004	4.99	0.000***	(1.013	1.031)
	final_nzdep_score_10	1.007	0.006	1.22	0.222	(0.996	1.018)
	risk_cat2	0.587	0.139	-2.26	0.024**	(0.369	0.932)
	risk_cat3	1.509	1.129	0.55	0.582	(0.348	6.538)
	risk_cat4	0.919	0.254	-0.3	0.760	(0.535	1.580)
	risk_cat5	0.711	0.108	-2.25	0.025**	(0.528	0.957)
	risk_cat6	0.408	0.127	-2.88	0.004***	(0.221	0.751)
	climate_zone_dummy_2	0.913	0.126	-0.66	0.509	(0.697	1.196)
	climate_zone_dummy_3	0.973	0.143	-0.19	0.852	(0.730	1.298)

* p<0.1, ** p<0.05, *** p<0.01

Table 87: Full regression output for model presented in Table 48

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	-3.186	0.694	-4.59	0.000***	(-4.545	-1.826)	R-sq: within = 0.0006 R-sq: between = 0.0165 R-sq: overall = 0.0012	F(7,29000) = 60.61 Prob > F = 0.0000
	Heating only	-0.186	2.236	-0.08	0.934	(-4.569	4.197)		
	Insulation and heating	-2.963	1.511	-1.96	0.050*	(-5.926	-0.001)		
	Occ_diff	37.667	1.995	18.88	0.000***	(33.757	41.577)		
	Age_diff	0.641	0.115	5.55	0.000***	(0.415	0.867)		
	Monthly mean temp	0.331	0.111	2.99	0.003***	(0.114	0.549)		
	Winter	-0.617	0.798	-0.77	0.439	(-2.182	0.948)		

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter and monthly average temperature

Table 88: Full regression output for model presented in Table 49

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
Difference in all-cause hospitalisation costs	Insulation only	-3.503	0.791	-4.43	0.000	(-5.053	-1.953)	R-sq: within = 0.0006 R-sq: between = 0.0165 R-sq: overall = 0.0012	F(7,29000) = 42.52 Prob > F = 0.0000
	Heating only	-0.018	2.461	-0.01	0.994	(-4.841	4.805)		
	Insulation and heating	-3.739	1.651	-2.26	0.024	(-6.974	-0.503)		
	Insulation only* winter	0.943	1.185	0.790	0.427	(-1.381	3.267)		
	Heating only * winter	-0.473	3.095	-0.150	0.879	(-6.540	5.594)		
	Insulation and heating * winter	2.275	2.067	1.100	0.271	(-1.776	6.327)		
	occ_diff	37.669	1.995	18.88	0.000	(33.759	41.579)		
	age_diff	0.641	0.115	5.56	0.000	(0.415	0.867)		
	Monthly mean_temp	0.331	0.111	2.98	0.003	(0.113	0.548)		
	1.winter	-1.147	0.959	-1.2	0.231	(-3.027	0.732)		

Table 89: Full regression output for models presented in Table 61

Dependent variable modelled	Variable name	Coeffic ient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
All cause pharmaceutic al use cost difference	Insulation only	-2.008	0.250	-8.03	0.000***	(-2.498	-1.518)	R-sq: within = 0.0113 R- sq: between = 0.0674 R-sq: overall = 0.0403	F(7,28968) = 164.55 Prob > F = 0.0000
	Heating only	-3.008	0.807	-3.73	0.000***	(-4.589	-1.427)		
	Insulation and heating	-1.195	0.568	-2.1	0.036**	(-2.309	-0.081)		
	occ_diff	36.608	1.117	32.78	0.000***	(34.419	38.797)		
	age_diff	3.309	0.158	20.99	0.000***	(3.000	3.618)		
	Monthly mean_temp	0.028	0.025	1.15	0.249	(-0.020	0.077)		
	Winter	0.013	0.190	0.07	0.945	(-0.360	0.386)		
Circulatory illness-related pharmaceutic al use cost difference	Insulation only	-0.390	0.068	-5.75	0.000***	(-0.523	-0.257)	R-sq: within = 0.0037 R-sq: between = 0.1007 R-sq: overall = 0.0472	F(7,29054) = 67.00 Prob > F = 0.0000
	Heating only	-0.578	0.220	-2.63	0.009***	(-1.009	-0.147)		
	Insulation and heating	-0.234	0.154	-1.52	0.129	(-0.536	0.068)		
	occ_diff	6.688	0.310	21.59	0.000***	(6.081	7.295)		
	age_diff	0.755	0.041	18.27	0.000***	(0.674	0.836)		
	Monthly mean_temp	0.006	0.008	0.73	0.463	(-0.010	0.022)		
	Winter	-0.029	0.068	-0.43	0.666	(-0.162	0.104)		
Respiratory illness-related pharmaceutic al use cost difference	Insulation only	-0.034	0.038	-0.88	0.379	(-0.109	0.041)	R-sq: within = 0.0012 R-sq: between = 0.0124 R-sq: overall = 0.0061	F(7,28946) = 19.52 Prob > F = 0.0000
	Heating only	-0.042	0.124	-0.34	0.732	(-0.286	0.201)		
	Insulation and heating	-0.016	0.090	-0.18	0.857	(-0.194	0.161)		
	occ_diff	1.757	0.155	11.35	0.000***	(1.454	2.060)		
	age_diff	0.186	0.021	8.81	0.000***	(0.145	0.228)		
	Monthly mean_temp	-0.002	0.004	-0.48	0.631	(-0.010	0.006)		
	Winter	-0.035	0.031	-1.15	0.251	(-0.095	0.025)		

* p<0.1, ** p<0.05, *** p<0.01

Models control for occupant_diff, age_diff, winter and monthly average temperature

Table 90: Full regression output for models presented in Table 62

Dependent variable modelled	Variable name	Coefficient	Robust Standard error	t	P>|t|	95% Confidence intervals		R^2 within and R^2 overall	F-statistic
All cause pharmaceutical use cost difference	Insulation only	-1.962	0.278	-7.06	0.000***	(-2.507	-1.418)	R-sq: within = 0.0113 R- sq: between = 0.0674 R-sq: overall = 0.0403	F(10, 28968) = 115.26 Prob > F = 0.0000
	Heating only	-2.855	0.878	-3.25	0.001***	(-4.576	-1.133)		
	Insulation and heating	-0.951	0.639	-1.49	0.137	(-2.203	0.301)		
	Insulation only* winter	-0.113	0.307	-0.37	0.712	(-0.715	0.489)		
	Heating only * winter	-0.356	0.877	-0.41	0.685	(-2.076	1.363)		
	Insulation and heating * winter	-0.511	0.623	-0.82	0.413	(-1.732	0.711)		
	occ_diff	36.608	1.117	32.78	0.000***	(34.419	38.797)		
	age_diff	3.309	0.158	20.99	0.000***	(3.000	3.618)		
	Monthly mean_temp	0.029	0.025	1.18	0.240	(-0.019	0.077)		
	Winter	0.073	0.205	0.36	0.722	(-0.329	0.474)		
Circulatory illness-related pharmaceutical use cost difference	Insulation only	-0.397	0.077	-5.16	0.000***	(-0.547	-0.246)	R-sq: within = 0.0037 R- sq: between = 0.1007 R-sq: overall = 0.0472	F(10, 29054) Prob > F = 0.0000
	Heating only	-0.579	0.233	-2.48	0.013**	(-1.035	-0.122)		
	Insulation and heating	-0.201	0.174	-1.16	0.247	(-0.542	0.139)		
	Insulation only* winter	0.014	0.084	0.17	0.868	(-0.151	0.179)		
	Heating only * winter	0.002	0.221	0.01	0.994	(-0.432	0.435)		
	Insulation and heating * winter	-0.065	0.179	-0.36	0.716	(-0.417	0.286)		
	occ_diff	6.688	0.310	21.59	0.000***	(6.081	7.295)		
	age_diff	0.755	0.041	18.27	0.000***	(0.674	0.836)		
	Monthly mean_temp	0.006	0.008	0.73	0.464	(-0.010	0.022)		
	Winter	-0.030	0.073	-0.41	0.683	(-0.173	0.113)		
Respiratory illness-related pharmaceutical use cost difference	Insulation only	-0.012	0.045	-0.26	0.795	(-0.099	0.076)	R-sq: within = 0.012 R- sq: between = 0.0124 R-sq: overall = 0.0061	F(10, 28946) =14.35, Prob > F = 0.0000
	Heating only	0.143	0.145	0.98	0.326	(-0.142	0.428)		
	Insulation and heating	-0.017	0.104	-0.16	0.874	(-0.221	0.187)		
	Insulation only* winter	-0.053	0.054	-0.97	0.332	(-0.159	0.054)		
	Heating only * winter	-0.419	0.154	-2.72	0.007***	(-0.721	-0.117)		
	Insulation and heating * winter	-0.006	0.111	-0.06	0.954	(-0.224	0.211)		
	occ_diff	1.756	0.155	11.35	0.000***	(1.453	2.060)		
	age_diff	0.186	0.021	8.81	0.000***	(0.145	0.228)		
	Monthly mean_temp	-0.002	0.004	-0.42	0.674	(-0.010	0.006)		
	Winter	-0.011	0.033	-0.35	0.730	(-0.076	0.053)		

* p<0.1, ** p<0.05, *** p<0.01

Model controls for occupant_diff, age_diff, winter,monthly average temperature, interaction terms for treatment and winter

APPENDIX 3. Evaluation of Warm Up New Zealand: Heat Smart Extension

An extended analysis of the impacts of retrofitted insulation and new heaters on health services utilisation and costs, and mortality

Evaluation of Warm Up New Zealand: Heat Smart Extension

Nick Preval
He Kainga Oranga/Housing and Health Research Programme,
University of Otago, Wellington

October 2012

Author contact details

Nick Preval
He Kainga Oranga/Housing and Health Research Programme
University of Otago, Wellington School of Medicine and Health Science
nicholas.preval@otago.ac.nz

Acknowledgements

We thank the Ministry of Economic Development (MED) for commissioning and funding this study, and the following agencies for considerable assistance in obtaining and interpreting the requisite data: the Energy Efficiency and Conservation Authority (EECA), Quotable Value New Zealand (QVNZ), and the Ministry of Health (MoH). Particular thanks go to Chris Lewis (MoH).

We thank Phillipa Howden-Chapman (University of Otago Wellington School of Medicine and Health Science), Arthur Grimes (Motu Economic and Public Policy Research) and Richard Arnold (Victoria University of Wellington) for peer reviewing this report.

We also acknowledge the work and support of Lucy Telfar-Barnard (University of Otago Wellington School of Medicine and Health Science), whose work formed the basis for the creation of the dataset that this report is based on.

CONTENTS

List of Tables 3

List of Figures 4

Executive Summary 5

Background and introduction 8

Aim 10

Methodology 11

- Method 11
- Data description 11
  - Data sources and sequence 12
  - Dataset creation 15

Data analysis 17

- Key characteristics 17

Model selection 17

- Mortality 17
- Hospitalisation: Cost data 19

Methodological limitations 21

Results 21

- Mortality 21
- Mortality: cost data 22
- Hospitalisation: Cost data 24
- Summary data for cost benefit analysis 26

Conclusion 27

References 28

LIST OF TABLES

Table 1. Type and reasons for excluding hospitalisation cost data 16

Table 2. Age distribution of treatment, control and Census 2006 populations 17

Table 3. Effect of treatment on mortality rates in people aged 65+ hospitalised with circulatory illness prior to treatment month (original results with additional calculation) 18

Table 4. Effect of treatment on mortality rates in people aged 65+ hospitalised with circulatory illness prior to treatment month 22

Table 5. Change in monthly total hospitalisation costs per household..... 24

Table 6. Summary of annual health-related benefits (savings) per household..... 26

Table 7. Summary of annual health-related benefits (savings) per household from original report for comparison..... 27

LIST OF FIGURES

Figure 1: Flow diagram of method of calculating difference between treatment and control household total monthly hospitalisation costs..... 20

EXECUTIVE SUMMARY

Background

- This report is an extension of an evaluation of changes in the incidence and costs of health services, pharmaceutical usage and mortality in the first 46,655 houses retrofitted under the Warm Up New Zealand: Heat Smart (WUNZ:HS) programme. The initial report was entitled *The impact of retrofitted insulation and new heaters on health services utilisation and costs, pharmaceutical costs and mortality: Evaluation of Warm Up New Zealand: Heat Smart* Telfar Barnard, Preval [1].
- The original evaluation found statistically significant changes in mortality rates for people aged 65 and over who had had a circulatory system related hospitalisation prior to treatment under the WUNZ:HS programme.
- Other key results included a statistically significant (but small) reduction in total monthly hospitalisation costs at a household level.
- These results were consistent with previous clinical and public health research, including the results of two community trials, the Housing, Insulation and Health Study and the Housing, Heating and Health Study that have shown that both respiratory and circulatory symptoms are affected by indoor temperature and relative humidity.
- This report presents the results of adding an additional year of post retrofit hospitalisation and demographic data for the key mortality rate and total hospitalisation cost analyses presented in [1].
- The mortality and total hospitalisation cost benefits estimated in [1] comprised the majority of the benefits reported in the report *Cost Benefit Analysis of the Warm Up New Zealand: Heat Smart Programme* [2] which calculated that the programme had a benefit to cost ratio of approximately 3.9:1.

Method

- The original evaluation was a retrospective cohort study. QV matched EECA addresses for 46,655 treated dwellings to addresses in their database; 37,163 of the treatment addresses were able to be matched to a QV property listing, a match rate of 79.7%.
- The cohort was selected by matching dwellings that received insulation or heating retrofits (treatment homes), by address, to up to 10 similar (control) homes in the same Census Area Unit (CAU); 31,423 treatment homes (67.4% of all treatment dwellings) were able to be matched to at least one control address.
- Subsequently, via an anonymisation process, we identified records listed on the New Zealand National Health Index (NHI) as resident at those treatment and control addresses. The original health cohort finally comprised 973,710 individuals from 255,672 homes.
- As part of the extension EECA supplied anonymised details of the retrofits that it carried out during the period Dec 2010 – Dec 2011 which allowed us to correctly characterise cohort households that had received additional retrofits during this time period, and to identify

cohort addresses that were no longer valid due to participating in WUNZ:HS during this time period. Removing no longer eligible control households eventually reduced the cohort to 944,059 individuals from 247,444 homes.

- In order to extend the analysis, in early 2012 an additional year of hospitalisation and demographic data was obtained for the original cohort members. Matching issues with the additional NHI data reduced the cohort to 943,323 individuals from 247,388 homes.
- Our analysis followed the method set out in the original evaluation but was applied to the larger data set.
- Our mortality analysis used a sub-cohort of the study group, comprising those aged 65 and over who were hospitalised with a circulatory condition, but not deceased, prior to treatment date. We then compared survival time for treatment group and control group individuals using a Cox proportional hazards model, and valued the difference.
- Our analysis of total hospitalisation costs followed the 'difference in difference' approach used in the original report. We compared the difference between each treatment group household's monthly total hospitalisation costs and the mean of its matched control group household's monthly hospitalisation costs both before and after the intervention.
- Methodological limitations included imprecision in assigning NHI records to addresses and the possibility that control group households may have installed insulation or heating before or during the study period outside of WUNZ:HS.

Results

- Among those in the mortality sub-cohort, those in the treatment group had a statistically significant lower chance (hazard rate) of mortality after treatment than those in the control group (HR 0.68, 95%CI 0.55-0.84, p<0.000) where HR is the hazard rate.
- As with the original evaluation, we valued this statistically significant reduction in mortality based on the demographic structure of the treatment group as at July 2009, and estimated that there would be an annual reduction of 0.908 deaths per 1000 households. The life years gained can be conservatively valued at $468.40 per year per treated household. The benefit per year was valued at $662.58 for households that received treatment as Community Services Card Holders and $228.07 for those who did not, reflecting different proportions of vulnerable occupants.
- At the household level we found that there was no statistically significant change in total hospitalisation costs as a result of receiving insulation. This result was inconsistent with our previous findings. We speculate that this result may reflect the impact of reduced mortality for the treatment cohort.
- Limiting analysis either to those households who received WUNZ:HS funding as Community Services Card (CSC) holders or those who did not also did not demonstrate any statistically significant change in hospitalisation costs as a result of insulation.
- As with the original analysis there were no statistically significant changes in total hospitalisation costs at a household level as a result of receiving heating under WUNZ:HS.

- Combining the changes in circulatory mortality related costs and total hospitalisation costs calculated in this report with the reduction in pharmaceutical costs calculated in our original report and the imputed benefits resulting from reduced GP visits, time off work and school, reduced caregiver costs calculated in our original report we estimate an on-going annual benefit per household of $527.19 (95% CI: $292.99, $717.54) as a result of receiving insulation under WUNZ:HS. We estimate a greater annual benefit for households that participated in the programme as Community Service card holders at $758.07 (95% CI: $426.78, $1027.25) and a lesser annual benefit of $239.11 (95% CI: $125.07, $331.76) for households that did not participate in the programme as Community Service card holders.
- We found that receiving heating under WUNZ:HS produced annual health benefits of only $4.64 per household, $9.27 per household that participated in the programme as a Community Service card holder, and no benefit for those who did not participate in the programme as a Community Service card holder. These small annual benefits were entirely imputed from previous work.
- It is important to note that these benefits do not include any improvements in comfort, which are separate from health-related benefits, and so these calculated benefits should be treated as conservative.

Conclusion

- From a policy perspective the key result from this report is the replication of the reduction in mortality found in the original report, which was the primary benefit that was included in our cost benefit analysis [2].
- Further research will be required to explore the difference between the total hospitalisation cost savings from insulation found in this report and those in the original report.
- The reduced health benefits calculated in this report can be applied to the calculations in the original cost benefit analysis in order to estimate the likely reduction in the benefit cost ratio. We calculate that the benefit cost ratio of 3.9:1 presented in the original cost benefit analysis[2], under the favoured set of assumptions, drops to3.65:1 based on the extended data.
- This figure is conservative given that WUNZ:HS has recently exceeded the number of retrofits that the original cost benefit analysis was based on.

BACKGROUND AND INTRODUCTION

The Warm Up New Zealand: Heat Smart (WUNZ:HS) programme is a large scale government intervention that part-funds insulation and clean heat retrofits for New Zealand homes administered by the Energy Efficiency and Conservation Authority (EECA). It began on the 1st of July 2009 as a four year programme with a $340 million budget and has subsequently been extended for an additional year (2013-2014) which will be funded from the original budget, the programme having exceeded targets relative to cost expectations.

Under WUNZ:HS home owners with homes built before 2000 can receive funding towards a retrofitted ceiling insulation and/or underfloor insulation or moisture barrier, a clean heating device (if floor and ceiling insulation requirements are met), and a range of other improvements including draught proofing, hot water cylinder wraps, and pipe lagging.

WUNZ:HS represents a massive expansion of previous programmes administered by EECA such as the EnergyWise Home Grants programme both in terms of size and scope. It was motivated by the twin goals of improving energy efficiency in New Zealand and improving the health of New Zealanders. New Zealand homes are notoriously cold and damp [3, 4] for a variety of historical reasons and improved insulation and heating were expected to partially address these issues by improving indoor temperatures and reducing damp and mould. New Zealand currently experiences 1600 excess winter deaths per year[5], primarily from respiratory and circulatory causes, representing a higher rate than many much colder OECD nations. It is hypothesised that cold exposure due to poor housing/ heating is one of the explanatory factors.

Cold exposure is linked to respiratory and circulatory health through a number of plausible causal routes. With regard to circulatory health, as quoted in our original report, "the increases in platelets, red cells, and viscosity associated with normal thermoregulatory adjustments to mild surface cooling provide a probable explanation for rapid increases in coronary and cerebral thrombosis in cold weather"[6]. For respiratory health, cold exposure inhibits various respiratory defences against infection in both the upper and lower respiratory tracts. [7-9]. In addition to respiratory infections, cold exposure can induce asthma in asthmatics, and colder houses are mouldier and have higher levels of dust mite allergens, both of which can produce allergic reactions in sensitised asthmatics. Finally, unflued gas heaters and open fires release harmful particulates and gases into the indoor environment, thus improvements in heating may improve respiratory health not only by increasing temperature but by reducing indoor concentrations of these pollutants.

New Zealand randomised controlled trials have clearly demonstrated the benefits of improving temperature on respiratory outcomes such as wheezing and GP visits for asthmatic children [10, 11]. A recent systematicl review of housing intervention studies concluded that "[h]ousing improvements, especially warmth improvements, can generate health improvements; there is little evidence of detrimental health impacts. The potential for health benefits may depend on baseline housing conditions and careful targeting of the intervention" [12].

In 2010 the Ministry of Economic Development commissioned an evaluation of WUNZ:HS. A team consisting of researchers from the University of Otago (Wellington), Motu Economic and Public Policy Research, and Victoria University of Wellington produced a series of reports evaluating the impact of WUNZ:HS on industry [13], energy use [14] and health[1]. Key results from these reports were then combined to produce a final cost benefit analysis [2].

The central scenario presented in the cost benefit analysis estimated a ratio of benefits to costs of approximately 3.9:1. This outcome was driven largely by favourable findings from the original health report [1], in particular the finding that receiving treatment under the programme reduced mortality for people aged 65 and over who had had a circulatory related hospitalisation during the period between the start of the study and receiving treatment under the programme. This was estimated to result in an on-going benefit of $439.95 per year per household.

In May 2012 the Ministry of Economic Development commissioned an extension of the health research presented in [1]. The goal of this extension was to explore the impact of an additional year of health and demographic data on the key hospitalisation and mortality related analyses presented in that report. Given the impact of the mortality result on the cost benefit analysis presented in [2] our major focus is on this aspect of the prior study.

AIM

The aim of this report is to extend the hospitalisation and mortality related analyses carried out in our previous report [1] by obtaining and analysing an additional year of post-intervention data for the cohort.

In that report we calculated statistically significant reductions in mortality rates for people aged 65 and over who had undergone circulatory hospitalisation during the period from the start of the study up to and including the month before that individual received treatment. We also calculated statistically significant reductions in monthly total hospitalisation costs at a household level as the result of receiving insulation under the WUNZ:HS programme.

The current report will extend these key analyses using the expanded data set and combine the results with previously calculated outcomes for pharmaceutical costs and imputed costs that were presented in [1]. This will enable us to present revised total health benefit per household figures which can then be applied to the cost benefit analysis presented in [2].

METHODOLOGY

METHOD

The study design was set out in detail in [1]. A very brief summary of the relevant details is presented here.

The study is a retrospective cohort study. EECA provided us with addresses and details of the work done under WUNZ:HS during the period 1st July 2009 – 31st May 2010. These addresses were provided to Quotable Value (QV), who matched them with up to ten control houses that were selected to be as physically similar as possible using criteria such as age, location, size and condition.

Via a complex anonymised process, a cohort of individuals who were identified as living at these treatment and matched control addresses as at the 15th of October 2010 was then constructed based on the National Health Index data held by the Ministry of Health (MoH).

Hospitalisation and prescription data were originally collected for the cohort for the period 1st January 2008 to 31st December 2010 (we only used hospitalisation data up to 30th September 2010 due to data reliability concerns). This extension involved obtaining an additional year of hospitalisation and demographic data for the cohort, extending the study period to the 31st of Dec 2011.

Mortality data was divided into before and after periods based on the date that the first WUNZ:HS treatment occurred for treatment houses. The before period was defined as the time between the 1st of January 2008 and the last day of the month preceding the month in which the first treatment occurred. The after period began on the 1st of the month following the initial treatment and ended on the 31st of December 2011.

Mortality analysis was limited to all causes mortality for cohort members who were aged 65 and over and who had been hospitalised with a circulatory condition during the before period. We compared survival time after treatment for treatment and control group individuals.

Our analysis of hospitalisation costs took place at the household level and was based on a 'difference in difference' approach. In the original report we compared the average monthly hospitalisation costs (total, respiratory and circulatory) of treatment group households with their matched controls in the periods before and after treatment occurred. In the extension we repeat this process using the larger data set.

DATA DESCRIPTION

The original data were fully described in Telfar-Barnard et al. (2011). A very brief summary is given here. Two key adjustments to the cohort are described in full. Firstly, cohort households that

received treatment from EECA during the extension period (1st Dec 2010 – 31st Dec 2011) are identified, with treatment households having their details updated and control houses being removed from the cohort. Secondly a matching process is carried out for the extension demographic data, with individuals whose extension data does not conform to the original data being dropped if details are not sufficiently similar.

DATA SOURCES AND SEQUENCE

EECA

EECA provided a list of 46,655 addresses for homes which had been treated under the WUNZ:HS programme between the 1st of July 2009 and the 31st of May 2010. The list also included detailed information on the treatments carried out and a variety of information including tenure type and whether the treatment was funded based on the home having an occupant eligible for a Community Services Card.

QV

QV matched EECA addresses for treated homes to addresses in their database. 37,163 of the treatment addresses were able to be matched to a QV property listing, a match rate of 79.7%. QV then used a dwelling match protocol based on relevant home characteristics (see Telfar-Barnard et al. 2011 pp 15-16) to identify up to 10 control addresses for each dwelling.

31,423 treatment homes (67.4% of all treatment homes) were initially able to be matched to at least one control address and 269,110 control homes were selected. This created an initial cohort of 300,533 homes.

QV DATA FOR COHORT DWELLINGS

QV provided detailed data for the 300,533 homes including: house type, size, construction materials, location and condition. This data was not ultimately used in the original report or our current analysis but location data did inform our analysis of energy use [14].

EECA

EECA later provided us with information about treatments that were carried out during the period 1st June 2010 to 31st of November 2010[1]. We dropped 7,141 homes from the control group which EECA identified as having received treatment under the WUNZ:HS programme during that time.

[1] Telfar-Barnard et al. (2011) did not incorporate the additional treatments that treatment homes received during this time. This oversight did not impact individual level hospitalisation and mortality data but had the potential to introduce a minor misclassification bias to hospitalisation cost analyses. The analysis presented in this report adjusts for this.

Dropping these 7,141 homes resulted in dropping a further 22 treatment group homes that no longer had at least one matched control.

MINISTRY OF HEALTH DATA

At this point in the data collection process the Ministry of Health (MoH) provided addresses for every NHI record, with a unique identifier, to QV. QV matched these addresses to its list of those records still in the study cohort (31,401 treatment dwellings and 261,969 control dwellings).

We then excluded dwellings that did not have an occupant according to the NHI data provided[2]. This resulted in the exclusion of 1,492 treatment group dwellings and 26,724 control group dwellings. Removing these dwellings resulted in the further removal of 164 treatment group dwellings that no longer had at least one matched control group dwelling, and 9,318 control group dwellings that no longer had a matched treatment dwelling.

The final cohort for our initial study was 973,710 individuals from 255,672 homes. Once we had identified our final study cohort QV returned this list to MoH, who then provided us with NHI-sourced demographic data, hospitalisation (NMDS) data, and prescription data (not used in extension) for all individuals occupying a cohort home.

From the data supplied we used the following fields in the current study:

- Demographic data[15]: date of birth, death (as listed in NHI), NZ residency status, gender, ethnicity.
- Hospitalisation(NMDS) data[15]: date of admission, admission type, ICD-10 code for primary diagnosis, cost weight (calculated value designed to weight a base rate payment), and cost weight code.

EECA

In May 2012 EECA provided us with details of treatments under WUNZ:HS during the period 1st December 2010 to 31st Dec 2011, following the commissioning of this extension. 53,433 homes with distinct addresses received treatment according to our analysis of address data.

QV

[2] It is important to note that dwellings without an occupant according to NHI records are unlikely to be unoccupied in reality. NHI address records are updated regularly from public hospital databases or sweeps of primary healthcare provider databases; however interactions with hospitals are rare events, so it is possible that the current occupants of an "unoccupied" dwelling have not interacted with a hospital or primary healthcare provider while living at that address, while previous occupants have had their address updated due to an interaction with either hospital or primary healthcare provider (See Methodological Limitations section for further discussion).

We provided addresses for these 53,433 homes to QV who matched them with their own dataset. 32,703 could be suitably matched, a 61% match rate.

We linked these successfully matched addresses with the original dataset from EECA, and, using a version of the EECA dataset which was formatted as per our original data we coded the work that was done for each home. After various data cleaning processes we had 32,483 unique usable home records with accurate QV identifiers.

We then matched these 32,483 QV identified records with our cohort of 255,672 homes. We found that 439 treatment homes had had additional treatment during the December 2010 –December 2011 period and that 8,193 control group homes had received treatment thus making them ineligible for inclusion in the study. This in turn meant that the 29,513 individuals occupying these homes were no longer eligible were removed from the study. Removal of these 8,193 control homes also left 35 treatment homes without at least one matched control home. Dropping these 35 treatment homes mean removing 138 treatment individuals from the cohort.

The new cohort contained 247,444 homes and 944,059 occupants.

MINISTRY OF HEALTH

The Ministry of Health provided us with NHI-sourced demographic data (current as at July 2012), and hospitalisation (NMDS) data for 2011. We compared demographic data for the 944,059 individuals remaining in the study.

Table 1. Matching based on MoH demographic data

	Date of birth	Birth protocol	Gender	Ethnicity
Match	941,939 (99.78%)	943,668 (99.96%)	943,771 (99.97%)	917,047 (97.14%)
Don't match	2,120 (0.22%)	391 (0.04%)	288 (0.03%)	27,012 (2.86%)
Total	944,059	944,059	944,059	944,059

The very minor discrepancies in the data are a natural artefact of the processes whereby the NHI dataset is updated. Individual contact with hospitals and regular electronic "sweeps" of primary health databases are opportunities for newer data to replace previously held data and for data entry errors and mistakes to enter the system (and for inaccuracies to be corrected). Internal MoH algorithms regularly resolve issues with multiple NHIs assigned to a single individual and in some cases several people may knowingly use the same name in order to access healthcare that might not otherwise be legally available, resulting in discrepancies in NHI data.

Given that we have no independent way to assess whether the original demographic data or the more recently supplied data are more accurate, we chose, for pragmatic reasons, to base the current dataset on our original dataset, which meant formulating rules for the degree of discrepancy with the new demographic data that seemed acceptable before an individual had to be removed from the cohort. We accepted that a perfect match on date of birth was probably too strict a criterion, but

gender seemed too important to accept a mismatch. We thus removed 288 individuals who were mismatched on gender (note that gender in the NHI context refers to birth sex and would not change following gender reassignment surgery). After consultation with MoH staff we established a protocol to ensure that the date of birth field was similar enough for a legitimate match. We eliminated records if date of birth was more than five years different from the original record. We also eliminated records if more than one field (day, month, or year) was different from the original. This ensured that we did not eliminate records that were different as a result of a single typing error. We dropped 382 further individuals using this method (9 having already been dropped due to gender mismatch).

Ethnicity is a more complex variable (evidenced by a higher percentage of mismatches) with numerous coding options and more potential variability in the way that individuals may self-identify over time and thus potential for the variable to differ. After consultation with MoH we concluded that it would be too difficult to formulate a meaningful matching rule for this variable and thus did not remove any cohort members due to discrepancies.

This left the cohort at 943,389 individuals from 247,413 homes.

Finally we had to test whether these removals left any homes "empty" and thus any treatment homes without an occupied matched control home or vice versa. 24 control homes with 61 occupants were dropped due to their matched treatment home now having no occupants and 1 treatment home with 5 occupants had to be dropped because it had no matched occupied controls.

The final adjusted cohort was 943,323 individuals from 247,388 homes, of whom 110,048 individuals were in 29,706 treated homes and the remainder were in control homes.

DATASET CREATION

EXPOSURE TIME

As noted in our original study, days were assigned as "before" if they occurred during the period from the start of the study to the last day of the month before the month in which the first treatment occurred. The days of the month in which treatment occurred were excluded to account for uncertainty around the actual date of installation. Exposure days were assigned "after" if they occurred from the first day of the month following treatment to the end of the study to the 31st of December 2011. Individual exposure was limited by birth date and death date if either occurred during the period of the study. Exposure based measures inform both the analysis of mortality and household level hospitalisation costs.

HOSPITALISATIONS

EXCLUSIONS

The exclusions in our present study for household level cost analysis are identical to the exclusions in the original study.

Table 1. Type and reasons for excluding hospitalisation cost data

Exclusion	NMDS description	Reason for exclusion
Waiting list admissions	Admission type="WN"	Potential for systematic bias: being on a waitlist may make a person more likely to be identified as suitable for WUNZ:HS participation, thus potentially introducing "return to the mean" type biases i.e. unwell people typically improve in health after hospitalisation which could make the hospitalisation rates of the treatment group appear to improve post treatment.
Birth events	Event type="BT"	Not adverse health events
Transfers	Admission source="T"	Included (see below)
Readmissions	Admission date within 30 days of previous discharge date	Included (see below)
Non-New Zealand residents	New Zealand resident indicator="N"	Not relevant to study
ICD-10 Chapter 15 (Pregnancy, childbirth and the perineum)	Clinical code starting with "O"	Majority not adverse health events; distribution of hospitalisations driven by birth rate and events nine months prior.
ICD-10 Chapter 16 (Certain conditions originating in the perinatal period)	Clinical code starting with "P"	Majority non-relevant health events; distribution of hospitalisations driven by birth rate and events nine months prior.
ICD-10 Chapter 17 (Congenital malformations, deformations and chromosomal abnormalities)	Clinical code starting with "Q"	Distribution of hospitalisations driven by birth rate and events nine months prior.
ICD-10 Chapter 21 (Factors influencing health status and contact with health services)	Clinical code starting with "Z"	Non-relevant health events; most do not represent conditions subject to change

We included transfers and readmissions in our analysis of cost data because, while not new health events, transfers and readmissions represent real costs, and needed to be captured in order to accurately capture costs. Our analysis focuses on total hospitalisations and includes all hospitalisation events apart from those excluded above.

DATA ANALYSIS

KEY CHARACTERISTICS

As with our original study, the current study is observational, rather than experimental, and this leads to the possibility for confounding where the self-selected treatment group differs systematically from the matched control group. The original report compared characteristics such as age, gender, ethnicity and deprivation for the treatment group, control group and the New Zealand population based on 2006 Census data. The original report concluded that while there inevitably were statistically significant differences in these variables (due to the high numbers involved) the only variable that was deemed different enough to warrant inclusion in the analyses of individual level data was age.

Table 2 below includes age data for the updated dataset of 943,323 individuals, the original dataset and the 2006 Census. No obvious changes have resulted from the reduction of the dataset.

Table 2. Age distribution of treatment, control and Census 2006 populations

	Treatment (original)		Control (original)		Census 2006	
Age group	n	%	n	%	n	%
0-4 years	7,286	6.9	53,500	6.3	275,034	6.8
5-14 years	13,126	12.4	117,262	13.7	592,365	14.7
15-24 years	12,563	11.9	124,064	14.5	570,960	14.2
25-44 years	32,493	30.7	265,619	31.0	1,133,739	28.2
45-59 years	17,823	16.8	162,098	18.9	779,226	19.4
60+ years	22,523	21.3	135,916	15.6	671,718	16.7
	Treatment (following exclusions)		Control (following exclusions)		Census 2006	
0-4 years	7,315	6.91	50,204	6.22	275,034	6.8
5-14 years	13,096	12.37	110,582	13.71	592,365	14.7
15-24 years	12,517	11.82	117,593	14.58	570,960	14.2
25-44 years	32,411	30.60	249,016	30.87	1,133,739	28.2
45-59 years	17,769	16.78	153,062	18.98	779,226	19.4
60+ years	22,803	21.53	126,125	15.64	671,718	16.7

At the household level, the original report noted statistically significant differences in the distribution of dwelling types based on a dwelling typology developed by Lucy Telfar-Barnard [1]. These differences appeared too small to have clinical significance. We follow the original report in excluding a dwelling type variable from analyses.

MODEL SELECTION

MORTALITY

This report focuses on mortality in individuals aged 65 and over who had had a circulatory hospital admission during the before period, as this was the group which showed a statistically significant reduction in mortality as a result of participating in the programme in our original report and this result in turn was the primary benefit included in our cost benefit analysis[2].

It is not possible to do a before-after analysis for mortality, since people dying in the "before" period could not participate in WUNZ:HS, meaning that mortality rates in the treatment group would be lower than otherwise expected in this period.

The original model used was a standard Poisson model for which the dependent variable was "d-after", a binary variable that was 0 if an individual did not die during the post-treatment period and 1 if they did. The model adjusted for the cost of previous circulatory hospitalisations, as a marker for severity of illness using a variable called "d-cost". Potential exposure time for a given individual was defined as the time period beginning on the first day of the month after the treatment month and ending on the 31st of December 2010 (the "after" period).

The current analysis uses a Cox proportional hazards model. The change in model addresses two concerns. Firstly, we concluded that the use of a binary dependant variable ("d-after") was not ideal for a Poisson model, particularly one that included a continuous variable, although the outcome was unlikely to be much different using a Cox proportional hazards model [16]. Secondly, we concluded that we needed to modify how we adjusted for hospitalisation cost. We replaced "d-cost" with a variable called "annual_circulatory_hosps" because "d-cost" was calculated in such a way as to introduce endogeneity. The replacement "annual_circulatory_hosps" variable is calculated by summing the cost of all circulatory hospitalisations that an individual underwent during the "before" period, calculating a cost per year figure and finally taking the log of this figure.

The Cox proportional hazards model defines "failure" as death during the "after" period, being the period from the end of the treatment month for a given individual to either the end of the study (31st of December 2011) or the date of death if a person died during the post-treatment period.

Applying this model to the original dataset, the Cox proportional hazards based model produces outcomes similar to those produced using the original Poisson model when the "dcost" variable was excluded, as suggested by Laird and Olivier [16]. Table 3 sets this result out.

Table 3. Effect of treatment on mortality rates in people aged 65+ hospitalised with circulatory illness prior to treatment month (original results with additional calculation)

	Hospitalised before treatment month(number)	Deaths after treatment month (number)	Mortality rate per 1000 people per year	RR (95% CI) p-value	Hazard Ratio (HR) calculated using Cox proportional hazards model including "annual_circulatory_cost" variable
Treatment	958	51	69.8 (51.2-88.3)	0.62 (0.45-0.84) p=0.002	0.61 (0.45-0.81) p=0.002
Control	2231	184	112.7 (97.2-128.2)		

HOSPITALISATION: COST DATA

Analysis of hospitalisation cost data followed that in the original report and was based on a difference in difference approach at the household level. The model compared the difference between each treatment group household's monthly hospitalisation costs with the mean of its matched control group households' monthly hospitalisation costs both before and after the intervention.

The data was cleaned by removing values lower than the 1st percentile and higher than the 99th percentile – for reasons of consistency and balance we removed all observations associated with a household cluster (before and after) if either observation was removed. One of the key reasons for cleaning the data in this way was reduced skew. Figure 1 details the data cleaning process.

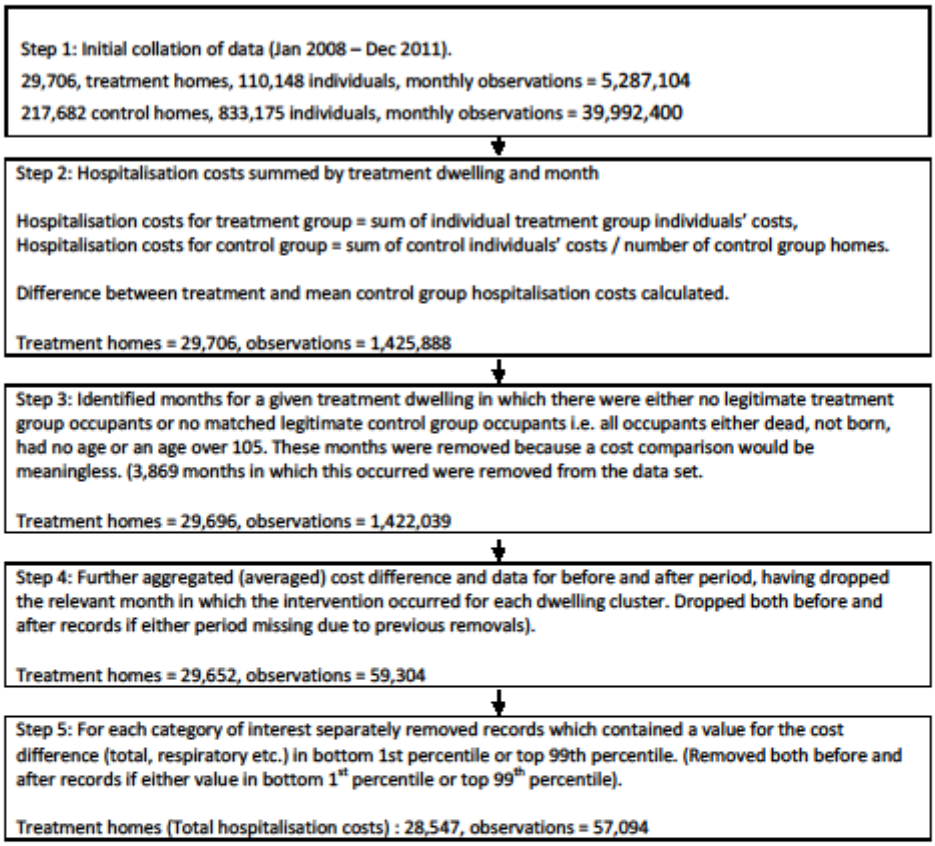
We use a fixed effects OLS estimator with standard errors clustered by treatment/control pairings to analyse the cost data. The equation and explanation following are taken directly from the original report [1](p 31).

$$Hospitalisation\ Cost\ Diff_{it} = \alpha_i + \beta_1 insulation_{it} + \beta_2 heating_{it} + \varepsilon_{it}$$

Hospitalisation Cost Diff$_{it}$ represents the difference between the averaged monthly hospitalisation costs of treatment house *i* and its control group houses over time *t* (*t* is either before or after). α_i represents the unobserved individual house fixed effect of house *i* relative to its control houses, *insulation*$_{it}$ is a dummy variable, that is 1 if house *i* [was assumed to have] retrofitted insulation (ceiling, floor or both) for the entirety of period *t* (i.e. the after period), heating$_{it}$ is a dummy variable that is 1 if house *i* [was assumed to have] retrofitted heating for the entirety of period *t:* the coefficients of the insulation and heating dummies, if statistically significant, indicate the size and direction of any change in the difference in average monthly hospitalisation costs as a result of treatment . ε_{it} is the residual term, which is correlated between periods within houses, but independent between houses. Note that all covariates which are constant over time for a given household (e.g. region, deprivation, ethnicity of occupants (assuming no changes to household composition)) are absorbed into the fixed effect α_i and do not need to be explicitly included in the model. We initially considered including an additional interaction term between insulation and heating. However exploratory analyses found that the coefficient of this term was not significantly different from zero, so we did not include it in the final model set out above.

The analysis focuses on exploring the impact of insulation and heating on total hospitalisation costs, a key result which informed the cost benefit analysis presented in [2]. We also carry out a further sub-analysis looking at the results of the intervention for those households who qualified for the WUNZ:HS programme as Community Services Card holders, and those who did not. For each fitted model we check the validity of our conclusions by utilising a bootstrap calculation of the standard errors with 500 repetitions to address issues noted in the original analysis around the fact that residuals were highly concentrated at zero.

Figure 1: Flow diagram of method of calculating difference between treatment and control household total monthly hospitalisation costs



METHODOLOGICAL LIMITATIONS

The methodological limitations of the current study are largely those of the original study discussed in the original report [1] (p 35). The following is a brief summary:

The accuracy of the NHI dataset addresses for individuals, which were extracted on the 15th of Oct 2010, is a primary source of uncertainty. The address field in the NHI dataset reflects the last update of the field, which may occur following contact by an individual with a hospital or with a primary health care provider if they have provided address details. As we understand it, there has never been any research carried out to independently verify the accuracy of the NHI address field at a given point in time.

It seems likely that the treatment group addresses are more accurate because the group are less healthy on average and thus more likely to have contact with treatment providers and thus have updated address fields, and also because the decision to undertake treatment under WUNZ:HS suggests a commitment to one's current location (at least for some individuals). It is not clear which direction the likely greater accuracy of treatment group addresses may bias the analysis.

The study is, however, inevitably biased towards the null because some control households will already have the same "treatment" i.e. recent insulation or a heat pump as their matched controls, which they have purchased prior to the study period and separate from the WUNZ:HS programme.

In addition, during the period of the study some control households/ individuals will purchase insulation and/or heating outside of WUNZ:HS. These factors mean that the study is not comparing treated with untreated homes, but rather "all treated" with "some treated". This is particularly plausible in the case of heat pumps given their recent popularity.

A limitation of the original study was the limited post-treatment exposure time for the cohort. This study addresses this limitation, increasing the amount of post-treatment exposure time and thus giving the study greater power.

RESULTS

MORTALITY

The key result of our original analysis was a statistically significant reduction in mortality for people aged 65 and over who had been hospitalised with circulatory conditions (ICD-10 chapter IX) as a result of receiving treatment under the programme. The analysis of the extended data set which included mortality data for 2011 was largely consistent with the previous analysis. Individuals aged 65 and over in the treatment group had a significantly lower mortality rate than those in the control group (HR 0.68, 95%CI 0.55-0.84, p<0.000) after adjusting for cost of hospitalisation using the annual_circulatory_cost variable. Table 4 presents the results.

Table 4. Effect of treatment on mortality rates in people aged 65+ hospitalised with circulatory illness prior to treatment month

	Hospitalised before treatment month (number)	Deaths after treatment month (number)	Mortality rate per 1000 people per year	Hazard ratio calculated using Cox proportional hazards model including "annual_circulatory_cost" variable
Treatment	885	136	75.57	HR 0.68 (0.55 – 0.84) p<0.000
Control	2036	413	101.32	

The new results are largely consistent with the original results calculated using the Cox proportional hazards model , although the Hazard Ratio calculated is slightly less favourable (the HR calculated using the original data was 0.61 (0.45-0.81) p=0.002 using the annual_circulatory_cost variable). This may reflect normal statistical variation or it may suggest that the protective effect of treatment under the study is strongest soon after a circulatory illness, which seems plausible.

MORTALITY: COST DATA

We follow the analysis presented in the original report [1] in order to calculate the value of the reduction in mortality rates presented above. Firstly we estimate the change in the number of deaths per household in the year following treatment:

- We chose a point in time, July 2009, the start of the WUNZ:HS programme, as the basis for our calculations.
- We counted up the number of living individuals with valid ages (age less than 105, age not missing) in our total treatment cohort at that time (107,165), the number of dwellings they lived in (29,661) and the average number of individuals per dwelling (3.61).
- We counted the number of individuals who were over 64 and who had been hospitalised for circulatory illness in the prior 18 months within this cohort (834).
- We calculated the average number of such individuals living in a treatment dwelling *=834 /29,661 = 0.028 (3 d.p.)*

Assuming that, in the absence of treatment, the mortality rate for such individuals would be 101.32 deaths per 1000 individuals in the year after treatment (i.e. the estimate for the control group in Table 4), and given that there is a 32% reduction in mortality (see Table 4), we estimate that in the year following the installation of insulation the following number of deaths would be avoided:
*= 0.028 (no. of individuals) * (101.32/1000) (control group mortality rate) *0.32 (32% reduction in mortality)*
= 0.908 deaths avoided per 1000 households (95% CI: 0.454, 1.277)

As with the original report it is necessary to estimate how much additional life an individual might gain as a result of reduced mortality due to treatment under the programme, and to assign a value to the additional life gained by these individuals. We follow the original report in assigning a value of $150,000 per life year gained, irrespective of age or health status based on an influential 1991 report

by Miller and Guria [17] which estimates the value of a preventable fatality at $3,352,448 ($NZ 2008).

We predicted that 0.908 deaths per thousand households would be avoided in the year following treatment. We value the reduced mortality using a figure of $150,000 for the value of a life year. We expect that, on average, a person that would have died would have died halfway through the year: this means in the year following treatment the value of reduced mortality per *dwelling* is:

= (0.908/1000) (reduced mortality) * $150,000 (value of a life year) *0.5 (to account for half year)
=$ 68.10 (95% CI: $34.05, $95.78)

A complication is that, for the 0.908 individuals per thousand dwellings who did not die in the first year, additional benefits will accrue for each year that they live beyond the first: based on an average age of 76.63 we conservatively predict they will live an additional 11 years based on figures reported in Statistics New Zealand's Life Tables (www.statistics.govt.nz).. Note the saving from reduced mortality per household in year one is $53.18 because avoided mortality is assumed to happen halfway through the year , which means that the additional savings in future years are valued at 2*$53.18=$106.36. At a discount rate of 5% the saving per dwelling would be:

$$= \$68.10 + \sum_{n=1}^{n=10} (\$136.20)/((1+0.05)\text{^}n)$$
$$= \$1119.80 \text{ (95\%CI: \$559.90,\$1574.87)}$$

We have modelled the probability of a vulnerable person avoiding mortality as a result of the intervention. The probability of this is (101.32/1000)*0.32= 0.032 (2.5%). As we explained in the original report we treat avoidance of mortality by treatment in each year as independent events and thus the multi-year benefit calculated above would accrue based on the life years gained as a result of deaths avoided in year one. We expect that these benefits will accrue in year two for different vulnerable individuals (aged 65 and over with a cardiovascular related hospitalisation in previous 18 months), and for different individuals again in every subsequent year that the treatment continues to have an effect. This means an on-going stream of benefits of $1119.80 per year.

We follow the original report in reducing the figure of $1119.80 per dwelling per year because it is based on the untenable assumption that an elderly person who has had a circulatory hospitalisation will continue to live for an average length of time predicted for a person of their age. We instead assume that they will live for five additional years. At a discount rate of 5% the on-going annual savings are:

$$= \$68.10 + \sum_{n=1}^{n=4} (\$136.20)/((1+0.05)\text{^}n)$$
$$= \$551.06 (95\%CI: \$275.53,\$775.00)$$

As we describe in the original report, we reduce this on-going annual figure by a conservative 15% in order to account for individuals who avoid death due to the treatment in multiple years, thus avoiding over-counting.

=$551.06*0.85

=$468.40 (95% CI: $234.20, $658.75)

Finally, we note that, unsurprisingly, the proportion of vulnerable individuals was different depending on whether or not they lived in a dwelling that received treatment under the programme as a Community Services Card holder. In our treatment cohort as at July 2009 there were 60,365 CSC individuals, and 655 were in the vulnerable group. There were 46,797 non-CSC individuals and of them 179 were in the vulnerable group. This means that, assuming a household with 3.61 occupants and carrying out the calculations described above, we would expect the following on-going annual savings given our assumptions:

CSC dwelling = **$ 662.58 (95% CI: $331.29, $931.76)**

Non CSC dwelling = **$ 228.07 (95% CI: $114.03, $320.72)**

HOSPITALISATION: COST DATA

We estimated each of our hospitalisation cost models using a fixed-effect OLS estimator with standard errors clustered by house. Table 5 presents the results of our primary analysis of hospitalisation costs. Our conclusions were tested using a bootstrap estimation of standard errors involving 500 repetitions. Bootstrapping produced similar standard errors and p-values, confirming the validity of the model.

Table 5. Change in monthly total hospitalisation costs per household

	Total hospitalisation	**Total hospitalisation (CSC holders)**	**Total hospitalisation (non-CSC holders)**
Insulation	**$1.67**	**$1.31**	**$1.96**
Robust standard errors	(2.007)	(3.002)	(2.552)
95% confidence interval	(-2.26, 5.61)	(-4.57,7.19)	(-3.04,6.97)
Heating	**$5.93**	**$11.38***	**-$1.15**
Robust standard errors	(4.028)	(5.908)	(5.222)
95% confidence interval	(-1.96, 13.83)	(-0.20,22.96)	(-11.36, 9.08)
Observations	57,148	31650	25498
Number of households	28,574	15825	12749
R-Squared (within)	0.0002	0.0003	0.0000
R-Squared (adjusted)	0.0001	0.0002	0.0000
*** p<0.01, ** p<0.05,* p<0.1			

The results presented in Table 5 indicate a non-statistically significant *increase* of approximately $1.67 in total hospitalisation costs per month for a household that received some combination of ceiling or floor insulation under WUNZ:HS. The results also show that total hospitalisation costs were increased for households who received heating under the programme by about $5.93 although this result was also non-significant.

For households that received the intervention as Community Service card holders there was a non-statistically significant increase in hospitalisation costs of $1.31 per month for those who received insulation, and a larger increase in hospitalisation costs of $11.38 that was statistically significant at the 10% level for those who received heating.

For households that received the intervention as non-Community Service card holders there was a non-statistically significant increase in hospitalisation costs of $1.96 per month for those who received insulation, and a non-statistically significant reduction in hospitalisation costs of $1.15 for those who received heating.

These results appear to be inconsistent with those obtained under the original analysis, which found a statistically significant (albeit small) reduction in total hospitalisation costs as a result of receiving insulation under WUNZ:HS. We speculate that these less favourable results may reflect the impact of reduced mortality resulting from participation in WUNZ:HS. Over a greater period of time the impact of having more vulnerable elderly people continuing to live in the treatment households – and therefore continuing to need hospitalisation – could increasingly outweigh any reduction in hospitalisation costs resulting from improved occupant health.

SUMMARY DATA FOR COST BENEFIT ANALYSIS

Here we take the benefits and costs imputed from previous studies and the changes in pharmaceutical costs that were calculated in the original report and combine them with the mortality and total hospitalisation cost results presented in this report in Table 6. For details on how these figures were calculated consult the original report[1]. We also include comparable results from the original report in Table 7.

Table 6. Summary of annual health-related benefits (savings) per household

Household type	Benefits	Insulation	Heating
All households	Total Hospitalisation related benefits calculated in present Study	$0.00	$0.00
	Total pharmaceutical use related benefits calculated in original study	$11.04	$0.00
	Additional benefits imputed from previous studies	$47.75	$4.64
	Value of reduced mortality in present Study	$468.40 (95% CI: $234.20, $658.75)	$0.00
	SUM OF HEALTH BENEFITS	$527.19 (95% CI: $292.99, $717.54)	$4.64
Households that participated in WUNZ:HS programme as Community Services Card holders	Total Hospitalisation related benefits calculated in present Study	$0.00	$0.00
	Total pharmaceutical use related benefits calculated in original study	$0.00	$0.00
	Additional benefits imputed from previous studies	$95.49	$9.27
	Value of reduced mortality in present Study	$ 662.58 (95% CI: $331.29, $931.76)	$0.00
	SUM OF HEALTH BENEFITS	$758.07 (95% CI: $426.78, $1027.25)	$9.27
Households that participated in WUNZ:HS programme as non-Community Services Card holders	Total Hospitalisation related benefits calculated in present Study	$0.00	$0.00
	Total pharmaceutical use related benefits calculated in original study	$11.04	$0.00
	Additional benefits imputed from previous studies	$0.00	$0.00
	Value of reduced mortality in present Study	$228.07 (95% CI: $114.03, $320.72)	$0.00
	SUM OF HEALTH BENEFITS	$239.11 (95% CI: $125.07, $331.76)	$0.0

Table 7. Summary of annual health-related benefits (savings) per household from original report for comparison

Household type	Benefits	Insulation	Heating
All households	Hospitalisation and pharmaceutical use related benefits calculated in present Study	$64.44 (total hospitalisations) + $11.04 (total pharmaceuticals) = $75.48	$0.00
	Additional benefits imputed from previous Studies	$47.75	$4.64
	Value of reduced mortality	$439.95 (95% CI $0.00 – $765.84)	$0.00
	SUM OF HEALTH BENEFITS	$563.18 (95%CI $123.23 - $889.07)	$4.64
Households that participated in WUNZ:HS programme as Community Services Card holders	Hospitalisation and pharmaceutical use related benefits calculated in present Study	$109.80 (total hospitalisations)	$0.00
	Additional benefits imputed from previous Studies	$95.49	$9.27
	Value of reduced mortality	$613.05 (95% CI $0.00 - $1,067.16)	$0.00
	SUM OF HEALTH BENEFITS	$818.34 (95% CI $205.29, $1,272.45)	$9.27
Households that participated in WUNZ:HS programme as non-Community Services Card holders	Hospitalisation and pharmaceutical use related benefits calculated in present Study	$11.04 (total pharmaceuticals)	$0.00
	Additional benefits imputed from previous Studies	$0.00	$0.00
	Value of reduced mortality	$216.38 (95%CI $0.00 - $376.66)	$0.00
	SUM OF HEALTH BENEFITS	$227.42 (95% CI $11.04 - $387.70)	$0.00

CONCLUSION

This report confirmed the key finding in our original report, while challenging others. The key benefit reported in the original report, reduced mortality for individuals aged 65 and over who had a circulatory hospitalisation in the "before" period of the study, was also found in the present report.

The statistically significant reductions in total hospitalisation costs as a result of receiving insulation under WUNZ:HS found in the original report were not found in the current report. We are not certain of the reason for this discrepancy, but speculate that any reductions in terms of total hospitalisation costs due to insulation or heating are being offset by increases in hospitalisation costs for treatment households due to lower mortality rates for those aged 65 who had recently had a circulatory hospitalisation in the treatment cohort. This possibility will be further explored by Nicholas Preval as part of his Doctoral Thesis.

From a policy perspective the key result from this report is the replication of the reduction in mortality found in the original report, which was the primary benefit that was included in the cost benefit analysis presented in [2].

The cost benefit analysis states that "the central estimate of gross benefits for the programme is $1.28 billion compared with resource costs of $0.33 billion, a net benefit of $0.95"[2], a benefit to cost ratio of 3.9:1. We estimate that the lower health benefits per household calculated in this report will reduce the benefit cost ratio presented in [2] from 3.9:1 to approximately 3.65:1. It is important to note that as WUNZ:HS has recently exceeded the target number of retrofits that the cost benefit analysis was based on this figure should be seen as conservative.

REFERENCES

1. Telfar Barnard, L., Preval, N., Howden-Chapman, P., Arnold, R., Young, C., Grimes, A., and Denne, T., *The impact of retrofitted insulation and new heaters on health services utilisation and costs, pharmaceutical costs and mortality: Evaluation of Warm Up New Zealand: Heat Smart*, 2011.

2. Grimes, A., Denne, T., Howden-Chapman, P., Arnold, R., Telfar-Barnard, L., Preval, N., and Young, C., *Cost Benefit Analysis of the Warm Up New Zealand: Heat Smart Programme*, 2011.

3. Telfar-Barnard, L., *Home truths and cold admissions: New Zealand housing attributes and excess winter hospitalisations*, in *Public Health Department*2010, University of Otago: Wellington.

4. Howden-Chapman, P., Viggers, H., Chapman, R., O'Dea, D., Free, S., and O'Sullivan, K., *Warm homes: drivers of the demand for heating in the residential sector in New Zealand.* Energy Policy, 2009. **37**(9): p. 3387-3399.

5. Davie, G., Baker, M., Hales, S., and Carlin, J., *Trends and determinants of excess winter mortality in New Zealand: 1980 to 2000.* BMC Public Health, 2007. **7**(1): p. 263.

6. Keatinge, W.R., Coleshaw, S.R.K., Cotter, F., Mattock, M., Murphy, M., and Chelliah, R., *Increases in platelet and red cell counts, blood viscosity, and arterial pressure during mild surface cooling: factors in mortality from coronary and cerebral thrombosis in winter.* British Medical Journal Clinical Research Ed, 1984. **289**(6456): p. 1405-8.

7. Hajat, S., Bird, W., and Haines, A., *Cold Weather and GP consultations for respiratory conditions by elderly people in 16 locations in the UK.* European Journal of Epidemiology, 2004. **19**(10): p. 959-968.

8. Mäkinen, T.M., Juvonen, R., Jokelainen, J., Harju, T.H., Peitso, A., Bloigu, A., Silvennoinen-Kassinen, S., Leinonen, M., and Hassi, J., *Cold temperature and low humidity are associated with increased occurrence of respiratory tract infections.* Respiratory Medicine, 2009. **103**(3): p. 456-462.

9. Eccles, R., *An explanation for the seasonality of acute upper respiratory tract viral infections.* Acta Oto-laryngologica, 2002. **122**(2): p. 183-191.

10. Howden-Chapman, P., Matheson, A., Crane, J., Viggers, H., Cunningham, M., Blakely, T., Cunningham, C., Woodward, A., Saville-Smith, K., and O'Dea, D., *Effect of insulating existing houses on health inequality: cluster randomised study in the community.* BMJ, 2007. **334**(7591): p. 460-469.

11. Howden-Chapman, P., Pierse, N., Nicholls, S., Gillespie-Bennett, J., Viggers, H., Cunningham, M., Phipps, R., Boulic, M., Fjallstrom, P., and Free, S., *Effects of improved home heating on asthma in community dwelling children: randomised controlled trial.* BMJ, 2008. **337**(a1411): p. 848-862.

12. Thomson, H., Thomas, S., Sellstrom, E., and Petticrew, M., *The health impacts of housing improvement: a systematic review of intervention studies from 1887 to 2007.* American Journal of Public Health, 2009. **99**(Supplement 3): p. S681-S692.

13. Denne, T. and Bond-Smith, S., *Impacts of the NZ Insulation Fund on Industry & Employment*, 2011.

14. Grimes, A., Young, C., Arnold, R., Denne, T., Howden-Chapman, P., Preval, N., and Telfar-Barnard, L., *Warming Up New Zealand: Impacts of the New Zealand Insulation Fund on Metered Household Energy Use*. 2011.

15. National Health Board Business Unit, *National Minimum Dataset (Hospital Events) Data dictionary*, 2010, Ministry of Health: Wellington.

16. Laird, N. and Olivier, D., *Covariance analysis of censored survival data using log-linear analysis techniques.* Journal of the American Statistical Association, 1981. **76**(374): p. 231-240.

17. Miller, T.R. and Guria, J., *The Value of Statistical Life in New Zealand: Market Research on Road Safety*, 1991: Wellington, New Zealand.