Energy Epidemiology: an epidemiological approach to empirically-based population-level energy demand research

By

Ian George Hamilton

UCL Energy Institute

The Bartlett School of Environment, Energy and Resources

University College London

A dissertation submitted for the degree of

Doctor of Philosophy

Declaration

I, Ian George Hamilton confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgments

In completing this work I would like to thank a number of people who provided guidance and support along the way, without whom this ambitious task would not have been completed.

From UCL, I would like to sincerely thank my supervisors: Prof. Tadj Oreszczyn, who provided sage advice and support whenever I needed it; Prof. Philip Steadman, who offered thoughtful analysis and considered advice, and grammatical and editorial assistance; and Dr. Alex Summerfield, who provided the initial ideas behind this work, lots of advice on methods, and plenty of stimulating conversations. Also, to the late Harry Bruhns whose thoughtful advice and instruction prepared me for undertaking this research.

In addition to my outstanding supervisory team, I would like to thank other UCL colleagues who spent time in supporting my research, including: Prof. Mike Davies, who tempted me back to UCL, Prof. Bob Lowe, and Dr. David Shipworth.

My PhD was carried out part-time and I would like to acknowledge the support from the Engineering and Physical Science Research Council (EPSRC) under the research grants 'New Empirically-Based Models on Energy Demand in Buildings' [EP/I038810/1] and 'RCUK Centre for Energy Epidemiology' [EP/K011839/1]

I would like to thank my family who played such an important role in helping me achieve this degree: my parents, Judith and Bob Hamilton, who have supported me throughout my career by providing me advice and encouragement, and who gave me all the love and opportunities in life; my Mississauga family, Jo-Anne and George Osterbauer, who offered love, support, interest and a place to stay; and my grandfather, Ian Matthew, who has been a lifelong (and long-lived) source of inspiration, along with my late grandmother, Sheila Matthew, who missed 'having a doctor in the family'.

I would also like to specially thank Liz Gill, my London champion and family while away from home, whose generosity, encouragement and enthusiasm have been dearly appreciated.

Finally, I would like to thank my wonderful wife Dawn, whose companionship has given me the strength to pursue this task and whose encouragement and patience made this possible.

Abstract

The shift to a low carbon economy and the need to address energy demand priorities will involve the retrofit of millions of buildings resulting in changes in energy demand services at the national and international scale. Studying energy demand in buildings at a population level is different than in individual or small samples because of population heterogeneity. Evaluating policies and determining the effect of technologies *in situ* in millions of buildings means using techniques that support that level of analysis and use empirically derived data that can represent complex real-world conditions. Health epidemiology, which studies the distribution and determinants of population health outcomes, offers a compelling framework for studying population level energy demand.

The aim of this thesis is to determine whether the adaption of the conceptual and methodological framework of epidemiology can support the study of energy, people and buildings. This thesis tests this hypothesis by examining relevant epidemiological concepts and its methodological framework along with three studies that adapt and apply epidemiological methods to energy demand and energy efficiency retrofits in UK houses. The method studies use a database of over 13 million dwellings to study energy efficiency retrofit uptake and their impact on energy demand.

The method study findings support the case that an epidemiological approach to energy demand provides an appropriate and plausible conceptual and methodological framework for determining population-level evidence to inform modelling and policy development and evaluation. Adapting the epidemiological approach is not a panacea to dealing with the challenges facing the field of research in energy demand in buildings. However, it does provide a set of concepts, methods and analysis tools that are capable of supporting an empirically-based population-level research approach, identified as a necessary step towards to developing a robust foundation of evidence.

Table of contents

List of figures

List of tables

Chapter 1 Introduction

"A new approach to studying population-level end-use energy demand"

Chapter Introduction

Globally, over the past decade, there has been a strong call from researchers and policymakers for action to address end-use energy demand. These calls are predicated on the need to address a range of issues, including: global climate change and national greenhouse gas (GHG) abatement targets, energy security, price stability, economic productivity, and consumer access to health and well-being. Achieving a sustained change in energy demand requires a deep understanding of the fundamental drivers across the population, accounting for the environmental and socio-technical interactions.

The purpose of this chapter is to introduce the thesis topic and to set out the path this research will follow. This chapter introduces the context around research in energy demand, outlines the problem that this thesis will focus on, sets out the aims of the research, and provides an overview of the components of the thesis.

1.1 Context

Over 70% of global GHG emissions are related to fossil fuel use (IPCC, 2014). Allocated to end users, buildings comprise ~18% of total GHG emissions, transport ~14%, and industry ~29%, with waste, flaring and other energy ~14%. Agriculture and land use comprises the remaining ~25%. The majority of these emissions are related to the built environment (i.e. buildings, transport and industry), providing services such as cooling and heat in buildings, power for lights, appliances, electronics and computing, and motive power for moving to and within largely urbanised places.

In 2012, global energy demand was approximately 154 PWh (EIA, 2013). It represents one of the single most important components of the global economy and is an essential contributor to supporting quality of life, maintaining health and wellbeing, transporting people and goods, extracting materials and creating products, and communicating, amongst many others. Global energy demand estimates by final user are dominated by the industrial sector (51%), followed by transportation (20%), residential (18%) and commercial (11%) (EIA, 2013). Total energy demand in buildings (i.e. non-industrial processes or assembly) for services (e.g. heating, cooling, lighting, appliances) account for around 20% of total global energy use (EIA, 2013).

Energy efficiency in buildings has been identified as a major potential source of GHG savings but, to exploit this resource, current efforts require a significant increase in programmes and policies for energy efficiency in buildings than is currently applied (IPCC, 2014). The International Energy Agency (IEA) has identified both a growth in demand for energy globally and, simultaneously has expressed caution about future traditional (fossil) fuel supplies (IEA, 2012). They have advocated that energy savings should be viewed as the *first fuel* resource to exploit (IEA, 2013), and that these could deliver savings in 2035 of up to 20% of total demand in 2010 (IEA, 2012). However, the IPCC have highlighted that uncertainties in meeting emission reductions are related to the complex and highly contextual requirements around energy demand and the application of efficiency programmes.

Globally, these energy savings need to occur alongside a projected continuing growth in low-carbon energy demand and the shift towards low-carbon fuels. Energy demand in the poorest households must increase to improve their livelihoods and health (OECD/IEA, 2010). This tension between the energy demand convergence of developed and developing consumers and the need to contract overall carbon-intensive energy use means that achieving an equitable and sustainable change in demand is fundamentally complex.

1.2 Problem statement

The inter-relationship between people, energy and buildings is complex and brings together many interacting factors and activities. Energy policy for residential energy demand is focused on a heterogeneous populations of dwellings and households. The complexity of the housing stock, the importance of houses in people's lives, and the wide spectrum of interacting agents all make energy and housing an important area of "policy resistance".

Energy policy seeks to manage and shift energy demand in buildings through changes in occupant practices and improvements in building performance through technological improvements. However, for the most part, basic information about energy demand in buildings, e.g. trends and patterns along with simple descriptions of population and stock segmentations is limited or simply lacking (Skea, 2012; Summerfield and Lowe, 2012). Without even basic descriptions and agreed metrics of energy demand in buildings, developing a policy framework to achieve change in demand is undermined by the general lack of a strong evidence base and a misunderstanding of consequential drivers. Historically, this lack of evidence is related to prioritisation of funding, the transient nature of academic research, and a dearth of observed data and therefore reliance on models that are often poorly informed or out-dated (Lowe and Oreszczyn, 2008; Skea, 2012; Summerfield and Lowe, 2012).

The shift to a low carbon economy and the need to address energy demand related priorities, such as fuel access and affordability, are drivers of change in the way energy is used and demanded in buildings at the national and international scale. In the UK, such change will involve the retrofit of millions of buildings and changes in energy demand services. Studying energy demand in buildings at a population level is different than in individual or small samples of buildings or users because of the heterogeneity present among populations for which studies need to account. Essential to the development of a strong evidence base is the use of empirically derived data from large populations that can represent the real-world conditions of a complex building stock and population. Evaluating policies and determining the effect of technologies *in situ* in millions of buildings means using techniques that support that level of analysis.

Energy policy has thus far failed to achieve a multitude of objectives, including long-term reduction in energy demand, limited take-up of energy efficiency retrofits for those not in receipt of direct government support, regressive regulation around energy markets and risk of profiteering; increases in households vulnerable to fuel poverty alongside reduced fuel affordability, and regressive changes in values and attitudes towards protective action (e.g. fuel poverty and climate). This general failure of policy against numerous defined objectives can be related to the many interacting stakeholder tensions played out in the political sphere. The failure of energy policies endangers actions for social good, such as alleviating socio-economic energy phenomena (e.g. fuel poverty) and climate change commitments. Further, policy failures may also result in unintended consequences caused by poorly designed and targeted policies and programmes (Davies and Oreszczyn, 2012).

Numerous experts have called for a strong foundation of evidence-based policies and strategies to achieve targets for energy demand, climate change, and other socio-economic goals (Lomas, 2010, 2009; Lowe and Oreszczyn, 2008; Oreszczyn and Lowe, 2010; Skea, 2012; Summerfield and Lowe, 2012; Whitesides and Crabtree, 2007). These same authors outline that the evidence must be made up of the latest best-practice information drawn from relevant research that is properly designed, conducted, interpreted and presented; and drawn from inter-disciplinary activities that address the complex, contextually distinct and politically diverse nature of energy demand. At present, however, much of the research is either too focused on small samples or single cases or is hindered by lack of funding for large empirically-driven survey or monitoring projects (Summerfield and Lowe, 2012; Whitesides and Crabtree, 2007).

The research problem is that this piecemeal approach prevalent in energy demand studies has meant that a methodological framework that captures the complex interactions between people, energy and the built environment has not fully emerged or advanced from the field. This has, therefore, severely limited the development of a foundation of evidence that is able to address the pressing issues related to climate change, socio-economic imperatives and the deployment of technological advances. For these reasons, there remains a need to better understand how and why energy is used in buildings, along with how and why energy demand various across the population.

The study of energy demand in buildings is now at a point where a methodological framework is needed. With the emergence of interval meter data, ubiquitous sensors and large data frameworks of people energy and buildings data, a methodological framework will help to structure existing research practices and help guide interdisciplinary research.

A strong methodological framework would broaden the approaches used to study energy demand to better handle complex socio-economic and technical interactions, to guide, frame and contextualize a growing number of field trials and surveys, and to provide strong methods for examining trends and account for individual variations within a population. This methodological approach should be prepared to draw together those perspectives that play a role in the demand and use of energy within buildings, including engineering, physics, planning, geography, social sciences, economics, and health.

1.3 Aim and scope

This dissertation focuses on developing a methodological framework comprising methods for the study of energy demand among populations while accounting for drivers of variation, methods that can address uncertainty and bias. The framework should also include methods that examine population level trends, evaluate technologies, and guide policy development with evidence that is drawn from well-designed empirical studies.

An approach that offers a compelling framework from which the growing field of energy demand studies could draw is epidemiology. Health epidemiology, which studies the distribution and determinants of health outcomes among a population to address health issues, offers a compelling framework for studying population level energy demand. Although there are various branches to epidemiology, there exists a well-defined methodological structure and foundation that offers tools and study designs, common definitions and standard approaches to analysis. Further, the approach has well-established methods for dealing with (and unravelling) the inter-relationships of socio-technical, environmental and geographical factors.

The aim of this thesis is to determine whether the adaption of the conceptual and methodological framework of epidemiology can support the study of energy, people and buildings. This thesis tests this hypothesis by examining relevant epidemiological concepts and methodological framework along with three studies that adapt and apply epidemiological methods to energy demand and energy efficiency retrofits UK houses.

1.4 Significance of the study

This thesis, therefore, is concerned with the development, application and results from an epidemiological approach to energy demand. It focuses on the adaption of the epidemiological method to the study of energy demand. The thesis aims to define a relevant methodological framework for an empirically-based population-level study of energy demand, with a focus on residential buildings. The application of the epidemiological approach to the study of energy demand in buildings is illustrated using the UK housing stock. Whilst elements of epidemiological techniques have been applied to the study of energy demand in buildings, such as the case-control trials in Twin Rivers by Socolow (1977) and the Pennyland studies by Chapman (1985), they have not been rigorously or explicitly been applied.

This thesis provides a series of methodological studies of energy demand in the UK housing stock that illustrate the adaption and application of epidemiological concepts and methods. The method studies include in-depth discussions on the suitability of epidemiological methods as applied to energy, as well as relevant findings from the studies themselves. In addition to a discussion of the methods, new results on energy demand in UK dwellings are presented in this thesis.

In terms of the wider applicability of this research, it is anticipated that this framework will be capable of offering a stronger methodological foundation for developing evidence of energy demand in buildings that is based on empirical data, includes variation and distributions, can determine factors of causation, and can be used to target and predict changes due to interventions. Furthermore, an evidence base that relies on consistency and transparency will mean that energy policies can be developed in a more informed manner and tackling complex problems can be more achievable.

1.5 Overview of the study

This research comprises eight main parts that address the research problem and the research aim:

> Chapter 2 provides background on research in energy demand. It covers key drivers of interest, dominant research motivations and research paradigms, and the existing approach to population-level studies in energy demand and its problems. It will also review several population-level research approaches.
>
> Chapter 3 introduces the epidemiological approach in medicine, offering a brief history of the field, the key concepts and their basis. This chapter is meant to provide the reader with a sufficient understanding of epidemiology so as to better judge its adaption to energy demand research.
>
> Chapter 4 develops the case for an epidemiological approach to empirically-based population-level energy demand research, focusing on residential buildings. The chapter goes through the process of examining the key concepts introduced in Chapter 3 and assessing whether they are appropriate for the study of energy demand, whether the epidemiological conceptual basis is maintained when applying them, and whether they advance the study of energy demand.
>
> Chapter 5 outlines the epidemiology methodological framework in terms of its key principles, methods and analysis techniques. It then outlines the method study approach that will be used to investigate the application of epidemiological methods to practical problems in energy demand research.
>
> Chapters 6, 7 and 8 present applications of selected epidemiological approaches to practical research problems in energy demand research in residential buildings in the UK. The method studies use selected study designs and highlight key epidemiological concepts. The results from each method study provide interesting

findings for UK residential energy use research. Each method study is critically discussed in terms of its application and suitability to the research problem.

Chapter 9 provides an overall discussion of the research in terms of its ability to address the research problem and the research aim, the adaption of the epidemiological approach to energy demand research, and the findings from the critical discussion of the method studies. It also proposes ways forward from the research findings.

Chapter 10 provides a discussion of a selection of findings from the method studies as they pertain to energy demand in UK dwellings.

Chapter 11 provides concluding thoughts on the research carried out in this thesis.

Portions of this thesis have been published in a number of peer-reviewed journals over the past four years. The co-authors of the journal articles provided suggestions and comments on content; however, the main text and results were drawn from this thesis and is the work of the author.

The articles are:

Hamilton, I. G., Summerfield, A. J., Lowe, R., Ruyssevelt, P., Elwell, C. A., & Oreszczyn, T. (2013). Energy epidemiology: a new approach to end-use energy demand research. *Building Research & Information*, 41(4), 482–497. doi:10.1080/09613218.2013.798142

Hamilton, I. G., Shipworth, D., Summerfield, A. J., Steadman, P., Oreszczyn, T., & Lowe, R. (2014). Uptake of energy efficiency interventions in English dwellings. *Building Research & Information*, 42(3), 255–275. doi:10.1080/09613218.2014.867643

Hamilton, I. G., Steadman, P. J., Bruhns, H., Summerfield, A. J., & Lowe, R. (2013). Energy efficiency in the British housing stock: Energy demand and the Homes Energy Efficiency Database. *Energy Policy*, 60, 462–480. doi:10.1016/j.enpol.2013.04.004

Chapter 2 Literature Review

The study of energy demand: "Current practice in energy demand research"

Chapter Introduction

Policies focused on energy demand in buildings are developed in a complex environment of crosscutting multi-objective and interacting issues of climate change, prices and affordability, energy supply, market regulation, and health and wellbeing. To date, however, energy policy has not adequately recognised or been able to respond to this complexity, which has meant that policies have failed to deliver or adequately address many of these complex, socio-technical challenges in a timely manner. Energy policy is focused at the population scale, but current research is largely carried out at the individual unit level (e.g. building, person, household) and small-scale, driven by single discipline perspectives. Beyond policy, the building industry and technology manufacturers create products that are focused at populations. These industries rely on population data to understand their market whilst also carrying out technology field trials to determine product potential. However, the limited availability of detailed empirical data on energy demand in buildings makes it difficult to understand the market potential and impact of widely installed technologies. This has meant that deeper insights into problems around energy demand, their presence and persistence across the population, are severely limited, which in turn undermines effective policy, product development and deployment.

The aim of Chapter 2 is to describe the research context of the study of energy demand in buildings. In developing an approach to studying the socio-technical interactions of people, buildings and energy demand at a population level, it is essential to understand the interests driving current research and their paradigms. This examination will provide some insight into current research approaches, their strengths and weaknesses, and therefore what form a new approach might take.

This chapter begins (Part A) by briefly introducing several key areas of the study of energy demand and existing research paradigms. The second section (Part B) of the chapter focuses on the approach used to studying energy demand in dwellings at a population level and provides a brief review of current study methods and designs. The final section (Part C) sets out strengths and weaknesses of the current approach to studying end-use energy demand in dwellings at a population level. The section concludes by making the case that a more clear research methodology is needed in order to address the general and specific aims and objectives of studying population-level end-use energy demand.

Section A

2.1 Introduction to the study of energy demand in buildings

Over the last two decades the role of energy demand associated with buildings has gained prominence in the international effort to contain GHG emissions and limit climate change to a 2 °C rise in global average temperature (Jollands et al., 2010). Given its national importance and contribution to total GHG emissions, energy demand reduction in buildings has been identified as needing to make a substantial contribution to GHG mitigation across the globe. For example, nations in the EU-27 have committed to reducing emissions from existing buildings by 80-95% by 2050 (European Commission, 2011a), with all new stock to comprise of 'nearly zero-energy' buildings by 2020 (European Commission, 2010). In the US, recent legislation has mandated increased energy efficiency measures in buildings (Dixon et al., 2010), but targets for lowering energy demand are being led by individual states (Lutsey and Sperling, 2008). California aims to reduce state level emissions 30% by 2020 from 1990 levels and has proposed to reduce energy consumption in existing homes by 40% during this period (CPUC, 2008), saving 19.5 $MTCO_2e$ (California Air Resources Board, 2008).

There has been considerable investment in energy efficient technologies and a host of policies adopted in developed countries related to energy efficiency in buildings since the 1970's (Geller et al., 2006; Noailly, 2012). This attention to energy use at the economy-wide level has resulted in per capita demand holding in these countries whilst expanding energy-using services (Laitner, 2009) and increasing total demand (OECD/IEA, 2012); while many developing countries have seen gains in per capita energy demand and increases in total demand (OECD/IEA, 2012). Energy use in buildings in developing nations has increased most quickly in middle and high-income urban areas and in industrial demand (IEA, 2010). Energy demand in buildings, whether for economic or quality of life-related reasons or for meeting abatement targets, remains high on the international agenda (Jollands et al., 2010).

2.1.1 End uses and energy demand

Energy demand is a by-product of an end use or some desired service. Energy used inside a home consists of a form of fuel (e.g. electricity, gas, oil, and solid fuels) used for a particular service (e.g. space heating, hot water, cooking, appliances, lighting, and cooling). The amount of fuel used will be a product of the type of service, minus the losses through the system used, the environment and context conditions that the service is demanded within, and the frequency and duration of the demand made for that service, and the choices of the user or system designer. As a wider system, end-use energy demand is characterised by the interaction of physical, technical and social components, which act together in a chaotic manner to have a varying effect on energy demand. The physical components are those features of the system that are physical processes (e.g. thermal flow through a wall). The technical components may be characterized as the engineered systems (e.g. the space heating system or its sub-systems). Social practices relate to the way that an occupant will make use of the systems. These practices are themselves driven by values, culture, institutions, preferences, understanding and beliefs. This might be the desire to have a 'comfortable' temperature while at home in the evenings before regular sleep times. The interaction of the physical processes, engineered systems, and practices result in an energy demand for a given end-use. An end-use may involve numerous components of the system (e.g. comfort could include the thermal, auditory and lit environments) and will be driven by the occupant's practices (or where a system is automated, the pre-specified practice). Characterising the physical and engineered systems is limited by the ability to collect information on their features and functioning condition. Characterising social practices is subject to an even greater level of uncertainty and these are likely to vary considerably between occupants and for an occupant over time.

2.2 Study areas of interest

This section describes the areas of energy policy that have driven energy demand research over the past 40 years. There are broadly six main areas that have been the subject of end-use energy demand research since the mid-1970's. These are: energy market regulation, energy security, fuel access and poverty, energy efficiency and climate change. These areas have resulted in a great deal of research into energy demand and have tended to be associated with certain disciplines, e.g. economics, engineering and physics, and social sciences. The following section provides a brief discussion (largely focused on the built environment) of these study areas as background to a following discussion on the current approach to studying energy demand.

2.2.1 Energy market regulation

Following the privatization of the UK retail energy market in the mid 1990's, the energy market was set up to freely allow consumers to choose suppliers (Helm, 2005; Price, 2005). The objective of the open market was to encourage competition in supply and to reduce prices for consumers (Price, 2005). Research in this area has largely focused in the following areas: market competition (e.g. prices, customer portfolio and services), fuel security, efficiency of supply and demand, and access to supply (Helm, 2005, 2003; Perrels et al., 2006; Price, 2005).

In terms of energy research, the impact of energy market liberalization has been clear. The investment in research and development between 1985 to 1995 fell in the UK by more than 80% (Dooley, 1998). This shift in investment strategy in research has further declined following that period, with figures from the UK Energy Research Council (UKERC) and the Engineering and Physical Sciences Research Council (EPSRC) showing a considerable decline from that point onward, see Figure 1 (Skea et al., 2013). This steep decline in public research investment followed the deregulation of the energy markets in the early 1990s. Beginning in the early 2000s, the Research Council UK (RCUK) investment in end-use energy demand has gradually increased. In 2014, the EPSRC spending on energy efficiency (including end-use energy demand) is £72.4 million (EPSRC, 2014). This increase in funding coincided with numerous national and international climate abatement policy activities.

However, there is still very little energy demand research being carried out in the private sector. There have been similar trends of under-investment in the US (R. M Margolis and Kammen, 1999; Nemet and Kammen, 2007).

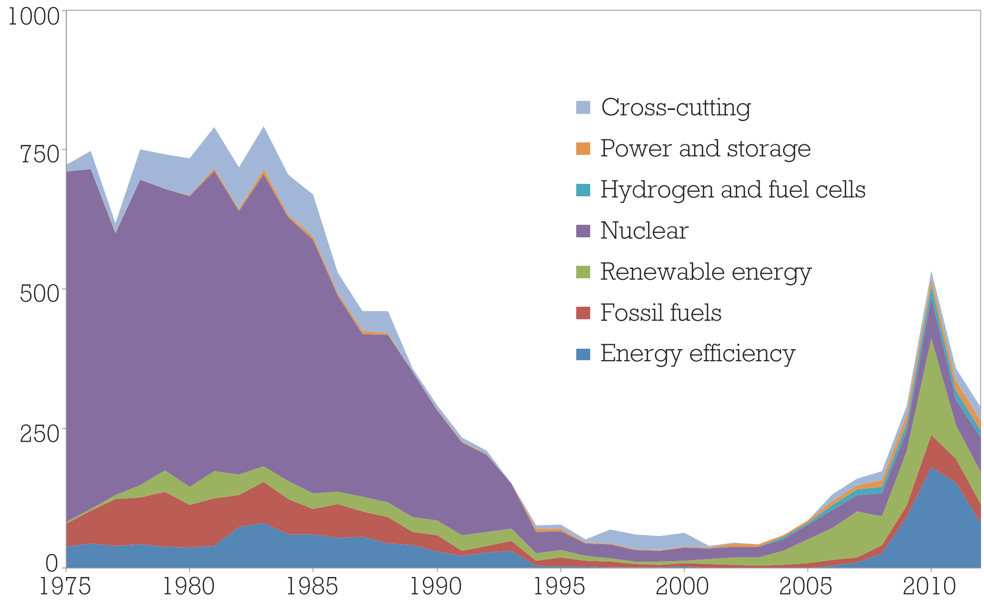


Figure 1 - UK annual public sector research, development and demonstration budgets (£m, 2012 money) (Skea et al., 2013)

Deregulation and the decline in investment has also marked a shift towards short-term research and development projects, with the average project term falling from 5-7 years to less than 3 years, and where long-term innovative technologies are less likely to be funded under competitive deregulated markets (Dooley, 1998). Nemet and Kammen (2007) identify that government funding in R&D can help support early-stage technology and send signals to markets in terms of commitment to wider potential expansion of technology (Nemet and Kammen, 2007). They go on to emphasise that the drop in private-sector investment in research limits an 'innovation' based strategy to achieving energy demand objectives. Without longer-term investment in research, there a much reduced capacity for innovation and also the ability to shape and develop policies that are able to address current and future pressing problems (Whitesides and Crabtree, 2007).

2.2.2 Energy security

Energy security is primarily seen through a lens of ensuring reliability and affordability (Hughes, 2009). The concern over energy security was a primary driver of research in efficiency of energy demand from the early 1970's onward, although the explicit prominence of security has since varied (Bielecki, 2002). The drive for energy security following the OPEC crisis led to numerous energy efficiency policies, a number of which subsequently attached research-focused evaluations to these policies (Schipper, 1987). In the UK, the crisis instigated several initiatives that focused on improving energy efficiency with the prospect of preserving reserves through reducing demand. The initiatives included grants for energy surveys and (notably) a number of demonstration schemes that sought to illustrate techniques and technologies to improve efficiency (Mallaburn and Eyre, 2013). An example of such a scheme was the Pennyland project, which was an estate built in 1976 to examine the potential for passive solar design of dwellings in the UK (Chapman et al., 1985).

Presently, the UK government is focusing on energy security as part of its on-going commitment to economic growth, market liberalization and reducing risk from severe weather, terrorism, technical failure and labour actions (HM Government, 2013). In terms of directing research, a recent House of Commons Energy and Climate Change report called for the government to undertake more research in terms of understanding public attitudes and behaviours, modelling scenarios, along with the impact of demand-side response (House of Commons, 2011). The UK ERC has suggested that energy security indicators should cover resilience of the energy supply, energy infrastructure and energy users. Indicators for this latter group could include: the level of energy demand, energy intensity, energy costs and availability and type of reserve and back up systems.

2.2.3 Energy access and affordability

The term 'fuel poverty' was introduced in the early 1980's as a means of describing households who spend an inordinate proportion of their income on energy use (Hutton, 1984). Boardman (1991) defined these households as those who spent more than 10% of their income on all energy needs (Boardman, 1991). As a concept, fuel poverty has had considerable traction in research terms and has been the subject of a number of UK government policies. The first Fuel Poverty Strategy was produced in 2001 and aimed to end fuel poverty in vulnerable households by 2010 (DTI and DEFRA, 2001). A number of subsequent strategies were published along with statistical outputs that mapped the prevalence of fuel poverty throughout the UK. Such targets have been revised numerous times due to the difficult nature of defining and tackling fuel poverty (DECC, 2013a; Moore, 2012). The *Hill* report, which reviewed the concept of fuel poverty, suggested that a preferable measure would focus on households whose income is below the poverty line and which have higher than typical energy costs, known as the 'low-income, high costs' definition (Hills, 2012).

In terms of research, a number of studies have looked at the occurrence of fuel poverty throughout the UK (Boardman, 2004; Milne and Boardman, 2000; Walker et al., 2013) and other countries (Healy and Clinch, 2002; Howden-Chapman et al., 2009). The concept has also been used in health research that has looked at the impact of living in cold homes in terms of excess winter deaths (Healy, 2003; Walker et al., 2006; Wilkinson et al., 2001) and mental health (Green and Gilbertson, 2008; Harrington et al., 2005; Shortt et al., 2007).

2.2.4 Energy efficiency

As an issue, efficiency of energy demand has been closely associated with both energy security and affordability (Saidel and Alves, 2003) , and more recently climate change

(Mallaburn and Eyre, 2013). While deregulation of energy markets has seen a major drop in energy research, energy efficiency has remained a key area of funding for many governments (Saidel and Alves, 2003). Energy efficiency has had a considerable impact on overall energy demand in advanced economies since the 1970's. Geller et al (2006) estimate that energy demand could have been up to 49% higher in OECD countries without the introduction of energy efficiency. They state that these savings are attributable to the policies and research and development investment in efficiency of buildings, appliances, vehicles and industrial operations (Geller et al., 2006). For example, since the 1990's the US government has expanded its research initiatives into the improvement of energy efficiency in buildings, industry, electric power and transportation (Dixon et al., 2010).

In the UK, research into improving energy efficiency was brought alongside market deregulation through the introduction of various energy suppliers obligations in the early 1990's that aimed at improving energy efficiency for domestic and non-domestic customers (Ofgem and Energy Saving Trust, 2003). The Energy Efficiency Standards of Performance (EESOP) scheme (1994-2002) sought to improve energy performance of supplier customers and had focused on assisting 'disadvantaged' customers. The target energy savings for EESOP approximately 18.5 TWh (lifetime savings) (Rosenow, 2012), and also sought to achieve a number of social goals, determine supplier capability and provide environmental benefits. The follow-on Energy Efficiency Commitment (EEC) (2002-2008) had a similar rationale in terms of achieving multiple benefits, but greatly increased the target energy savings to approximately 192 TWh (lifetime savings) (Rosenow, 2012). Subsequent programmes continued to focus on both carbon emission reductions in line with climate change mitigation goals, such as under the Carbon Emission Reduction Target, and energy efficiency, as under the Community Energy Saving Programme, Green Deal and Energy Company Obligation. In terms of directing research, Rosenow (2012) identified that the research investment (i.e. R&D and energy monitoring) associated with these programmes has dropped substantially over the period, with: 3% (£750k$_{1994}$) of scheme budget under EeSOP 1, 0.42% (£100k$_{1998}$) under EeSOP 2, 0.75% (£400k$_{2000}$) under EeSOP 3 and then no allowance for research and development. More recent programmes (i.e. CESP, Green Deal and ECO) also had some allowance for research in terms of undertaking process and impact evaluation, but no explicit amount of funding is stated (DECC, 2012a).

More recently, the UK Government has made a call for more evidence to support energy policy and have made a commitment to work more closely with research councils (DECC, 2014a). Investment in 2013 by the UK research councils of approximately £39 million over five years (or £7.8 million per year) for end-use energy demand research helps to improve the commitment to funding (EPSRC, 2013), but this is still a long way from the earlier levels of investment seen in the 1970's.

2.2.5 Climate change

In the UK, the threat of climate change has been the major driver of energy demand research in the last 10 years, primarily as the result of government actions that have directed both funding and policy development to this area (Mallaburn and Eyre, 2013). Prior to this period, climate change provided a platform for a number of policies that aimed to reduce the impact of UK GHG emissions on the global climate. Following the signing of the UN Framework Convention on Climate Change in 1992, the first major piece of UK climate legislation was the 1997 commitment to the Kyoto protocol. This commitment held that the UK would reduce its CO_2 equivalent emissions by 12.5% of 1990 levels by 2008-2012 (Pearce, 2006). In 2003, the UK government made a further commitment to reduce emissions by 60% of 1990 levels by 2050. These earlier commitments were superseded by the Climate Change Act of 2008 that set out an at least 80% reduction in UK CO_2 equivalent emissions from 1990 levels by 2050 (OPSI, 2008). In terms of directing research, the Climate Change Act provided the basis for an independent committee on climate change that would have the powers to "gather information and carry out research". Beginning in 2008, the UK Committee on Climate Change (CCC) published recommendations for the first three Carbon Budgets that set out the pathways to achieving the reductions through sector-specific mitigation measures covering the period up to 2022 (UK CCC, 2008). A subsequent fourth budget recommendation was released in 2010 that covered the period to 2027 (UK CCC, 2010). For buildings, the target reduction in carbon was to be near to zero carbon by 2050, which is an ambitious task for a sector that has consistently underperformed in achieving energy efficiency and effectively technology deployment (Oreszczyn and Lowe, 2010).

The CCC compiled a technical evidence base, and for energy demand and housing this was primarily derived from available surveys and the use of modelling. The residential sector used a stock level implementation of BREDEM (BREHOMES) that used to underpin much of UK government energy efficiency policy (Shorrock and Dunster, 1997). However, this model has been criticized for not being transparent or clear in its underlying assumptions, making it difficult to determine the accuracy of any predictions made (Firth et al., 2010; Palmer et al., 2013).

In 2011, the Carbon Strategy set out the various objectives that would need to be achieved to meet the UK's GHG reduction commitment. For buildings, the strategy stated that the government was on track to meet the first three recommended carbon budgets but that a further 25 to 75 $MtCO_2e$ needed to be reduced during the fourth carbon budget period – a considerable reduction. The strategy maintained the need to "complete the cost effective 'easy wins'" during the current decade (HM Government, 2011, p. 33). While the strategy emphasized the need for more research in many sectors, the term 'research' was strangely absent from the section dedicated to buildings (HM Government, 2011). However, in the subsequent 2012 *Energy Efficiency Strategy*, the government acknowledged that research would take a central role in helping to achieve the emissions reductions related to energy efficiency, stating:

> [The Government is] working to strengthen the evidence base through: commissioning research into the potential of advanced heating controls; working with the IEA to explore all benefits of energy efficiency; setting out a future DECC Evidence Strategy; and coordinating with Research Councils UK and others, to support the development of a knowledge hub for the refurbishment of existing

homes as well new Energy Demand Research Centres, announced with this strategy. (DECC, 2012, p. 6)

This document provided the clearest directive calling for research related to energy demand and energy efficiency in existing buildings (DECC, 2012b).

2.3 Existing research paradigms for studying energy demand

Like the drivers of energy demand research, there are a number of prevailing research approaches and activities that are used in the examination of energy demand in dwellings. These conceptualisations of energy demand research follow the predominant theories of the active disciplines within the field. Known as research paradigms, these activities drive the manner by which energy demand is both viewed and researched within those disciplines. A paradigm (as conceptualised by Kuhn (1962)) is described as being "a coherent pattern of research organized around commonly shared theoretical propositions and models" (Helm, 2005). A paradigm is not so much developed but instead identified through the actions of the research community, the identified gaps that need covering, and the wider scientific zeitgeist within which research is carried out. Energy demand paradigms provide the lens through which to examine and interpret pressing issues and act as a gravitational pull around which research collects. A number of energy demand research paradigms exist that are most easily classed in themes of economics, social sciences, engineering and environment and sustainability. Although classed in this manner, there is overlap in the paradigms between disciplines and difference that exist between those paradigms within disciplines. As with research area interests, paradigms are important for understanding current research approaches and their limitations.

2.3.1 Economic

Economic paradigms have focused on the economic efficiency of the energy demand (and supply) system. Energy efficiency has different meanings depending on the discipline. Economic efficiency of energy focuses on rational or optimal use of resources to maximise the benefits for society (Gillingham et al., 2009). The efficiencies of the liberalised market through privatization attempted to release the creative economic activities of the open market (Reddy, 2002). In the UK, this approach dominated much of the early policies set out around energy supply and had an impact on energy demand through a market-based approach to energy prices and eventually to deregulation of energy suppliers (Helm, 2005; Mallaburn and Eyre, 2013). This paradigm has also directed the approach to investment in the energy supply and demand systems. For the most part, investments in both these systems have been the victims of the open market approach advocated under this paradigm (Robert M Margolis and Kammen, 1999). In the UK, US and elsewhere, energy companies purchased publicly built infrastructure and have subsequently exploited these for profit but lacked the re-investment and upgrades necessary to maintain efficiency. Energy efficiency, in these terms, is seen as a rational and beneficial investment to reduce costs of the energy used for demanded services (Laitner, 2009). For energy demand, the investment in buildings efficiency has been the subject of public regulation (i.e. building regulations) and programmes directed through the energy suppliers (Mallaburn and Eyre, 2013). However, outside of those programmes this paradigm views energy demand as a rational choice, whereby the consumer would increase or decrease their demand in accordance with their perceived benefits or dis-benefits. It also assumes that consumers will invest in energy efficiency when it becomes economically beneficial to do so (Gillingham et al., 2006; Hassett and Metcalf, 1995, 1993).

The economic paradigm also assumes that through market liberalisation consumers will benefit from more choice in terms of services and prices and will make decisions to achieve efficiency in their supply by 'switching' suppliers. The reality is that while many consumers do switch suppliers, many are either unable or uninformed and as a result are faced with higher prices (compared to those who do switch) (Price, 2005).

The 'rebound' paradigm has also been a major area of research in energy demand. The 'rebound' effect has been defined as the increase in energy use services that corresponds to a decrease in price (or cost of demand) (Greening et al., 2000). This approach crosses economic and technical viewpoints by bringing together the technological efficiency of changing features of the energy demand system (e.g. energy efficiency retrofits) and rational expectations (Sorrell et al., 2009). The concept has focused on explaining why changes to structural features of the energy demand system have not resulted in predicted savings. A number of technical, social and environmental contributors are identified as being important factors in the occurrence and degree of rebound (Greening et al., 2000; van den Bergh, 2010). This approach has been criticised on the basis that there is insufficient understanding of the system that could predict changes in energy demand (Herring and Roy, 2007).

A further paradigm being applied to energy demand and economic efficiency is through an econometric lense. Econometrics applies mathematical and statistical methods to explore empirical economic relationships in order to develop and put expand theory. Econometrics as a tool is being applied to understand a range of energy demand problems, including energy efficiency, energy investment and other (Alberini et al., 2013; Meier and Rehdanz, 2010). The paradigm focuses on using empirical methods through the application of econometrics to examine economic relationships of energy demand.

2.3.2 Technological

Energy demand research from an engineering and physics perspective has grown greatly over the past two decades (Schweber and Leiringer, 2012). The technical paradigm has placed buildings and their accompanying technologies and physical processes at the centre of describing energy demand – often from a disaggregated (i.e. building bottom-up) level (Baker and Rylatt, 2008). This approach is generally technologically focused, even when including occupants as part of the system (Palmer et al., 2012). At its core is the thermodynamic evaluation of the energy system with demands being met by energy systems converting primary energy to delivered energy and useful energy. Energy services are delivered and limited by the thermodynamic efficiencies of conversion. These approaches are often used to examine specific technical problems and rely on models in the absence of empirical data (Kelly, 2011; Palmer and Cooper, 2013). The approach attempts to incorporate users through further parameterisation of uncertainty (Hughes et al., 2013a), but often lacks detail on the range of user practices that would capture these unknown traits. The failure of this approach in being able to identify and understand the key drivers of energy demand has led to various conceptual ‘fixes’ that attempt to mitigate the expectations of technical understanding (Sunikka-Blank and Galvin, 2012).

The transition to a more energy-efficient building stock has been cited as an active paradigm where technologies are applied to improve the efficiency of energy demand for services (Boden, 1996). This paradigm views the delivery of technologically more efficient products as the key means to improve energy performance in buildings, thereby reducing demand for those services. The primary goal is to achieve a much higher degree of performance through the integration of advanced architectural and engineering building design, commissioning, and operation. The research in this field, which includes intelligent building design, has focused on integration of building services through communication and adaptive technologies (e.g. adapting to changes in weather or occupants) (Wong et al., 2005).

There have also been paradigms that attempt to merge the technical and the social elements through various approaches. Under a technical-social paradigm, more focus is placed on understanding the uncertainty around the user as part of the technological and physical system, for example by widening the sensitivity of the user inputs (Booth and Choudhary, 2013). This paradigm has attempted to understand the impact that users have on influencing energy demand and energy savings. This area has come to the fore in an attempt to explain why technical models that attempt to account for user interactions have been unable to predict ‘actual’ demand. The practice of using ‘normative’ models has been criticised as being too crude and inflexible (Cayre et al., 2011). Recently, the socio-technical perspective has been argued as a more constructive and inclusive approach to study energy use in buildings that acknowledges the mutually co-dependant nature of social actors and technologies and the enmeshing of social activities and technological artefacts (Chiu et al., 2014).

2.3.3 Social research

The economic and technological approaches have been criticised for treating the occupant as either an autonomous rational being or as a functioning component in a system rather than attempting to reflect his or her social practices as fundamental drivers of energy demand (Wilhite et al., 2000a). Under this paradigm, energy demand practices include the behaviours and norms, personal beliefs and values, and understanding of social institutions. Where engineered systems historically viewed users as passive and with little formal feedback, the social model sees users as actively and unintentionally interacting with the energy system to ‘demand’ services. This paradigm sees energy consumption as being derived from the services demanded by users and reflecting their values, motivations and social influences. Whereas the technical approach seeks to explain energy demand through understanding the technical components in a building or dwelling, under an occupant practices paradigm, occupant behaviour or social practices are the main focus (Yohanis, 2012). Factors such as education levels, household type (e.g. age and dependents), income levels and ownership have all been identified as significant contributors to energy demand (Mills and Schleich, 2012; Pelenur and Cruickshank, 2012). It has been noted, however, that energy use practices are subject to complex social and technical and personal interactions that often lead to difficulties in interpreting individual actions, making conclusions difficult to generalise or quantify (Attari et al., 2010; Dietz, 2010). This realm of research has looked at these individual factors in order to better understand how and why energy is used and therefore how better to promote efficiency and changes in behaviour (Steg, 2008). There is a growing field of research in this 'non-technical' area of energy demand, which has primarily focused on occupants and has somewhat neglected institutions and policy (Schweber and Leiringer, 2012).

2.3.4 Environmental and Sustainability issues

The energy demand paradigm seen through environmental and sustainable development lenses is concerned with holistic energy demand in the context of other pressing global issues (Reddy, 2002). This paradigm seeks to place energy demand within a wider context and to focus on service demand, for which energy is an enabler, and away from total consumption. The concerns of this paradigm are: the disconnect between economic drivers and energy consumption (versus demand), the recognition of changes to lifestyles, universality of access to affordable energy for basic services, expanding decentralised energy systems, modernizing rural energy system to improve quality of life, and humanising energy demand (Reddy, 2002). The arguments around sustainability and energy demand focus on needing to make changes to the way energy is supplied and used, for the purposes of tackling broader issues, in which its use is seen in an ethical context, rather than a technical one (Herring, 2006).

The low-carbon society is one of the key paradigms (possibly the dominant paradigm) to have emerged in the last two decades. The 'low-carbon' commitment paradigm has focused on examining energy demand and its contribution to anthropogenic GHG emissions (Helm, 2005; Lomas, 2010). In this paradigm, the need to avoid catastrophic climate change provides the basis for a number of research activities that seek to reduce the rate and total amount of greenhouse emissions; the case for such actions is supported by evidence provided by the International Panel on Climate Change (IPCC) (IPCC Secretariat et al., 2007a). This paradigm has driven energy demand research in the last decade aimed at seeking the social good of mitigating the negative impacts of climate change (Skea, 2012). In the UK, this paradigm has been embraced primarily through economic terms (Stern, 2007), which helped make the case for policymakers to set out strategies to curb emissions. In the UK residential sector, government policy has been focusing on reducing the GHG contribution of new dwellings (i.e. through building regulations with the 'zero carbon dwellings'), improving standards for rental properties (i.e. enforcing incremental minimum energy rating standards for lettings), and through improving the energy performance of existing buildings (i.e. through Green Deal and ECO) (DECC, 2012b).

2.3.5 The dominant paradigm of energy demand

The low-carbon society paradigm has become the dominant paradigm within which research in the energy demand field is taking place. The argument around a low-carbon society is that the benefits are multiplicative beyond more efficient use of energy and decarbonised energy, but extend to encompass social connections, health and wellbeing, and economic growth (Haines et al., 2009; Seyfang, 2010; Skea and Nishioka, 2008). This persuasive argument has permeated into and altered most other existing paradigms. Culminating in the Stern review (Stern, 2007), the economic rationale is changing to incorporate a triple-bottom line approach (economy, environment and society). Meanwhile the socio-behavioural paradigms have largely become concerned with understanding the drivers around making behaviour changes to take action towards mitigating climate change (Dietz et al., 2009). It has been argued that a shift into a more holistic framing of energy demand would lead to a move away from a model of balancing the economic, social and environmental demands to one that is focused more on resilience and adaptation, which holds benefits for initiating change amongst stakeholders (du Plessis and Cole, 2011).

The implication of the diverse nature of research on energy demand issues coming from a number of disciplines is that the strengths of the different disciplinary approaches did not emerges over time at the same rate. The dearth of funding in energy demand forced a lot of research activities away from empirical data collection and analyses towards theoretical research, this latter approach being facilitated by the availability of cheap computing power (Oreszczyn and Lowe, 2010).

Section B

2.4 Current approach to studying energy demand in dwellings and households

Under the dominant paradigm of the low-carbon society, the call for action to transform the built environment and respond to the threat of climate change, its scale, scope and urgency, is clear and the reasoning behind such action sound (European Commission, 2011b; IPCC Secretariat et al., 2007b; Stern, 2007). To address this call, development of plans and strategies that are able to direct effort and resources to achieve these changes in the most effective manner, while building support for both their investment and expansion across multiple sectors of the built environment. Yet to date policymakers have not been presented with conclusive evidence for how some policies have delivered intended outcomes, and importantly the reasons why others have not (Skea, 2012). Although buildings may represent one of the single largest sectors for potential CO_2 emission reductions (UNEP, 2007), the ability to achieve these reductions is limited by a knowledge gap that means it is not possible to fully identify and describe the drivers of energy demand in the built environment upon which strategies can be built, and the success of policies judged (Oreszczyn and Lowe, 2010).

Though ambitious plans and policies are being developed to tackle the climate change and energy challenges, research on energy demand in buildings has been criticised for being largely incapable of describing even the most basic conditions around energy demand (Oreszczyn and Lowe, 2010). In examining the estimate for mitigation potential in buildings from the IPCC, Oreszcyzn & Lowe say:

> [The estimate] prompts a number of concerns. The most important of these is the poor quality of the data available to support the [estimate]. This concern in turn stems from weak links between the policy research community responsible for producing the [estimate], the building science research community responsible for producing the underlying data and analysis, and the communities of practice responsible for translating the speculations presented above into reality.

They go on to discuss the implication that this paucity of data means for addressing energy demand in dwellings, seen as 'low-hanging fruit' by policymakers, but which is vastly more complex and difficult than imagined. They challenge policy orthodoxy that states that interventions in the building stock aimed at reducing energy demand are 'quick wins' and cost-effective.

The simplistic approach to the understanding of energy demand and the built environment is a key risk for why policies may not deliver expected carbon savings. Despite many programmes and policies targeting the UK's energy use in the housing stock, changes in real energy demand have been limited and generally poorly explained (Summerfield and Lowe, 2012). Summerfield and Lowe point out how the lack of good quality information on energy use, buildings and technology and households and their practices severely limit researchers' and governments' ability to address the decarbonisation challenge across a heterogeneous and complex building stock and its occupants. They go on to set out the challenges for achieving the transition to a low carbon society which include: the scale of emission reductions needed, the rate of change needed in emissions and transformation to the energy system, the scope of the sectors and actors interacting in the built environment, and finally the trans-disciplinary approach needed. One of the key features of their approach to addressing these challenges is the role of empirical evidence and high quality data.

If the targets to avoid catastrophic climate change are to be met, changes to energy demand practices and the energy performance of buildings need to happen faster, to many more buildings, and drawing in many more households. Because energy demand in dwellings is highly complex and subject to many interacting factors, a 'one-size fits all' approach would be completely inappropriate. The approach must be capable of dealing with the key challenge of complexity and must be evidence-based. The empirical evidence must be able to describe the basic conditions and features of buildings and occupants and should be sufficiently comprehensive to represent the population to which policies and actions are targeted.

Given the scale of change and the heterogeneous nature of energy use in buildings and the systemic lack of evidence, an empirical population-level approach to studying energy demand is essential. The following section examines the existing approaches to studying empirically-based population-level energy demand research in residential buildings.

2.4.1 Existing approach to population-level energy demand research

Over the past decade there has been a growing amount of research focused on energy demand in buildings that has been empirically based. Schweber and Leiringer (2012) found in a review of academic journal databases that over the past decade that there has been an upswing in the number of articles with the terms 'energy' and 'building' in the abstract or titles. Articles focusing explicitly on research objectives defined as 'energy demand' comprise around 20% of all articles (Schweber and Leiringer, 2012). They claim that of those articles, 48% marked as 'construction research' were empirical-statistical in nature and nearly 80% were approached from a positivist stance, compared to an interpretivist position. This highlights that energy demand is studied among disciplines using a range of approaches, from theoretical or experimental, and techniques such as quantitative and qualitative. This thesis focuses on population-level approaches from an empirical perspective using quantitative analysis techniques.

Empirical population-level data is directly observed from a population of individuals. In energy and buildings research, the individuals that comprise the population can be buildings, people, or some other units, but can be collected together into a coherent group. There have been only a handful of studies that have analysed energy demand at a population level, and they have used a wide range of study designs and analysis approaches. There is a wider number of studies that have analysed factors related to energy demand, many of which have used population-level approaches. Among these studies a number of different designs are used, including: individual case studies, ecological (i.e. aggregate and area-based studies), several case-control studies, a number of cross-sectional and longitudinal studies, and various intervention studies.

Case series approaches typically take a cluster of observed individuals within a defined place or time, that are very often selected on availability or circumstances. In their study of heat loss through party walls, Lowe and Wingfield (2007) used a sample of 700 newly constructed dwellings near Manchester. The population comprised surveyed and monitored dwellings studied over a period of two years. The study found that mid-terraced dwellings were losing 60% more heat than would have been expected using notional calculations (Lowe et al., 2007). Another example of a case series study is earlier work by Bell and Lowe (2000) who used three sets of dwelling groups with varying levels of monitoring and surveyed information around the city of York to provide a case study of the potential for CO_2 emission reductions through modern building practices (Bell and Lowe, 2000). Baker and Rylatt (2008) used a survey of 42 respondents from Leicester and Sheffield on their energy use and features of their home and household characteristics to examine the presence of clusters of high and low users (Baker and Rylatt, 2008). In a study of perceptions of energy demand and savings related to household activities, Attari et al. (2010) used a nationally drawn survey of 500 participants in the US to show that there was a considerable gap between actual and predicted demand/savings (Attari et al., 2010). A major drawback of case series studies is that they do not provide a representative picture of the issue being studied and as a result are at risk of being over-interpreted. Case series are typically used to investigate particular issues and to generate hypotheses. They may also be used to define or establish particular conditions (e.g. as in the party-wall heat loss effect). Case series studies comprise the largest group of study types in energy demand, reflecting the small-scale funding approach that tends not to support large-scale studies.

Several studies have been published using aggregated population-level data on energy demand, known as ecological studies. Aroonruengsawat et al (2012) looked at the impact of state-level building codes in the US on per capita electricity use in dwellings using a database of annual total electricity collected for 48 states between 1960 and 2006, along with state level details on building codes, per capita income, and weather data (Aroonruengsawat et al., 2012). Druckman and Jackson (2008) used census output area statistics combined with cross-sectional data to estimate energy demand at the 100-dwelling level (Druckman and Jackson, 2008). The ecological study is ideal for making use of aggregated data such as census or national registries (i.e. home sales) and to generate hypotheses that can be tested in more detailed studies.

Following case series, cross-sectional studies are another common approach for studying energy demand. This is due to the greater availability of national cross-sectional surveys of buildings. In the UK, the English Housing Survey (EHS) (formerly the English House Condition Survey (EHCS)) has been undertaken in some form since 1968. It has only collected information on energy demand periodically through bespoke follow-up surveys, in

1991, 1996, 2001 and recently in 2011. Brechling and Smith (1994) used the 1986 EHCS to examine the take-up of energy efficiency measures. Tovar (2012) also used successive EHS's from 1996 to 2008 as a longitudinal dataset to examine how investment in energy efficiency levels changed over time. Kelly (2011) used the 1996 EHCS along with the Food and Expenditure Survey (FES) to examine whether more efficient homes used less energy. In the U.S., the Residential Energy Consumption Survey (RECS) has been used to examine differences in energy demand between sub-groups. Poyer et al. (1993; 1997) examined energy demand and minority or ethnicity status (Poyer and Williams, 1993; Poyer et al., 1997). More recent studies using the RECS have looked at identifying the influence that household characteristics have on quantiles of energy demand, showing that dwelling and household factors (e.g. income levels, ownership, rurality and dwelling age) explain varying amounts of energy use in different quantiles (Kaza, 2010). Similar approaches to studying energy demand using national cross-sectional data have been undertaken in Ireland, Greece, Holland, India (Guerra Santin et al., 2009; Leahy and Lyons, 2010; Nair et al., 2010; Pachauri, 2004). National cross-sectional surveys have the advantage of being carried out so as to be representative of the dwelling stock and households therein, and are more likely to have undergone more rigorous tests for sampling error and bias. The disadvantage of national surveys is that they are less flexible regarding the variables being surveyed, and in the case of the EHS completely omit energy demand (or fuel use) in most years. Other more bespoke surveys have been carried out to look at energy demand using a cross-sectional approach. Shipworth et al (2010; 2011) undertook a survey (drawn from a stratified random sample of England's postcodes) of home heating use practices in 427 dwelling, focusing on thermostat settings and heating patterns (Shipworth, 2011; Shipworth et al., 2010), along with subsequent studies of indoor temperatures (Huebner et al., 2013; Kelly et al., 2013). The benefit of a bespoke survey is the ability to collect specific details of interest; however, the disadvantage is that these surveys tend to be smaller and more open to various forms of bias and measurement error.

Several studies have used longitudinal data to examine trends in energy demand over time within the same houses or households. Meier and Rehdanz (2010) used a longitudinal panel survey to identify household and physical dwelling factors that explained expenditure on space heating in Britain (Meier and Rehdanz, 2010). An early study in the US, known as the Twin Rivers study, examined several samples of houses in different details over the course of approximately 6 years from 1972 (Socolow, 1978; Sonderegger, 1978). This particular study was a landmark in identifying the significant effect of occupant behaviour on energy demand in dwellings that were effectively the same. Summerfield and Lowe point out that the advantage of high-quality longitudinal data is the ability to examine how different features present and persist within buildings and how they relate to energy demand and its change over time. Longitudinal and cohort studies are generally difficult to implement due to costs and loss of participants in follow up.

Compared to cross-sectional and case studies, there have been fewer intervention studies on energy demand in dwellings, possibly due to their expensive nature and the limited funding in this area. In England, the Warm Front studies used details of a public energy efficiency programme aimed at reducing energy expenditure for low-income households and examined the relationships between dwelling energy efficiency and household characteristics, and indoor temperature, mould growth and thermal comfort (Hong et al., 2006; Oreszczyn et al., 2006a, 2006b). The research used data collected under the Warm Front scheme for approximately 2000 households that received a number of different interventions. These Warm Front studies employed several methods including a before and after approach and a limited control group. In the late 1970's and early 1980's a number of intervention programmes were carried out in the US, which found a variety of energy savings for different intervention types (Goldman, 1985; Wall et al., 1983). A recent study applied a pragmatic randomised control trial (RCT) to study the impact of energy efficiency retrofits on indoor temperature. The study provided a randomly allocated intervention to approximately 129 households in fuel poverty with another 108 houses used as a control group (Heyman et al., 2011). Intervention trials offer the most reliable method of testing and determining the effect of a particular control or programme on energy demand within a well-defined population.

Population-level studies of energy demand have primarily been undertaken using small samples that have largely been selected ad-hoc and potentially suffer from a range of biases, hence they may not be truly representative of the target population being examined. This risks over-interpretation of the results and can result in misguided policies or poorly informed models. There are a number of studies that have used national cross-sectional surveys and these are more able to provide a representative focus on the issue being examined and are ideal for describing the presence of factors of interest. However, these surveys may be unsuitable for tackling specific or timely issues and may not necessarily provide the needed data. Longitudinal studies are more able to examine trends in end-uses and energy demand and a wide number of factors that affect changes, but they are expensive to run and are few and far between. Intervention studies are also used to examine particular issues in-depth but, like case series, run the risk of being based on small samples and, without larger precursor studies, it is difficult to contextualise their findings. This leads to a number of problems with the current approach, which are discussed in the following section.

2.5 Problems with the current approach

In their paper, Lowe and Oreszczyn (2010) lament the lack of a framework for studying energy demand and the resulting limits on the understanding of trends and determinant factors of energy demand. From the point of view of carrying out research, this lack of framework has meant there is little guidance on how and what data should be collected and what tools should be used to analyse relationships and provide feedback to policymakers, industry and researchers. A number of authors have highlighted the problems with the current approach to studying energy demand in buildings, described in further detail below, namely that the built environment is too complex for simple plans, there is a lack of integration between disciplines and study approaches, there is a lack of high-quality empirical data and evidence, and the evaluation of practices and programmes is limited (Firth et al., 2010; Oreszczyn and Lowe, 2010; Skea, 2012; Summerfield and Lowe, 2012).

2.5.1 Too complex for simple plans

Although the UK Government has made a commitment to an extensive decarbonisation of the existing housing stock through improved efficiency of new and existing houses and through an almost zero-carbon grid, there remains little research or evidence to suggest how this programme will be carried out and what effect this real-world experiment may have on a host of household indicators (Kelly, 2009; Skea, 2012). Further, the lack of quantitative and descriptive foundation means that contextualising intervention studies and field trials that aim to support this decarbonisation programme, along with understanding short and long-term trends in energy demand, is severely limited. It has been suggested that there could be a range of unintended consequences that may present themselves through the decarbonisation of the stock, for example, those related to health impacts, or behavioural or economic effects (Davies and Oreszczyn, 2012).

Unravelling the relationships between factors that influence intervention outcomes and trends related to energy demand in buildings is extremely difficult and often subject to highly speculative assumptions and caveats. Empirical collection of data on building characteristics and energy demand has historically been ad-hoc or subject to interruptions and there has been little tradition of reporting data in a formal sense, thus undermining any concerted advances in the research (Summerfield and Lowe, 2012). When data is collected on an ad-hoc basis it very often lacks key features that allow for cross-comparison or linking to a broader foundation and therefore risks misunderstanding and limited evaluation. The overall effect of this lack of data collection has meant that models that attempt to describe energy demand in buildings have been seriously limited and often rely on unconfirmed theory rather than empirical observations. For example, the party-wall effect described by Lowe et al, in which heat was being lost through the gap in the party wall and directly conducted to the un-insulated roof, was the effect of regulatory and construction practices that policy, regulations and models simply did not account for (Lowe et al., 2007). Further, the use of conservatories as an extension of the living space, especially those with double-glazing, brought about significant increases in heat-related energy demand (Oreszczyn, 1993); this phenomena was similarly missed until identified through survey work. Worryingly, these phenomena have not been explored within the wider UK housing stock and have effectively remained discovered but understudied.

The complex process surrounding the policy, regulation, design, development, construction and operation of dwellings means that these types of unintended effects have been too easy to miss. They also suggest that there is the possibility of much wider and equally troubling issues related to energy demand waiting to be found. This also means that without the combined ability of in-depth study along with wider (and consistent) survey work, trends in energy demand may arise without there being a conceptual basis with which to examine the problem.

2.5.2 Lack of integration in approach

As an approach, the engineering/physics paradigms tend to tackle mechanisms and systems. They take as their base the laws of thermodynamics supported by laboratory and physical measurements. The challenge occurs when applying these to the complex socio-technical system of people, buildings and services. The limitation of this approach is the assumptions and simplifications in describing the system. In addition, the assumption that changes to the technical system has little or no impact on the wider system other than the direct feature being changed. This approach has trouble accounting for variation due to the unexpected or unobserved. The engineering approach typically includes socio-technical factors through parameterisation and standardisation. Otherwise the influence of the occupant happens outside the system defined by the model domain. Although based on experimentation and physical relationships these technological paradigms are severely limited when used to consider effects of occupant practices on energy demand and cannot support the modelling of broad and complex policies.

Sociological and economic approaches are concerned with understanding behavioural mechanisms and norms, and value decisions that are associated with a condition or level of energy demand, but are limited in the description of the physical processes surrounding energy demand, although causal models may be established to explain these. Social scientists who study energy demand rightly point out that addressing only the physical problems and focusing on the technical potential offers little insight into practices, behavioural factors and attitudes that influence energy demand (Guy and Shove, 2000). Recent research has pointed out that economic sociologists have considered the built environment in terms of behaviour and psychology, energy technology and housing, sociotechnical systems and technology-shaping processes, but that the built environment and energy use are understudied (Biggart and Lutzenhiser, 2007) and that the nature of energy demand in particular is little studied (Wilhite et al., 2000a). Without an appreciation of occupant practices it is likely that insight into patterns and trends in energy demand will be shallow (Guy and Shove, 2000; Parnell, 2005; Shove, 2003).

It is vital to move beyond narrow disciplinary approaches to study energy demand phenomena across a population. This shift must be driven by a need to understand the mechanisms that cause changes in patterns of energy demand and also the consideration of the differences in these changes among groups and sub-populations. Ultimately, this requires empirical data from cross-disciplinary studies analysed within a common research framework to disentangle the dynamic and interrelated effects of environmental, socio-cultural, lifestyle, and economic factors that influence occupant practices and energy demand (Attari et al., 2010; Dietz, 2010; Sorrell, 2007; Wilhite et al., 2000a).

Perceptions and practices of disciplines can be barriers to interdisciplinary research (Lutzenhiser and Shove, 1999). Therefore, a strong (independent) methodological framework, along with definitions and detailed and consistent studies, can help to encourage a sense of collective understanding and foster an environment of interaction.

2.5.3 Lack of data and evidence

In the UK, most of the statistics on trends in domestic stock level energy demand over time are largely estimates based on engineering-physics model outputs of stock models based on BREDEM. That the national level trends in energy demand are derived from models rather than observed data inherently weakens the value of the figures and makes it very difficult to analyse behavioural influences and technological changes from other observed datasets.

There are several statistically representative stratified surveys of actual fuel and energy demands that can be used to observe trends in energy demand in UK housing (i.e. English House Survey (EHS), and the British Household Panel Survey (BHPS)). However, there are limitations that make it very difficult to use these datasets for analysing trends and focussing on particular segments of the house and household stock, such as frequency of collection and measured variables. This makes observing trends in energy demand in the UK housing stock possible but limited in time and drivers.

The more recently developed national energy efficiency data-framework (NEED) provides a new source for analysis. NEED is a collection of government administrative data (i.e. Valuation Office Agency (VOA)), energy efficiency intervention data (drawn from the Home Energy Efficiency Database (HEED)) and energy supplier (gas and electricity) data, and has been selected by DECC as a representative sample of the housing stock of Great Britain. However, although this dataset offers a significant advance in data pertaining to energy demand and physical characteristics, there is no available information on households, thus making it very difficult to establish what effect occupants have had on changes in demand over time.

The absence or limited access to high-quality people and buildings data and high-resolution energy data of the statistical and methodological quality that other disciplines would consider a prerequisite for the pursuit of good science and robust conclusions is a major challenge. Along with this, a limited research capacity currently exists to organize or archive data, despite significant sums of money invested to collect data through individual projects. It will be essential that research data are captured along with detailed meta-data allowing for use by other researchers and held in a suitably accessible repository for future analysis and connection. Without this detailed and comprehensive data collection there is little basis for systematic reviews of research findings to support project-by-project learning.

2.5.4 Limited evaluation of past practices

Energy demand research in the built environment remains characterised by piecemeal studies and fragmented discipline-specific methods and perspectives that can limit, rather than expand, the broader relevance of findings (Lowe and Oreszczyn, 2008; Whitesides and Crabtree, 2007).

To date no intervention to reduce energy demand in buildings has been subjected to comprehensive empirical evaluation in the UK or any other major European or developed country of the type or scale being identified in many carbon abatement plans. As a result the data, tools, systems and models to support the design, implementation and evaluation of such interventions are absent or un-calibrated. This therefore leaves a significant gap in understanding of what the direct and indirect effect of a policy or technology has in practice as compared to assumptions drawn from models.

The theoretical understanding of social and technical factors that influence energy demand remains underdeveloped (Schweber and Leiringer, 2012). This is due to the inevitable lack of information that can be used to test hypotheses in real world situations and under operating conditions. The prevailing approach struggles to identify associations and establish underlying causes behind outcomes and variations in energy demand seen within a population. As an example, although there has been some decline in average UK household delivered energy in recent years (DECC, 2012c), it remains unclear whether this is attributable to improvements in building fabric and energy systems, occupant behaviour in response to increases in energy prices, the global financial crisis or other factors.

2.5.5 The need for a new approach

The sheer scale of interventions being proposed to reduce building-related energy demand requires an approach that is capable of dealing with population-level observations and interventions, while supporting and learning from other disciplines and strands of work, including field and case studies. The evidence base must be capable of supporting the development and application of well-targeted abatement measures that identify critical areas for investment and transformation while also being able to evaluate historic efficiency programmes and activities. Evaluating past policies and programmes used to address specific energy demand or related issues is essential to developing effective plans that are able to meet the set objectives. There is real risk that without high quality and representative data, free of (or with limited) bias, policies cannot be informed by past practices. The lack of data and evaluation also makes it harder to identify and understand the causes of unintended consequences or features of policies that did or did not work, and why.

The energy research community is lagging behind the evolving policy agenda, and in a number of cases, practice, to reduce energy demand from the built environment. The existing approach to studying technologies, socio-cultural practices, and the deployment of technologies and other interventions in the field, limits both the generalizability of findings, and the range of challenges to which existing models and theories can be subjected. It leaves policymakers and other stakeholders without insight, and limits the ability to achieve effective change. Many of the problems or limitations faced reflect the underlying disconnect between the different disciplines involved, from engineering and building physics to social sciences, economics and health. A host of factors, such as the low prioritization for funding and limited empirical data (Gupta and Gregg, 2012; Kelly, 2009; Schweber and Leiringer, 2012), have led to an overreliance on models that are often poorly informed or outdated (Laurent et al., 2013; Lowe et al., 2007). Overall, this has meant that a methodological framework that captures the complex interactions between people, energy and the built environment is only just beginning to emerge from the field.

Different research approaches are needed to help tackle the complex issues surrounding energy demand in buildings that can focus on providing insight into how and why different levels of energy demand occur among different groups of buildings, occupants and environments. In turn this will help to provide an understanding of the impact interventions aimed at reducing demand or increasing efficiency will have, what is an adequate or profligate level of energy demand, and what are the causes and critical factors related to system performance and behavioural interactions.

Section C

2.6 An integrated population level approach to studying energy demand

As governments move towards implementing large scale energy-related intervention programmes, a far more comprehensive evidence base is needed to support the development, implementation, and ongoing evaluation of energy demand policy (Clery, 2007; Oreszczyn and Lowe, 2010; Skea, 2012; UK CCC, 2010; Whitesides and Crabtree, 2007; Wilkinson et al., 2007). Delivering such a transformation in the way energy is used in buildings will require a deeper level of understanding of the underlying relationships between energy use and socio-cultural practices, engineered systems, physical processes and environment so that effective technologies, practices and behavioural changes can be adopted and supported through evolving policies (Dietz, 2010; UNEP, 2011, 2007). Appraising evidence and providing feedback to policymakers on the evolution of applied policies, while addressing the complex, contextual environment, will help to identify the key determinants of successful interventions and policy mechanisms. Appraisal will also provide insight into which unsuccessful policies were the result of poor delivery and/or flawed measures and controls (Lowe and Oreszczyn, 2008; Rychetnik et al., 2002; Skea, 2012; Sorrell, 2006).

2.6.1 Population studies in other disciplines

There are several disciplines that focus on the study of populations. Demographers and geographers study people and places, in order to understand changes within society and their impact on places (Rowland, 2003). These disciplines uses information on populations and their activities to examine trends and occurrences and features within a larger context. Demographic methods are used in geography, sociology, economics, marketing and planning and political sciences, among others. Demographic research is also informed by findings from other disciplines, e.g. health, economics, and psychological and behavioural studies, to help understand trends in society. Demographers use a number of population research designs, such as cross-sectional, ecological or cohort studies. The focus on the study of population and context offers an interesting parallel. However, for the purposes of examining energy demand in residential buildings, demographic approaches are unlikely to provide a strong methodological framework within which to draw together detailed studies on mechanisms with the wider population context. The reason is that while demographics examine individuals and their activities, e.g. birth rates, the focus is very much at the population level. This is not to say that those detailed mechanisms that are drivers of population activities are not integrated into population sciences, but rather that the mechanisms themselves are not primary objects of study. Though there are inevitably strong parallels between demographics and any population study approaches, the demographic approach may not satisfy the need to balance the study of detailed mechanisms with population-level effects.

In the economic realm, actuarial sciences study risk among the population and use details of behaviours and activities from samples and groups to determine potential risks of outcomes. The findings of these smaller groups are used to inform models that allocate these behaviours and risks of outcomes to a broader population. Actuarial science's main tools are mathematics, economics and statistics and it seeks to project and predict requirements for coverage or provision of services. For the purposes of energy demand and building research, actuarial sciences have more in common with the current model-driven approach. Further, there is very little scientific literature on actuarial research methods or conceptual approaches and therefore this approach is unlikely to offer a meaningful analogy.

A further area of study that holds considerable relevance to studying empirical observations of populations is econometrics, which is the application of empirical methods of analysis to economic relationships in order to develop, test and expand economic theory (Geweke et al., 2008). Econometrics use statistical methods applied to observational data to explore patterns and project estimations through forecasting. The focus on analysis of cross-sectional and panel data is a particular strength of econometric methods. The approach offers many potential advantages for studying energy demand among populations, which has been illustrated by the numerous energy demand studies in economics described above. However, whilst econometrics offers the marriage of statistical and observation, it may not necessarily offer a strong methodological framework for drawing in detailed studies of underlying mechanism or laboratory testing. There may also be a wider conceptual framework limitation around approaches to study design and meta-analysis studies, though these may emerge from the field.

The many research approaches used in the health sciences examines a range of systems including the detailed mechanisms of human systems (e.g. physiology, immunology), though processes, emotions and decision-making (e.g. psychology, psychiatry) and population health (e.g. epidemiology, primary care, health services). Health science research has a long experience of dealing with complex problems and bringing together evidence from a host of disciplines with a focus on improving both individual and population-level health. The inter-disciplinary transactions around any health outcome range from clinical observations, laboratory testing and measurements, to group-level trials and surveys to overall population health at a global level. The approach used in health science research, in particular epidemiology, offers a compelling framework for undertaking research in end-use energy demand. Epidemiology, which studies the distribution and determinants of health in a population, is well suited to complexity and for dealing with physical, socio-economic, environmental and geographical factors. Although there are various branches to epidemiology, there exists a well-defined methodological structure and foundation that offer tools and study designs, common definitions and standard approaches to analysis, which are the focus of this thesis.

Epidemiology, from its Greek roots, is literally defined as 'the study of what is among the people'. Described as "the study of how often diseases occur in different groups of people and why?" (Rose and Barker, 1978), epidemiology is primarily concerned with the description of health-related events or conditions within a population, and with applying findings to the control of health problems. Epidemiology makes use of a set of tools in the form of common study designs and analysis, definitions and metrics, and the collection and comparison of data that are considered under a wide range of theories and hypotheses in studying disease aetiology. These tools also support a longitudinal approach to trends, using data collected over many years in a standardised manner (where possible) to inform and develop hypotheses. There is also a strong methodological tradition for large-scale studies and the relationship between the clinical level and population level, with a great deal of effort employed to ensure that findings can be relevant outside of the individual case. Furthermore, epidemiology has made significant strides in an evidence-based approach to problem-solving and decision-making (Jenicek, 1997). In particular, epidemiology's strong methodological foundation and study designs support the gathering and assessment of evidence needed for such an approach. Overall, this suggests that the epidemiological approach is adept at addressing issues of scale, definition, and differing viewpoints.

2.6.2 A transformative approach

The study of energy demand in buildings is now at a point where a methodological framework that broadens the approach to include complex socio-economic and technical interactions is needed; one that is able to guide, frame and contextualize a growing number of field trials and surveys, examine trends, and identify the impact of behaviours and interventions on energy demand. It is proposed here that epidemiology offers a compelling framework from which the growing field of energy demand studies.

The epidemiological approach offers both a set of tools and a methodological framework within which to undertake analysis in search of causal factors associated with a condition and to frame results and findings. Unlike energy demand, there are a number of bodies and organizations that undertake epidemiological health studies, with many focusing on a select set of issues with the primary aim of designing, conducting, interpreting, and presenting relevant and timely research.

Using an epidemiological approach to tackle the complex issues surrounding energy demand in buildings could provide the means for research to focus on providing insight into how and why different levels of energy demand occur among different groups of buildings, occupants and environments. In turn this could, for example, help to provide a deeper understanding of what effect interventions aimed at reducing demand or increasing efficiency will have, what might be an adequate or profligate level of demand, and what could be critical factors related to system performance and occupant interactions. This methodological approach should include the various perspectives that play a role in the demand for and use of energy within buildings, and in doing so support and draw from the engineering and physical sciences, sociology and economics, studies of health and wellbeing, environmental sciences and geo-spatial studies.

2.7 A case for an epidemiological approach to energy demand

The buildings and energy demand research field could build on lessons learnt in epidemiology with respect to research concepts, methods and study designs, used to examine both detailed mechanisms and their manifestation at a population level. For example, the availability of national health records has led to the development of health epidemiology as a disruptive discipline for understanding human health. There is the potential that an epidemiological approach to energy demand can strengthen the energy and building research approach and improve the evidence base in comparable fashion. Through methods founded on a strong conceptual framework, consistent and detailed definitions, well-designed studies that are subject to evaluation and critical analysis, along with the availability of individual and sub-meter high frequency data, deeper insights into energy demand can be established and better policies developed.

Although the initial collection of data has begun, the conceptual and methodological framework required to use data effectively, develop causal associations that are coherent and defensible, and make wider comparisons is still largely lacking. The study of energy demand needs to broaden its approach to include a wider range of issues, a concept of causality, and the interactions between the 'user' and the 'system', which need to be supported within a methodological framework of analysis. Epidemiology provides a strong research approach and policy setting emphasis, and offers a host of concepts, methods and analysis tools within an integrated framework.

The allocation of economic resources and effort in managing energy demand in the built environment means that decisions regarding implementing a policy or changing a practice must be sufficiently well informed so as to deliver desired results and minimize risk of unintended outcomes (Davies and Oreszczyn, 2012). Doing so means gathering evidence of the potential benefits, harms and costs, along with their magnitude and accuracy, so as to compare possible outcomes. The scale of changes required to decarbonize the building stock underlines the importance of gathering evidence through a methodological framework that allows for common definitions and a set of study designs applicable to both individuals and population levels, along with collaboration between disciplines. Adapting the epidemiological approach to end-use energy demand studies may provide the means to describe the trends and patterns of demand and begin to establish causal factors that lead to outcome events. It may also provide the means to undertake and contextualise intervention studies. The benefits of such an approach would be to strengthen the empirical foundation from which evidence is drawn to inform policy decisions and evaluate past intervention programmes or regulatory actions while also acknowledging the complex environment within which the studies occur.

The development and application of an epidemiological type approach could be used to study and describe the mechanisms of energy demand and determinants of conditions that lead to levels of demand. A greater understanding of the mechanisms and determinants of energy demand will ultimately extend domain knowledge. Through the recognition of conditions (or 'disease') it would be logical to move towards 'treatments' and management of problems. Having sufficiently large populations for study it will become possible to identify relevant populations (e.g. people, households, buildings, etc), to assess and determine causes of conditions and the effect of interventions, so as to establish 'cures' or effective treatments. Mechanisms can be tested in a much richer context, where social scientists can engage with the engineers and physicists and beyond to the fields of planning, economics and health through an integrated framework.

The breakthroughs required for the step-change increase in efficiency and reduction in energy demand will be facilitated by the unprecedented availability of and access to new energy and buildings data. For example, through the installation of high frequency metering and sensors, a huge amount of information can be accessed to describe patterns of demand, manage peak loads, and allow consumers to interact with the supply system and be charged an accurate and fair price. These changes in technology make an epidemiological approach more feasible; data collection is becoming cheaper and more accessible than ever before. In the near future, minute-to-minute data from high-frequency meters could become more widely available. This data will need to be subject to high levels of protection for privacy; however, with the development of suitable controls and under the aegis of the government, access to anonymised data could be extended across the research community to create an unprecedented, open environment for empirical testing of theory, policy and technology. The buildings and energy demand field must build on the lessons learnt around data access and protection in the health research field. Just as linking patient records to the use of health services has led to the development of epidemiology as an indispensable part of public health policy, the availability and use of individual and sub-meter high frequency data and collection of building and occupant data through coherent research designs can support an epidemiological approach, essential for the development of policy for evidence-based energy demand.

This thesis therefore is concerned with the development and application of an epidemiological approach to energy demand. It will focus on the adaption of epidemiological concepts and methods to the study of energy demand in the built environment, as applied to the UK housing stock.

2.8 Summary

The study of energy demand has largely been undertaken in five main areas of interest: market reform, security, access and affordability, efficiency and climate change. These interest areas have been approached under a range of research paradigms driven by the different themes, i.e. economic, technological, sociological, and environment and sustainability. The dominant paradigm through which energy demand in buildings is currently being researched is the sustainability paradigm, namely the 'low-carbon society', which, at its core, focuses on bringing about a major shift in how energy is used in buildings. The current approach to studying energy demand in buildings has been domain- and discipline-specific, drawing on those disciplines' own technical and conceptual understanding of characteristics of energy demand in buildings. Economic and social approaches have tended to focus on occupant practices and decision-making, but have neglected the technical nature of building systems and their interaction with occupants. On the other hand, the technological approach has focused too much on the engineered system and physical processes, leaving occupant practices as a component of the 'energy demand system'. The tensions of this discipline-focused approach to study energy demand has left the field with a number of problems, namely: that the interactions between environment, buildings and systems and occupant practices and social institutions that interact around energy demand are complex but have been addressed with simplistic models; there is little integration between the different disciplines researching energy demand and limiting potential insights; a lack of high-quality empirical data on people, buildings and energy demand makes it difficult to contextualise detailed studies and examine trends among the population; and there is limited evaluation of past practices leading to poorly informed policies. To address these issues, an integrated approach is needed that draws together the technical and sociological dimensions of energy demand within the guidance of a strong methodological framework.

Epidemiology is adept at drawing together detailed mechanisms and complex social issues, and contextualising and studying their occurrence among a population for the development of policies. Epidemiology provides a strong conceptual and methodological framework within which to develop and test hypothesis using empirical population level information. It is proposed here that an epidemiology approach to energy demand can assist in strengthening the evidence base needed to bring about a step change in understanding of the drivers behind energy demand to support the transition to a low-carbon society.

In the following chapter, a brief summary of the field of epidemiology and its key concepts are introduced and discussed for the purpose of informing the process of adapting the approach to the study of people, buildings and energy demand.

Chapter 3 Epidemiology

The epidemiological approach: “What epidemiology can offer to the study of population-level energy demand”

Chapter Introduction

> “Unlike medicine which, despite its many specializations has always held together and has always maintained the tradition of linking practice, research and teaching, architecture and the other professions in the construction, design and urban planning [field], with few exceptions, have not developed a tradition of practice-based, user research [.]” Frank Duffy, *BR&I 36(6)*, 2008, p.657

Research of end-use energy demand at a population level requires a framework that is able to capture the heterogeneous and complex nature of populations. Further, such a framework should be capable of offering a structure that allows for interaction between the multiple disciplines that impinge on energy demand. The empirically-based population-level research approach typified by the epidemiology offers a compelling framework within which to examine problems of energy demand.

The aim of this chapter is to describe key concepts of health epidemiology and the research framework that surrounds population-level health study, i.e. the study of health outcomes within a population, the antecedents of their expression and differences.

The chapter begins with an introduction to the epidemiological research paradigm along with a brief description of its evolution into modern practice. Then a number of concepts of epidemiology are introduced that cover issues such as defining outcomes and populations, variation, and risk. The epidemiological approach of causality and causal models is described along with attributes and fundamental features that support the approach. Following this, a brief description of the epidemiological analysis framework is outlined, with more details for selected study types provided in Chapter 5. Finally, the chapter concludes with a critique of the epidemiological approach.

3.1 Introduction to epidemiology

Epidemiology, derived from its Greek etymology, is literally defined as "the study of what is upon the people". However, modern health epidemiology is currently primarily concerned with

> "The study of the occurrence and distribution of health-related states or events in specified populations, including the study of the determinants influencing such states, and the application of this knowledge to control health problems" (Porta et al., 2008).

In its initial stages, epidemiology mainly addressed pressing diseases affecting public health, such as cholera and influenza, which were especially prevalent among urban populations. The focus was on the prevention of disease along with a focus on socioeconomic and structural factors that were linked to differences in health (Pearce, 1996). In more recent decades much of modern epidemiological research has focussed on non-communicable diseases and issues related to public health, such as cardiovascular disease and smoking (Horton, 2012). Epidemiology provides robust and evolving scientific methods through which to study health outcomes and disease, providing a foundation from which to inform health care practitioners, health policymakers, health researchers, and the general public (Bailey et al., 2005). Epidemiology, being focused on studying patterns and causes of disease between different groups, is an important part of the overall study of health. It is not able on its own to identify and describe the mechanisms that cause disease; rather it is complementary to other areas of medicine, such as clinical pathology, genetics, microbiology, or immunology (Rose and Barker, 1978). Further, epidemiology has been described as a "biosocial-environmental" science due to its consideration of context and practices, as well as the nature and causes of disease (Bhopal, 2008). The tools used in epidemiology studies have been (and continue to be) developed to focus on addressing population-level health issues that take into account differences in human practices and experienced environments and rely on detailed and specific studies of, say, disease mechanisms.

A main goal of health epidemiology research is to use qualitative and quantitative evidence along with systematic reviews (reviews that assess the results of primary studies against demanding criteria) to inform public health policy development and evaluate past practices (Brownson et al., 2009). For policymakers, using this evidence requires a clear and timely message that can inform political debate (Petticrew et al., 2004), while researchers require that evidence is detailed and adheres to standard study practices allowing for thorough evaluation (Whitehead et al., 2004).

3.1.1 Theory and the central paradigm of epidemiology

Epidemiology is a science primarily based on empirical observation (Bhopal, 2008). This approach is fundamentally based on the philosophy of science known as induction, i.e. deriving knowledge from fact. In epidemiology, this means making statements about health and disease states from observed populations (i.e. sample populations) and applying these to unobserved populations (i.e. target populations) (Karhausen, 1995). There is nevertheless debate about this conceptualisation that follows the more general reflection and evaluation of the philosophical thinking and discourse relating to modern science, not just epidemiology itself (Greenland, 1998; Rothman et al., 2008).

This philosophical foundation of empiricism provides the basis for the current central paradigm (i.e. shared ideas through which scientists communicate and collectively judge knowledge (Kuhn, 1996)) of epidemiology, namely:

> "That patterns of ill health and disease in population may be analysed systematically to understand their causes and to improve health." (Bhopal, 2008, p. 3)

However, epidemiology has experienced an evolution in the paradigms through which knowledge and understanding have developed. Susser and Susser (1996a) outlined three of these paradigms. These are: 1) sanitary statistics and the paradigm of miasma; 2) infectious disease epidemiology and the paradigm of germ theory; and, 3) chronic disease epidemiology and the paradigm of the 'black box' (see Table 2 below). These changes in accepted thinking reflected the scientific understanding, philosophical and political discourse of the times, and the evolution of disease itself (Susser and Susser, 1996a).

Epidemiology has been criticised as being atheoretical and more fundamentally concerned with understanding patterns of health and disease through the development and use of a set of tools for study design and statistical analysis (Bhopal, 2008). However, epidemiology does have a number of basic theories that are used in current practice (see Table 1). These theories help to recognize that not all classes of disease, as experienced in populations, are alike despite the fundamental pathological or genetic similarities. These theories provide a basis for examining populations and health variation and help to explain why many diseases are so difficult to eradicate in practice.

Table 1 - Epidemiological theories and principles (from Bhopal, 2008 p. 16)

Theory:
Disease in population is more than the sum of disease in individuals
Populations differ in their disease experience
Disease experiences within populations differ in subgroups of the population
Principles arising:
Disease variations can be described, and their causes explored, by assessing whether exposure variables are associated with disease patterns
Knowledge about health and disease in human populations can be applied to individuals and vice versa
Health policies and plans and clinical cases can be enriched by understanding of disease patterns in populations

3.1.2 Aims and objectives of epidemiology

The primary aim of epidemiology is to describe the health of a specified population and to investigate causes of key factors and their effects on health outcomes. These outcomes may be disease states (e.g. healthy, sick) or measures of health state outcomes (e.g. underweight, normal, obese, morbidly obese). Using information on the causes of disease, the primary objective of epidemiology is to apply findings to the control of health problems (Bailey et al., 2005). Epidemiological studies seek to consider the complex interactions between the physical environment, socio-economic features, and individual interactions and practices, along with biological mechanisms. Epidemiology is also concerned with the broader context and provides an environment within which individual studies may be contextualised and systematically assessed.

Epidemiology is concerned with the description of a health-related event or condition within a population. The epidemiological approach offers both a set of tools and a methodological framework within which to undertake analysis and frame results and findings. The approach is based broadly on three main functions (Rose and Barker, 1978), which are to:

1. Describe and measure the distribution of a condition or health event;
2. Explain that distribution by its determinant factors (biological, environmental, social, behavioural and economic)
3. Predict the changes expected in that distribution from interventions and control measures.

Epidemiology extends its interest beyond disease mechanisms to include social factors, environmental features, physiological conditions, and more. The approach is informed by studies from genetics to global patterns of disease in order to identify mechanisms for effective intervention and the control of disease or health conditions. The development of results is based on comparison, with a considerable attention to ensuring that causes and their components are able to describe disease prevalence (i.e. proportion of the population with a disease) and incidence (i.e. number of additional cases of disease over time) within a population (Rose and Barker, 1978).

3.2 The evolution of epidemiology to modern practice

The need to address public health came to the forefront of the collective imagination and political will during the industrialisation of most currently advanced western economies (Susser and Susser, 1996a). The close association between population health and the anthropogenic environment has existed for many centuries (Perdiguero, 2001), and the approach to treating serious epidemics of infectious disease, such as cholera, relied on the development of this understanding. Epidemiology has changed from traditional epidemiology, which focused on addressing public health and preventing disease through understanding of causes, to modern epidemiology, which focuses on identifying an individual's risk factors, using statistical techniques and study designs, and exposure measurement (Pearce, 1996). The paradigms through which modern epidemiology has emerged provides a lens through which to examine this evolution (see Table 2). The modern epidemiological approach itself is being pushed in a direction that encompasses the understanding of 'determinants and outcomes at different levels of organisation' (Susser and Susser, 1996b). In the following section, the paradigms of epidemiology are used as a basis for a brief history of the approach.

Table 2 – Eras in the evolution of modern epidemiology (adapted from Susser & Susser, 1996b)

Era	Paradigm	Analytic Approach	Preventative Approach
Sanitary statistics (first half of 19th century)	Miasma: poisoning by foul emanations from soil, air, and water	Demonstrate clustering of morbidity and mortality	Drainage, sewage, sanitation
Infectious disease epidemiology (late 19th century through first half of 20th century)	Germ theory: single agents relate one-to-one to specific diseases	Laboratory isolation and culture from disease sites, experimental transmission, and reproduction of lesions	Interrupt transmission (vaccines, isolation of the affected through quarantine and fever hospitals and ultimately antibiotics)
Chronic disease epidemiology (latter half of 20th century)	Black box: exposure related to outcome without necessity for consideration of intervening factors or pathogenesis	Risk ratio of exposure to outcome at individual level in populations	Control risk factors by modifying lifestyle (diet, exercise, etc.) or agent (guns, food, etc.) or environment (pollution, passive smoking etc.)
Eco-epidemiology (emerging)	Chinese boxes: relations within and between localised structures organised in a hierarchy of levels	Analysis of determinants and outcomes at different levels of organisation: within and across context (using new information systems) and in depth (using new biomedical techniques)	Apply both information and bio-medical technology to find leverage at efficacious levels, from contextual to molecular

3.2.1 A brief history of epidemiology

In its initial stages during the early 19th century, epidemiology was driven forward by the need to solve health issues that threatened the lives of many urban populations, such as cholera, dysentery and the spread of influenza. These issues centred on developing sanitary practices and the theory of their incidence held of its core a belief that ill-health was caused by poor environmental conditions, fundamentally related to the growth of cities (Susser and Susser, 1996a). The theory of *miasma*, popular under the sanitary approach, attempted to link disease to noxious pollutants in the air (and later in soil and water) caused by contaminants, such as 'rotting corpses, the exhalations of other people already infected, sewage, or even rotting vegetation' (Halliday, 2001). In Britain, this theory was advocated by leading scientists and practitioners at the time, including Edwin Chadwick and Florence Nightingale (who 'attributed scarlet fever, measles and smallpox to the practice of building houses with drains beneath them from which odours could escape and infect [others]' (Halliday, 2001)). The theory was used to advocate major infrastructure changes to cities in an attempt to reduce the morbidity and mortality of diseases caused by miasma, including drainage, sewerage, clean water supplies and sanitation practices (Susser and Susser, 1996a). The approach to understanding disease was based on observation and mapping cases in an attempt to identify sources of contaminants.

While the improvement of the urban infrastructure was brought about by the miasma theory, advances in the study of microbiology and the development of germ theory soon eclipsed the idea. Identifying the principal causes of disease as single agents, such as bacterial infection, allowed for interventions to be appropriately directed to those causes of disease (unlike miasma) (Susser and Susser, 1996a), for example through the use of antibiotics and immunization. Early epidemiologists, such as Dr John Snow and Jakob Henle, developed theories of disease spread through pathogens and their vectors, known as the 'germ theory' of disease. In the famous case of the Broad Street water pump, John Snow sought to identify and eliminate the source of a cholera outbreak in London by mapping the illness and deaths due to cholera around the Broad Street area[1] (Cameron and Jones, 1983). He hypothesised that the disease was being transmitted by contaminated drinking water (Snow, 1857), thus developing a model of the relationship of single agent to specific disease. The mythologised events surrounding the outbreak have Snow (or local authorities) removing the handle to the pump and thus ending the outbreak and proving his theory of water-borne transmission (McLeod, 2000). Snow's causal theory of cholera being water-borne was not quickly accepted, for example the 'Great Stink' that prompted Bazalgette's main sewer across London was initially explained on the basis of air being a vector for cholera. However, William Farr, a statistician to the registrar general, brought wide acceptance of the theory through his analysis of London's last cholera outbreak in Whitechapel in 1866, which showed that contamination of the water supply had been the principal vector of transmission (Whitechapel was not yet linked to the sewer) (Halliday, 2001). The germ theory focused on controlling the agents of infectious disease, the identification and treatment of which were advanced in laboratory settings by, for example, notable figures such as Louis Pasteur (Susser and Susser, 1996a). During this period up to the early 20th century, the occurrence of communicative diseases dropped substantially and

[1] Incidentally, some of Snow's work was published in a magazine called *The Builder, 1855: an Illustrated Weekly Magazine for the Architect, Engineer, Archeologist, Constructor, Sanitary–Reformer and Art Lover.* London. The builder is now known as *Building* (building.co.uk).

most principal agents of disease were identified, leading to further awareness of social context in determining disease occurrence.

In the developed world from the mid-20th century onward, as length of life increased, the challenges associated with chronic diseases exceeded those of infectious diseases, pushing diseases such as cancers, heart disease, and strokes to the forefront of epidemiological research (Susser and Susser, 1996a). However, unlike the germ theory, the factors behind these diseases were virtually unknown and as such, laboratory experiments were unable to provide the necessary explanatory factors for their occurrence among the population. The antecedents of these chronic diseases were multitudinous and complex and required a more systematic approach to describe them. Thus began the development of more rigorous study designs and data analysis in order to deal with issues of confounding, bias and misclassification, and risk factors (Susser and Susser, 1996a). Understanding causality in these situations required an approach that took into account complex interactions and also the various factors that when combined would contribute to increased disease risk, if not a direct outcome (Bhopal, 2008). The studies undertaken by Doll and Hill (1950) examining the relationship between smoking and lung cancer are often held up as an exemplar of the epidemiological approach to chronic disease. Their research illustrated how cases of lung cancer needed to be contextualised in order to expand on the theory of causation linking smoking and lung cancer. This was done by selecting controls without lung cancer for comparison to those cases with lung cancer that were similar in age, sex and from the same area (Doll and Hill, 1950). Analysis of a survey of the participants' habits, including whether the cases and controls smoked, provided the evidence to show that smoking was not only related to the occurrence of lung cancer but that the more people smoked the higher the risk. The current form of epidemiology is now closely associated with understanding disease occurrence among a population and using this knowledge for prevention and control (Bhopal, 2008). However, as disease occurrence and priorities change across the globe and new diseases present themselves and technological advances are made, the approach to studying disease through epidemiology is also changing.

3.2.2 Current concerns of modern epidemiology

The main focus of recent epidemiological research, particularly in developed countries, is on non-communicable outcomes related to public health and social conditions (Horton, 2013), highlighting the importance of social factors in affecting health. Susser and Susser (1996b) have described this emerging era of epidemiology as being viewed through a 'Chinese boxes' paradigm, which attempts to take into account the 'causal pathways at the societal level' along with the 'pathogenesis and causality at the molecular level' (Susser and Susser, 1996b). They use an example of HIV infection, which includes (at least) four levels, from: 1) molecular level, 2) individual practices level, 3) population prevalence and characteristics level, and 4) the global level of societal interconnections. With each level holding determinants for the occurrence of HIV, the study of the disease is tackled through the analyses of these determinants and outcomes within a system of levels and degrees of complexity.

In recognising these levels, there have been calls to develop an approach to epidemiology that integrates socio-biological factors within complex systems (Diez Roux, 2007). This approach accepts that social and biological factors are interrelated within a system of many parts and interactions, with nonlinear relationships and feedbacks. For example, a complex approach recognises that gene expression differs depending on the surrounding environment, but also that these environments can affect the 'structure and function of biological systems' (Diez Roux, 2007). This move towards complex systems epidemiology is recognised from within the field: as techniques and technology change so too must the paradigm.

Modern epidemiology still represents an empirically-driven and systems approach that provides a framework to add value, rather than displace, the evidence from other research methods (Bhopal, 2008). Findings from population studies of risk factors through randomised control trials (RCTs) and laboratory studies of genetic markers and diseases are informed by insights gained from relationships identified at the population level from observational studies, which further combine to underpin public policy development and advancement of medical theory and practice.

3.3 Fundamental concepts in epidemiology

As stated above, current epidemiological research aims to identify trends and causes of health outcomes and their differences among a population. In describing these health outcomes, a number of concepts are fundamental to the study and description of the patterns among individuals.

In their concise introduction to epidemiology, Bailey *et al.* (2005) describes the epidemiological approach using an ordering of the five W's, i.e. *what*, *who*, *where*, *when* and *why*. 'What?' uses the definition of an outcome of interest and how cases with the outcome are included or excluded; 'who?' focuses on both the enumeration of those with and without outcomes of interest, and also the difference in attributes and characteristics of those individuals; 'where?' identifies the place that the outcomes and individuals are drawn from; 'when?' describe the trends over time of the outcome; and finally, 'why?' seeks to identify the determinant factors of an outcome by quantifying their association and testing hypotheses on causality using unbiased methods of comparison (Bailey et al., 2005).

The following section provides a basic overview of these epidemiological concepts. The key concepts include: defining disease outcomes, measuring frequency, populations, variation, risk factors and causality, measuring association, bias and confounding, study interpretation, and standards and reviews. These concepts are provided here in order to draw upon these insights in the following Chapter 4 and also throughout the remaining chapters.

3.3.1 Defining disease outcomes

The definition, description and classification of diseases or health outcomes are essential to support epidemiological research into causes and patterns of health outcomes (Bhopal, 2008). Having a clear understanding of what a disease consists of, along with standardised benchmark definitions, ensures that the condition being considered is widely comparable, allowing for pattern detection between studies and the ability to monitor outcomes in different times and locations. The *International Classification of Disease* (ICD) is maintained by the World Health Organization (WHO) and provides a system of classification using established criteria (WHO, 2010); the criteria themselves were first proposed in their basic form by the statistician William Farr in 1856. The ICD is a 'standard diagnostic tool' for disease research and ensures consistent and comparable disease reporting, allowing for analysis, interpretation and comparison of disease outcomes in different regions and populations. These classification systems rely on the definition of disease. Such definitions are produced through understanding causality and the disease mechanisms through which, for example, laboratory identified biological markers, their pathology and risk factors are manifested within an individual. However, in practice, when the causes of disease are complex and multitudinous, or unknown, a more pragmatic approach that takes into account the limitations of identifying and understanding the causes of disease has been advocated (Severinsen, 2001). Severinsen (2001) offers two principles that embody a pragmatic approach to disease definition:

> "The disease entity must be applicable to concrete cases of disease with reasonable reliability, within reasonable economical limits, and with limited discomfort to patients" (Severinsen, 2001, p. 327)

And,

> "Disease entities should be defined so as to secure a reasonably high correlation between, on the one hand, the disease entity and, on the other hand, certain severe consequences and complications of diseases, certain success rates of various treatments or – if the disease is undesired in itself – certain causes that are relevant to its prevention" (Severinsen, 2001, p. 328)

The definition and classification of disease also provides a means of defining the occurrence of disease that in turn allows for the definition of the unit of observation and its distribution among a population (more on this below). Epidemiological studies are able to focus on describing disease that originates from an individual person (or sometimes a group). This common unit of analysis makes describing disease status or contributing factors within a study more straightforward, allowing for comparisons to be made. However, while epidemiology has the distinction of focusing on an individual within a given population, there must also be definitions of the decisions made about what measurement will be the proxy for features that may be a cause or antecedent of a disease, the diagnostic tests used, and the time periods within which to frame the study must also be defined. Bhopal (2008, p. 6) states that the ideal epidemiological variable should:

- "Impact on health status in individual and populations
- Be measured accurately
- Differentiate population in terms of patterns or status

- Differentiate population by relevant characteristics
- Generate testable causal hypotheses, or develop health policy, or deliver intervention programs, or help prevention and control"

Definition and classification also allow for the identification of putative or speculative causal factors around the occurrence of the disease outcomes – a basic approach in the process of establishing strong associations from epidemiological studies that can lead to an understanding of causality.

3.3.2 Measuring frequency of outcomes

Analysis of the patterns and distribution of a sufficiently-defined disease among a population, aiming to lead towards building causal relationships and thus control it, must rely on some form of frequency quantification. Disease occurrence varies between populations and over time. As such epidemiological studies have two basic forms of frequency measurement – incidence and prevalence. The type of frequency being measured will affect the periods considered, the type of study, the approach to analysis, the description of the events and the interpretation of the results.

Prevalence rate is defined as 'the total number of individuals who have an attribute or disease at a particular time divided by the population at risk of having the attribute or disease at a specified point in time' (Porta et al., 2008). The prevalence rate can measure the proportion of disease in those at risk from the disease over three forms of time measurements: point, period and lifetime. The definition provided above is for point prevalence, i.e. the number of cases at a defined time point. Commonly, this is the mid-year point, reflecting the commonly used demographic variable of the mid-year population. Period prevalence is the proportion of the disease in those at risk over a (usually) short defined period. In both cases, the concept is meant to express the proportion of a population. Lifetime prevalence is an extension of the period concept and is a count of the occurrence of a disease (or factor of interest) at any point during the person's lifetime. The point, period and lifetime prevalence can be described as (Bhopal, 2008):

Equation 1

Point prevalence rate = *All cases of disease at time point/ Population at risk at time point*

Equation 2

Period prevalence rate = *All cases (old and new) of disease during time period/ Average population during time period*

Equation 3

Lifetime prevalence rate = *All cases who ever had the disease during time period/ Population at risk at beginning of time period*

Incidence is defined as 'the number of *new* events or cases in a defined population within a specified period of time and may be measured as a count, rate or proportion' (Porta et al., 2008). Incidence can be measured in two ways: risk or cumulative incidence and person-time incidence. *Risk* (or simply *incidence*) is the new occurrences of outcome for a period of time. When expressed as a proportion over the population at risk (i.e. number of disease-free individuals at the start of the time period) over the period of time, it is known as *cumulative incidence*. Risk provides a measure of the probability of an event occurring within a defined population. *Person-time incidence* accounts for the potential issue of persons dropping out of a study and uses the total amount of time that the study population was disease-free. This concept is known as the 'person-time at risk' and the denominator is the total sum of the time at risk (in years): see Figure 2 for example. These terms are described by (Bhopal, 2008):

Equation 4

Cumulative incidence rate = *New occurrences of outcome over a period of time/ Population at risk over the period of time.*

Equation 5

Person-time incidence rate = *New occurrences of outcome over a period of time/ Time spent by the study population at risk over the period of time.*

Often incidence rates are expressed as a rate per standardised population size, for example: occurrence per 1000 persons. Describing incidence in these terms allows for a more practical comparison that is independent of the original sample size. Further, it may be that the cumulative incidence rate is adjusted to take into account death in the population at risk over the period, in which case it becomes very similar to person-time incidence (Bhopal, 2008). This is typically done where a disease under study is common.

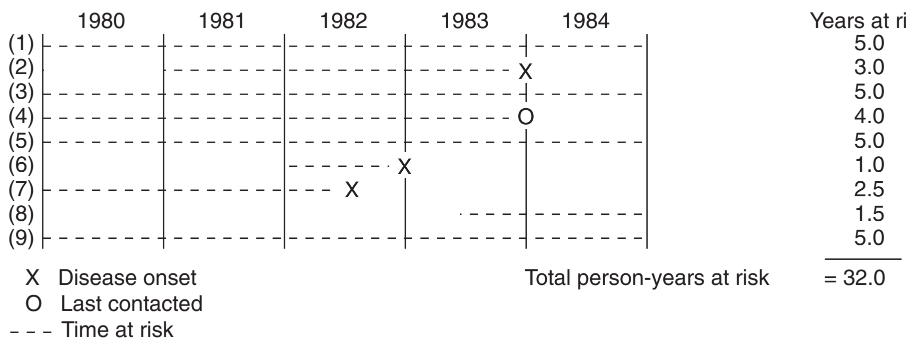


Figure 2 – Example of person time incidence rate (dos Santos Silva, 1999)

Incidence and prevalence are clearly related to each other. An example used to describe the relationship is the 'bathtub model'. In this model, incidence is the incoming flow of cases and prevalence is the existing cases (Bhopal, 2008). This model can be expanded to consider dynamic population flows by adding a reservoir upstream of births and immigration from which the incidence cases flow and where the recoveries of the disease feed. The prevalence tub would then contain two drains, one for recoveries and the other for emigration, un-measured cases and death. Figure 3 illustrates these two related concepts.

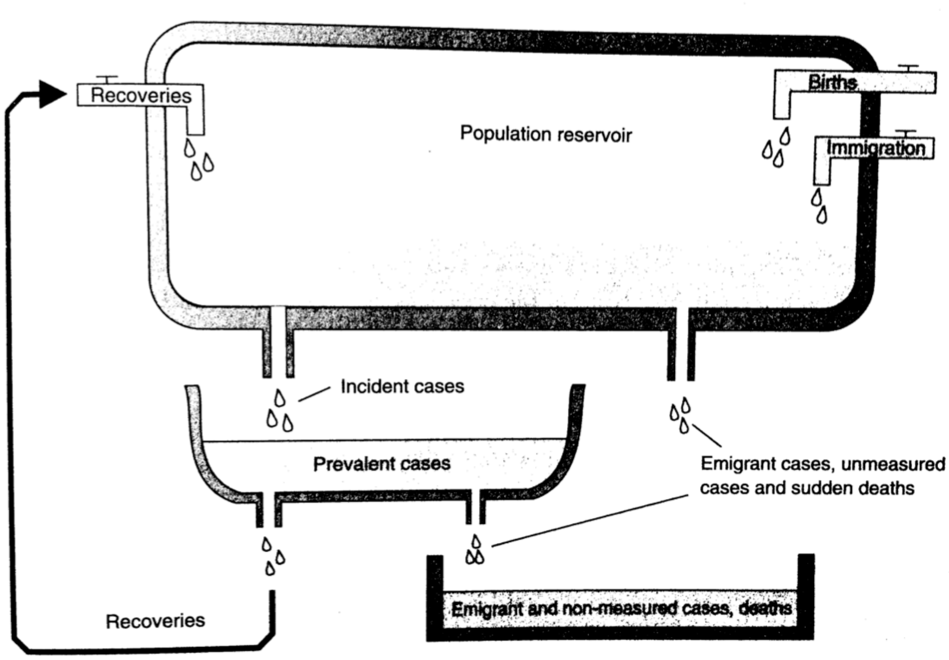


Figure 3 - Incidence and prevalence in a natural population: the bathtub model (Bhopal, 2008, p. 225)

3.3.3 Populations

Describing disease frequency among a population clearly requires a strong understanding of what a population is, as it relates to epidemiological studies. Epidemiology is fundamentally a 'population' science, exploring disease patterns in groups using information on individuals through which to develop appropriate responses to health priorities. The focus is often to identify causal factors and establish attributable risk or describe conditions and underlying issues within a target population (e.g. the risk of hypertension in working age women) from which a sample population is selected that can reasonably approximate the target group (e.g. a random selection of working age women from clinics in 5 U.S. states with risk markers of hypertension) and results extrapolated (Woodward, 2004).

Distributions of population health outcomes are a fundamental part of epidemiology and are used to express the spread and shape of disease occurrence or key factors of interest. This concept suggests that those with health problems are part of a continuum (i.e. from diseased to healthy), existing in the extremes of the spread. This can be extended to a strategy of controlling disease by 'shifting' the distribution towards the healthy end of the spectrum, thus defining a new 'normal', see Figure 4 for an example. The causes of differences between health and unhealthy populations are then the focus of epidemiological research and thus preventative actions should focus on the 'causes of causes' (Rose et al., 2008).

Research on obesity is of particular relevance in that while obesity is not a disease it is a variable condition that acts as a strong risk factor for numerous chronic diseases and adverse health conditions in later life. Moreover, studies on obesity have increasingly recognised the role of wide-ranging and inter-related factors in the social and physical environment that influence its prevalence, from dietary patterns influenced by the availability and supply of foods to the role of the urban environment that acts to discourage or limit physical activity (this parallel is discussed further in Chapter 4). In comparison are those preventative strategies that focus on the edges of these distributions whereby those with extreme risk are the subject of controls or actions. This is known as a 'high-risk' prevention strategy.

The ideal population would include all members at risk of a particular outcome, but this is likely to be difficult to ascertain and therefore samples are drawn so as to be as representative of the target population as possible (Bhopal, 2008). Describing the population requires knowledge of people's demographic and health features, often drawn from registries or surveys. As a result, large population surveys, such as the census, or appropriately sampled surveys, such as the Health Survey of England, provide a contextual resource when drawing much smaller samples for specific studies. These registries provide vital knowledge on the size and composition of the target population and their characteristics. When developing studies, these resources are used to inform the selection of individuals, and the physical and social boundaries from within which to conduct surveys. Populations, of course, are not static entities and as such the characteristic descriptions will change to reflect changes in groups and individuals, for example changes in age structure (and therefore risk of disease), or through migration and immigration (and therefore changes in background occurrence of disease).

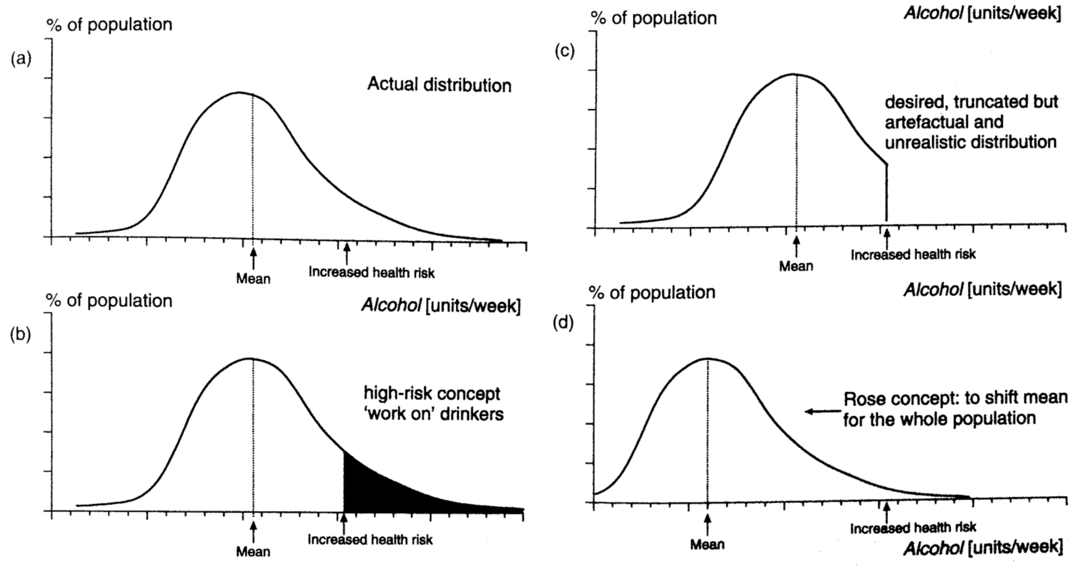


Figure 4 - Rose's distributions (Bhopal, 2008, p. 31)

3.3.4 Variation in outcomes

The study of the occurrence of disease within a defined population, its prevalence and incidence, provides the means through which to investigate changes in disease. Changes in health outcomes may be the result of many interacting factors, the understanding of which will provide a better means of their control. Policy can be developed to undertake actions to bring disease occurrence under control by using knowledge of the time trends of the disease and its putative or established causes. Often, patterns of health outcomes and their changes will be the result of inter-related social and biological factors, such as changes in employment or environmental exposure. In epidemiological studies, it is often the goal to understand factors that cause variation in health outcomes amongst a particular population or group. However, to identify the causes of variation is difficult and requires the use of well-structured analyses to identify potential causal factors. In understanding variation in disease occurrence, artefacts of data collection or measurement must be excluded, and hypotheses of association tested and interpreted in a causal framework (Bhopal, 2008).

Variation can be tested by measuring changes in health outcomes over time and associating these with changes in putative risk factors to determine an association. Association may be based on first principles or through observation. While association may imply causation it is insufficient to act as evidence of a causal link. Instead, epidemiological studies seek "to quantify the association, to assess whether the association is causal, and, if so, to explain how" (Bhopal, 2008 p. 61). This is done through deeper and more specialised study of the putative causal risk factors, to understand the mechanism through which they affect the disease.

3.3.5 Risk factors and disease mechanisms

Epidemiology provides research on the occurrence of disease in populations that have been exposed to a risk factor, compared to those that are unexposed, in order to identify putative associations between these factors and disease outcomes, in support of causality, and thus the application of appropriate interventions (Bailey et al., 2005). Establishing disease aetiology (the causes of disease) by identifying ‘attributable’ risk factors (i.e. those excess cases of disease occurrence that are attributable to a risk factor) is a central goal of epidemiology. A risk factor is defined as:

> “1. An aspect of personal behaviour or lifestyle, an environmental exposure, or an inborn or inherited characteristic that, on the basis of scientific evidence, is known to be associated with meaningful health-related condition(s);
>
> “2. An attribute or exposure that is associated with an increased probability of a specified outcome, such as the occurrence of a disease. Not necessarily a causal factor: it may be a risk marker;
>
> “3. A determinant that can be modified by intervention, thereby reducing the probability of occurrence of disease or other outcomes. It may be referred to as a modifiable risk factor, and logically must be a cause of the disease.” (Porta *et al*, 2008, p. 218)

These definitions specify that a risk factor may be a directly observed attribute, such as smoking, or may instead by a proxy for practices that together contribute to a higher disease risk, such as low income, or that may be the subject of some sort of modification or intervention. Identifying disease risk factors requires a thorough understanding of the contextual features that surround and interact with them as much as those biological factors that are inherited by them. Defining risk factors also requires some concept of the putative causes of a disease or health outcome. For example, while Doll and Hill (1950) identified a number of putative factors associated with lung cancer, smoking (and the amount smoked) was shown to have a strong effect. Later, Doll *et al.* (2004) highlighted how social institutions also played a strong role in smoking practices, by showing a difference in smoking habits between cohorts (Doll et al., 2004). The concept of putative risk factors should be developed within a causal framework that can provide the basis for the development of hypothesis from which to test aetiological relationships.

3.3.6 Causality

Understanding the cause(s) of health outcomes is essential to their control. The causal mechanism, however, is inherently complex and often subject to limited understanding for reasons of inadequate definition or measurement. Simply, “a factor is a cause of an event if its operation increases the frequency of the event” (Elwood, 2007). A causal factor therefore leads to an effect through the understood causal mechanism, which may be singular in its nature or multitudinous and complex.

The causal model has provided a means of understanding causal mechanisms. These models have evolved along with the understanding of diseases and their manifestation.

There are generally four types of causal model, moving from simple to complex: 1) Line, 2) Triangle, 3) Wheel and 4) Web (Bailey et al., 2005; Bhopal, 2008) – see Figure 5 below.

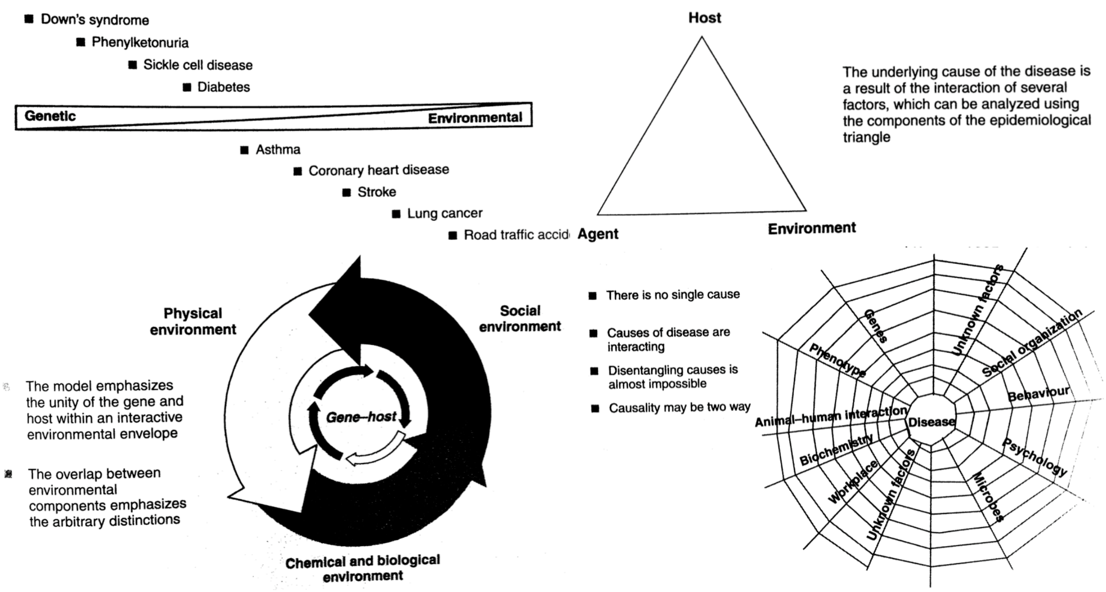


Figure 5 - Epidemiological causal models (Bhopal, 2008, p. 128, 131, 135, 137)

The most basic causal model is the 'line of causation', where causes lie along a spectrum of influence, from those factors that are inherited (e.g. genetic) to those that are environmental (e.g. contextual). This model emphasises the relative degree to which factors are specifically related to an outcome compared to those that are the result of many coinciding factors. The 'triangle' model captures the concept of interactive causal relationships within the 'agent-host-environment' formation. In this concept, agents may be those factors that by themselves have an active effect (e.g. a bacillus such as streptococcus); hosts factors are the features and characteristics of individuals that may have a different effect from the agent factors (e.g. age or weight); and environment factors are those that are socially and contextually driven (e.g. indoor air pollutants). The 'wheel' model emphasises the role of interacting factors that overlap between the physical, social and biological environment. The 'causal web' acknowledges the complex interactions between a range of physical, social and biological factors. This approach suggests that there is no single and specific cause, but instead there are many causal pathways to a health outcome. This last approach has characterised the past two decades of epidemiological study, emphasising the need to search for multiple causes and effects (Krieger, 1994).

An extension to the concept of multi-causal factors is the idea of 'sufficient and necessary' causal factors. Sufficient causes are a 'minimum' set of features, conditions or factors that together produce an outcome, while a necessary cause is a factor whose presence is required for an outcome (Rothman and Greenland, 2005). This concept suggests that no specific putative causal factor is by itself a sufficient cause of a health outcome but that the interaction of a number of causal factors may contribute to an outcome. Figure 6 shows an example of a collection of components (i.e. wedges) as a causal mechanism of disease. Using lung cancer as an example of the disease, in I) a number of components could come together, such as smoking, diet, genetics, workplace, etc. to result in lung cancer (i.e. a causal mechanism. In II), exposure to asbestos (G) may be a single component cause of lung cancer but remains influenced by other components. While III) still results in lung cancer but with a largely different set of components. The point Rothman makes is that more than a single causal mechanism can cause a given disease.

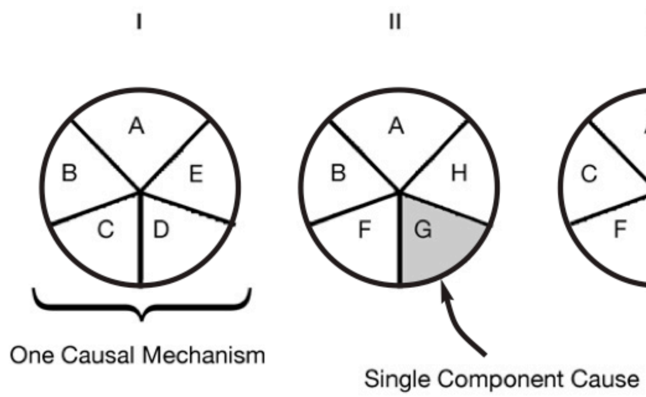


Figure 6 - Components of disease (Rothman, 2005, pp. S145)

The means by which causal inferences can be made are not straightforward and will often depend on the background of those developing and interpreting studies seeking aetiology. Epidemiological studies in themselves are not sufficient to establish causal mechanisms (except in a probabilistic manner); rather they seek to identify strong associations between health outcomes and putative risk factors that can be supported through complementary investigations. This often means that epidemiology relies, if possible, on other sciences to understand the function of the causal mechanism. In epidemiological studies, causal criteria are often used to differentiate between causal and non-causal factors. Commonly accepted criteria are the ‘viewpoints’ proposed by Bradford Hill (1960), who sets out a series of ‘tests’ that attempt to provide a process by which to establish a causal link. The criteria rely on the reasoning of John Stuart Mill’s canons using the complementary methods of concomitant variation, agreement, difference and residues. Hill’s guidelines are summarised in Table 3 below.

Each guideline should be taken not as a prescriptive test that must be met in order to establish causality but instead as a marker and recommendation for approaching a causal inference. These guidelines are subject to their own caveats and limitations that mean if one particular guideline is not met then this does not necessarily mean that the factor is causal or not. Rothman & Greenland (2005) offer a number of caveats against each guideline. For example, weak associations may more likely be the result of some form of bias; however it may still be the case that weak effects are causally linked, while some very strongly associated exposures may not be causal (Rothman and Greenland, 2005). These guidelines are an important part of epidemiological studies used to determine the strength of the investigated relationship in leading towards causality. These guidelines are referred to throughout this thesis.

Table 3 - Hill's Causal guidelines, modified from (Rothman & Greenland, 2005, pp.S148 and Bhopal, 2008, p. 146)

Guideline	Concept
i. Strength	Strong associations are more likely to be causal than weak associations. If other factors could explain the effect they would need to be even stronger than the observed association.
ii. Consistency	An observed association is consistently shown in different populations and under different circumstances and study designs.
iii. Specificity	A cause leads to a single effect.
iv. Temporality	A cause must precede an effect in time.
v. Biological gradient	A gradient, or dose-response, should exist.
vi. Plausibility	The biological causal mechanism of the exposure should be plausible, given current scientific knowledge.
vii. Coherence	A cause and effect interpretation for an association does not conflict with what is known of the natural history and biology of the disease.
viii. Experimental evidence	That changes in the level of exposure, through prevention or intervention, cause changes in disease experience.

3.3.7 Measuring association

Statistical analysis is the primary method through which associations between risk factors and disease outcomes are measured. This complements the observational nature of epidemiology and the approach to study design, whereby sampled populations are used for inference to larger target populations. For reasons related to the (often) complex nature of hypothesized causal pathways, the association between exposures and effects can be difficult to establish. Studies measuring 'real-world' effects do not have the same level of control available to laboratory studies and therefore the tests used for measuring association must be carefully applied.

Epidemiological studies primarily seek to test causal hypotheses between risk factor exposure and health outcomes using comparisons between exposed and unexposed populations. The tests of association should measure the impact while accounting for differences between the groups. At an individual level, the exposure to a risk factor may be the subject of multiple forces acting to increase or mitigate the effect. The quantification of association often relies on relative measures of the size of the impact and the likelihood of the health outcome among the population being studied. Absolute measures of association are also used and provide an estimate of the frequency of disease between exposed and unexposed groups. They can be used to measure how strongly an exposure is associated with a particular disease and are often used to determine the number affected by a disease.

Epidemiological studies that rely on between-group comparisons for the measurement of associations will often make use of ratios between numbers exposed and unexposed. Similar to the measurement of disease incidence, there are three types of association ratio: risk ratio, rate ratio, and odds ratio. Studies that assess the absolute frequency of disease measure total changes, known as the attributable (or excess) risk, which can also be measured as the 'population attributable risk' or 'fraction'.

The *relative risk ratio* (known as either relative risk or risk ratio) is the 'ratio between the cumulative incidence [i.e. new occurrences of outcome over a period of time as a proportion of the population] in the exposed group and the cumulative incidence in the unexposed group' (Bailey et al., 2005). The unexposed group may also be a reference population whereby the exposure is the population average and provides a baseline for a group with a higher exposure. Where the comparison is a simple dichotomy between exposed/unexposed, the comparison of risk is often presented in a 2x2 table – see Table 4 and Equation 6. The relative risk is the main measure of the size of the effect of a risk factor and strength of association and is used as a measure of aetiological strength (dos Santos Silva, 1999). The relative risk can be derived from those studies that provide incidence data, such as cohort studies, registry studies, trials and cross-sectional surveys. The risk ratio is a unitless measure, whereby a value of 1.0 indicates that there is no difference between the risk rates of the exposed and unexposed groups. A value greater than 1.0 suggests a positive association and an increased risk of disease among the exposed group, while a value less than 1.0 suggests a negative association and therefore a decreased risk among those exposed. A decreased risk associated with a factor is often termed as 'protective' against the disease.

Table 4 - Epidemiological 2 x 2 table

	Clinical outcome		
Risk factor/ exposure	**Diseased**	**Not diseased**	**Total**
Present (exposed)	*a*	*b*	*a + b*
Absent (not exposed)	*c*	*d*	*c + d*
Total	**a + c**	**b + d**	**a + b + c + d**

Equation 6

Relative risk ratio *= Cumulative incidence rate in those with risk factor/ Cumulative incidence rate in those without risk factor.*

or

Relative risk $= (a/(a+b))/(c/(c+d))$

The *odds ratio* is a measure of association that relies on the "chance of being exposed as opposed to not being exposed" (Bhopal, 2008). The odds ratio is measured as the number of outcomes (or non-outcomes) for those without a factor over those with a factor. It measures the number of times the outcome occurs relative to the number of times it does not. It describes "how many more times likely the cases are to have been exposed to the factor under study compared with the controls" (dos Santos Silva, 1999). The odds ratio (OR) and risk ratio (RR) are read in the same manner and are generally equivalent, i.e. when OR=1 so does RR, meaning the estimated effect is the same. Odds ratios are used when incidence of disease is unknown, or for studies where participants are selected on the basis of disease status (i.e. retrospective). The odds ratio can be expressed (using values in Table 4) as:

Equation 7

Odds ratio = *Odds of disease in those with risk factor/ Odds of disease in those without risk factor.*

or

Odds ratio = $(a/b)/(c/d)$

Although they have strong numerical similarities, note that in medical literature, a distinction is made between ORs and RRs. In case-control and cohort studies, where an outcome is rare (e.g. <10% in the unexposed population), the OR is a reasonable approximation of RR. However, when the outcome is common, the OR will exaggerate the RR. This is because the OR is symmetrical regarding the outcome, while the RR is not. Figure 7 provides an illustration of this issue, where RR and OR are not equivalent when comparing between the Low and High vehicle crash speed.

Table 3. Hypothetical Cohort Study of Seat Belt Use and Death in a Traffic Crash

Vehicle Crash Speed	Seat Belt Used	Outcome, No. Died	Outcome, No. Survived	Risk	Risk Ratio	Odds	Odds Ratio
Low	Yes	25	4975	.005	0.50	0.005	0.50
	No	50	4950	.010		0.010	
High	Yes	125	375	.250	0.50	0.333	0.33
	No	250	250	.500		1.000	
Total		450	10 550	.041		0.043	

Figure 7 - Example of differences in symmetry of ORs and RRs

The *rate ratio* is similar in nature to the risk ratio, but instead of using cumulative incidence it uses the person-time incidence rate (i.e. new occurrences of outcome over a period of time). By accounting for the amount of time spent in a study, it is often used when studying common outcomes because it allows for people entering and leaving the study. The rate ratio can be expressed as:

Equation 8

Rate ratio *= Person-time incidence rate in those with risk factor/ Person-time incidence rate in those without risk factor.*

Attributable risk provides a measure of the excess frequency of disease occurrence in an exposed group over that in the unexposed group. This measure provides an indication of the impact of disease through the total increase in numbers or percentage for that population being affected by the exposure. The attributable risk provides a method of considering alternative actions by identifying the *preventable fraction* associated with a disease, and by applying knowledge of preventative actions that mitigate or reduce the occurrence of the health outcome. In doing so, the risk factor under assessment should have strong evidence of its contribution to the cause of disease, along with knowledge of the occurrence of the risk factor amongst the population, the overall burden of disease (i.e. occurrence compared to other health outcomes), and knowledge of effective preventative actions (Bhopal, 2008) pp. 254. The attributable risk, attributable risk percentage, and preventable fraction can be expressed as:

Equation 9

Attributable risk *= Incidence rate in exposed group – incidence rate in unexposed group.*

or

Attributable risk per cent *= (Incidence rate in exposed group – incidence rate in unexposed group) / Incidence rate in exposed group*

or

Preventable fraction *= (Risk rate in exposed group – risk rate in unexposed group) / risk in unexposed group*

The *population attributable risk* is an extension of the attributable risk but is a measure for the target population (e.g. persons of a certain age). The measure uses knowledge of the incidence of a health outcome in the whole population along with the incidence of the outcome in the non-exposed population.

Equation 10

Population-attributable risk *= (Risk rate in whole population – risk rate in unexposed population) / Risk rate in total population*

In many cases the total incidence of the outcome will not be known for the whole population. An alternative method of estimating this incidence is via a representative sample of the population, such as a cross-sectional survey, that can provide estimates on the prevalence of exposure. These are combined with the relative risks ratio from cohort studies or the odds ratio from case-control studies. This can be expressed as:

Equation 11

Population-attributable risk *= Prevalence of risk factor in population * (relative risk ratio – 1) /*
*1 + prevalence of risk factor in population * (relative risk ratio -1)*
or,
Population-attributable risk *= Prevalence of risk factor in population * (odds ratio – 1) /*
*1 + prevalence of risk factor in population * (odds risk ratio -1)*

The above measures of association are used to determine the strength of association between a putative risk factor and a health outcome. The use of the different measures will depend on the study being performed in addition to available information. Each has its own requirements that must be satisfied when used.

3.3.8 Bias and confounding

While 'the aim of many epidemiological studies is to establish causal association between exposures and outcomes' (Bailey et al., 2005), knowing whether an effect or association is real requires that all other possible reasons that could lead to the relationship are excluded. The reasons include whether there is error in the way the study was undertaken, referred to as bias, or differences between groups that were not accounted for (i.e. confounding), or chance. A study with internal validity is one that does not suffer from bias, confounders or chance.

Error is defined as a "false or mistaken result of a measurement" (Porta et al., 2008) and may take the form of random error or systematic error. Random error, or chance, is unconnected to other measured variables, while systematic error, or bias, occurs unequally and usually in a particular direction. In epidemiological studies, bias is often categorised as *selection* or *information* bias. These forms of bias can be further sub-divided to identify possible sources of bias. Table 5 lists these sources of bias along with some details. Selection bias is most likely to seriously undermine the comparison between groups of case-control or intervention studies through differences in the group characteristics. For cross-sectional or cohort studies, selection bias can cause the study group to be unrepresentative of the target group (i.e. population to which inference is made). Information bias, also known as differential misclassification, is where one group of subjects is more likely to be misclassified. The effect of this type of bias is that subjects are wrongly classified as having been exposed to a risk factor or having a disease, leading to incorrect estimates of their risk or altering the strength of association.

Avoiding biases requires that their potential is addressed and assessed throughout the study, from the design phase, through data collection and analysis, and finally to results and especially interpretation. Methods to address selection bias could be through randomization (depending on the study type) and the careful selection of subjects and matching to a baseline population. Information bias can be avoided through careful attention to the development, collection and analysis of variables. Examples of these could be avoiding poor quality sources of information, such as surveys that rely on memory, or ensuring that questionnaires are validated through cross-referenced 'gold standards', i.e. tested and defined survey instruments and variables, such as the General Health Questionnaire (GHQ). Often used in intervention or case-control studies, a further method of reducing bias related to the research is to use a 'blind' or 'double-blind' approach. Blind studies are where the participants do not know which intervention they are receiving, while double-blind are those where both the participant and observer are blind to the intervention or allocation (Elwood, 2007). A number of organisations exist that provide guidance and protocols for carrying out studies and avoiding bias.

Table 5 – Forms of epidemiological study error and bias, from Bhopal, 2008 pp. 87 and Bailey et al, 2004, pp. 98

General form	Specific form	Source or cause
Selection bias		
	Population bias	Population chosen is not representative of the target population.
	Berkson's bias	Both the exposure and the disease under study affect the selection group due to increased probability of hospitalisation, causing cases and controls to be systematically different in hospital studies.
	Response bias	Unequal participation, interest or motivation in certain groups
	Follow-up bias	Difference in maintaining contact with subjects
	Drop-out bias	Difference in continuation of participation in study
Information bias		
	Measurement error	Mismeasurement of disease and risk factors arising from improper collection process, technique or calibration
	Recall bias	Unequal recall of occurrence or risk factors among subjects
	Work-up bias	Difference in collection ability among comparison groups
	Interview bias	Difference in information collected from groups due to interviewer

A confounder can be defined as “the distortion of a measure of the effect of an exposure on an outcome due to the association of the exposure with other factors that influence the occurrence of the outcome” (Porta et al., 2008). A confounder must be related to, but not a consequence of, the health outcome or disease; and it must be related to, but not a consequence of, the risk factor (Woodward, 2004). Common confounders in epidemiological studies are age and gender, both of which have been shown to have an effect on disease outcomes. Figure 8 provides an illustration of confounding. A confounder (C) may be causally related to both the risk factor and the disease (a), or a confounder may be non-causally related to both the risk factor and disease (b), or the confounder may be causally related to the disease but the risk factor is non-causally related to both the disease and confounder (c). In the latter case of diagram C, it may be that the risk factor is a proxy for other unobserved factors. A perfect confounder occurs when the risk of an outcome for the different levels of confounding variables are the same, although this rarely occurs in practice. A practical example would be a relationship between private rental housing and low energy demand. Whilst the two may be related, simply being a privately rented dwelling does not cause low energy demand; rather household socio-economic status is a confounder of the relationship. In this case, being of a low socio-economic status makes it more likely to live in private rented accommodation and also less likely to be able to afford much energy.

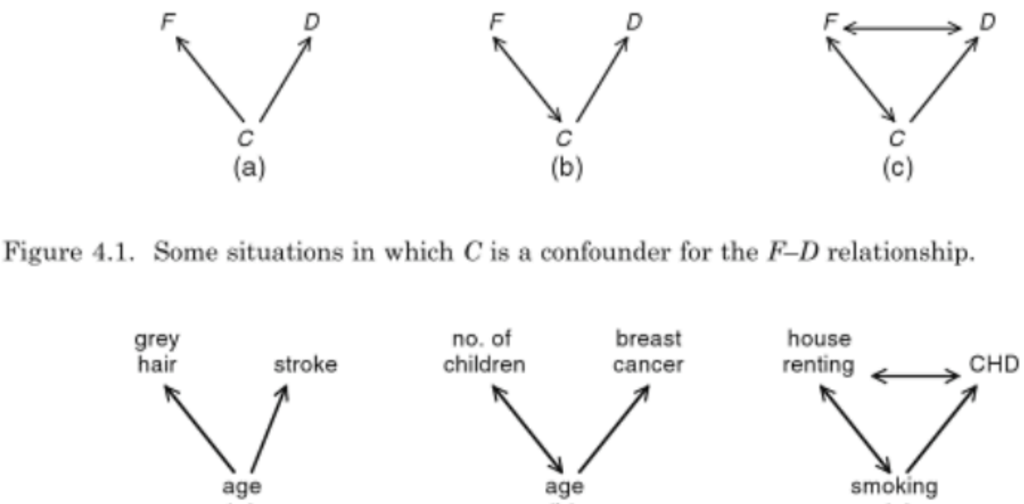


Figure 8 – Example of causal diagrams (Woodward, 2004, p. 169); CHD is coronary heart disease.

In epidemiology, an unadjusted relationship or pattern is denoted with the term ‘crude’, which means that the relationship is the overall actual value and reflects the real effect for that population. However, a crude rate can mislead by not accounting for difference or potential confounders that would affect the relationship found. Identifying confounders relies on the observed data and also *a priori* knowledge of the risk factors and outcomes under study. In practice, it is better to treat factors suspected of confounding and assess whether they have a real influence by estimating the effect of the risk factor on the risk of health outcome with and without controlling for the confounder. This approach often estimates the effect by strata of the confounder (one or more levels) and is termed as being 'adjusted'. Adjusting provides a means for ensuring that the relationship found can be compared to other studies looking at a similar issue. Other approaches adjust for confounding by looking for an independent effect that occurs over and above any adjustment, i.e. where the relationship between the factor and disease is still (statistically) significant after adjustment. Often, where several confounding variables exist, it is necessary to adjust for all the factors and compare each separately and then with adjustment for all confounding factors (i.e. multiple adjustment). This approach provides the most effective way of determining if confounding has occurred.

Chance variation is most commonly addressed through statistical analysis that attempts to determine the probability of an effect occurring due to chance. There are various forms of statistical error that are linked to significance levels set for statistical tests (i.e. Type I and Type II errors). A Type I error occurs by falsely rejecting the hypothesis when it is true and a Type II error occurs by falsely accepting the hypothesis when an alternative hypothesis is true. These forms of statistical error are closely associated with the size of the study. For epidemiology, the sample size will be determined by the research question being asked, the magnitude of the effect being assessed and the minimum size of any difference being compared (Bhopal, 2008; Woodward, 2004).

3.3.9 Interpretation

Epidemiological studies often involve clinical or medical researchers whose interest primarily lies in biological mechanisms for disease, but their expertise can ensure that appropriate methods can be applied in the data collection or the analysis relevant to the mechanism.

The epidemiological approach offers the specialist:

- The opportunity to use hypotheses generated through observation of the population for their own studies
- A method of testing hypotheses in their research by comparing their own prediction with that in the population
- The ability to ground their work in a defined sample of the population
- A theoretical framework and environment within which to study specific causal relationships and risk factor mechanisms

The epidemiological approach offers both a set of tools and a methodological framework within which to undertake analysis in search of socio-technical aetiology of models and to frame results and findings (Rose and Barker, 1978). The approach is based on four main functions, which are to:

i. Describe and measure the distribution of a condition or adverse outcome;
ii. Explain that distribution by its determinant factors (e.g. biological, environmental, social, behavioural);
iii. Predict changes expected in that distribution from interventions and control measures; and,
iv. Evaluate and shape policies to improve population health (Rose and Barker, 1978).

3.3.10 Data and data collection

In health research, data is commonly collected with the aim of developing datasets that are comparable while avoiding sources of bias. Data is collected through disease registries for individuals, often in clinical or hospital settings, or new studies using the designs mentioned above. Data used in epidemiology studies that come from aggregated collections (known as 'routine' data) is used to describe multiple individuals within a an area and time period and is often used to study the prevalence and to suggest hypotheses on the basis of which more complex study designs are applied. To be valid, specialised designs are used to identify associations and links between causal factors, and these need to tackle issues of sample population, size, variables, and ethical considerations to name a few. As a result, studies of disease are often subject to high degrees of justification and scrutiny from expert panels prior to research funds being granted. This has had the effect of formalising the data collection approach and instilling rigour in the science. Collecting established data over an extended period provides the opportunity to undertake longitudinal health studies, an important element in tracking changes in disease patterns and evaluating policy.

3.3.11 Standards and reviews

The evaluation of the evidence collected under epidemiological studies relies on the development of standards to govern and guide research methods and enable comparison of findings. For instance the Cochrane Collaboration represents an international network of researchers that over two decades has established a set of protocols for conducting systematic reviews that benchmark current findings, for the purpose of providing evidence-based healthcare (Cochrane Collaboration, 2013). Systematic reviews seek to collect all the evidence relevant to a specific pre-defined eligibility criteria to address a specific research question. They use explicit and systematic methods to minimise bias when interpreting the strength of the evidence, therefore offering more reliable findings to inform decision-making. While these reviews primarily focus on findings from appropriate randomised control trials (RCTs), as the gold standard for intervention studies to form evidence-based policy, other collaborations have formulated guidelines for observational studies (von Elm et al., 2007).

Systematic review in health sciences provides a method of assessing the validity of studies pertaining to an issue or disease. Doing so requires that quantitative results be presented along with confidence intervals (and other necessary statistics) in order to judge precision and assess the soundness of the study. The approach requires a focus, selection and synthesis of relevant studies, and that results be judged and conclusions drawn (Oxman, 1994). These reviews are an important activity within a discipline as they offer others in the field a consistency that could not be achieved by any single reviewer. Systematic reviews provide decision-makers with a robust resource of the latest evidence by identifying appraising and synthesizing a large amount of information and evidence and interpret the strengths and weaknesses (Higgins and Green, 2008). Without such a rigorous standard or operating process, the outputs of research activities may be critically faulted and judged as being unable to support an evidence basis.

3.4 Summary

This chapter provided a brief introduction to the field of epidemiology. It offered a short history of the evolution of epidemiology in terms of its theories and central paradigms, along with its present aims and objectives. Epidemiology is fundamentally an empirical science that seeks to explain distribution of health outcomes by their determinant factors to better manage and improve health. Modern epidemiological theory is concerned with understanding the determinants of health outcomes under differing social, environmental, biological, and genetic contexts.

The epidemiological conceptual framework includes defining health outcomes, frequency, variation, risk factors, disease mechanisms, causality, associations, and bias and confounding. The concepts are briefly discussed to give an overview of their basis and characteristics and in preparation for the following Chapter 4, which will examine their application to the study of energy demand.

Chapter 4 Conceptual Framework

Energy epidemiology: "a conceptual framework for population-level energy demand studies"

Chapter Introduction

The epidemiological approach has evolved over the past century to a discipline focused on understanding determinants of health outcomes and the influence of surrounding environments and underlying mechanisms on those outcomes. The approach is built around a conceptual framework that provides the basis for examining the occurrence of outcomes and their variation among populations.

This chapter explores the opportunity for applying an epidemiological approach to problems of end-use energy demand, in terms of its applicability. A number of the key epidemiological concepts identified from the literature are discussed in terms of how they might apply to the population-level study of end-use energy demand in buildings. The purpose of this chapter is to set out how these concepts could be adopted, adapted and integrated.

The chapter begins with an outline of the case for applying an epidemiological approach to studying energy demand at a population level. It goes on to examine and evaluate the epidemiological concepts and their applicability to problems of energy demand. The evaluation of concepts and their use in studying energy demand problems comprise three questions: Is the concept appropriate and applicable? Is the basis of the epidemiological concept maintained? And, can the concept advance the study of energy demand? The chapter concludes with a proposal for aims and principles for an energy epidemiological approach and a supporting conceptual framework.

4.1 Energy demand and the epidemiological approach

The lack of a strong evidence base is a fundamental part of the argument underpinning the identified need for a greater understanding of the factors that determine energy demand for services at a population level (Kelly, 2009; Oreszczyn and Lowe, 2010; Summerfield and Lowe, 2012). The case for more empirical studies of a sufficient size and quality able to identify and explain factors that relate to energy demand is made numerous times throughout the literature (Lomas, 2009; Lowe and Oreszczyn, 2008; Skea, 2012). However while a call has been made for more evidence it is not clear from the literature what conceptual framework such studies might be conceived within. A conceptual framework is important because it sets out the concepts through which to interpret and understand phenomena while also reflecting the research paradigm (Jabareen, 2009). These concepts contribute to, and are informed by, the over-arching aims and objectives present within a research paradigm. Having a clear conceptual framework provides a structure on which to develop a methodological framework, i.e. the methods (e.g. procedures and techniques) that support and test the concepts held within the paradigm. A methodological framework consists of methods that are: consistent in their approach, internally valid and free of bias, capable of generating and testing hypotheses, can identify and test relationships, can be used to examine groups within a defined context, and whose evidence can be used to inform policy. This thesis takes as its initiating rationale that the call for a more coherent approach to studying population-level energy demand can be addressed through a well-defined conceptual framework and the application of a strong methodological framework.

It is posited here that an epidemiological approach offers both a suitable conceptual framework and provides a methodological framework that can support the energy demand studies at a population-level. While the epidemiological approach is most commonly applied to health studies[2], it is argued here that many of its concepts are relevant to studies that seek to describe the distribution of energy demand outcomes or phenomena and their influencing factors.

The research questions examined here are related to the above assertion, and seek to test whether an epidemiological approach is suitable for the study of end-use energy demand, specifically in the residential sector. First, the energy demand research paradigms introduced in Chapter 2 are synthesized within a new definition around which population-level research could go on. Second, an exploration of whether the aims and principles of an epidemiological study approach supports an empirically-based approach to the study of energy demand at the population-level is provided. Third, key epidemiological concepts are examined for their relevance to energy demand and are assessed for whether they can be reasonably adapted to the study of population-level energy demand.

An exploration and discussion of the relevance of epidemiological concepts applied to the study of population-level energy demand goes directly to answering whether epidemiology's methodological framework would be appropriate. The question of the

[2] Epidemiological approaches have been applied conceptually to non-health studies, though rarely, as in research undertaken on problems of indoor air quality (Andersson, 1998)

nature of an appropriate methodological framework for population-level energy demand research is explored in Chapter 5. The exploration is carried out against the backdrop of the literature of empirical studies on end-use energy demand, and whether gaps in the evidence could be addressed through the epidemiological approach.

4.1.1 Applying an epidemiological approach

The identified need to study end-use energy demand at a population-level raises questions of whether the epidemiological approach provides a basis for establishing a new research paradigm, and also if such an approach would align with the aims and objectives of population-level energy demand. This provides the first two research questions (RQ) that will be addressed, namely:

> RQ1: Can the epidemiological approach support the paradigm of population-level study of end-use energy demand (focused here on the residential sector)? And,
>
> RQ2: What are the aims and objectives of an epidemiological approach, as applied to energy demand, in support of this paradigm?

These research questions are explored through a consideration of the paradigm, or fundamental basis, behind the population-level study of energy demand and how an epidemiological approach would support the paradigm through aligned aims and objectives.

4.1.2 Population-level study of end-use energy demand paradigm

As stated in Chapter 2, a paradigm is “a coherent pattern of research organized around commonly shared theoretical propositions and models”. A paradigm is not so much developed but instead is identified through the actions of the research community, the identified gaps, and the wider scientific beliefs through which research is carried out.

The central paradigm of epidemiology is that the health of the population can be improved through the systematic analysis of patterns of ill health and diseases to better understand their causes and differences (Bhopal, 2008). Also, that the causes of variation in diseases differ between populations, and that strategies that seek to control these ‘causes of causes’ are different between populations. The focus is then on shifting the distribution of risk factors in the population rather than the individual (Rose, 2001). This complements the overarching goal that improving the health of the population is a social good, a concept that is embodied in the ethos of the global health community (Horton, 2012).

A paradigm for the *population-level* study of end-use energy demand has not been explicitly stated within the literature, although there are a number of authors that have begun to set out the features of a new paradigm focused at a population level. Most have emphasized the need for more evidence derived from empirical information in order to meet the pressing need for changing energy demand patterns through intervention or information for the purpose of greenhouse gas abatement, and fuel security and fuel access. For example, in setting out the challenges facing the energy and building research community, Oreszczyn and Lowe (2010) state:

> "What has been largely absent from the debate, to date, has been comprehensive and high-quality empirical evidence on the actual performance of low-energy housing" (Oreszczyn & Lowe, 2010, p. 110)

These sentiments are echoed by Schweber and Leiringer (2012) who conclude from their overview of energy and buildings research over the past four decades that advances towards the overall aim of mitigating climate change through a low-carbon and low-energy built environment are dependent on bringing together theory and empirical research, which should be informed by both quantitative and qualitative evidence. They state:

> "In the rapid and unpredictable development of energy and buildings, there is a need for research that examines the processes, understandings, and motivations which produce observed patterns and systems." (Schweber & Leiringer, 2012, pp. 490-491)

The need to expand research approaches is closely related to an earlier call for shifts in research practices to focus on the 'consumer of energy services' that seek to better understand "how and why, and for what purposes" consumers use energy (Wilhite et al., 2000b). This earlier call was for more qualitative research with an emphasis on how the actors, institutions and the networks that individuals exist within affect the consumer's demand for services, placing this demand within a social, cultural and technical context. With respect to research, Wilhite *et al.*'s aim was to 're-tool research' by borrowing and adapting appropriate approaches, saying:

> "[A review of evidence suggests] there are existing tools and resources which can be used to tackle the broader agenda of energy demand. It also suggests [there is a] need to develop, extend and adapt many of these ideas and approaches. […] [And] if questions of demand are to be tackled head on, […], it will be necessary to draw a new population of social researchers into the field." (Wilhite *et al.*, 2000, p. 13)

The above calls for a paradigm shift are also reflected in the work by Hartenberger *et al.* (2013) who claim that built environment professionals lack an inclusive and shared identify for an inter-disciplinary approach. For such an approach, they looked at the health profession as a guide for training and education of built environment professionals who would benefit from a closed 'practice-research-education-training' loop. Evolving the community towards this approach would help to provide built environment professionals with higher levels of qualification, personal identification, and motivation and empowerment (Hartenberger et al., 2013).

Some authors have suggested that the prevailing paradigm through which energy demand research is happening is the 'low-carbon society' paradigm (Lomas, 2010). While this paradigm provides a basis for research activities, it does not reflect the full nature of the pattern (or type) of research that is taking places or the gaps in research, such as the need for more empirical analysis of the drivers of energy demand. Further, it does not define the theoretical basis of the research that is (or needs to be) taking place, even though this is to a degree implied by referencing 'low-carbon' and therefore the social good of climate change mitigation. The low-carbon paradigm is useful in its introduction to one of the wider aims of energy demand research, but it is not sufficiently defined to capture the needs and activities of research in population-level end-use energy demand.

Another important goal of end-use energy demand research is that of fuel security and access. The fuel poverty paradigm takes as its basis that improving access to and use of fuel by reducing a household's spending a disproportionate level of their income will have benefits to their overall health and well-being (Moore, 2012). This paradigm includes research that has sought to better understand the connections between fuel poverty and health (Liddell and Morris, 2010), the effect that energy efficiency interventions have had on the environment experienced by fuel-poor households (Milne and Boardman, 2000), and the evaluation of policies that seek to target fuel-poor households (Sefton, 2002). The research in this field has also included an examination of the 'fuel poverty' concept and its definitions (Moore, 2012; Waddams Price et al., 2012) along with a comprehensive review of the theoretical basis of the concept (Hills, 2012). Fuel poverty provides an excellent example of a research paradigm that is defined, self-critical, debated and evolving.

This thesis seeks to interpret the above research calls for a defined paradigm that focuses on the population-level and takes an empirical approach. It should be able to develop relevant evidence that can help guide the shift in the physical and technological components of the built environment. It should be able to deal with the shift in consumers' energy-related service demand practices, and reflect the low-carbon commitment and fuel access paradigms. It should also be able to include qualitative or mixed methods. Using the epidemiological approach as a template for this definition, the central paradigm for a population-level approach to the energy demand research is therefore:

> That the shift to a low-carbon society, along with the alleviation of energy-related social and environmental phenomena, such as fuel poverty and discomfort, can be improved through population-based methods that analyse patterns and systems of energy demand end uses, in order to better understand the practices, drivers and differences of energy demand outcomes.

It is within this paradigm that an epidemiological approach to energy demand will be examined.

4.1.3 Interpreting the epidemiological approach to energy demand

Energy demand research uses many different methods that are largely drawn from the disciplines within which any given energy-related issue is being studied. These approaches, identified in Chapter 2, are broadly the engineering/ physics-based approach that tackles mechanisms and engineered systems, and sociological and economic approaches that investigate effects related to social activity and practices. Each approach draws research designs and analysis techniques from their respective founding disciplines with the purpose of offering insight into domain-specific questions, for example building façade performance, temperature in houses, the value of social networks in the diffusion of energy efficiency practices, or the price elasticity of energy demand. At present there are no unifying methodological frameworks within which to undertake inter-disciplinary studies of energy demand. Therefore existing research activities tend often to be disparate rather than concerted, creating isolated pillars of understanding that risk being undermined by limited knowledge of their relevance or uncertainty within the broader energy demand context. It is proposed that the epidemiological research structure can be re-interpreted in order to found a coherent research and analysis framework from which to address the issues surrounding energy demand.

The major advances in health sciences research seen over the past century-and-a-half have been the product of both individual disciplines and their interactions within an integrated model of research focused on health. This health research system includes a series of models and practices that draw together a number of concepts, methods and disciplines, for example: the *biomedical* model (e.g. pathology, biochemistry and physiology), the s*ocio-behavioural* model (e.g. psychology and sociology), the *genomic* model (e.g. genes and social/environmental 'switches'), along with the *epidemiological* model (e.g. population studies). A number of features are in place to support this integrated research system including: laboratory research and testing, clinical diagnostics, surveys and registries, systematic reviews and meta-analyses, and a range of long-term and in-depth data collection exercises. There is interdependence between these different research models (or systems) whereby each provides insights or evidence to be used by another. For epidemiology, Bhopal (2008) says:

> "The science of epidemiology, therefore, combines elements of biology, social sciences and ecology: a biosocial-environmental science focusing on disease in population. By its nature epidemiology is multidisciplinary." (Bhopal, 2008, p.5)

The proposed structure for an epidemiological approach to end-use energy demand includes at least three parts that reflects the different disciplines working within the research field (Table 6 and Figure 9). These include: *end-use energy processes and systems* (i.e. engineering and physical sciences), *end-use energy practices* (i.e. socio-behavioural interactions) and the *end-use energy context* (i.e. structure and conditions of systems and practices). This structure can be seen as being interdependent, with findings from the various models being shared and built upon. Although the population level 'epidemiological model' encompasses the other three models, this does not imply a hierarchy. Rather, it means to emphasis how an epidemiological approach uses theory, insights, understanding, aetiology and findings to drive forward the detection of patterns related to energy demand at the population level.

'*Physical processes and engineered systems*' related to energy demand, for example the engineered and designed system of the end-use (or service), building or built environment, are the physical processes through which the environment and user interact. Studies within this model may be mostly quantitative and observational and seek to describe functions and boundaries of a system. This physical and engineering model could be likened to a model of the cardiovascular system, circulating and oxygenating blood through the heart, or to a heating system that circulates hot water for space heating via radiators through a boiler. In both these examples, a detailed understanding of the mechanisms that operate the system, the tolerances and optimal operating conditions, and the context that the system is operating within should be well understood. '*End-use energy practices*' include the practices embodied in end-use energy through the behaviours and norms, personal beliefs and values, and communication of social institutions. For example, research into social practices could seek to understand how cultural norms affects the amount of resource used, such as the intake of calories in obesity studies, or the use of gas related to temperature in houses. These studies may straddle both quantitative and qualitative methods and play an important role when exploring new areas, developing theory, and describing experiences and insight into practices. The '*end-use energy context*' model describes the societal, political and environmental features that define and determine the engineered structures (e.g. regulation and standards) and the context within which user interactions take place. Whilst both the physical and social models shape context, this area of research focuses on the environment, not only the natural environmental conditions but also on 'place' and on social and political structures. For example, this model might consider how regulatory frameworks such as the application of building regulations affect the energy performance of buildings. The '*population level end-use*' model (or epidemiological model) should focus on describing and explaining end-use energy demand patterns and using this information to develop policies to address problems or modify physical and institutional structures to effect change. It must rely on the insights provided by the other models to inform the development of the interacting pathways or identify putative factors that might affect the outcome of interest.

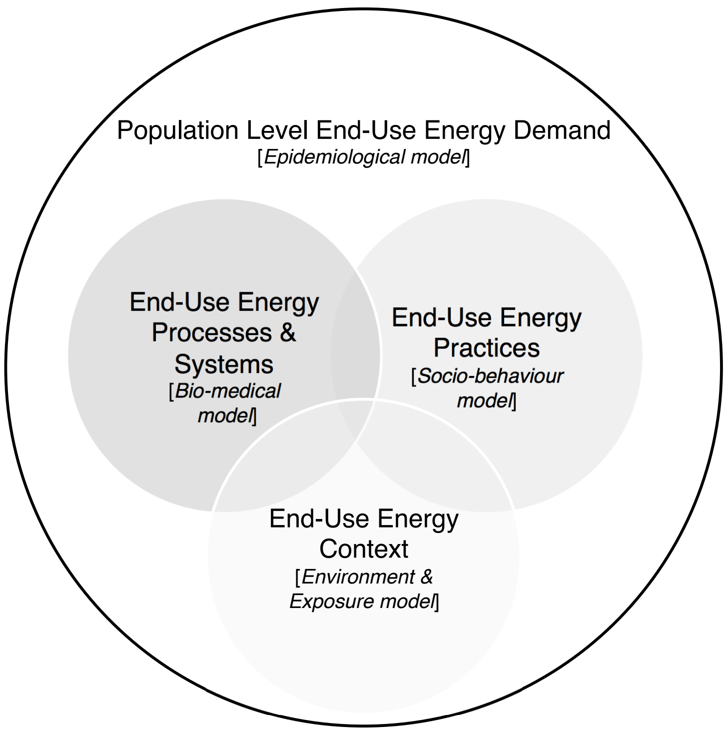


Figure 9 – Research concept for research in end-use energy demand epidemiology

Table 6 – Research approach for research in end-use energy demand epidemiology

Domain	Features
End-use energy physical processes and systems	The *physical systems* (e.g. thermo/fluid physics and engineered/technological systems) devised for service *demand*, within a *context* **Focus:** Study the physical processes and technological systems and mechanisms that support the use of energy within a given context
End-use energy practices	The *interactions* of users with a *physical system* for a service, within a *context.* **Focus:** Examine the motivations, values and reasons through which to interpret the relationship between physical mechanisms and social-cultural practices that contribute to the development of phenomena in energy demand
End-use energy context	The given *context* of both *practices* and *physical processes and systems* **Focus:** Examine the structure and conditions of the physical processes and systems, socio-cultural practices in context with a wide range of factors that act on the complex energy demand structure.

This proposed research structure makes use of existing expertise in buildings and energy research, and by applying an approach using existing energy demand concepts, and adapting relevant epidemiological concepts, can develop a common framework. The next section considers the conceptual similarities and differences between end-use energy demand and medical epidemiology.

4.1.4 Similarities and differences between epidemiology and energy demand

Beyond first glances there are a number of similarities between issues studied in health epidemiology and end-use energy demand that support the assertion that an epidemiological approach could be adapted to energy demand research. These include the multi/ inter-disciplinary nature of the respective spheres; the complex interactions of physical systems, environment, and socio-behavioural actions and their heterogeneous nature; a concept of a range of conditions, from 'normal' to 'abnormal'; and a surrounding support and delivery infrastructure to enable or prevent undesirable outcomes. There are also differences in the subject matters being addressed, e.g. that the unit of observation in epidemiology is more readily defined (i.e. a person) and is subject, as an entity, to less variation than say a dwelling (though that may depend on how variation is defined). However, as will be discussed, these differences may be seen as superficial. This work focuses on the approach, methods and techniques and not the specifics of any epidemiological studies.

For end-use energy demand, the need to control energy use for reasons of climate change abatement and socio-economic issues is similar in nature to the need to prevent and control adverse health outcomes. As with health studies focusing on improving public health, the process of developing appropriate interventions for a population or building stock will require that the detailed findings from research measuring physical processes and monitoring engineered systems are integrated with a knowledge of the social practices that affect the demand for energy. However, few studies make the connection between systems, practices, and context that would fit the concepts identified under an epidemiological approach. These are specifically, identifying putative factors (social, technical, physical, environmental, or some combination) that might influence the outcome of an intervention, and undertaking a systematic process of investigating and measuring the strength of their associations. The interaction of energy demand and buildings is a socio-technical phenomenon by nature and it needs to be investigated under an approach that is capable of dealing with this complexity.

Epidemiology has been defined as a study of health-related states or events and their occurrence among a population, along with the study of the determinants that influence those states or events (Porta et al., 2008). There are *many* different health-related states or events, for example infectious, contagious, mental disorder, or lifestyle, with many appearing and disappearing and at any given time affecting numerous people around the globe at a range of stages and outcomes. For example, the International Classification of Disease published by the WHO provides contains over 120,000 codes for diseases, along with a number of more specialist classifications (WHO, 2010).

Many epidemiological studies aim to describe disease, illness or sickness, their prevalence and incidence in a population, and the application of treatment leading to a health outcome. Disease can defined as:

> “a disorder of function or structure, especially one that produces specific signs or symptoms or that affects a specific location" (Oxford English Dictionary Online, 2010).

In epidemiology, disease takes the ‘biological dimension’, as compared to illness that reflects the ‘subjective or psychological state’ of being unwell, and sickness that reflects the ‘social condition’ of being ill with disease (Porta et al., 2008). Recent arguments in the health field have suggested that the term ‘disease’ should reflect a more inclusive concept. For example, Bircher (2005) suggests that:

> “Health is a dynamic state of wellbeing characterized by a physical, mental and social potential, which satisfies the demands of a life commensurate with age, culture, and personal responsibility. If the potential is insufficient to satisfy these demands the state is disease. This term includes sickness, illness, ill health, and malady. (Bircher, 2005)

What is important to note from the above definitions is that disease, illness, sickness, etc… do not imply final states, but rather allow for conditions to be defined within a complex context and a range of states. How do such concepts of disease or health-states apply to end-use energy demand?

End-use energy demand can be described as the "total energy supplied to the final consumer for energy-related services" (International Energy Agency (IEA), 2012). This definition suggests that the energy used (in the form of a fuel) is sought (i.e. demanded) and is related to both a service (i.e. those that use energy) and also a user (i.e. they that demand energy). Energy use is a means to a desired service (e.g. heating a living space or a tank of water to a desired temperature). As a research field its definition is less clear but, broadly speaking, it is defined here as seeking to describe the drivers of the demand for energy, its sources and fuels, services and uses, practices and norms, across interacting sectors and actors, with a focus on the built environment.

It is proposed here that end-use energy demand can be likened to a condition (or state) with an outcome that can be compared to other members of a population for a defined period or point in time. End-use energy demand is a proxy for the fuels demanded for a given service by a consumer (or a system), and the amount used within a period provides an indication of the underlying conditions that create variation across what could be a seemingly similar service. For example, the amount of space heating in a dwelling will vary as a result of a number of physical and environmental characteristics and consumer practices. The amount of end-use energy demanded by consumers could provide a basis for describing a range of outcomes related to the physical condition of the engineered systems within dwellings (e.g. the fabric heat loss), the practices of consumers for particular services (e.g. weekend heating patterns), or changes in the social context (e.g. price increases) or experienced environments (e.g. warm winters).

In this work, excessive end-use energy demand is likened to excess weight and obesity, itself a growing global epidemic threat whose study is fraught with complex interactions (Caballero, 2007). As a condition, obesity does not represent a biological disease; rather it is a strong risk factor for subsequent adverse health outcomes in later life. There is debate in the health field as to whether obesity is a disease at all or is instead a marker and precursor of diseases (Formiguera and Cantón, 2004). The very condition itself is difficult to accurately define. Indeed it is hard to define what is a 'normal' weight range for individuals (Canoy and Buchan, 2007), despite its ready depiction. Obesity is subject to variation in how it is defined and measured. For example, the Body Mass Index is a commonly used metric but has been shown to be limited and inaccurate when applied across a population (Canoy and Buchan, 2007), though it may have merit for identifying risks for individuals in clinical settings (Mooney et al., 2013). Other forms of measurement are used with similar effect (Hu, 2008). Obesity is a state that lies along a spectrum of symptoms and features that may contribute to ill health but is not itself a biological disease (it is a metabolic condition or disease), although it is a factor in biological outcomes such as heart disease or high blood pressure. The causes of the condition of obesity relate to individual activities such as diet, activity levels, physiology; but it is also a disease of complex interactions and feedbacks. This complexity is characterised by the concept of the 'obesogenic' environment whose conditions contribute to increased weight outside the direct control of the individual, e.g. through cheap access to high-fat diets, communities planned primarily for vehicle use, reduced physical activity, and social image pressures that reinforce particular weight norms (Egger and Swinburn, 1997). The concept has been defined as:

> “the sum of influences that the surroundings, opportunities, or conditions of life have on promoting obesity in individuals or populations” (Swinburn and Egger, 2002, p. 292)

The study of obesity continues to examine it as a 'disease' but has acknowledged the complex social and environmental factors that play a part, along with genetic and biological characteristics, in bringing about adverse outcomes.

Energy demand, like obesity, can be described along a spectrum with a host of interacting factors leading to a particular defined and measured outcome. While individual features of dwellings or consumer practices can highly influence the level of energy demand, defined with a given metric, the combined knowledge and exploration of these key determinants can offer insight into certain types of outcomes, such as the causes of excessive use or underuse of energy for a given population.

Further, the concept of the ‘obesogenic environment’ makes an interesting framework through which to examine trends in energy use. An ‘energy-genic environment’ might be one whose influences such as the physical condition of properties (e.g. building age), social opportunities (e.g. income levels or employment patterns) and conditions (e.g. colder climates) interact to contribute to a particular energy demand outcome (e.g. comparatively high or low consumption of energy services) for consumers within a given area.

The approach applied to the study of obesity offers a compelling model for the study of energy demand. Like obesity, the use of energy is complex and highly related to the characteristics and circumstances of the individual, such as fuels available, services, physical conditions, environmental conditions, social norms and individual practices quite as much as the structure of the technology and surrounding physical environment.

4.2 Concepts of energy epidemiology

In this work, a conceptual framework is defined as a set of related concepts that provide a comprehensive means of understanding phenomena (Jabareen, 2009). Here, the inter-related epidemiological concepts are used to understand and interpret the paradigm of population-level end-use energy demand and provide a basis for the third research question, namely:

> RQ3: Are the key epidemiological concepts appropriate and applicable for the population-level study of energy demand?

In the following section, this research question is addressed through a translation of the key epidemiological concepts presented in Chapter 3. These concepts make up the components of the conceptual framework and provide a means for proposing a set of energy epidemiology objectives. The concepts are adapted for energy demand and discussed in terms of the overall suitability of the epidemiological concept as it applies to the general study of energy demand. The exploration of these concepts contributes to a conceptual framework for energy epidemiology as proposed below.

The key concepts examined are:

- Cases, outcomes, conditions and events
- Measurement and definition
- Populations and sampling
- Change and variation
- Interventions and control
- Data collection
- Risk factors
- Causal pathways and causality
- Association
- Bias and confounding
- Policy and evaluation

The adaption of these concepts uses the information provided in Chapter 3 as a basis from which to discuss the practical application of an epidemiological approach to population-level end-use energy demand. The aim of this discussion is to identify the applicability of the concept to determine whether the concept reasonably applies to the study of energy demand.

In the following sections, each concept's main idea or theory is briefly summarized. Then, a discussion of its applicability to population-level end-use energy demand is presented using the following criteria to judge the appropriateness of applying the concept:

1. Is the concept appropriate to end-use energy demand?
2. Is the basis of the epidemiological concept maintained?
3. Can the concept advance the study of energy demand?

4.2.1 Epidemiological cases: outcomes, conditions and events

One of the main purposes of epidemiology is to study the occurrence of health-related outcomes, conditions or events among a population and look for factors that explain variation or differences. Outcomes in population-level studies of energy demand have focused primarily on how much energy (or fuel) is being used for services among the population and how this has changed over time.

Investigating changes in outcomes requires having a metric or standard against which these changes in conditions can be compared. In their work on the impact that government programmes had in meeting the UK Government's stated objective of reducing fuel poverty, Walker et al. (2013) sought to evaluate the change in fuel poverty risk following energy efficiency interventions. They concluded that the policies had some impact in reducing household energy costs but that the approach was 'hit-or-miss', and offered suggestions for improving targeting of investment and therefore the outcome measure (i.e. reducing fuel poverty) (Walker et al., 2013). The outcome in this case was whether the objective of reducing the level of fuel poverty had been met. The assessment looked at the introduction of energy efficiency interventions (i.e. events) in dwellings to ascertain whether a condition (i.e. fuel poverty) had been changed, resulting in a new outcome.

In an analysis of energy use in dwellings in Milton Keynes, Summerfield et al. (2007) undertook a follow-up study of dwellings identified as 'low-energy'. The dwellings were classified into three energy demand groups, high, medium and low, reflecting the relative 'condition' of the level of energy demand and investigated factors that might be associated with different space heating conditions (Summerfield et al., 2007). They showed that energy demand in the historically 'high' energy demand dwellings (measured in kWh/m^2/year) increased by approximately 25% from 1990 to 2005. This included a 75% increase in electricity use for non-heating and a 20% increase in gas demand. The study also sought to examine differences in the change in demand due to other household and dwelling factors, showing that 'high' users were primarily living in larger semi-detached houses with higher income levels. The categorisation in the energy demand condition (i.e. high, medium and low users) provided a means of examining the putative factors that might otherwise have confounded associations, e.g. the differences in the change in electricity demand and temperature.

In a study of dwellings undergoing measures to improve their energy performance as part of the Warm Front scheme in England, Oreszcyzn et al (2006) examined the factors that affected the indoor winter temperatures before and after the measures. The event being examined was the introduction of an energy efficiency measure in the dwelling, which was used as a transition point from which to examine association between winter indoor temperature and determinant factors. Differences in indoor temperature were associated with physical dwelling characteristics (i.e. age, type and size) and household occupant characteristics (i.e. age of household head, number of persons) (Oreszczyn et al., 2006a). The event (i.e. retrofit) was also shown to associate with increases in internal temperature, providing initial empirical evidence of the temperature take-back effect.

In answering the first question, is the concept applicable to energy demand, it is proposed here that the examined concepts of an outcome, condition and event are readily applied to the study of energy demand. The selected studies are only a handful of examples that could have been used as illustrations. In terms of whether the basis of the epidemiological concepts are maintained, it is proposed that for the most part the circumstances in which they have been applied in end-use energy demand are generally in line with the epidemiological intent of identifying a case for systematic examination, although in an informal way. The Summerfield et al (2007) study provides a good example of identifying cases, i.e. historic energy demand levels, which were then used to examine factors that resulted in differences in their ‘current’ energy demand condition. Regarding the last point of whether the concept can advance the study of energy demand, the method of identifying outcomes, conditions and events has already been used to advance thinking in population-level end-use energy demand studies. The work by the Warm Front group examining the impact of energy efficiency interventions (i.e. events) has provided an evidence base for the government in their impact evaluation of the Warm Front scheme.

4.2.2 Measuring outcomes, conditions and events

Description and definition are essential to support research into cases and patterns in epidemiology. Having a clear description of a disease, along with a standard definition, ensures that the condition or outcome being considered is widely comparable, allows for pattern detection between studies and thus monitor outcomes by time and location.

Is the concept appropriate?

Defining an appropriate discrete entity for analysis in energy demand studies is challenging. The current approach is to measure units of energy, usually derived and converted from fuel use defined over a period of time and normalised by a unit of observation to allow for comparison (e.g. gas used for space heating per net internal area m^2 per year). It is entirely appropriate to have a variety of conventions for measuring and describing end-use energy demand. However, more consistency in defining the parameters can allow for broader comparison. This means stating and justifying the unit(s) of analysis used at the outset, along with the development of an energy demand taxonomy.

In making any comparison meaningful, it is essential that the items being compared are approximately similar. In health epidemiology, the self-contained unit is (often) a person. In making comparisons in energy demand among populations, it is not only essential to have a well-measured variable of energy demand but essential too that the unit of observation is equally well defined so that features that would have a differential effect can be accounted for. As with people, the self-contained energy demand unit will have numerous characteristics that will affect the measured energy use for a given period that should be controlled for when undertaking studies. Adjusting for differences is an essential part of examining trends among a population and their associations with suspected influencing factors.

Is the basis of the epidemiological concept maintained?

For the most part, studies examining energy demand among dwellings or households do adjust or control for variables that could affect a comparison, particularly when examining a heterogeneous population. For example, controlling for floor area is a common approach when making comparison between the amounts of energy used in dwellings. In Summerfield et al (2007), the change in electricity and gas demand between 1990 and 2005 was adjusted for heated floor area and number of occupants in the household, which were shown to have an effect on demand. In Oreszczyn et al’s (2006) study of indoor mould and relative humidity (RH) differences in the standardised RH (standardised to 5 °C outdoor temperature) between different levels of dwelling age, type, energy efficiency and deprivation were taken account of in the average values reported (Oreszczyn et al., 2006b). A similar approach was used to examine the determinants of indoor temperature in the same study of homes (Oreszczyn et al., 2006a). In a study that looked at differences between reported thermostat settings in English houses between 1984 and 2007, Shipworth (2011) showed that there was no statistical difference between the examined periods. The results accounted for differences from a range of dwelling stock features (i.e. household age, tenure, dwelling age, fabric air-tightness, double glazing). In reporting the findings, 'crude' mean thermostat settings were given at the outset with a further examination of the difference for each set of putative influencing variables. However, in these latter studies, the differences in the outcome variables being analysed were not brought together to develop a fully adjusted outcome variable that could be used for further comparison to other studies.

Bhopal (2005) suggests that a good epidemiological variable should have some discernable health impact among the population, be measured accurately, serve to differentiate populations by patterns and underlying characteristics, generate testable hypotheses, develop policy, and deliver interventions and controls. It is proposed here that the need for well-defined and measured outcome variables is applicable to the study of energy demand. Measurement and definition are the starting point for most studies of energy demand in the population. Further, it is also important that measured variables are well defined, such as the standardised temperature and relative humidity measurements used in the Warm Front studies. The overall intent of collecting additional contextual information is so that differences can be accounted for when an outcome is systematically analysed and compared to other studies. An example of the importance of measurement in energy and buildings research is some recent evidence on the U-values of solid walled dwellings, which showed that the in situ measured values were a third lower than the normative assumptions and that only the distributions came near the standard values (Biddulph et al., 2014; Li et al., 2014). The risk, as Li shows is that the expectations of impact of solid wall insulation will result in a considerable shift in the energy rating of British dwellings.

Can the concept advance the study of energy demand?

While a number of studies accounted for difference examining a variable, the idea of using this information to develop an adjusted variable accounting for the difference has not been extended. However for the most part many putative influencing factors are identified within the studies that would be necessary conditions for extending comparisons to other studies.

4.2.3 Populations and sampling

The main intent of the epidemiological concept of a population is to characterise differences that would be seen within a specific target population and to describe the occurrence or trends of an outcome, condition or event. Undertaking studies with populations provides a means of generating generalizable results by accounting for individuals' differences in the population.

Is the concept appropriate?

There are a number of studies of energy demand (and related areas) that use populations of individuals to examine trends or factors that are associated with differences in an outcome or condition. These range from studies that look at groups with a common feature (e.g. Warm Front) and those that examine broader target populations, such as the study of the practices around home thermostats by Shipworth (2007). Studies that attempt to characterise or examine trends for a national level target population have been less frequent due to limits on empirical data availability.

In the UK, several studies have examined trends and differences using a sample that seeks to represent the national residential housing stock or households. Wyatt (2013) used a sample population from the National Energy Efficiency Database (NEED) to investigate the physical and socio-economic drivers of energy consumption and to assess the impact of energy efficiency on energy consumption in the English residential stock. The sample population was drawn from records held in a council tax database that covered all residential properties to represent English dwellings; a 2.5 million sample properties were matched to their energy meters for the years 2005 to 2008 (Wyatt, 2013). The matching process removed approximately 37% of the initial sample (mostly flats), which meant that the analysis was predominantly focuses houses. The analysis provided both 'crude' (i.e. unadjusted) energy demand rates (i.e. kWh/year) and also assessed the differences in demand due to putative influencing characteristics. The impact of energy efficiency measures was evaluated by comparing the change in energy demand for dwellings with efficiency retrofits against a 'control' group without retrofits. Although the concept of a case-control approach was valid, the populations sampled for 'like-for-like' comparison and consequently the results needed to be treated with caution, an issue that Wyatt acknowledged. The studies by Shipworth (2010; 2011) used a dataset of 358 households from an initial survey size of 1134 (selected for a spread of geographic and socio-demographic features), which were drawn to be representative of English households. Shipworth compared several physical dwelling characteristics against a national cross-sectional survey in order to examine differences that should be accounted for in the analysis comparing reported thermostat settings (Shipworth, 2011; Shipworth et al., 2010). As in the Wyatt study, Shipworth (2011) provides both 'crude' (i.e. unadjusted) reported thermostat settings and also those for putative influencing factors (i.e. dwelling type, age, region, efficiency features, type of thermostat control and its location).

Energy demand-related studies have also drawn samples to reflect particular groups or features present within a population. For example, the studies by Oreszczyn (2006a; 2006b) and Hong (2006; 2009) used a sample of approximately 1200-1600 dwellings from the Warm Front scheme, depending on the study, to assess the determinant factors associated with differences in indoor temperature, relative humidity, space heating, and thermal comfort. The samples were selected to represent the environment, housing types and climate of England where the Warm Front scheme was carried out. The studies provided description of the variables being examined (i.e. temperature, energy, mould growth risk) by influencing characteristics and also their distribution within the sample, though crude rates were not reported (Hong et al., 2009, 2006; Oreszczyn et al., 2006a, 2006b). In these Warm Front studies, the target population was not the national stock but rather those that received Warm Front assistance, i.e. households that applied and met the benefit eligibility criteria. The studies did not provide a comparison of the features examined within the sample to its target population (i.e. Warm Front), which meant that the findings were less easily generalizable to the wider Warm Front population (though there are expected to be similarities).

The above studies provide examples where the population approach has been applied to describe the features and differences in energy demand outcomes (and related variables) and also where findings from the sample were used to give insight into a larger target population. The concept is shown to be applicable to the study of energy demand, although its application is imperfect. A weakness with the current approach to the concept is that a number of studies did not formally compare the characteristics of the population sample to their target population, which meant it was unclear if the results were generalizable. In addition, while most of the above studies presented the variables being analysed by putative influencing factors and also crude values, they did not attempt to adjust these to account for these differences for wider comparison.

End-use energy demand research does not widely use the population as a basis for studying issues of energy demand. Although population-level studies exist commonly in social and economic analyses, they are less common in engineering or physics-based studies. The concept of developing comparable groups and seeking differences and associations in order to create testable hypothesis is under-utilised in energy demand research. This is attributable to the fact that energy demand research was originally seen as physics based study and is combined with the difficulty of accessing population-based energy data following privatisation of the utilities. This can mean that studies which seek to explain behaviours seen at the individual level (e.g. person, household, premises, building, etc.) lack a broader context within which to compare results, determine possible biases, and describe the target population (e.g. elderly households in old houses). For example, in their examination of the impact that more energy-efficient building practices had on energy demand in a selection of new housing developments in York, UK, Bell and Lowe (2000) compared the energy demand for spacing heating and hot water to mean values derived in other studies, including the national stock, all dwellings in York and a selection of other low-energy projects (Bell and Lowe, 2000). Ideally, the comparisons would have been made for equivalent types of developments as those examined in the study, or else the comparison could have used values adjusted for known influencing features such as dwelling size, age, climate and household features. However, at the time there was little information on energy demand in England that could be used to contextualise the impact that these building practices had.

An extension of the population approach is to make use of distributions. Distributions provide a means to express the spread and shape of the occurrence of outcomes, conditions or key factors of interest. In health, this approach suggests that those with health problems are part of a population with variation in the outcome of interest (i.e. from diseased to healthy). Phenomena related to energy demand can also be considered in a similar manner; for example, the uptake of low-energy light bulbs had for some time been very low (despite the obvious economic rationale) for a host of reasons related to perceived issues of quality, access, and desirability.

Is the basis of the epidemiological concept maintained?

The basis of the epidemiological concept of a population is: that differences in outcomes and the impact of influencing factors can be better understood by accounting for differences in population characteristics (Rose and Barker, 1978). Population samples should be representative of the target population so that inferences can be made, and should be of sufficient size to reduce the chance of error in the measured effect. As shown above, a number of energy demand studies at the population level have examined determinants of energy demand (or related) outcomes with the intent of determining the association of factors and a change in condition. However, the current limited application of the concept may be overcome as more information on energy demand in the population becomes available for comparison. Applying the concept of distributions to energy demand is not new but is certainly rare in practical research and policy. Using the distribution from a population provides a basis for investigating what contributing factors may lead to a given outcome in energy demand level and pattern over time.

Can the concept advance the study of energy demand?

Many European and advanced nation governments have committed to an extensive decarbonisation of the existing housing stock through improved efficiency of new and existing houses and through an almost zero carbon grid. Policies that aim to achieve these decarbonisation goals need to be applied to a complex building stock and diverse population. Such actions therefore require an understanding of the heterogeneous features and variation exhibited in the stock. Using studies based on population samples can provide a reasonable representation of the target population (e.g. national stock or households) for which to develop policies. Further, population studies provide a basis for contextualising small-scale intervention studies (i.e. in small and specific populations) and field trials that aim to support this decarbonisation programme.

4.2.4 Incidence, prevalence and variation

Epidemiological studies examine the occurrence of cases using defined and measured variables from representative samples of identified target populations, in order to move towards understanding aetiology for a particular outcome. The occurrence or frequency of cases among the population provides the means of examining how putative influencing factors affect outcomes. Prevalence studies provide the contextual background for more detailed studies, without which it is difficult to determine the level of the problem being investigated.

Is the concept appropriate?

Several energy demand studies examined the prevalence of conditions among the population. Shipworth's (2011) examination of set-point temperatures aimed to determine what household practices were with respect to the set-point of their thermostat during winter heating periods. Their analysis provided a basis for comparison to an earlier prevalence study that examined the same phenomena. Strictly speaking, the Shipworth study did not give a measure of prevalence (e.g. *period prevalence rate = all cases/ population at risk*) but instead provided a frequency occurrence among the sample survey. The variables can be converted into estimates of period prevalence using the frequency of the thermostat cases divided by the average population who are eligible to experience the condition (i.e. all dwellings in the target population). In a study by Healy & Clinch (2004) examining the occurrence of households that spend over 10% of their income on home heating, they make use of survey data to determine the percentage of cases from the population at risk. From a cross-sectional survey, they estimate that 17.4% of Irish households were reported as being in some level of fuel poverty, which equates to approximately 227,000 households (Healy and Clinch, 2004). Determining the prevalence among the household population of Ireland then allowed for deeper analysis to understand the association between reported fuel poverty and other potential risk factors. For example, 39.2% of fuel-poor households experienced condensation and 60% of those households had some difficulties heating their home adequately. For studies that focus directly on residential energy demand (or fuel use) the use of the prevalence rate is less frequent. In Wyatt (2013) the sample was drawn to be representative of English dwellings, but there was no formal comparison of the sample to the target and little data was available for the comparison of energy demand to the target population.

Incidence is more readily applied to features that act on energy demand, such as the installation of energy efficiency retrofits or frequency of different levels of demand. Incidence studies are typically used to examine whether an outcome of interest is changing over time. Several studies have examined the change in energy demand across years, but most do not consider the change beyond the population mean and do not attempt to examine the change in the frequency or distribution of the occurrence. In a study on the determinants of residential space heating expenditure in Great Britain, Meier and Rehdanz (2010) provided price and income elasticity estimates for heating fuel demand over the period 1991 to 2005. Although they were able to identify factors that were associated with changes in demand (focusing on owners and renters) such as the type of dwelling and the size and number of occupants, they did not give any figures on how different classes of energy users have changed during that period (Meier and Rehdanz, 2010). The Summerfield et al. (2007) study looking at Milton Keynes Energy Park used incidence of the change in different levels of energy demand (low, medium and high) to determine an association between demand levels with dwelling and household characteristics. The study provided a good example of how the change in a factor of interest (energy demand) and its relationship with putative factors can offer a broader insight into the drivers of change.

Is the basis of the epidemiological concept maintained?

The concept of prevalence is certainly used in energy demand studies, but not always formally. However, the need to identify severity of conditions among a population of interest and to contextualise case studies means that the concept is increasingly being used in the literature. By comparison, the study of incidence and its use to examine the frequency and distribution of 'new' energy demand, change in demand or related conditions is not widespread. This may in part be due to the lack of available longitudinal data. With policies seeking to improve the energy performance of dwellings and reduce demand, it will be important to examine the change in demand (i.e. incidence) and its relationship with the introduction of retrofits, for example to determine the spread of the change across the population, along with the size of the shift.

Can the concept advance the study of energy demand?

Regarding the epidemiological basis, both prevalence and incidence are attempting to describe the occurrence and frequency of some condition of interest and the factors that are associated with it among an identified population 'at risk' – or population who could experience the condition. This concept is presented in the Healy and Clinch (2004) study of fuel poverty in Ireland. It is also present in the study by Summerfield et al. (2007). However, the other studies discussed do not explicitly make links to the 'at risk' population. By identifying this population, the findings and associations can be made more generalizable but also their limitations can be made more explicit. Further, with an examination of the incidence it is difficult to identify where the condition occurs and whom the problem most affects. This is of particular importance for studies concerned with changes in high and low users of energy over time and the relationship of putative factors associated with the condition. The related concepts of prevalence and incidence can advance the study of energy demand by providing more contextual information about a condition of interest and also how it changes over time. Without measures of incidence under an epidemiological approach, it would not be possible to move towards understanding aetiology (i.e. causes of a condition) and factors that affect the outcome.

4.2.5 Putative influencing (risk) factors

Putative influencing factors, or 'risk' factors, are those features that are not only strongly associated with an outcome, condition or occurrence of an event among the population, after adjustment for potential confounds, but may comprise components of the 'causal pathway' (Greenland et al., 2004). Identifying those putative influencing factors provides a route through which to prevent or control the condition of interest.

Is the concept appropriate?

The vast majority of energy demand studies of populations consider 'explanatory' variables and the degree to which they predict an outcome. These variables are generally selected on the basis that they are expected to have some influence on the measured outcome variable. For example, the Hong et al (2006) study of the impact of energy efficiency on space heating demand in Warm Front dwellings examined the influence that putative factors would have on the change in modelled energy demand. The factors considered to be likely to have an effect (i.e. in the causal pathway) were dwelling age, type, insulation level, primary heating and standard assessment procedure (SAP) rating (Hong et al., 2006). In Tovar's (2012) study of expenditure on energy efficiency in English dwellings a number of factors putatively related to expenditure on retrofits were tested, for example that low-income households with dependent children living in privately rented dwellings were less likely to adopt measures. The statistical approach that Tovar took to examining the factors that are associated with different levels of energy efficiency is not uncommon in energy demand studies (especially those with an economic approach). In the Milton Keynes Energy Park study, Summerfield et al (2006) examined factors that were considered important in determining changes in temperature within the sample and also between the two points in time being examined.

Is the basis of the epidemiological concept maintained?

Risk factors are currently applied to energy demand studies as explanatory variables, which may be statistically tested as in Tovar (2012) or descriptive as in Summerfield et al. (2006). The purpose is to identify factors that are antecedents of (i.e. that precede) the condition or outcome. For studies that look at interventions, as in Hong et al.'s (2006) study of retrofits and space heating, it is more clear that the intervention is within the causal pathway towards changing the energy used for heating. However, when studies are attempting to examine a number of factors, such as in Tovar (2013), it needs to be clear that the factors being considered are part of the causal pathway or are suspected to act on the outcome of interest and are not artefacts of the data or statistical associations. For the most part, the concept of risk factors is applied in population energy demand studies and do attempt to determine the attributable proportion when there are a number of potential factors acting on the outcome of interest.

Can the concept advance the study of energy demand?

The concept of risk factors could be more widely applied in order to reduce potential errors of 'causal' association. This more cautious approach would emphasise the use of causal pathways to identify those putative influencing factors that could be tested in statistical and analytical terms. By extending the concept of attributable factors that act on the condition, mechanisms aimed at controlling or alleviating the condition could be more effectively employed. For example, in the Warm Front study examining the impact of interventions on wintertime temperatures (Oreszczyn et al. 2006), an increase in temperatures was observed for those dwellings that were most inefficient prior to the retrofit. One potential risk factor of energy savings is the desire of the occupants to raise the temperature following a heating retrofit in order to improve comfort. A further benefit to including putative influencing factors is the preceding analytical process of establishing the causal pathway, which is an essential part of understanding whether a factor has a real effect on the outcome.

4.2.6 Causality and causal pathways

The methods used in epidemiology are primarily focused on observing and quantifying associations. Epidemiology is not able to demonstrate disease mechanisms, but rather infers the presence of these mechanisms and relies on the work in related disciplines to identify and describe them in detail (Joffe et al., 2012). It does this by acknowledging the interplay of the ‘host, agent and environment’ that interact to form the components of the causal model, i.e. the components or factors that interact jointly to determine an effect.

Is the concept appropriate?

The causality concept is certainly presented in population-based energy demand studies; often the term itself is referred to in building towards a description of the variables of interest, a statement of the findings or as part of the interpretation process. In a study of the spectrum of residential energy consumption in the US, Kaza (2010) examined the factors that were expected to have an effect on residential energy consumption and explained the mechanisms that were known to have an effect on levels of energy use but cautioned against assuming their causal nature. The association’s between higher levels of energy demand and a number of tested putative explanatory variables are discussed that could contribute to different levels of demand (e.g. more appliances or higher incomes) (Kaza, 2010). This cautious use of the term causality is seen in a number of population-level studies (Cayla et al., 2011; Moran et al., 2012). For the most part, these types of study identify associations that provide means of generating hypotheses for other studies that will examine the theorised causal mechanisms. For example, in the description of the party-wall thermal bypass related to the interpretation of buildings regulations, Lowe et al (2007) identified a causal factor that increased the dwelling heat losses. Had a study been undertaken looking at a population of terraced dwelling types it might have found an association between higher space heating energy demand and a number of factors that are known to have an effect on demand. One of those factors could have shown a positive association of new build and higher demand (compared to some other groups) that reflected party-wall bypass.

Implicit in most population studies that examine factors that have an effect on energy demand is that they include a number of ‘causal’ components that act on a given outcome. However, it is not always clear that these studies maintain the causal concept. A concept that would offer considerable benefit to the study of variation in energy demand is that of multi-causality, which sets out that a condition can be caused by a number of mechanisms. This is because there is often no single factor driving energy demand in buildings. Multi-causality means that most individual causal factors are not sufficient or necessary to produce an outcome. It then follows that not every component must be addressed or identified to affect the condition. For example, the condition of low energy demand for space heating may be caused by a number of interacting factors, such as low fabric heat loss, high performance heating system, low thermostat settings, low fuel expenditure, and so on. By altering one of those factors, it would be possible to have an effect on the low energy demand. In the Warm Front study looking at the impact of energy efficiency retrofits on space heating fuel use, Hong et al (2006) showed that the heating fuel demand reduced by only 1% (118 to 117 Wh/K/m^2/day). The objective of the Warm Front scheme was to reduce heat fuel expenditure (rather than energy demand) for ‘vulnerable’ households (i.e. those on a number of priority benefits) by installing a range of retrofits. The study suggests that poor energy efficiency, while certainly a ‘cause’ of improving energy efficiency and therefore demand, was not a sufficient cause, but rather a component of a number of inter-related factors that acted on demand. In a related study, Oreszcyzn et al (2006) identified that there was an increase in internal wintertime temperature following a number of retrofits, which they hypothesised reflected the low comfort levels and therefore a 'temperature take-back' effect (Oreszczyn et al., 2006a). Hong et al (2009) concluded that this temperature take-back in post-intervention Warm Front dwellings reflected improved comfort conditions (Hong et al., 2009).

The causal pathway (or causal model) is an important concept that is rarely used explicitly in energy demand studies. The concept provides a framework that sets out how those putative influencing and interacting factors can inform the analysis and interpretation of results. As in the work around the Warm Front study, the interacting causes of changes in energy demand, indoor temperature and moisture conditions and comfort could be drawn into a pathway or model framework to provide a better understanding of how changing components, such as energy efficiency, might contribute to reducing heating expenditure. Figure 10 shows this causal system as applied to the Warm Front interventions.

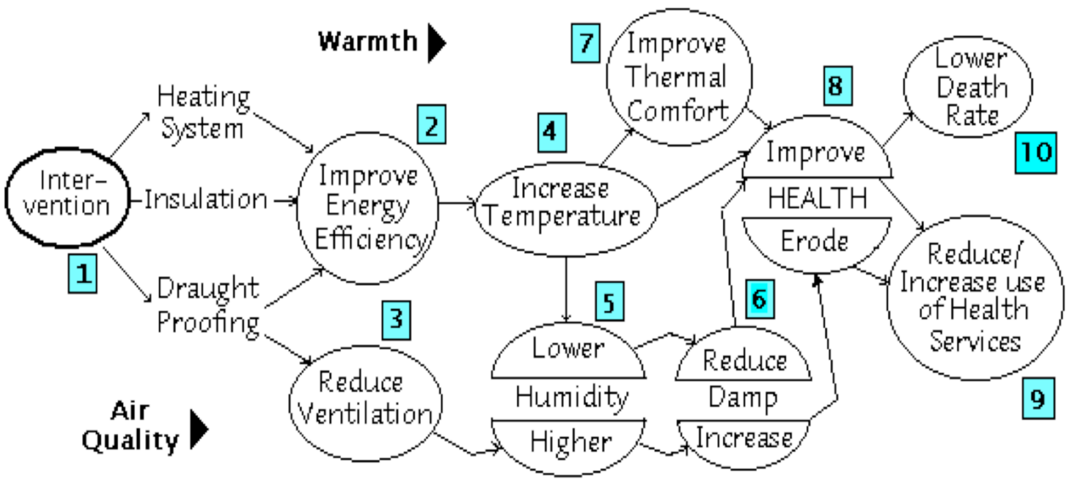


Figure 10 - Warm Front pathway diagram showing relationship between energy efficiency retrofit and health outcomes (Wilkinson *et al.* 2008)

Is the basis of the epidemiological concept maintained?

The basis of the concept of a causal factor (sufficient, necessary and component) is that the factor (or more likely a collection of factors) act to produce an effect. There is much debate in the epidemiological world around how to define and illustrate the concept of causality. Hill's criteria set out what might help to 'prove' causality through analysis (i.e. Hill's criteria: strength, consistency, specificity, temporality, gradient, plausibility, coherence, experimental evidence and analogy) but these are only given as steps towards causality. Energy demand studies of the population should continue to identify and test associations of factors thought to have an impact on energy demand and work towards seeking to establish causality through more detailed studies.

Can the concept advance the study of energy demand?

Applying the concept of causality to the study of end-use energy demand can provide the opportunity to study and describe the key mechanisms related to conditions around demand and the determinants associated with levels of use. Through the recognition of conditions (or 'disease') and identifying the multitude of factors that act on energy demand, and by describing in further detail the causal pathways that lead to an outcome, it is logical to move towards treatments. For example, it would be possible to identify risk factors to energy savings for an energy efficiency refurbishment or education programme. Further, by having sufficiently large populations for study it would become possible to identify relevant populations (e.g. people, households, buildings, etc) to assess and determine component causes of conditions and the effect of interventions to establish effective treatments.

4.2.7 Determining association

In epidemiological studies, association is simply the relationship (usually statistical) between the occurrence of an event or characteristic that varies along with the occurrence of other events or characteristics (Porta et al., 2008). As with other fields of study, an association does not necessarily imply a causal relationship. Epidemiological associations are primarily measured through the *relative risk ratio*, *odds ratio*, and *rate ratio*; the use of each depends on the study design and the outcome under examination. These measures can be extended to examine the *attributable risk*, or the risk that is associated with a given exposure or set of factors.

Is the concept appropriate?

In population-level energy demand studies, the outcome of interest is generally the fuel demand for given end uses. As discussed above, the concepts of prevalence and incidence are not widely used, although there are examples of each in energy demand. However, the aim of determining the relationship between supposed influencing factors and outcomes of interest is common in studies of energy demand. In Meier and Rehdanz's (2010) examination of the determinants of space heating expenditure in British households, a number of social and dwelling characteristics were tested for their relationship to fuel use. For example, they found strong associations between total fuel expenditure and household income and that expenditure (per room) decreases with more rooms, but increases with more people (Meier and Rehdanz, 2010). One particular issue they identify was that heating expenditure increases with age until a peak of around 68 to 79 years, when it then declines. They speculate that this could be related to older persons (more likely retired) occupying and heating fewer rooms, or to the fact that they may be more likely to be in fuel poverty (although their study finds that the average expenditure for these households on heating fuel is ~4% of annual income). It is not possible to ascribe a causal relationship between age and heating fuel expenditure (i.e. that being old causes lower expenditure) without further study. It is possible, as Meier and Rehdanz have shown, to present the relationships and generate a number of hypotheses for further examination.

Measures of association in energy demand (and related) studies where 'incidence' is of interest sometimes make use of odds ratios in order to determine the likelihood of particular outcomes. In a study examining the link between demographics and barriers to adopting energy efficiency measures in UK households, Pelenur and Cruickshank (2012) used the odds ratio to determine the strength of association between adoption and selected barriers. Using the odds ratio, they showed that, for example, marital status was strongly associated with barriers to adoption of energy efficiency related to the landlords and also the characteristics of the property. More specifically they determined that single households were at least twice more likely to cite that their 'landlord' was a barrier to adoption than married households (Pelenur and Cruickshank, 2012). Rather strangely, although they used the odds ratio of association to determine the strength, they do not show the confidence intervals, thus making it difficult to determine the uncertainty of the estimate. In a study examining energy use and appliance ownership in Ireland, Leahy and Lyons (2010) determined the association, and its strength, of the incidence that a household would own an appliance versus a reference category (also with no confidence intervals). For example, they determined that urban households were 1.25 times more likely to have double-glazing than rural households, that renting households were 52% less likely to have double glazing than owner-occupied dwellings, and that dwellings built post-2000 were 9.5 times more likely to have double glazing than those building between 1918 and 1960 (Leahy and Lyons, 2010). The odds ratios in this study were used to describe the incidence of the ownership and were used to inform the model development. When used in this way, the odds ratio provides a means of examining the likelihood of the incidence among a population. In Oreszczyn et al. (2006), odds ratios were used to examine a dwelling's likelihood of having a mould severity index >1 (i.e. moderate and above). For example, they identified that dwellings in the highest standardised relative humidity quartile (compared to the lowest quartile) had an OR of 3.57 (95% CI 1.84, 6.94) of having a mould severity index >1 (Oreszczyn et al., 2006b). In this study, the odds ratios were also used to examine the incidence of mould severity and a number of contributing factors, such as energy performance, energy efficiency retrofits and humidity of the dwelling on cold days. In Oreszczyn et al. (2006) the odds ratios were accompanied with sufficient information to determine the strength of association along with their significance. In addition, the ORs were adjusted for year and location. Broadly, the use of odds ratios is not particularly common in energy demand studies, but there is potential for its use when examining both prevalence and incidence. Measures such as relative risk have not been found in energy demand studies.

Is the basis of the epidemiological concept maintained?

Most energy demand studies do not use commonly applied epidemiological measures of association, such as relative risk ratio or odds ratio, for the simple reason that most studies are looking at continuous variables, e.g. energy use. Often there is no particular reason to bin a continuous outcome such as fuel demand for the reason that the unit increase is sufficiently large (i.e. kWh) and meaningful for interpretation. This often implies the use of regression techniques in examining the association between the energy demand outcome and a selection of influencing factors, as seen in most of the energy demand studies discussed thus far. However, there is recognition that a household or dwelling's relative placement along the energy demand distribution is itself an important consideration in determining the association of energy demand with putative influencing factors. In a study using a cross-sectional survey of US residential energy use, Kaza (2013) examined the differential effect of the relationship between selection of dwelling and household features and energy demand. Kaza found that the marginal effects across the quantile distribution provide a more refined examination than an ordinary least squares linear regression approach (Kaza, 2010). For example, dwellings in the upper quantile (i.e. 80th percentile and above) of heating demand reduced demand by 3 MWh/yr for every 20 years of dwelling age, compared to only 0.4 MWh/yr for the lowest quantile of energy demand (i.e. 20th percentile and below).

The concepts of relative risk and the odds ratio are certainly appropriate to energy demand, but depend very much on the study design and its aim. Most studies focus on the absolute change in fuel use across the stock and therefore concentrate on various regression techniques that complement continuous variables. Where the outcome of interest is an incidence or some form of change between discrete, or appropriately discretised continuous, measurements, the use of odds ratios is appropriate.

Can the concept advance the study of energy demand?

Regardless of which measure of association is used, the primary interest is to explore these associations within the framework of causal thinking (Bhopal, 2008). What is important is that the relationships examined and the strength of the association found is not mistaken for causality. Measures such as relative risk and odds ratio are often associated with more guarded language around their meaning. Their use may have advantages in certain circumstances, for example where not all putative factors are known and therefore the study of associations of any selected factor is exploratory. Most population-level energy demand studies will suffer from incomplete knowledge of the factors around which fuels are used and must rely on proxy or surrogate measures. For example, dwelling age is often used to represent energy efficiency because in practice dwelling building standards have targeted thermal and ventilation performance and heating system efficiency over the past few decades. Household income might be used as a proxy for the ability to spend on heating fuel. For these reasons, it may be preferable to have a measure of association that does not rely on absolute levels, but instead compares the association of levels of energy demand against different levels of the putative factors. Where the outcome of interest is incidence or prevalence among a population, there could also be interest in the effect that having certain characteristics have on the likelihood of the incidence or its distribution among the population, for which relative risks or odds ratios are well suited. An important focus of any measure of association however is whether the effect is simply due to chance, or if it is the result of some form of bias or confounding.

4.2.8 Bias and confounding

Epidemiological studies are concerned with the 'in situ' or 'real world' occurrence of disease, rather than the controlled environment of a laboratory – although laboratory testing is invaluable to understanding disease mechanisms. A true and statistically sound relationship between an exposure and an outcome is the product of internally valid (i.e. lacking bias or confounders) and externally valid (i.e. generalizable and replicable) studies.

Is the concept appropriate?

The term bias is commonly used in population energy demand studies of the residential sector. Leahy and Lyons (2010) refer to potential bias in understanding determinants of fuel demand in Irish dwellings when appliance ownership is not taken into account. This might be considered a form of measurement error, rather than a specific form of bias, for the reason that the condition (i.e. level of fuels used) can be further understood by modifying factors such as number of appliances owned. In making a case for using quantile regression techniques to study determinants of residential fuel use in the US, Kaza (2013) referred to those studies that focused on particular types of households (i.e. high energy users) as being affected by sampling bias. In making the case for further research and evidence requirements to support the decarbonisation of the built environment, Skea (2012) refers to the problem of selection bias in findings that are related to participation in energy efficiency retrofit programmes (Skea, 2012). An example is provided by findings from the Warm Front scheme that focuses on households on various forms of benefits and the generalizability of those relationships found to all other English households. In the studies by Oreszczyn et al. (2006a, 2006b) and Hong (2006, 2009), the sample being analysed consisted of homes drawn from the Warm Front scheme that received measures over the winters of 2001-02 and 2002-03. The sample was of those dwelling receiving Warm Front measures during that period, and the target population was of all those households that were eligible for Warm Front measures by virtue of being on benefits. Therefore, these findings are not necessarily generalizable to English households without further comparison. A different form of bias was referred to by Schweber and Leiringer (2012) related to the reporting of certain results, known as publication bias, which is a form of selection bias that affects the generalizability of the findings (Schweber and Leiringer, 2012). The concept of bias is certainly applicable in population-level energy demand studies but it could be more formally discussed in designing and undertaking analysis. This may be through a more thorough description and comparison of the study sample against the target population and being explicit as to how the findings are generalizable. An example of this is seen in Shipworth et al paper (2010) examining and comparing the thermostat settings in English houses between two survey years and the English Housing Survey (Shipworth et al., 2010).

Confounders (i.e. distortion due to another 'hidden' effect) in energy demand study are occasionally identified as such, although they can sometimes be confused for 'effect modifiers'. In a study that evaluated the targeting of the Warm Homes scheme that provided retrofits to owner-occupied or privately rented dwellings in Northern Ireland, Walker et al (2013) controlled for the potential confounding effect related to areas with high levels of social housing. Including these areas could have resulted in an incorrect under-targeting in areas with high a proportion of (socially) rented dwellings (Walker et al., 2013).

Effect modification occurs when a putative influencing factor related to the outcome of interest varies across levels of a group. An example would be where household income and per capita energy demand are being investigated and higher incomes are shown to use more energy. However it is necessary to control for the type of dwelling these households occupy (e.g. perhaps high-income households are more likely to occupy detached properties) and also the age of the property (e.g. perhaps more affluent areas are characterised by older homes). The factors of both the type and age of the property would interact to result in differences in the relationship between income and energy demand.

Is the basis of the epidemiological concept maintained?

The basis of the epidemiological concepts of bias and confounding relate to the need to ensure that any associations or putative causal relationships found are real and not the product of some form of error. It is unlikely that any energy demand study would knowingly ignore bias; however it is possible that because the evidence is not subject to the same rigorous reviews as in epidemiology they could occur unknowingly.

Can the concept advance the study of energy demand?

By recognizing various forms of bias and confounding in energy demand studies it would be possible to improve the certainty around the associations found and, in particular, the generalizability of the study findings beyond the sample. For those studies that make comparisons between different sample populations, examining potential confounders that could cause spurious results will improve the internal validity of the study. In doing so, the findings from such studies will be more open to being reproduced and can provide a much stronger evidence base for policymakers. At the core of mitigating bias and confounds is clear study designs and also the collection, analysis and interpretation of data. Having a strong foundation for data collection methods and resources for energy demand will assist these investigations by setting out guidance on how to increase the validity of studies.

4.2.9 Data and data collection

In health research, a goal of data is collection is to develop datasets that are comparable and avoid sources of bias. Routine data, i.e. from wide-sale standard collection processes, can be used to study prevalence and to pose hypotheses on which more complex study designs are based. Associations and potential causal factors are studied through specialised designs and must address sample population, sample size, variables, and ethical considerations to name a few. Collecting data over an extended period provides the opportunity to undertake longitudinal health studies, an important element in tracking changes in disease patterns and evaluating policy.

Is the concept appropriate and applicable?

There has been a chronic lack of access to good quality energy and building data (Summerfield and Lowe, 2012). In the UK, the empirical collection of data on building characteristics and energy demand has historically been ad-hoc or subject to interruptions and there has been little tradition of reporting data in a formal sense, thus undermining any concerted advances in research. For example, the Energy Follow-Up Survey (EFUS) was performed in 1991, 1996, in 2001 and most recently in 2011. While the 1996 provided a sample of approximately 2300 dwellings to represent the 22 million English dwelling stock for an in-depth collection of fuel use and factors thought to affect demand (e.g. temperatures and heating patterns), the subsequent 2001 survey was never publicly released due to problems with the sampling methodology. Despite the premise of the follow-up studies being to determine the relationship between dwelling energy performance and energy technologies and energy use, the surveys have used very different measurement techniques that make longitudinal comparison very difficult. The US, by comparison, has undertaken 13 cross-sectional surveys of national residential energy consumption from 1978 to 2009. There have been several other monitoring studies of residential energy use in the UK that were collected to be representative of English houses. The Carbon Reduction in Buildings (CaRB) Home Energy Survey (HES) sought to survey national home energy use from a sample of 427 households (Shipworth et al., 2010). The UK's statistical releases of sub-regional gas and electricity supplier meter data provide a form of registry data, which can be used in ecological style studies. The collection of different forms of energy demand data for population groups occurs, but is not systematic.

Is the basis of the epidemiological concept maintained?

Routine data collection provides the basis for drawing samples for further epidemiological studies. It also provides a means for contextualising studies to help improve study designs and possible confounding or modifying factors. Recently in the UK, the Department of Energy and Climate Change (DECC) have begun to link various energy, building and socio-economic datasets. The National Energy Efficiency Data-framework (NEED) links annualised meter data for gas and electricity from energy suppliers, administrative data on dwelling characteristics, information on energy efficiency retrofits undertaken, and also a number of modelled socio-economic variables (DECC, 2012d). This data framework differs from health registries of (for example) death records in that the underlying datasets are not ordinarily available for researchers to access. It is also different in that many disease or death record registries will use a classification system that has been developed to be systematic in its definitions.

Can the concept advance the study of energy demand?

When data is collected on an intermittent or inconsistent basis it can lack key features that allow for cross-comparison or provide the means to contextualise smaller bespoke study designs. The overall effect of this lack of data collection has meant that models that attempt to describe energy demand in buildings have been severely limited and often rely on unconfirmed theory rather than empirical observations. More widely available data on the actual use of fuels among the UK dwellings and households would provide a basis to examine trends and differences in energy use based on empirical information. It would also provide a means for linking smaller and more bespoke studies and to contextualise others.

At present it is difficult to make generalizable findings from studies that focus on energy demand in dwellings because there is only limited data available to describe the target (i.e. wider) population.

4.2.10 Describing results

In epidemiological studies, results are often drawn from studies of comparison between populations that have been subject to some exposure, treatment or difference in characteristics. For these results to be accepted, they are presented with in-depth descriptions of the data and the collection process and possible sources of bias, and a range of statistical tests is applied to assess quality of the data and result and thus the validity of the study. The formalised process of presenting and discussing results provides the opportunity for researchers elsewhere to accept or refute the results and equally important, to undertake reviews of the data in a systematic manner. Further, systematic reviews provide a means of assessing the validity of studies pertaining to an issue or disease. Doing so requires that quantitative results be presented along with measures of uncertainty (e.g. confidence intervals) in order to judge precision. The description of the study population, the development of the causal pathway and inclusion of important risk factors, along with a description of and control of potential confounders are required to assess the soundness of studies. The systematic review process requires a focused question, selection criteria for inclusion and exclusion, appraising the quality of the findings, synthesis of the results, and disseminating the conclusions (Oxman, 1994). The benefits of a well-defined systematic review are to provide a method of appraising and interpreting evidence and presenting it to a wider audience in a clear and accessible format.

End-use energy research rarely offers the same degree of formalised results. There are however common outputs from studies that use a defined metric and assure the validity of any statistical models developed. However these details could be further structured to create consistent summaries of evidence, i.e. showing the robustness of results expressed with confidence intervals along with population parameters to avoid bias and misinterpretation. Presenting results in this manner would allow for the evidence from energy demand studies results to be more effectively judged and assessed. This means thoroughly describing the data and its source and summarising the data so that results can be reviewed. Several authors provide a useful hierarchical framework for quantitative and qualitative studies in epidemiology that can usefully be applied to energy, with the lowest order being single-case studies, through descriptive, conceptual and, the most important, generalizable studies (Daly et al., 2007; Oxman, 1994). The concept of the systematic review has already been extended to energy demand and is described as being able to improve the evidence base for developing energy policy by improving the definition of key issues, increasing the number of studies considered, identifying gaps and pinpointing uncorroborated results (Sorrell, 2006). However Sorrell (2006) also notes that some forms of systematic review may be too narrow and not sufficiently flexible to address some of the complex problems and policies involved in studies of energy demand (Sorrell, 2006). The combined effect of a weak methodological framework and lack of good quality data has meant there is little opportunity for systematic reviews of research findings, and little basis for project-by-project learning. This is in stark contrast with the systematic reviews undertaken to assess the strength of evidence on diseases used to inform policymakers and thus improve health.

4.3 Conceptual framework for energy epidemiology

The epidemiological approach in medicine has evolved through systematic application and evaluation undertaken by a wide number of disciplines working towards a better understanding of disease for the improvement of health. Epidemiological concepts have co-evolved to address the study of detailed mechanisms and practices and their manifestation among populations. Epidemiology is also well suited to dealing with uncertainty through the use of methodological tools and analysis techniques that examine, support and strengthen those concepts. These methodological tools include: common definitions and metrics, population selection techniques, study designs for data collection, comparison and analysis, approaches to dealing with bias and confounding factors, guidelines for working towards identifying causal relationships, and systematic approaches to reviewing evidence. These tools also support a longitudinal focus that provides a means of examining the trends and changes in health phenomena using data collected over many years in a standardised manner to determine response functions, examine drivers and evaluate past policy or practices. There is also a strong methodological tradition for large-scale studies and the relationship between the clinical level and population level, with a great deal of effort employed to ensure that findings can be relevant outside of the individual case. These concepts and tools are used for a wide range of theories and hypotheses in working towards improving population health and controlling health phenomena. Epidemiology has made significant strides in an evidence-based approach to problem-solving and decision-making (Jenicek, 1997). In doing so, the epidemiological, medical, and policy-making professions have had to come together to prioritize health research, examine and scrutinize the findings, disseminate them widely and openly, and implement programmes through well-informed policy.

The main questions being examined in this section of the thesis are: whether the epidemiological approach is appropriate and applicable to the study of energy demand at a population scale; whether the basis of the epidemiological concepts are maintained; and whether the application of those concepts can advance to studies of population-scale energy demand.

In the previous sections, a number of key epidemiological concepts were tested against these criteria by examining where and how the current energy demand research community is applying them. The findings suggest that, by and large, these concepts are applicable and appropriate. Many concepts such as outcomes, population selection, influencing factors and bias are commonly applied in energy demand studies. Several other concepts, such as incidence and prevalence, causal pathways and confounding, are being applied but only in a limited form. The examination has suggested that many of the concepts found in medical epidemiology are applicable to the study of energy demand due to its catholic (i.e. wide-ranging) nature. Several concepts are directly applicable, such as drawing population samples and examining outcomes and measuring association. One test for determining whether these concepts are applicable to energy demand is whether they are appropriate. It is proposed here that indeed they are, as shown by their existing application.

A second test was whether the concepts are able to maintain their medical epidemiological basis, i.e. whether a concept can mean the same thing when studying energy demand. This judgement is more challenging for the reason that it relies on an interpretation of the concept as it is used in the study of energy demand. For the most part, the concepts examined here have evolved from a long practice of describing, comparing and testing differences and changes in outcomes of populations. For the most part, disease mechanisms and the factors that create them are still not fully understood. In part this is because the sheer variety is overwhelming and subject to biological, environmental, genetic, cultural, and societal changes and interactions that results in constant changes in disease itself. However, this might imply that there is no analytical approach to solving such problems. In studies of energy demand that relate to populations, there are many of the same limitations to developing a deeper understanding. Although the mechanisms and consumption of energy are, for the most part, the work of people, the structures, their systems and their operational life have strong parallels with disease and health outcomes. As with medical epidemiology, the mechanisms and factors that interact to create demands for energy are still not fully understood. The epidemiological concepts examined here are primarily focused on understanding how mechanisms and outcomes are manifested among populations. Achieving clarity through understanding the mechanisms that result in the demand for energy among dwellings and households that can assist in meeting the objectives of a low-carbon society. In the discussion of the selected concepts, many of the studies provided did offer, at least to some degree, a hint of what the epidemiological basis sought to answer. There were some concepts such as disease itself, that are difficult to resolve but nonetheless can be in part applied to the study of energy demand (e.g. the parallel between obesity and high levels of energy demand). In examining these concepts as they are currently being applied it is clear that for the most part the epidemiological basis is being maintained. This means that the motivation and objective behind the concept as applied in the energy demand literature is consistent with the intent behind the concept as applied in medical epidemiology.

The final test was whether the approach could be used to advance the study of energy demand at a population level. In this context it is proposed that many of these concepts, being already applied in a limited manner, could be expanded so as to found a more robust approach to population-level research. Examining energy demand phenomena at the population level requires an understanding of the technical details of the mechanisms within the energy demand system, the practices related to how energy is used, and the environment within which those practices and physical and engineered systems exist. The main concepts and tools for describing epidemiological outcomes would assist in this understanding of energy demand and would therefore be useful in advancing the scientific basis of the disciplines working in this area.

It is proposed here that, through the examination and discussion of the existing use of epidemiological concepts in energy demand research, a more formal adoption of the approach would contribute to the objectives of the paradigm of population-level energy demand study. Specifically, the application of epidemiological concepts can provide a means of understanding the distribution and differences of energy demand phenomena among a population. Through the application of a methodological framework that supports these concepts, the drivers and factors that create differences among the population can be identified, examined and better understood. Also, the approach can provide insight into how those drivers and factors can be manipulated to manage and control phenomena in order to improve quality of life and access to energy, and manage the transition to a low-carbon society.

Epidemiology and its concepts and tools are recognized as having a wider potential use than solely within the domain of the traditional public health field. In the introduction to the *Dictionary of Epidemiology*, Porta et al. (2008) describes how the concepts developed and refined within epidemiology are providing a lens through which to examine problems in other related fields, saying:

> Today research methods with strong epidemiological roots and properties are fruitfully applied 'within' and 'outside' epidemiology. A positive blurring of the boundaries of epidemiological research methods occurred in the last decades of the last century; e.g., the integration of population thinking and group comparison into clinical and public health research. The expansion of this influence toward other research areas remains a significant—and in my view highly attractive—challenge for many scientists. […] Today epidemiological thinking continues to create new approaches, research designs, strategies of analysis, and ways to assess causality for [other health] disciplines. Thus the influence of epidemiology continues. (Porta, 2008, pp. xi-xii)

Within the energy demand sector, there are several who have begun to lay the foundations for an epidemiological approach to energy demand studies. Summerfield et al (2005) provided an initial case for the use of an epidemiological approach in the analysis of longitudinal energy studies. Through the examination of the long-term patterns over the life-course of many buildings, a picture of how differences in design, construction, operation, maintenance and alterations can provide insight into how different factors can have effects on energy use (Summerfield et al., 2005). Sorrell (2007) has described how an evidence-based policy approach would provide a much needed level of rigour in developing energy policies that are driven by an understanding of the determinants of an outcome (Sorrell, 2006). In particular, epidemiology's strong methodological foundation and study designs support the gathering and assessment of evidence needed for such an approach.

Overall, this suggests that the epidemiological approach is adept at addressing issues of scale, definition, and viewpoints, which are greatly needed in end-use energy demand research.

4.3.1 Aims and principles of an energy epidemiology approach

Drawing from the proposed paradigm described earlier, the over-arching aim or purpose of empirically-based population-level research into end-use energy demand is to improve the understanding of variation and causes of difference among the energy-consuming population. The epidemiological approach provides an established set of concepts that have evolved through practice and debate for undertaking population-level studies.

The primary aim, therefore, of an energy epidemiology approach is to investigate causes and effects of key factors on energy demand within a population or sub-populations. It should consider the complex interactions between the physical and engineered systems, socio-economic conditions, and individual interactions and practices. It should also offer a description of the broader context and provide an environment within which individual studies can be contextualised and systematically assessed.

The transfer of the epidemiological approach to energy demand is not direct, but rather an adaption of those tools and methods that can best serve the study of the complex interactions between behavioural, physical, and environmental factors leading to an energy demand outcome. In meeting the primary aim of an epidemiological approach to energy demand, a number of objectives should be set out. In keeping with the basic approach of epidemiology (Rose and Barker, 1978), the objectives of an epidemiological study of energy demand are to:

i) Describe and measure the distributions of variable(s) of interest, e.g. energy demand per unit of observation;

ii) Explain the distribution by its determinant factors: physical, environmental, social, behavioural and economic;

iii) Support models that predict the changes expected in the distribution due to interventions, particularly energy efficiency and occupant practice measures; and

iv) Provide an evidence basis for informing policy related to end-use energy demand.

In the proposed epidemiological approach to studying energy demand, the emphasis is on understanding the occurrence of a phenomenon (i.e. end-use energy demand) among a population. The concept of a population in this context is described in further detail below, but the main concept is one that draws together a collection of common individual units of observation to provide a distribution to reflect variation. The second point sets out the need to identify factors that contribute to the variation, which is ultimately derived by differences between the observed units. There are likely to be many factors that interact and result in variation, even for the same unit over time. Deeper insight can be gained through a systematic approach to analysis and carefully consider confounding factors and bias. In doing so, associations between outcomes and putative influencing factors can be built, leading towards a causal understanding of how those factors affect an outcome.

The above objectives of an epidemiological approach to population-level energy demand research requires the development of a conceptual and methodological framework that consists of defined concepts and established analysis techniques and study designs. It is proposed here that these concepts are drawn from the epidemiological world, along with those from both the engineering and sociological spheres that are suited for adaption into this inter-disciplinary and complex approach.

4.3.2 Energy epidemiological conceptual framework

The energy epidemiology conceptual framework uses the concepts set out above. From these concepts, a set of objectives can be established that will enable an energy epidemiological approach to proceed. These objectives are to:

1. Apply the concept of causal models to describe (the conditions of) energy demand
2. Identify target populations of interest and select appropriate study designs to define samples and study groups.
3. Define and describe end-use energy demand variables and key confounding and potential causal factors in sample and study groups that can be extrapolated to the target population
4. Design and carry out formal studies, i.e. type, scale, method, data gathering, reporting
5. Describe key population characteristics, factors and distributions and establish causation.
6. Propose metrics to describe and define energy demand levels
7. Develop models, creating a formal conceptual framework for positing hypotheses
8. Translate results into policy and evaluate its implementation

Using the above approach, the purpose of the energy epidemiology approach is to:

- Describe the energy demand states within a population by a given metric (e.g. people, households or buildings);
- Understand endogenous and exogenous influences on energy demand;
- Understand the contributing factors that influence the use of energy within a population;

- Describe the longitudinal trends in energy demand; and as a result
- Provide a rigourous evidence base for directing energy demand policy.

The epidemiological approach may be applied to a host of issues within the energy demand field. In doing so, however, applying the concepts and methods and focusing on a clear set of aims are of particular importance for consistency and developing results that can be subject to systematic review. For example, when applied to energy efficiency, the approach should aim to:

- Detect causes that influence energy demand
- Quantify the association between energy demand and its key determinant factors (i.e. behavioural, structural, technological)
- Identify demand patterns, behaviours and trends that lead to more efficient energy use
- Test the efficacy of energy efficiency measures and interventions
- Monitor changes in energy demand over time

In meeting the above objectives, a methodological framework is needed that can support the energy epidemiology concepts. Such a methodological framework, comprising a set of tools, techniques and language is needed in order to establish, explore and support the concepts described above.

4.4 Summary

This chapter began by setting out a research paradigm within which an epidemiological approach would support empirically-based population-level energy demand research. The proposed paradigm is that the shift to a low-carbon society, along with the alleviation of energy-related social and environmental phenomena such as fuel poverty and lack of comfort, can be improved through population-based research methods that analyse patterns of demand in order to better understand the practices, drivers, causes and differences of energy demand outcomes. This provided the underlying theory on which an epidemiological approach to energy demand may be conceived.

The chapter went on to make a case that epidemiology is an appropriate approach with which to examine empirically-based population-level problems of energy demand in buildings. It made a case that the epidemiological approach is appropriate and that energy demand may be likened to a health outcome, such as obesity. Energy demand, like obesity, can be described along a spectrum with a host of interacting factors leading to a particular measured outcome.

The conceptual framework identified in Chapter 3 was examined and discussed in terms of its applicability to energy demand. The criterion was whether the concepts are appropriate for the study of energy demand, whether the epidemiological basis is maintained when applying these concepts, and whether they advance the study of energy demand. For the most part, the identified concepts are shown to be applicable, valid, and capable of advancing the understanding of energy demand research.

In Chapter 5, the methodological framework is described for energy epidemiology and the fourth research question is posed. The chapter goes on to set out the research approach used to exemplify the research questions, which leads to the remaining substantive chapters.

Chapter 5 Methodological Framework

Research approach: "Applying the epidemiological approach to population-level energy demand in residential buildings"

Chapter Introduction

In Chapter 4, a *conceptual framework* was proposed for an energy epidemiology approach to study population level energy demand. The conceptual framework provides the concepts that underpin the approach and respond directly to the *low-carbon society* research paradigm. The chapter also introduced the aims and principles of an energy epidemiology approach; epidemiology concepts are established, explored and supported through study designs and appropriate methods and analysis techniques. In this chapter, the epidemiology methodological framework is introduced and briefly described and its applicability to energy demand is considered.

The aim of this chapter is to introduce the research approach that will be used to illustrate how an epidemiological approach to population-level energy demand might function. The energy epidemiology approach will be explored through a series of method studies that will examine a selection of issues relating to energy demand in English dwellings. Each of the method studies applies a selected epidemiological study design and analysis technique in order to exemplify how the approach works in practical research. Each study concludes with a discussion of the approach and its strengths and weaknesses. The method studies provides a means of exploring the applicability and appropriateness of epidemiological concepts and methods as they apply to energy demand. It also provides an opportunity to explore relevant issues of energy demand in the UK residential sector.

This chapter consists of three main sections to describe the research approach that will be used to examine the selected research questions. The first section describes the basis of the methodological framework and how it supports the energy epidemiology approach. The second section describes the method study research approach, its aims and principles, and its purpose. The third section sets out the method studies that will be examined in the following chapters and the data sources that will be used. The three method studies selected are: an ecological study, a cross-sectional study, and a case-control study.

5.1 Methodological framework of energy epidemiology

The methods used by epidemiologists provide study designs and tools that support the central concepts to interpret and understand health-related phenomena. An energy epidemiology approach that studies phenomena of energy demand at the population level also needs appropriate methods. Therefore, the fourth research question being addressed in this work is related to the exploration of the key energy epidemiology concepts using common epidemiological methods. Namely:

> RQ4: Can epidemiology study designs and methods be applied to population level energy demand research problems using currently available people, energy and buildings data? And, what are their strengths and weaknesses when applied to energy demand?

In the following section, the epidemiological method (or methodological framework) of study designs is described and details are provided of when they are most typically used. This information will be used to address the above research question in order to inform how the identified methods could be used in the study of population-level energy demand.

5.2 Epidemiological study design framework

Epidemiology offers research on the occurrence of disease in populations that have been exposed to a risk factor, compared to those that are unexposed, in order to work towards a causal understanding and thus develop appropriate interventions (Bailey et al., 2005). Epidemiological studies in themselves do not prove a causal link. However, under the guidance set out by Hill[3], together with an understanding of disease mechanisms and other social and environmental phenomena, epidemiological studies provide some of the building blocks to do so.

Epidemiology is primarily an empirical science, relying on first-hand information and measurement in the 'real world', which has led to a framework broadly consisting of two spheres from which to study groups and individuals, observation and intervention (Bhopal, 2008). Observational assessments use study designs that primarily explore prevalence and incidence of a disease within a population due to the exposure to risk factors. Intervention studies assess the impact of interventions on outcomes. The methods used within these two spheres offer established study designs and data analysis practices capable of describing and determining association between key factors and outcomes for making inferences from population samples and sub-groups.

Developing causal associations is fundamentally problematic for observational studies because these study types lack several of Hill's proposed guidelines, namely: temporality. Analysis outcomes from observational studies more often identify associations between

[3] Note that Hill's guidelines were: strength of association, consistency among studies, specificity (single cause to single effect), temporality (cause precedes effect), gradient (dose-response), plausibility, coherency, and experimental evidence.

factors and outcome variables. Intervention type designs may provide stronger associations or evidence to suggest causality may be revealed through dose responses (for instance increased insulation associated with increased energy savings) and before-and-after treatment effects in longitudinal studies. For the most part, the populations being examined in observational studies tend to need to be larger in order to reduce error of the estimates.

5.2.1 Epidemiological study designs

In the process of conducting epidemiology studies, expertise has evolved to establish protocols for research methods and analytical techniques, for instance in systematically collecting data so as to minimise bias and measurement error and provide imputation methods for dealing with variables where data are not missing at random. Notably, emphasis is placed on reporting findings only after controlling for the effects of confounders; that is variables that are associated with both the risk factor and the outcome, so that the additional effect of a factor is identified.

The epidemiology methodological framework provides a series of analysis methods and tools that can be applied to a range of data. Figure 11 shows the common types of epidemiological study designs (Bailey et al., 2005). These include: ecological studies, case-control studies, cohort studies, intervention studies, randomised control trials, field trials, and variations thereof.

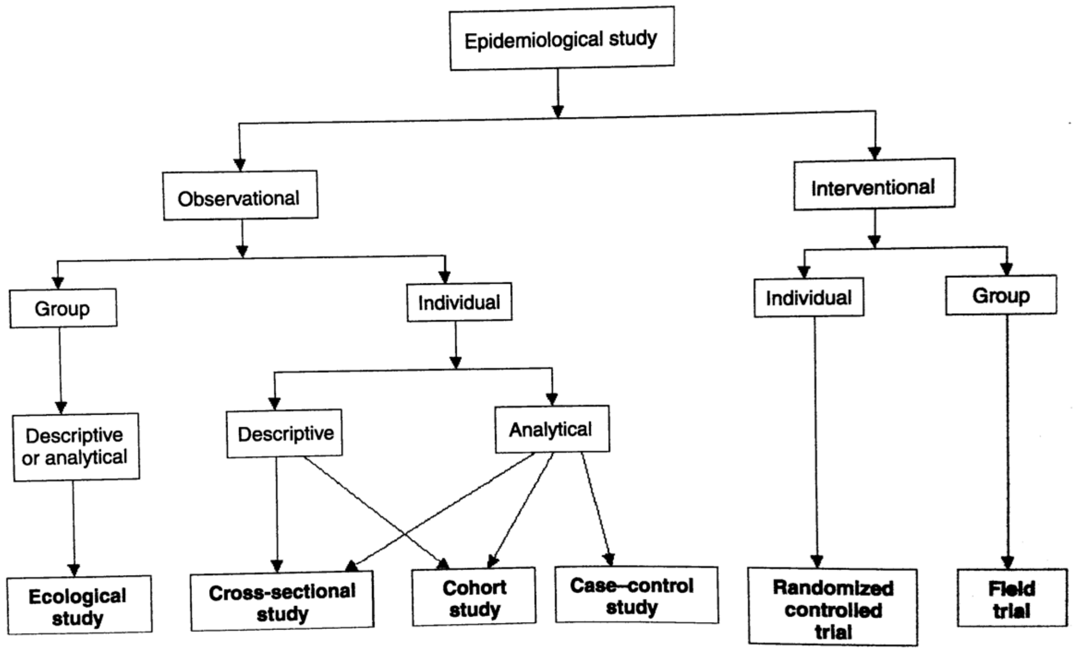


Figure 11 - Main types of epidemiological study design (Bailey et al, 2005. p.44)

The key features to consider are the nature of the study designs. Within the dichotomy of observational and interventional studies exists a further division of group vs individual.

This division is related to the unit of analysis and also the nature of what the study aims to do. Descriptive studies should simply describe outcomes or events in terms of their characteristics and distribution, looking for patterns and trends. Descriptive studies are often used in the initial stages of examining a phenomenon and should be used to pose hypotheses for examination in subsequent analytical studies. Analytical studies are used to test the hypotheses developed in descriptive studies and examine the relationship between an outcome and putative influencing variables.

Table 7 provides further details on the identified study designs. Each type of study will use certain forms of information on a population and examine several key concepts in building a whole picture of a phenomenon. The unique features of the study designs, such as the unit of analysis and its form (e.g. incidence, prevalence or outcome), the analysis techniques and measures of association, provide a means of examining a phenomena consistently. Also, each type of study is appropriate for certain types of collected data, e.g. aggregate, prospective or retrospective.

Table 7 - Epidemiological study designs and applications (from Bhopal, 2008, p. 15)

Study type	Study design	Essential concepts	Research purposes
Observational	*Case series and Population (or Ecological) case series*	Count cases (numerator) and relate to population data (denominator) to produce rates and analyse patterns. Look at characteristics of cases for causal hypotheses	Study signs and symptoms, and create disease definitions Surveillance of mortality / morbidity rates Seek associations Generate / test hypotheses Source of cases or foundation for other studies
	Cross-sectional	Study health and disease states in a population at a defined place and time Measure burden of disease and its causes	Measure prevalence (very rarely incidence) of disease and related factors Seek associations between disease and related factors Generate / test hypothesis Repeat studies (on different samples) to measure change and evaluate interventions
	Case-control	Look for differences and similarities between a series of cases and a control group	Seek associations Generate / test hypothesis Assess strength of association (odds ratio)
	Cohort	Follow up populations,	Study natural history of

		relating information on risk factor patterns and health states at baseline, to the outcomes of interest	disease Measure incidence of disease Link disease outcomes to possible disease causes, i.e. seek associations Generate/ test hypotheses
Interventional	*Intervention or Trial*	Intervene with some measure designed to improve health, then follow up people to see the effect.	Test understanding of causes Study how to influence natural history of disease Evaluation the effects (side-effects and benefits and costs of interventions)

Complexity is crucial to the epidemiological approach. Because epidemiological studies draw data from the 'real world' and examine how known mechanisms may manifest and distribute themselves among populations and sub-groups, study designs, measures of association and interpretations have been developed to cope with these limitations. The ability to examine phenomena under a number of different study designs is only possible with well-developed tools, such as: consistent study designs and analysis techniques, agreed definitions and common metrics, and the collection and comparison of data. Further, these tools also support a longitudinal focus for causality, using data collected over many years in a standardised manner (where possible) to inform and develop hypotheses. There is also a strong methodological tradition for large scale studies and the relationship between the clinical level and population level, with a great deal of effort employed to ensure that findings can be relevant outside of the individual case. This suggests that the epidemiological approach is adept at addressing issues of scale, definition, and viewpoints.

5.2.2 Selecting a study design

Undertaking an epidemiological study requires that a number of features are decided on so that the study can be effective and seek to attain internal validity (i.e. reducing the risk of bias). Bophal (2008) describes these steps as first setting the physical and temporal boundaries for the study, undertaking an enumeration of the population and describing its characteristics (e.g. socio-economic and risk factors), and developing a description of their health characteristics. The next step in the study design process depends on the research question being considered, i.e. whether the study will focus on questions of frequency and pattern, cause, or control – see Table 8. Understanding what type of study design is needed and the concept being examined is crucial to building an evidence base around a phenomenon and its control.

Table 8 - Data requirements and study designs for answering typical public health and epidemiological questions (Bophal, 2008, p. 40)

Public health and epidemiological questions	Data needs	Typical study design
(a) Frequency and pattern		
How common is this problem?	Case numbers for rare disease in stable population, otherwise population counts	Ecological Studies Disease registers Cross-sectional studies Cohort studies
Is the problem increasing, decreasing, or about the same?	Accurate numbers of cases and population counts	Repeat above studies/ analyses over time
Where does the problem occur most?	Case numbers and population counts by area	/ Analyse data by place
Who is affected most?	Case numbers and population counts by population characteristics	/ Analyse data by person
(b) Understanding the cause		
What are the causes of the problem?	Detailed information on the population and its social and environment context Data is collected to test hypothesis on causation	Case-control study Cohort study Trials (e.g. Randomised Control Trials)
(c) Control		
What strategy is needed to prevent or control the problem?	Understanding of what has worked elsewhere, the causal chain of the disease and of the resources available, together with understanding of the population in its context.	(Systematic) literature review
Are control measures working?	Case numbers and population data to monitor effectiveness of control measures; together with understanding of changes in population and environment	Evaluation using pragmatic designs, e.g. before and after analysis of disease Repeat cross-sectional studies Trials (e.g. Randomised Control Trials)

The above table provides not only a useful outline of the types of questions that might be asked and appropriate study types, but it also identifies the data that would be needed to address the problem. For example, much of the type of data needed relates to describing the population within which a problem is presenting itself or a control measure is being deployed or evaluated. Descriptive data is fundamental to most epidemiological studies for the simple reason that it provides a contextual basis to examine problems and evaluate controls. Further, several study designs are appropriate for particular question themes, for example cohort studies (if appropriately sampled) can be used to examine patterns and frequency, causal associations over time, and effectiveness of controls.

Epidemiology studies can also be categorised by their functional features (see Table 9); such as: whether they are descriptive or analytical in nature; whether they are followed over time or account for past events; whether they are observing a phenomenon or actively attempting to alter it; whether the outcome of interest has already occurred and which factors led to it or the presence of factors but where no outcome has yet occurred; and, whether comparison groups are required to make causal hypotheses.

Table 9 - Epidemiology study designs and classifications (Bhopal, 2008 p. 157)

Design	Descriptive / Analytic	Retrospective/ prospective	Observational / experimental	Beginning with disease/ causes of disease	Specific comparison group / no such group
Ecological Case series (individual and population	Descriptive	Retrospective	Observational	Disease	No
Cross-sectional	Descriptive	Retrospective	Observational	Both simultaneously	Usually not, but possible
Case-control	Analytic	Retrospective	Observational	Disease	Yes
Cohort	Analytic	Prospective and retrospective	Observational	Usually causes	Yes
Trial	Analytic	Prospective	Experimental	Usually disease, sometimes cause	Yes

Using 'real world' data, studies that seek to determine associations between putative risk factors and outcomes at a population level should provide comparison between groups of individuals and ensure that sources of bias are avoided (Woodward, 2004). Comparison provides a basis for determining whether the associations between risk factors and outcomes are consistent and whether other factors may modify the association. Study designs that make group comparisons need to be designed to compare 'like for like' to ensure that findings can be more broadly generalised. This means sampling appropriately and describing that sample so that controls can be applied if needed. For retrospective and prospective studies, comparison groups need to be as similar as possible in order to remove the potential or bias affecting an association. Retrospective study designs are particularly problematic due to recall bias, which sees those with an outcome more likely to recall certain events which they believe led to the outcome under study.

A further issue to consider in 'real world' data is that when using small samples, such as trials or case series, representativeness may be a problem. Case series, i.e. individual or highly selected groups, are generally not representative of any target group or may have unique underlying conditions that cannot effectively be controlled for. Making inferences from case series to wider populations should be avoided. However, these types of studies, often called clinical, can be very useful for examining supposed mechanisms because they can be highly controlled and can provide effects that can be examined in larger study types. Clinical studies are still considered 'real world' in that they will be controlled, but not to the degree that a laboratory study using test subjects or models (not theoretical) would be.

Other considerations for selecting study designs will be issues related to the length of time available for study and how frequent the outcome is. Case-control studies, for example, are able to examine a number of factors that could affect an outcome of interest, but they are subject to forms of selection bias due to the pre-allocation into groups. If they are retrospective they can be relatively quick to conduct, but require care in identifying events that have occurred in the past. Case-control studies are also more effective for examining rare events or outcomes that develop over a long period, for the reason that they can select the 'case' group and identify matching non-cases, or controls. However, because of recall and sample bias, this type of study is not ideal for studying aetiology. Cohort studies are better suited to examining causal associations because they are able to account for chronological events. They are less well suited to studying rare events or those that take a long time to manifest because follow-up over extended periods is less practical. Regardless of the selected design, there are many considerations that need to be taken into account when selecting and conducting an epidemiological style study.

5.2.3 Energy epidemiology methodological framework

The energy demand paradigm identified in Chapter 4 seeks to shift society to a low carbon state and alleviate energy-related social and environmental phenomena by examining trends, patterns and drivers at a population level. This paradigm was born from limitations of the past approach to studying energy demand through case study and modelling approaches that failed to grapple with and account for the complexity and uncertainty of how and why energy is used. However, by focusing at a population level, the issues being examined are more easily distorted and subject to variation than would be found in case studies. For this reason, the concepts identified in Chapter 3 and discussed in terms of their applicability to energy demand in Chapter 4 requires a method that can support the study of the underlying complex phenomena that influence energy demand and its related outcomes. The approach to data gathering and analysis will depend on the energy demand phenomena being examined, but using an epidemiological framework will help to ensure validity, consistency, comparability, and generalizability. The aim of this study is whether an epidemiological approach can provide insight into the major issues surrounding understanding trends, causal relationships and determinant factors of energy demand at a population level; and in addition, whether the approach can provide a strong foundation of evidence on which to build models, and to support policy to manage energy demand.

It is proposed that an energy epidemiology methodological framework can take as its base the methods framework used in epidemiology, described above, with the main purpose of providing a common set of tools, techniques and language, within which to carry out the following:

1. Identify energy demand phenomena and their manifestation and differences among a population;
2. Design and carry out formal studies to examine the phenomena for an identified target population, i.e. study type, size and sample, method, data gathering, analysis, and reporting;
3. Describe the distribution of the phenomena from the sampled population and identify patterns and trends
4. Generate hypotheses to explore and account for differences in the outcomes of interest and develop and undertake tests of the hypotheses in working towards causal associations;
5. Develop models based on the understanding of, and describe measures to control or manage, the energy demand phenomena.

The research question being addressed in this chapter relates to whether epidemiological methods can be applied to population-level problems of energy demand. The purpose of the methodological framework is to support the concepts outlined in Chapter 4 and to fit within the broader paradigm of population-level energy demand research. In the above section, the main purpose of the energy epidemiology methodological framework is to provide tools with which to examine energy demand phenomena. These tools must be applied in a manner such that findings can extend beyond the specific study and provide insights into broader issues being faced. In the following section, a research approach will be proposed that will test this research question.

5.3 Research approach

The research approach to be undertaken, which will comprise the remaining substantive parts of this thesis, is focused on addressing the fourth research question, i.e. can epidemiology study designs and methods be applied to population level energy demand research problems. The research approach uses a novel concept called a 'method study'.

A method study approach is similar to a case study approach, but aims to examine the method as applied to a particular problem. Thomas (2011) describes a case study as being focused on examining a particular 'subject' (e.g. "persons, events, decisions, periods, projects, policies, institutions, or other systems") that is part of a broader 'class of phenomena' using a selected method (or methods). The subject provides the basis for an 'analytical frame' or 'object', i.e. the idea or concept being examined (Thomas, 2011). In a method study the method is the *subject* that is being examined as being part of a broader class. The method study focuses on studying a given method, in this instance an epidemiological study design, to a defined problem and in doing so exemplifies underlying concepts and addresses the specifics of a selected problem. Like a case study, the selected problem being investigated should provide an example whose results can be extended to other similar issues. The analytical frame of the method study is the approach itself, again in this instance: epidemiology.

The method study approach is being used here in order to explore the practical application of epidemiological methods to energy demand phenomena at a population level. The focus of this work is on examining how appropriately the epidemiological research approach and its accompanying methods can be used to improve population-based study of energy demand. In developing the approach and methodological framework for the study of energy demand the research will:

i) Develop a methodological framework that supports the study of energy demand in buildings; and,

ii) Apply a selection of methods to selected issues within the energy demand in buildings research field.

The epidemiology methodological framework will be assessed for its applicability to provide suitable support and investigative tools in energy demand research. Each selected method will be assessed against a criterion with similar principles as the concepts, these are:

- Is the method appropriate to the issue being investigated?
- Does the method support the identified energy epidemiology concepts?
- Does the method expand the understanding of the issue at hand and is the approach generalizable to a broader set of related problems?

In the following section the research approach is expanded by first identifying several epidemiological methods to apply to common problems in energy demand research. Each research problem is introduced, leaving the detail to be further discussed within the chapters devoted to the selected method study. The section will briefly describe the research population and introduce the data sources that will be used for the method studies.

5.4 Method studies

Epidemiology seeks to identify a pressing problem or issue that needs to be addressed to improve population health. In establishing a research question and testable hypothesis, the first part of the process is to collect data on outcomes or events of interest and describe the frequency of the problem, its distribution among the population and differences among groups. The outcomes should be examined for patterns and trends among the population and what factors are related to any identified differences among groups. As described in Table 8, the methods used to examine the frequency of events are typically studies that describe incidence and prevalence, which could be through ecological, cross-sectional or cohort designs. These first steps are fundamental because they provide a means for generating hypotheses to test relationships between supposed influencing factors and the outcome of interest. Following this first part, generated hypotheses are more formally examined to determine whether the supposed factors have stronger associations with the outcomes or events of interest. This stage will rely on study designs that can support the elements of Hill's guidelines for determining causal associations. This latter part moves towards aetiology and potentially establishing a causal relationship. The final stage is the development of tests and mechanisms that are able to control or manage those outcomes (providing they are negative) or to expand treatment (for positive impacts). This stage also includes evaluating past practices, procedures and policies for their efficacy and their delivery.

5.4.1 Method study research problem

Energy efficiency improvements in UK homes are a major part of the Government's decarbonisation plans. The *Energy Efficiency Strategy* includes policies to reduce energy demand through improvements in building energy performance and aggressive decarbonisation of the energy supply system (DECC, 2012b). These proposals broadly seek a 20% reduction in heating-related energy demand in the existing housing stock by 2030. Under these plans, the UK government has set out pathways that will see millions of energy efficiency retrofits installed in houses by 2050 (as proposed by UK Committee on Climate Change (UK CCC, 2010)) at an estimated cost of £200 billion (BIS, 2010). As part of meeting these targets, the UK CCC mitigation pathways identify the need to insulate more than 7.5 million lofts, 4.6 million cavity walls and 3.3 million solid walls, and to complete 1.9 million double glazing upgrades by 2030 (UK CCC, 2010). Meeting these ambitious goals successfully will require a major shift in the delivery of these retrofit measures to UK dwellings, along with a better understanding of the impact on the use of energy in houses.

The method studies will be focused on the problem of:

> understanding the uptake of energy efficiency in UK dwellings and the impact these measures can have on energy demand.

This problem exemplifies the type of large-scale issues being tackled in the energy demand research field. It is a major concern among policymakers who are tasked with achieving multiple, and sometimes conflicting, objectives related to energy and housing. Further, gaining a deeper understanding of the problem requires building a context that describes the breadth and depth of the problem among the population of interest and also supports the development of testable hypotheses regarding impact.

5.4.2 Research population

The research population will vary slightly for each method study, but will broadly examine dwellings and households in England, Great Britain or the UK. The time period being considered will also depend on the study and issue being examined, but will again broadly be over the past-decade. This time period is chosen because it provides a relevant basis on which policies are being directed in order to undertake extensive retrofits for transition to a low-carbon society. There are a number of datasets that are available to examine this population which are described in further detail below.

In 2010 it is estimated there are 22.7 million dwellings in England, 1.3 million in Wales, 2.5 million in Scotland and 0.75 million in Northern Ireland (CLG, 2010a; DSDNI, 2011; Scottish Government, 2011; Welsh Assembly Government, 2011). The 2010 projections state there are 52.2 million residents in England, 3.0 million in Wales, 5.2 million in Scotland, and 1.8 million in Northern Ireland. The Census states that there are 22.0 million households in England, 1.3 million in Wales, 2.3 million in Scotland, and 0.70 million in Northern Ireland (NRS, 2013; ONS, 2012a, 2011).

5.4.3 Data sources

There are several sources of data available to describe dwellings and households in the UK, including national census output statistics and national cross-sectional surveys (e.g. the English Housing Survey (EHS)). In addition, there are various administrative datasets that provide information on dwellings and households collected by several government organisations. Chief among them is the Census that includes a great deal of information on both physical features of dwellings and the characteristics of the households. In addition, the Valuation Office Agency (VOA) collects dwelling information from local councils for the purposes of taxation. While there is a great deal of aggregated information on British dwellings and households most of this information is either aggregated (i.e. Census and VOA) or sampled (e.g. EHS). There is little information at a household record level available for analysis from which to draw study samples.

There are few sources of national level statistics on all retrofit measures in the UK. Different national UK governments collect information on government-back programmes, such as Warm Front (England) or Warm Homes (Scotland), and energy supplier obligations (e.g. Energy Efficiency Commitment). For measures that occur outside of these government-related schemes, the data availability is the domain of installer and accreditation organisations (e.g. Gas Safety). The most comprehensive information on all of these sources of data is the Homes Energy Efficiency Database (HEED), developed by the Energy Saving Trust (EST) and the Department of Energy and Climate Change (DECC).

Information on energy use at a national level (to Local Authority level) is available through the Digest of UK Energy Statistics (DUKES). DUKES covers all fuel use by sector but is developed from top-down estimates and is therefore not population-based. At a national level, DECC publish bottom-up energy demand data for all gas and electricity meters, both domestic and non-domestic. Energy demand from gas and electricity account for approximately 91% of total fuels used in the domestic sector (Palmer and Cooper, 2013).

The datasets used in this thesis are:

- Homes Energy Efficiency Database
- Energy supplier annualised gas and electricity meter data
- English Housing Survey
- Administrative Data (i.e. Census Outputs and VOA Council Tax Data).

The datasets are described in further detail below.

5.4.3.1 Homes Energy Efficiency Database

The Homes Energy Efficiency Database contains information on the characteristics and energy efficiency of approximately 16 million homes in England, Wales, Scotland and Northern Ireland. HEED covers over 50% of the UK housing stock (EST, 2010) and is drawn together from approximately 60 datasets and collected from approximately 20 organisations. The bulk of HEED data has been classified using the Reduced Standard Assessment Procedure (rdSAP) format, which attempts to categorise dwellings into common bands relevant to modelling energy demand (BRE and DECC, 2009). Where other forms were used, additional variables are added and are allocated to the best available class within rdSAP. EST undertook this data cleaning prior to HEED being made available for use in this study.

The data provided in HEED draws from survey data, plus data on specific measures installed under a variety of government-backed schemes, energy supplier obligations and installers. Table 10 provides a summary list of these data sources and Figure 12 shows a breakdown of the data sources from which the extract of HEED that was analysed included. Note that the variables collected under each source will vary and many sources for measures will include survey data. HEED comprises information at the individual dwelling level rather than by households or occupants. It contains no information on households or dwelling occupant, aside from the tenure, and thus socio-economic factors cannot be determined directly. It primarily contains information on the physical features of the dwelling as they pertain to the energy efficiency of the structure (i.e. fabric) and the heating system; see Table 11 for a summary of the survey and measures data. See Appendix A for more details on HEED.

Table 10 - Homes Energy Efficiency Database (HEED) data suppliers and programmes

Programme	Provider(s)	Survey/ Measures
Government schemes[a]	Warm Front & Warm Homes Scottish Central Heating Programme The Warm Deal	Survey and measures
Surveys	Home Energy Check National Registry of Social Housing Local Authorities	Survey
Installers[b]	Boiler installers Glazing installers Insulation installers Cavity wall installers Local Authorities Renewable installers	Measures
Energy Suppliers	Customer energy efficiency improvement schemes[c]	Measures

Note: [a] Government schemes are primarily targeting vulnerable groups, i.e. fuel poor or high indices of deprivation; [b] Installer data is collected as part of building regulation requirements, such as from Gas Safety and FENSA and also industry bodies such as the Cavity Insulation Guarantee Agency; [c] Energy supplier schemes target customers in fulfilment of carbon reduction targets.

Table 11 - Homes Energy Efficiency Database (HEED) example data

Data Type	Data Examples	
Survey Data	Property type Tenure No of Bedrooms Year of construction Space heating fuel Water heating fuel Loft insulation thickness External wall type	Window type Window frame type Levels of draught-proofing Main heating system Secondary heating system Hot water system Heating controls (various types) Energy rating (SAP) Hot water tank insulation
Measures Data	New/additional loft insulation and depth Cavity wall insulation Solid wall insulation/flexible linings Boiler replacements Heating control upgrades	Fuel switching Compact Florescent Lamps Renewable systems (e.g. solar thermal, solar PV, heat pumps)

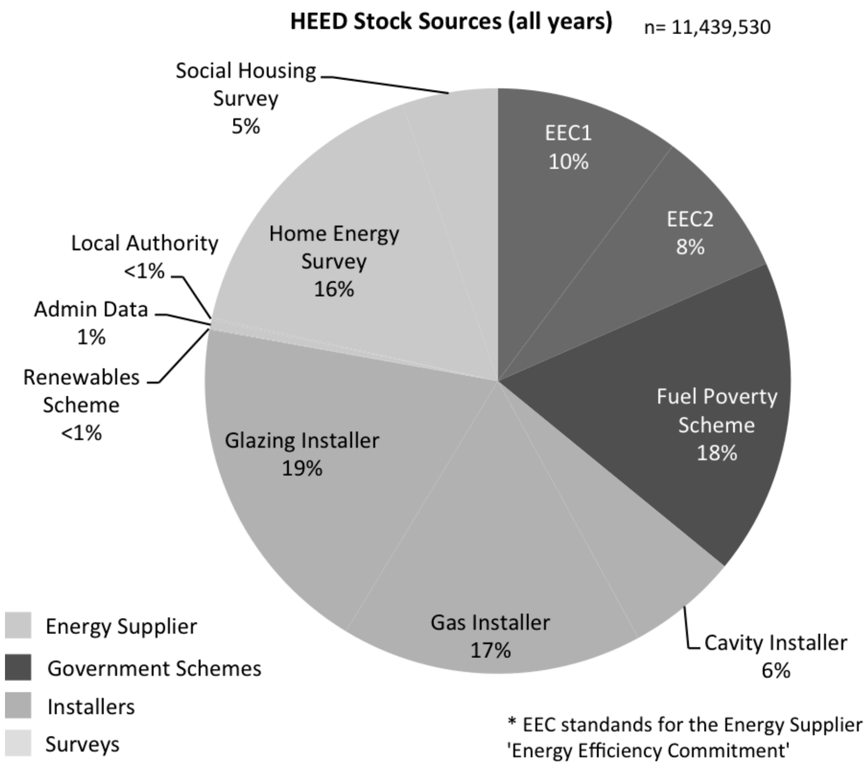


Figure 12 - HEED stock data sources

5.4.3.2 Energy supplier meter data

Energy suppliers collect and annualise final consumption gas and electricity data for individual meter points from customers which is then provided to DECC for the purpose of various statistical outputs; in 2008 there were approximately 22.6 million gas meters (22.3 million residential and 0.3 million non-domestic) and 29.1 million electricity meters (26.7 million residential meters and 2.4 million non-domestic meters) (DECC, 2008a). UK gas and electricity meters are classified into two types: daily (gas) or half-hourly (electricity), and non-daily (gas) or non-half hourly (electricity). Gas and electricity is thought to account for just over 90% of energy delivered to UK dwellings in 2010 (DECC, 2012e). Non- gas or electricity energy use, such as wood burning, is more difficult to meter and therefore little data is available on these fuel uses. The gas and electricity meter point data was provided by the Department of Energy and Climate Change (DECC) and covered the period of 2004 through 2007 and was linked to HEED by the Government for use in two EPSRC-funded projects that this dissertation draws on.

Gas non-daily meters are divided into categories based on their total expected annual load demand; although there is no user identification, 'residential' users are deemed to be those whose demand is less than 73.2 MWh/yr[4] and those above are commercial or industrial (DECC, 2009). The meter readings are converted into annual consumption values

[4] This figure, which has been inherited from the days before the privatisation of British Gas (it is the metric equivalent of 2500 Therms) is roughly 5 standard deviations above mean gas consumption. The high gas use threshold means that small non-domestic users are often included and (infrequently due to their numbers) very large dwellings are excluded.

using a common methodology with two meter readings at least six months apart (when no meter reading is available an estimate based on past demand is used in its place) and is corrected to a seasonal normal demand and an end-user climate sensitivity adjustment to derive a total annual demand (OFGEM, 2009). The purpose of the seasonal correction is to allow for inter-year comparisons that are weather independent. In terms of what the weather correction might mean for assessing the impact of energy efficiency interventions through the detection of changes in energy demand between years, it may be that long-term trends are more significant than year-on-year changes, but this will depend on the frequency of meter readings for which no information is available. The annual gas data period is 1 October to 30 September and covers a heating season.

Non-half hourly electricity meters are defined into classes representing likely demand profiles and are identified by user types. Residential electricity meters are classed into two types based on the meter: unrestricted electricity or economy 7. Economy 7 refers to meters that are on a time-charge tariff that offers cheaper electricity during off-peak hours, typically an 8-hour period, and are either time- or radio-switched (DECC, 2009); in dwellings, these meters are most often associated with electric heating, either space heating (e.g. storage heaters) or hot water, offering the customer the advantage of electricity bought at off-peak rates and stored as heat for daytime use; in this work economy 7 meters are kept distinct. Unrestricted meters are all other types of meters; these meters may be used for heating but are not time- or radio-switched. Electricity meters are annualised using actual meter readings or, if no readings are available, estimates are based on past use and historic usage patterns and are smoothed across an annual profile to derive a total annual demand in kWh (Elexon, 2009). The annualised electricity values are not corrected for weather. The annual period for electricity data is from 30 January to 29 January.

Both the gas and electricity data undergo a cleaning process to remove or identify potentially erroneous data points, such as negatives and dummy values (e.g. '1' values). For further details on the energy data see Appendix B.

5.4.3.3 English Housing Survey

The English Housing Survey reports on the overall condition of English dwellings and the households living in them (CLG, 2010a). The EHS uses an un-clustered stratified sample randomly drawn from a list of all addresses in England and has been updated and made available approximately every two years since 2002 (CLG, 2010a). The survey interviews approximately 17,500 households on the details of their home and household. A further physical survey of approximately 8,000 of those houses interviewed is undertaken. The EHS provides a statistically random sample of the English stock. The survey provides a factor with which to weight variables in order to represent houses or households in England; for this research the houses weighting was used.

5.4.3.4 Administrative data

The VOA's Council Tax Property Attributes tables are collected as part of the VOA's responsibility to group properties into the appropriate council tax bands (VOA, 2010). As part of the banding process, the VOA collects information on the characteristics of the property that may affect its value (including, age, type, area, and number of bedrooms). The VOA data is considered a 'global' listing, as it will contain information on all dwellings within an area (i.e. local council) and thus should have all dwellings in the stock. Note that the VOA only collects information for England and Wales. The VOA in maintaining a current valuation list revises the data annually; the data used in the comparison comes from an extract made in September 2010. The VOA dataset provides information at the Local Authority level for approximately 24.7 million residential properties.

5.4.4 Selected method studies

Three methods are selected to examine the above stated research problem sequentially. The methods consist of a study design along with the related energy epidemiological concepts that will be explored. The three methods are: an ecological study, a cross-sectional study, and a retrospective cohort study. The above research problem is structured so that each study will tackle a defined issue that can be built on in the following study. The study methods, the specific component of the problem they will address, the concepts examined, and their main aims and objectives are discussed below. Each method study is assessed in terms of its strength and weakness against the criterion set out in Section 5.3.

5.4.4.1 Ecological: Uptake of energy efficiency measures in England

The first method study examines the uptake of energy efficiency measures in England over the period 2000 to 2007. A longitudinal ecological study design was selected for this research, which is a means for investigating differences between populations, or for studying group-specific effects through the use of aggregated data.

Purpose:

Over the past two decades the installation of a range of energy efficiency measures such as cavity wall or loft insulation or efficient heating systems, has become increasingly widespread among dwellings in England. Beyond aggregate estimates of installations in the stock, however, little detailed evidence is currently available regarding the uptake rate or prevalence of interventions among specific household groups. This study uses the HEED to investigate both the combination of measures that have been installed, and in which dwellings, according to key neighbourhood socio-demographic variables, including income and tenure.

Method study research question:

It is well established that the drivers of energy efficiency in the UK encompass physical characteristics of the dwellings, household-level practices, institutional delivery mechanisms and policy priorities, along with the wider fiscal and environmental context (Dowson et al., 2012; Mallaburn and Eyre, 2013; Pelenur and Cruickshank, 2012; Tovar, 2012). While the analysis reported in the literature makes use of temporal survey data to assess the relationship of these drivers with energy efficiency uptake in English dwellings, what is missing is any geographical investigation of these effects, to show how local features may be associated with higher or lower levels of uptake. The research question asked in this study is therefore: What combinations of physical (dwelling), social (household), and environmental (geographical) characteristics and delivery mechanism are associated with the uptake of energy efficiency measures under a range of programmes during the last decade?

Concepts:

This study method focuses on two related metrics of uptake of energy efficiency retrofits in the dwelling stock: incidence and prevalence. Incidence is the rate at which new reported installations of an efficiency measure occur in the housing stock during a specified time period (e.g. new cases reported annually). Prevalence is the proportion of the housing stock with a reported efficiency feature at a given point in time (e.g. in 2007). The metrics, often associated with health-related studies, offer an approach for examining associations between socio-demographic and physical characteristics seen in the population and changes in levels of energy efficiency of the housing stock.

5.4.4.2 Cross-sectional study: Home energy efficiency and energy demand in UK dwellings

The second study method examines the relationship between dwelling characteristics, energy efficiency levels and gas and electricity demand. A cross-sectional study design was selected, which is a means of describing and analysing relationships for a specific population at a specific point in time. The study comprises a descriptive component that details the levels of energy efficiency characteristics for the sample population and their energy demand, and an analytical element to examine differences between the energy efficiency levels and energy demand. Cross-sectional studies are able to describe the prevalence of putative influencing factors and of outcomes and provide an overview of the 'current' status of energy efficiency and energy demand in the housing stock.

Purpose:

HEED records data on more than half of the UK housing stock, and contains information on a wide range of installed measures and building characteristics. In the first instance, it is necessary to determine whether datasets from a wide range of suppliers can be used effectively to describe energy efficiency, analyse energy performance and provide insight into parts of the UK housing stock and its energy demand behaviour. The aim of this method study is to 1) describe the HEED data, in particular to assess the overall representativeness of the houses in HEED as compared to other British housing stock data; 2) to describe the differences in energy demand (gas and electricity) of the HEED housing stock, segmented by built form characteristics and level of energy efficiency for a selected period (i.e. 2005 to 2007).

Method study research question:

Developing energy efficiency intervention programmes for the UK housing stock that are capable of achieving significant and sustained reduction in energy demand requires a clear description of the state of the existing stock. The details available on the energy performance of dwellings using HEED provides a means of examining who has received these measures at an individual dwelling level. The information collected by the government on the annualised gas and electricity use from energy suppliers provides a means of examining the relationship between those delivered energy performance measures and actual energy use. The two research questions being asked in this method study is: What differences are there between those who undertook energy efficiency retrofits and the overall dwelling population of England? And, what is the distribution of energy demand for differing dwelling characteristics and levels of energy efficiency?

Concepts:

This method study focuses on population samples and developing associations between putative influencing variables and outcomes. This method makes use of a large-scale sample and makes comparisons to other standard populations to determine differences and provide context for who received energy efficiency retrofits. The concept of establishing associations within sub-groups is also examined.

5.4.4.3 Retrospective cohort study: Energy efficiency retrofits and energy savings in English houses

The third method study examines the uptake of energy efficiency measures at the individual dwelling level and the impact that energy efficiency retrofits have had on gas demand in English houses. A population-based cohort study was selected to investigate the association between household and dwelling characteristics and the uptake of energy efficiency retrofits as well as changes in energy demand. A cohort study provides a means for examining trends and patterns of difference among a population that is designed to be representative of a broader group.

Purpose:

Between 2004 and 2008 gas and electricity demand in England fell by 3% and 2% per year (DECC, 2012c). The causes of the reduction in energy demand are incompletely understood. During this period, there have been a number of changes in factors that could affect gas and electricity use, from increases in domestic energy prices, changes in weather, changes in expenditure, changes in energy use practices, and a programme of energy efficiency interventions to name several. The uptake of energy efficiency measures in the housing stock has been identified as a potentially significant factor in the change in energy demand, and in particular is considered to have a strong effect on saving energy.

Method study research question:

A population-based cohort study of English dwellings selected from the Homes Energy Efficiency Database (HEED) was used to investigate the association between household and dwelling characteristics and the uptake of energy efficiency retrofits and also changes in energy demand. The study used a sample that was drawn to be representative of English houses, using the English Housing Survey (EHS) as a sample frame. The uptake of energy efficiency measures was examined from 2002 to 2007. The change in gas and electricity demand was followed from 2004 to 2007.

The primary research question being addressed in this study is: what differences exist in the take-up of energy efficiency measures and the relationship between dwelling and household characteristics and what impact have these measures had on gas demand between 2005 to 2007, accounting for variation among the sampled houses.

The aim of the study is to provide a better understanding of the uptake of energy efficiency retrofits and the resulting change in energy demand that accounts individual dwelling and household characteristics, adjusting for potentially confounding and interacting factors.

Concepts:

There are a number of important features and considerations when undertaking a cohort study. The first is to establish a working hypothesis to investigate for a target group. It is then a matter of defining the study group who will represent the target group. The premise of a retrospective cohort study is that the individuals in the cohort are 'retrospectively' followed in time by looking back in their history to identify factors of interest to study differences in present outcomes. A retrospective cohort study is limited by the available information that can be obtained during the study period, though they have the advantage of being able to examine a wide range of outcomes.

5.5 Summary

The thesis seeks to explore the applicability and appropriateness of epidemiological concepts by exploring and advancing them through the application of appropriate methods, comprised of study design and analysis techniques.

This chapter introduces the epidemiology methodological framework that supports the energy epidemiology conceptual framework. The application of epidemiological methods is dependent on the research question being examined and the methodological framework provides a series of analysis methods and tools that can be applied to study a range of problems. Common epidemiological concepts were introduced and their basic features discussed.

This chapter sets out the remaining investigations of the thesis, which is the application of selected epidemiological methods to the study of current energy demand research problems. A selection of datasets that will be used in the method studies are identified and briefly described. In the subsequent chapters, each selected method will be assessed to consider whether: a) the method is appropriate to the issue being investigated; b) the method supports the identified energy epidemiology concepts; c) the method expands the understanding of the issue at hand; and, c) the approach is generalizable to a broader set of related problems.

Chapter 6 Ecological Study

"Uptake of energy efficiency retrofits in England"

Chapter Introduction

Chapter 5 set out the methodological framework through which epidemiological concepts are supported, explored and advanced. The framework broadly consists of a number of study designs through which different classes of research questions are addressed, e.g. frequency and patterns, causes, and control strategies and their effectiveness. The application of these methods to problems of energy demand provides a process through which the epidemiological approach can be examined and appraised for its appropriateness and relevance.

The aim of this chapter is to illustrate the application of an epidemiological study design, along with its key concepts and selected analysis techniques, to an energy demand problem. The research problem being investigated is focused on the installation of energy efficiency retrofits in the UK. This method study uses the Home Energy Efficiency Database to investigate both the combination and location of retrofits that have been installed and in which dwellings, according to key neighbourhood socio-demographic variables, including income and tenure. This chapter covers the development of standard epidemiological metrics of rates - incidence and prevalence - looking at energy efficiency interventions in England over the period 2000 to 2007. The study uses a longitudinal ecological study design to examine the rate of uptake and its relationship with different population features.

The chapter comprises five parts: a) a brief introduction to the research problem along with relevant background and a description of the key concepts and features of the selected study design; b) a description of the features of the study design and the research problem being investigated; c) a presentation of the study results; d) a discussion and conclusions of the findings specific to the study; and e) a critical appraisal of the study approach, its strengths, weaknesses and limitations.

6.1 Research problem

The Government's decarbonisation plans call for millions of retrofits to take place in the housing stock over the coming two decades (UK CCC, 2010). To meet these targets, approximately 22 million homes need to be retrofitted by 2050, or approximately 1,600 per day. The retrofits include: loft insulation, cavity wall filling, solid wall insulation, boiler replacements, heating upgrades and controls, and draught-proofing (UK CCC, 2012).

6.1.1 Research question and aims

Barriers to the uptake of efficiency retrofits included beliefs and social norms, household practices and characteristics, upfront costs, perception of institutions (e.g. government or energy suppliers), the landlord-tenant split incentives, and the characteristics of the property itself (Brechling and Smith, 1994; Pelenur and Cruickshank, 2012; Tovar, 2012). Other factors, such as past government policies and targets, hidden costs, market shape and broader fiscal issues such as economic performance and taxation have also been cited as drivers influencing the uptake of retrofits in housing (Mallaburn and Eyre, 2013; Mills and Schleich, 2012).

Several studies have examined the uptake rate and associations with household characteristics (e.g. income, education, socio-economic class) and dwelling features (e.g. efficiency, size, type). These studies used temporal cross-sectional survey data to assess the relationship of these drivers with energy efficiency uptake in English dwellings. However, a key consideration is to understand the variation of the uptake rate across different geographic areas and how its differences may relate to features of those areas. To date, what is missing is any geographical investigation of these effects, to show how local features may be associated with higher or lower levels of uptake. The research question asked in this study is therefore: what combinations of physical (dwelling), social (household), and environmental (geographical) characteristics and delivery mechanism are associated with the uptake of energy efficiency retrofits in England during the last decade?

This study uses HEED to describe the level of uptake of retrofits in England between 2000 and 2007. In doing so, the coverage of the database is examined by comparing the number of reported retrofits with uptake estimates using English Housing Surveys over the same period. Then, the relationship between the uptake rate of energy efficiency retrofits and a selection of household features at the neighbourhood level (i.e. Lower Super Output Areas (LSOA)) is examined in order to describe potentially influencing variables. The LSOA is a geographical output area built for census purposes. Output areas are built from clusters of adjacent postcodes and are designed to have similar population sizes. LSOAs have between 400-1200 households (with an aim of 500 households).

The aim of this method study is to: a) describe the uptake of energy efficiency in (and across) English households between 2000 and 2007; and b) investigate the differences in uptake across England by programme type and measure by neighbourhood and household characteristics.

6.1.2 Context and background

During the period of interest, 2000 to 2007, the UK government had a number of programmes that aimed at providing energy efficiency retrofits to dwellings. In England, these were the *Energy Efficiency Commitment 1 and 2* (EEC 1 & 2, 2002 to 2008), *Warm Front* (launched in 2000 and continuing to the present day) and the *Energy Efficiency Standards of Performance 3* (EESoP 3, 1994 to 2002). The aim of Warm Front was to reduce the risk of ill health for vulnerable households due to cold, damp homes; that is, households with children, pregnant women, people with disabilities and long-term illness and elderly household on certain types of benefits (EAGA, 2004). The EESoP 3 scheme also focused on disadvantaged costumers (i.e. low-income, elderly or in debt) with a requirement of two-thirds of the targets for energy companies to invest in energy efficiency retrofits to be directed towards these households (Ofgem and Energy Saving Trust, 2003). EEC 1 & 2 also required that 50% of the energy efficiency measure delivery targets were focused on priority groups (i.e. those in receipt of particular benefits and tax credits). During this period there have also been revisions to the Building Regulations (2002 and 2006) that have required refurbishments to comply with higher fabric standards (walls, roofs, floors and windows) and meet higher ventilation and air-tightness standards (HMSO, 2006). The changes to the Building Regulations in 2002 required certification of installed double glazing through a regulatory body (i.e. Fenestration Self-Assessment Scheme (FENSA)) as a means of ensuring quality standards. Other certifying bodies were also set up to provide guarantees for retrofits, including cavity wall insulation and loft insulation. Government-backed schemes and supplier obligations required the use of accredited installers and certified products (Ofgem, 2002). Windows and cavity wall insulation retrofits have certification bodies that maintain standards of practice for registered installers. Other retrofits, such as loft insulation and draught proofing, are not subject to specific regulatory accreditation and are also easily undertaken by individual owners.

Various government organisations captured details on the uptake of energy efficiency retrofits as part of government schemes for reporting purposes and programme evaluation. Many of these sources of information were collected into HEED (EST, 2009). HEED is the most comprehensive database of reported energy efficiency retrofits available in the UK and will be used in this method study as a source of information on the uptake of efficiency retrofits over the period of interest for England.

6.2 Method

The method section describes the study design used in the method study, in term of its features and the concepts that it uses. This is followed by a brief description of data and the approach used to examine the stated research question.

6.2.1 Ecological study design and key concepts

A longitudinal ecological study design was selected to address the research question. The ecological study design is a means for investigating differences between populations, or for studying group-specific effects through the use of aggregated data. The ecological study is carried out using population level data (rather than individual level data) and is generally used in early examinations of an issue or problem. For this reason, ecological studies tend only to look at associations and are often descriptive and sometimes exploratory in nature. The key objects of interest in ecological studies are populations. These studies are used to examine differences between groups, rather than individuals, because it may be that certain features are associated with these differences in outcome (Bailey et al., 2005). For example, low-income households tend to be the target of more policy and therefore there maybe reasons to believe that these groups will have different outcomes than those that are not directly targeted. Other reasons to use ecological studies involve their cost and the availability of group level data. Because the study design relies on group level information certain biases may exist as a result (Pearce, 2011). Bias may exist due to differences in the way information is collected or differences at the individual level that may be concealed due to the boundaries used. A further issue is the inference from group level to individual level: this is known as ecological fallacy. This means that certain features may group together and may show an association but are not necessarily related. For this reason ecological level studies should not be used to infer information to the individual level.

This method study focuses on two related metrics of frequency of outcomes, using the uptake of energy efficiency retrofits in the dwelling stock. These are incidence and prevalence. Incidence is the rate at which new reported installations of an efficiency measure occur in the housing stock during a specified time period (e.g. new cases reported annually). Prevalence is the proportion of the housing stock with a reported efficiency feature at a given point in time (e.g. in 2007). The metrics provide a way of examining the relationship between socio-demographic and physical characteristics seen in the population and changes in energy efficiency levels of the housing stock.

This study uses the LSOA as the unit of analysis, the term LSOA is used to represent a 'neighbourhood' in terms of their size (i.e. ~500 households), though this is a simplification. There are approximately 32,480 LSOA within England. This study uses the LSOA level data to examine relationships between household characteristics and the uptake of energy efficiency by intervention programme (e.g. supplier obligation, government or other). Odds ratios (ORs) are derived from frequency data of both the outcome variable and also selected neighbourhood level household variables. The ORs are used to describe whether associations exist between presence/ absence of an outcome and the presence/ absence of those selected features. In this study, the outcome of interest is the uptake rate of energy efficiency retrofits over the period 2000 to 2007 within neighbourhoods. The main interest is whether there are differences in uptake between population sub-groups (e.g. areas with lower incomes or older populations). Here, the odds ratio represents the odds of an outcome (e.g. high uptake rate) in a group, given a particular feature (e.g. neighbourhood location) over the odds of not having an outcome (e.g. lowest uptake rate) given the same feature. If an outcome is associated with a feature, the odds of exposure in the outcome group will be higher than the corresponding odds in the non-outcome group (i.e. odds ratio >1). If an outcome shows no association with a feature the odds will be the same in both groups (i.e. odds ratio = 1). If an outcome is associated with a lack of a feature, the odds of exposure in the outcome group will be lower than the corresponding odds in the non-outcome group (i.e. odds ratio <1). For continuous variables (e.g. proportion of dwelling tenure) a unit of change is specified (e.g. additional 10%) in order to estimate the corresponding odds ratio for each unit of change. Most areas will have had some level of uptake over the period of interest and therefore there is no absolute threshold for what constitutes a high (vs low) level of uptake. Instead, it is possible to categorise these relatively and compare areas with the highest uptake rate of energy efficiency to areas with the lowest uptake rate, at the neighbourhood level.

6.2.2 Energy efficiency uptake study approach

This study of the historic installation of energy efficiency retrofits in England and their relationship with house and household characteristics comprised two main analysis components. The first was a description of the incidence of energy efficiency installation (i.e. reported) in the housing stock for a selection of retrofit retrofits during the period 2000 to 2007, along with an examination of the prevalence of energy efficiency features in 2007. The second was an analysis of the relationship between neighbourhood level[5] (i.e. LSOAs) household details and the uptake of efficiency retrofits during the period of interest.

To determine the representativeness of HEED for energy efficiency retrofits in England, HEED is compared with the *English Housing Survey*s made over the same period. Information on households was drawn from data at the LSOA level from the 2001 Census, Valuation Office Agency (VOA) dwelling attribute statistics, and inter-census administrative data. SAS 9.3 software was used in the data preparation and analysis of energy efficiency uptake incidence and neighbourhood level characteristics (SAS Institute Inc., 2011).

6.2.3 Data sources and variables

The Home Energy Efficiency Database is used to describe historic installation of energy efficiency retrofits in England from 2000 to 2007. Over the period of study (2000 to 2007) in England, 9.3 million homes were the subject of efficiency retrofits, covering approximately 40% of England's 22.6 million dwelling stock, with an average of 2.1 million intervention installations per year in 1.2 million dwellings per year. HEED was described in Chapter 5 and in further detail in Appendix A.

In this study HEED is considered to be a 'global' survey of the target population (i.e. England) or in other words a 'census' of reported efficiency retrofits. HEED is treated in this manner on the basis that it contained information from all government-backed schemes and the regulatory bodies and installers dealing with energy efficiency retrofits, along with extensive surveys of energy performance. HEED data are used as population level aggregate statistics (e.g. total number of reported retrofits in a given year) for comparison against other population level statistics. This assertion is tested through comparisons of selected reported efficiency retrofits to those reported in English Housing Surveys. However, it is

[5] For the purpose of this study the lower layer super output area (LSOA) is referred to as a 'neighbourhood'. The terms are used interchangeably when referring to the level of analysis. Whilst it is acknowledged that this is not necessarily a true description of a neighbourhood, it is a useful reference and aligns with the ONS description of available spatial statistics. On average, each LSOA has 500 dwellings.

acknowledged that there may be unreported or under-reported efficiency retrofits not included in HEED, which could be appreciable for some retrofits (e.g. DIY loft insulation). For these reasons, the retrofits of uptake were treated with a degree of caution, discussed further below.

HEED consists primarily of information on the physical characteristics of the dwelling, including: dwelling type, age, number of bedrooms[6], wall type and insulation level, loft insulation, glazing type, heat system, and fuel type. However, because the database is collected from a variety of sources the reporting of the physical description variables ranges from 67% for dwelling type to 59% for household tenure across HEED as a whole. Dwelling characteristic variables were not used for the purposes of description; instead neighbourhood level details from a variety of sources were used.

The database also included details on a number of installed energy efficiency retrofits. This study focused on: heat system replacement, glazing replacement, loft insulation and cavity insulation, although data on other retrofits are also shown for interest. Heat system replacement includes the installation of storage heaters, heat pumps, warm air systems, and boilers. Replacement of gas boilers (condensing and non-condensing) accounts for 98% of all heat system replacement data. Glazing replacement includes the installation of pre and post-2002[7] double glazing units and triple glazing. Loft insulation includes both 'top-ups' or laying additional loft insulation and 'virgin' or laying loft insulation where none existed; the data captures the level of insulation to a level of 250mm. Cavity insulation includes details on filling cavity walls built pre and post-1976, which marked the introduction of building regulations that required a wall U-value of at most 1.0 W/m^2K. For those retrofits installed under government programmes it was presumed that the regulatory guidance required was followed and where otherwise the retrofits were installed under best-practice or according to Building Regulations, and thus their reported presence was assumed as accurate.

The database provides details on the general location of the dwellings at the LSOA level and contains a date stamp associated with the details describing the installation of the retrofit or collected data. The date of the HEED detail is used to determine the uptake rate associated with retrofits installed in the England from 2000 to 2007. The location information was used to determine the rate of installation on an area-by-area basis and was also used to link neighbourhood level characteristics.

The EHS[8] was used as an independent comparison source of energy efficiency retrofits over the period of interest. The dwelling weighting gross factor was used for analysis here, as the study concentrates on dwellings as the unit of observation (rather than the households therein). The EHS contains details on selected efficiency features present in the dwelling (i.e. predominant type of window, loft insulation thickness, type of wall and insulation, and boiler type). Loft insulation thickness distinguishes 0mm, less than 100mm, 100mm to

[6] HEED does not contain dwelling floor area.
[7] The significance of pre- and post-2002 double glazing refers to a requirement introduced in the British Building Regulations of 2002 requiring that all windows (and replacement windows) conform to lower U-values.
[8] In 2008 the English Housing Survey replaced the former English House Condition Survey and Survey of English Housing.

150mm and 150mm or more. Windows include single and double glazed by casement material. Wall insulation included cavity with and without insulation and 'other' (e.g. solid wall, wood, pre-fabricated, etc…). Boilers include standard (i.e. non-condensing), back boilers, combination boilers, condensing boilers and condensing-combination boilers. The EHS efficiency variables were not directly comparable to HEED and therefore required careful grouping and selection for comparison; see Table 12.

Double glazing was not compared due to the differences in variable between HEED and the EHS. HEED reports the occurrence of a double glazing installation, which could be a single or multiple windows, while the EHS reports the predominant type. Glazing is presented for illustration only.

Additional data at the LSOA level were used to describe a range of household features. Data on median LSOA income and household type (following Mosaic classification) were drawn from Experian Mosaic Public Sector data (Experian, 2012). Data from the Office for National Statistics (ONS) was used for: age of population, level of central heating, number of benefit claims, and proportion of dwellings within council tax bands from the Neighbourhood Statistics service (ONS, 2012b). Data on dwellings counts by age and type were drawn from the Valuation Office Agency (VOA) property attribute datasets for 2010 (VOA, 2010). Also, data on the climate in 2005, measured in heat degree hours, was drawn from the UK Met office (UK Met Office, 2012). Appendix C contains further details on these LSOA level variables.

Table 12 – Description of energy efficiency variables used for comparison between Homes Energy Efficiency Database and the English Housing Survey

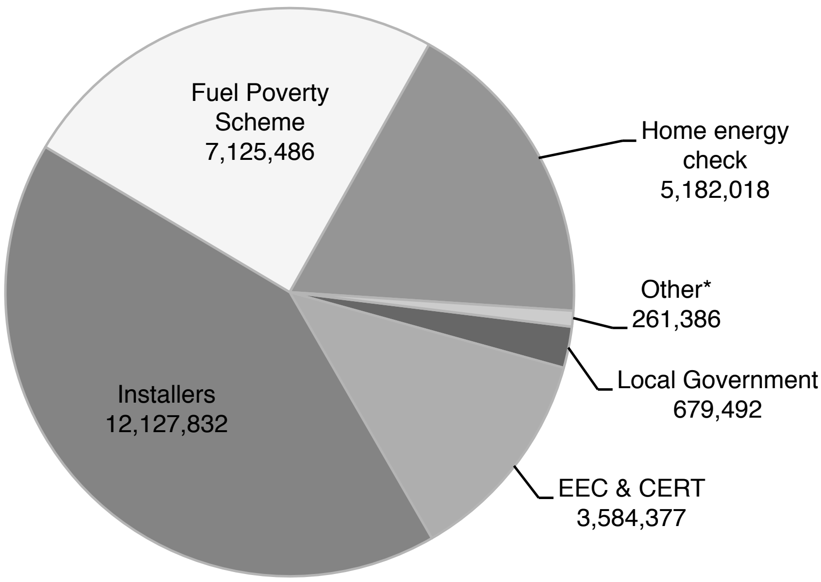
Measure	HEED Variables and "class"	EHS Variables and "class"
Loft insulation	Loft insulation measure "Insulation to 250mm" combined: "0-250mm", "25-250mm", "50-250mm", "75-250mm", "100-250mm", "150-250mm"	Loft {loftins4] "150 mm or more"
Cavity wall insulation	Cavity wall measure "Cavity wall insulation" combined: "Cavity wall insulation (pre-1976)", "Cavity wall insulation (post-1976)", "Cavity wall insulation (unknown age)"	Wall insulation type {wallinsx] "Cavity with insulation"
Boiler replace-ment	Boilers "Condensing Boilers" combined: "Condensing boiler", "Condensing boiler and controls", "Condensing-combination boiler", "Condensing-combination boiler and control"	Type of boiler [boiler] "Condensing boilers" combined: "Condensing boiler", "Condensing-combination boiler"
The following variables were not compared, but are presented for illustration		
Double glazing installation	Glazing measure "Double glazing replacement" combined: "Double pre-2002", "Double post-2002"	
Double glazing coverage		Predominant type of window [typewin] "double-glazed- wood", "double-glazed- UPVC", "double-glazed-

	metal”
	Extent of double glazing [dblglaz2] “unknown”, “less than 80% double glazed”, “80% or more double glazed”

6.2.4 Analysis approach

Longitudinal ecological studies use aggregated data to examine trends or changes over time for outcomes, in this case the uptake of a variety of energy efficiency interventions. This study examined whether there were any differences in the uptake across a selection of neighbourhood level features. The approach involved first developing the outcome variables of interest (i.e. the rate of uptake of efficiency retrofits), which were compared against national level statistics. It then examined whether any significant differences existed between the uptake rate and neighbourhood level features.

In the UK many energy efficiency retrofits are installed through government-supported schemes, while others are related to natural replacement or household decisions. Figure 13 shows the breakdown of the total number of retrofits (not houses) by programme or provider for England in HEED in 2007. HEED data was grouped into six categories according to the programme or data provider to analyse any change in uptake between government-backed programmes and industry or household efforts. These are: EEC, Installers, Fuel Poverty, Home Energy Surveys, Core Group (i.e. EEC, Installers and Fuel Poverty), and All sources.



*Other covers renewables installations

Figure 13 - HEED data by programme or data supplier in 2007 (sum of all retrofits since 2000)

Data on the number of installations (i.e. annual count) that occurred in HEED between 2000 and 2007 was used to examine both the uptake incidence (i.e. frequency) of energy efficiency retrofits and the prevalence of retrofits (the sum of incidence for the period) in 2007. The EHS from 1996 to 2008 was used to describe the prevalence of the selected features in the English housing stock. An estimate of the annual prevalence of efficiency features was made for each year not covered by the EHS through a linear interpolation between the known EHS survey years. From this the annual incidence for each of the selected retrofits was derived for comparison to HEED. A table showing the results of the interpolation is provided in Appendix D.

For further analysis at the neighbourhood level, the efficiency retrofits were grouped into three categories, one describing all retrofits, another describing only heat-related retrofits (i.e. condensing and standard boiler replacement, hot water cylinder replacement, and solar hot water systems) and a third describing fabric-related retrofits (i.e. loft and cavity insulation, glazing replacement, and draught proofing). The neighbourhood level (LSOA) energy efficiency uptake rate was determined using the uptake incidence data as a proportion of the total number of dwellings within the neighbourhood as of 2005, determined using Council Tax statistics (CLG, 2005) selected for being a recent count of English houses within the period covered. The energy efficiency uptake rate was divided into quintiles for the analysis, with 5 being the highest rate of uptake.

The neighbourhood level data were used to examine relationships between neighbourhood level household characteristics on the uptake of energy efficiency by programme. The SAS logistic regression procedure - *Proc Logistic* - was used to generate the odds ratios (ORs) for the association of selected household variables and the highest uptake rate of energy efficiency compared to the lowest at the neighbourhood level. The multiple logistic regression model is defined as:

$$log_e\left(\frac{\hat{r}}{1-\hat{r}}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

Where, $\hat{r}$ is the estimated likelihood of high vs low uptake and β are the explanatory variables. In this of this analysis, the explanatory variables for each LSOA are: rank of median income, proportion of owner occupied dwellings, proportion of flats, rank of heating degree days, proportion of council tax bands A-C, proportion of households on benefits, proportion of persons above 60, proportion of persons below 16. For proportional values, classes are generated as a 10% change from 0.

The odds ratio represents the odds of an outcome (e.g. high uptake rate) in a group given a particular feature (e.g. neighbourhood location), compared to the odds of not having an outcome (i.e. lowest uptake rate) given the same feature.

Because the analysis undertaken here is on a presumed ‘census’ of energy efficiency and also relies on actual English household census data, confidence intervals (i.e. the quantification of uncertainty of the estimates) are not required because the data represents the ‘true’ population. However, in practice missing observations of the population will occur and therefore the census represents a 'super-population'. As such, confidence intervals of the model are shown for reference.

6.3 Results

The annual uptake rate of energy efficiency retrofits England, using the data in HEED, increased between 2000 and 2007 for all retrofits, with the exception of draught proofing (Figure 14). On average, the uptake incidence rate for fabric retrofits over the eight year period was 0.87 per dwelling, and for heat system retrofits was 0.24 per dwelling for all houses in England. Loft insulation increased from approximately 29,000 installations per year to 700,000 installations per year by 2007, a 22-fold increase. Double glazing installations increased from 34,000 installations per year in 2000, to 2.8 million installations per year by 2007, a 81-fold increase. Cavity wall insulations increased from 99,270 installations in 2000 to 822,000 installations by 2007, a 6-fold increase. Condensing boiler replacement increased from 18,500 installations in 2000 to 1.1 million installations by 2007, a 59-fold increase. Most retrofits follow a relatively stable incidence trajectory, with the exception of double glazing installations which increased dramatically between 2003 and 2005, most likely as an artefact of the building regulation requirements. Fabric retrofits had the largest number of installations since 2000, approximately 18 million in total. The largest number of fabric installations by retrofits were: glazing (55%), cavity insulation (22%), loft insulation (17%) and draught proofing (6%). Approximately 5 million heating retrofits were installed, with condensing boilers making up 65% of the total.

The total cumulative number of installations over the 2000 to 2007 period was highest in the installers group at 42% of the total number of retrofits – see Table 13 below. The bulk of the installers' data was provided by gas and window installers, who were also responsible for the largest number of installation retrofits. Installations for government-backed schemes comprised 41% of the total installation from 2000 to 2007. Homeowner surveys provided 18% of the data on installations and local government the remaining 2% for the period.

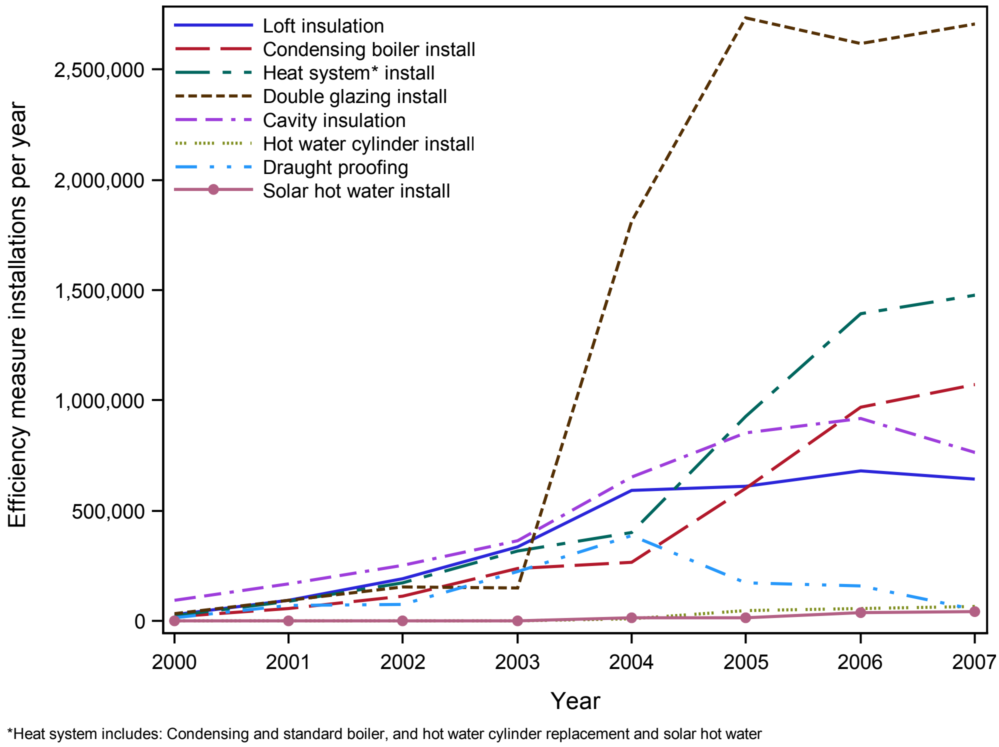


Figure 14 - Number of energy efficiency measure installations in England (non-cumulative) between 2000 and 2007 drawn from HEED (N= 9.3 M houses)

Table 13 - Energy efficiency installation uptake in England between 2000 and 2007 by data provider or programme, drawn from HEED (N= 9.3 M houses)

Efficiency installations in England (1000's)	Year								
Measure / Source	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**Total**
*Heat**									
Local Government	0.5	0.5	0.8	1.0	0.2	7	70	12	92
EEC 1 & 2	0.4	0.1	24	73	41	161	99	116	515
Installers	4	6	15	19	105	447	744	1,044	2,385
Fuel Poverty Scheme	8	48	81	211	194	280	280	-	1,102
Home energy check	12	35	56	16	86	92	296	409	1,002
Fabric†									
Local Government	6	5	8	8	1.5	35	369	94	527
EEC 1 & 2	3	0.6	80	207	334	936	417	401	2,378
Installers	54	44	60	36	985	1,781	1,969	2,604	7,534
Fuel Poverty Scheme	46	229	338	768	1,642	1,206	769	0	4,998
Home energy check	57	150	188	49	478	415	850	1,069	3,257

*Heat includes: Condensing and standard boiler, and hot water cylinder replacement and solar hot water; †Fabric includes: Loft, cavity insulation, draught proofing and double glazing

Across England, fabric and heat efficiency interventions were shown to be highest in the North East and North West regions and lowest across London and most of the southern regions. The uptake incidence rate (i.e. cumulative total number of installations for the period over the total number of dwellings in 2005) of fabric retrofits concentrates around midland and northern centres such as Leicester, Birmingham, Liverpool, Manchester, Leeds and Hull. Heat system installations also track the large urban centres in both the north and the south, with the exception of London.

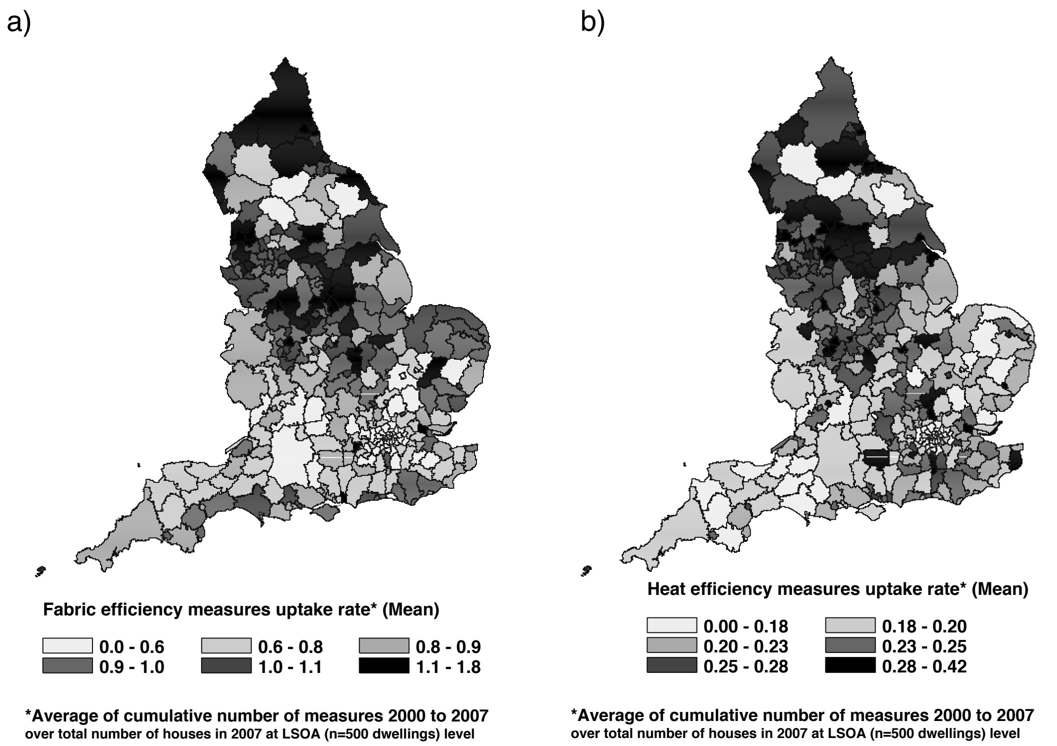


Figure 15 - Cumulative energy efficiency measure installations for a) fabric retrofits and b) heat system retrofits from 2000 to 2007 as ratio of number of dwellings in 2005, drawn from HEED (N = 9.3 million)

6.3.1 Comparison of uptake in England

Loft insulation, cavity wall filling and condensing boiler installation data in HEED are compared with the EHS and shows that boiler and loft retrofits track each other very closely, but that cavity wall filling diverge over the period of interest (see Figure 16). The installation of double glazing as reported in HEED, and the coverage of double glazing estimated in the EHS (Figure 17) are shown but not directly compared. The prevalence of retrofits in England for 2007 is also compared (Figure 18). Cavity wall filling had similar profiles but there were fewer installations reported in the EHS, where the discrepancy is approximately 1.1 million or 41% more reported in HEED than the EHS. The installation of condensing boilers according to the two sources was found to be very close; there was only a difference of 4% between HEED and the EHS, with 86,000 more installations reported in HEED. The double glazing installations profile in HEED shows a dramatic increase following 2002, which reflects the primary source of the glazing data, FENSA. FENSA was put in place in April 2002 and was set up to regulate the replacement of windows and doors as part of the introduction of the 2002 building regulations and as a result there were very few pre-2002 glazing installations reported in HEED. The EHS shows that the additional number of dwellings where the predominant window type was double glazing fell during the period. There were approximately 26% fewer loft insulation installations reported in HEED during the entire period; however, from 2003 onward there were only 7% fewer installations reported in HEED.

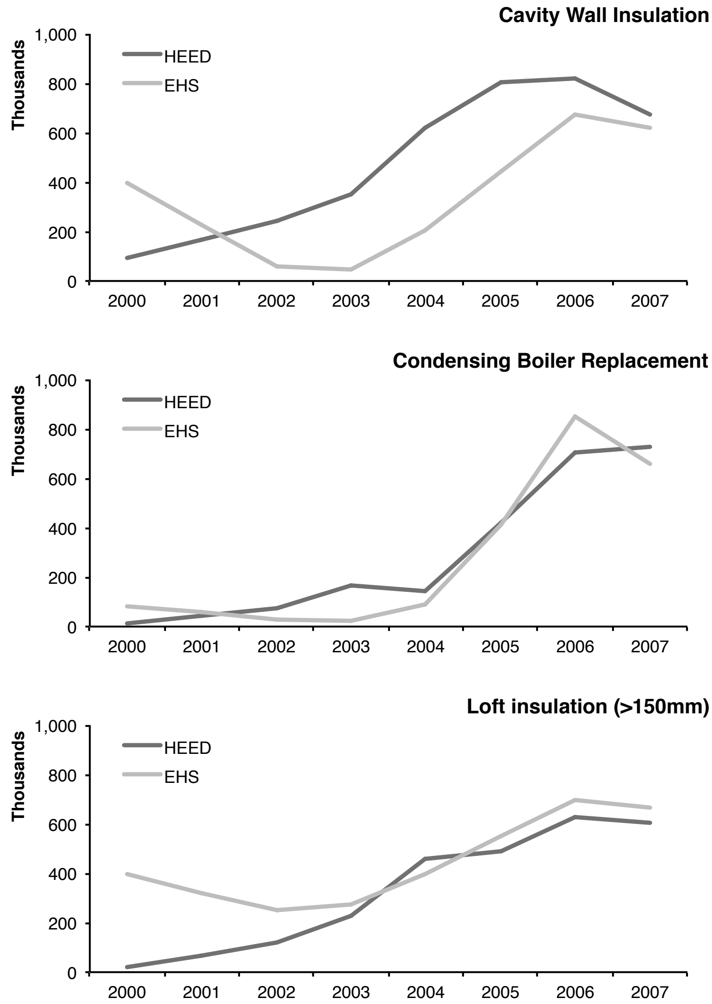


Figure 16 - Uptake incidence of energy efficiency retrofits in HEED compared to EHS for period covering 2000 to 2007

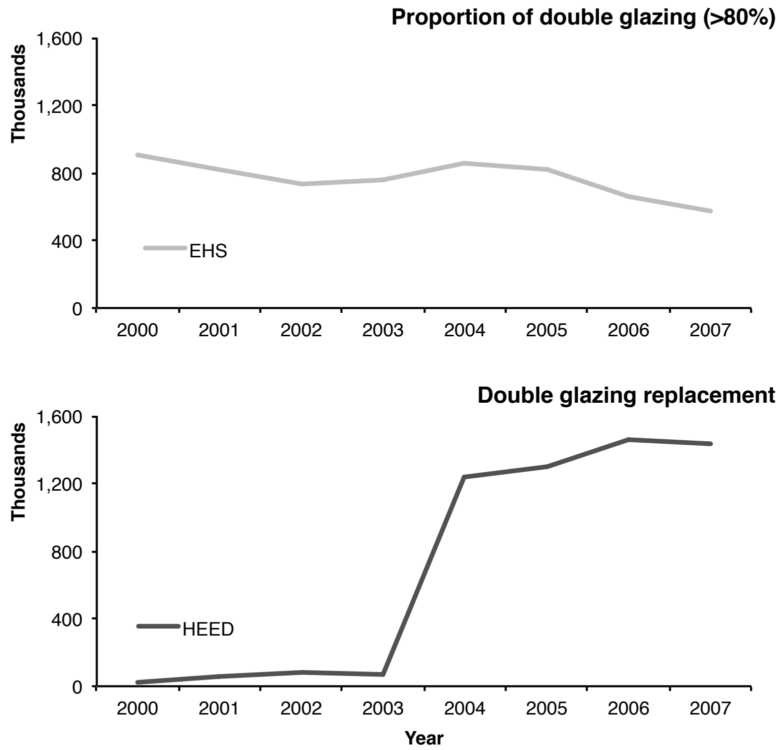


Figure 17 - Number of installed (HEED) and proportion of double glazing >80% (EHS) for period covering 2000 to 2007

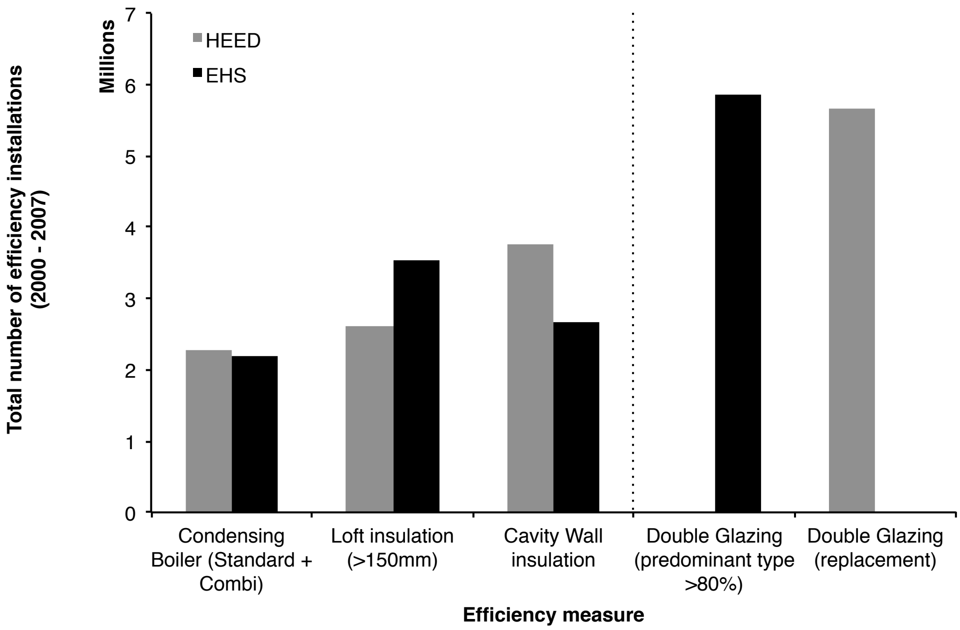


Figure 18 – Comparison of prevalence of energy efficiency retrofits in England between HEED and EHS for the period covering 2000 to 2007.

6.3.2 Household characteristics and uptake of efficiency retrofits

At the neighbourhood (LSOA) level, the uptake incidence rate of energy efficiency retrofits for the 2000 to 2007 period (measured in quintiles) was associated with income, tenure, the proportion of flats, climate and the proportion of the population with older adults and children (Table 14).

The energy efficiency uptake rate[v] is examined for three separate groups: Energy Efficiency Commitment (EEC 1 & 2), Installers, Fuel Poverty Schemes, and for all of HEED. For the EEC group, compared with neighbourhoods with households in the highest quintile of incomes (i.e. ≥£50,000), those neighbourhoods in the lowest income quintile were 2.71 times more likely to be in the highest uptake quintile for fabric efficiency retrofits and 2.05 times more likely for heating efficiency retrofits. Fabric measure uptake in the EEC group was higher for every 10% increase in the proportion of owner occupied dwellings, with an odds ratio (OR) of 1.27, and 1.11 for heat system retrofits. The OR of being in highest uptake of fabric retrofits for each additional 10% increase in flats was lower at 0.84 and also for heat retrofits, 0.91.

The uptake of energy efficiency retrofits drawn from the Installers group follow a similar trend as for the EEC group, although at greater magnitude. Neighbourhoods in the lowest median income quintile were 5.29 times more likely to have a high uptake of fabric retrofits and 4.11 times more likely for heating retrofits than neighbourhoods in the highest income quintile. For every 10% increase in households on benefits in a neighbourhood, the odds of being in the highest uptake group was 5% higher for fabric retrofits. In the Fuel Poverty Schemes group, a high uptake rate for fabric efficiency retrofits in neighbourhoods with low incomes was 10.74 times more likely and 6.5 times more likely for heating retrofits. The likelihood of a high uptake of both fabric and heating retrofits increased for every 10% increase in owner occupied dwellings, 1.88 and 1.69 respectively. The Fuel Poverty group was 20% less likely to be in the highest uptake as the proportion of flats increased. As the proportion of benefits increased, the likelihood of being in the highest fabric and heating efficiency uptake rate quintile increased by 63% and 50% respectively. The occurrence of being in the highest uptake group in the Home Energy Check group, i.e. those who self-reported energy efficiency retrofits, is similar magnitude and trend to the EEC group.

For all sources of efficiency retrofits, income was highly related for both fabric and heating retrofits, with those in the lowest income quintile being 12.23 times and 6.98 times more likely in the highest uptake rate quintile. The odds of being in the highest uptake rate quintile for fabric and heating measures for each 10% increase in the number of owner occupied dwellings was 1.73 and 1.30 and the number of homes with benefits was 1.33 times and 1.13 times respectively. Overall, the proportion of the neighbourhood population being older (i.e. ≥60 years) or having children increased the odds of high uptake only slightly. Weather was also shown to be somewhat related to the likelihood of being in the highest fabric and heating retrofits uptake quintile. Compared to the warmest neighbourhood areas,

[v] Reflecting the point on confidence intervals made in section 6.2.4, only the estimates are shown in the text. The confidence intervals are presented in the table for reference.

those in the coldest region were 1.96 and 1.73 times more likely to be in the highest uptake quintile for fabric and heating measures respectively. The ORs for the Core Programmes group are reported as a comparison to All Sources. It is found that the ORs are slightly higher in the Core group, but similar in trend.

Table 14 – Odds ratios of uptake incidence rate for energy efficiency (incidence frequency over dwellings in 2005) from 2000 to 2007 at t LSOA level and by energy efficiency programme in England

	Energy efficiency measures uptake incidence rate 2000 to 2007 (N= 32,480 LSOAs)							
	Energy Efficiency Commitment		Installers		Fuel Poverty Schemes		All sources	
Variable	Fabric	Heat	Fabric	Heat	Fabric	Heat	Fabric	Heat
Quintile of median income in 2005								
Q1 vs Q5	2.71** (2.39-3.06)	2.05** (1.81-2.31)	5.29** (4.67-6.00)	4.11** (3.64-4.64)	10.74** (9.42-12.24)	6.56** (5.78-7.45)	12.23** (10.72-13.95)	6.98** (6.16-7.92)
Q2 vs Q5	1.51** (1.38-1.65)	1.26** (1.16-1.38)	2.77** (2.53-3.04)	2.15** (1.97-2.35)	6.04** (5.49-6.65)	3.85** (3.51-4.22)	5.02** (4.57-5.52)	3.24** (2.96-3.55)
Q3 vs Q5	1.36** (1.26-1.46)	1.07* (1.00-1.15)	2.01** (1.87-2.16)	1.43** (1.34-1.54)	4.07** (3.77-4.39)	2.66** (2.47-2.87)	3.18** (2.95-3.43)	1.87** (1.74-2.01)
Q4 vs Q5	1.11** (1.04-1.19)	0.95 (0.89-1.01)	1.38** (1.30-1.47)	1.06* (0.99-1.13)	2.45** (2.29-2.63)	1.91** (1.79-2.04)	1.80** (1.69-1.92)	1.30** (1.22-1.38)
Tenure (proportion of dwellings)								
Owner occupied (units=10%)	1.27** (1.25-1.3)	1.11** (1.09-1.13)	1.43** (1.40-1.46)	1.18** (1.16-1.20)	1.88** (1.84-1.92)	1.69** (1.66-1.73)	1.73** (1.69-1.76)	1.30** (1.27-1.32)
Dwelling type (proportion of dwellings)								
Flats (units=10%)	0.84** (0.83-0.85)	0.91** (0.9-0.92)	0.78** (0.77-0.8)	0.87** (0.86-0.88)	0.76** (0.75-0.77)	0.88** (0.86-0.89)	0.70** (0.69-0.71)	0.80** (0.79-0.81)
Quintile of climate (heat degree days in 2005)								
Q2 vs Q1	1.56** (1.46-1.66)	1.55** (1.45-1.65)	1.75** (1.63-1.87)	1.52** (1.42-1.62)	0.80** (0.74-0.85)	0.76** (0.71-0.82)	1.55** (1.44-1.66)	1.30** (1.21-1.38)
Q3 vs Q1	2.10** (1.97-2.25)	1.86** (1.74-1.99)	2.20** (2.05-2.36)	1.77** (1.65-1.89)	0.72** (0.67-0.77)	0.72** (0.67-0.77)	1.97** (1.84-2.12)	1.64** (1.53-1.76)
Q4 vs Q1	2.20** (2.05-2.35)	1.89** (1.77-2.02)	2.56** (2.39-2.75)	1.95** (1.82-2.08)	0.68** (0.64-0.73)	0.62** (0.58-0.67)	2.21** (2.05-2.37)	1.67** (1.56-1.79)
Q5 vs Q1	1.99** (1.86-2.14)	1.71** (1.59-1.83)	2.00** (1.87-2.15)	1.70** (1.58-1.82)	0.79** (0.74-0.85)	0.83** (0.77-0.89)	1.96** (1.82-2.10)	1.73** (1.62-1.86)
Council Tax Band (proportion of dwellings)								
Band A&B (units=10%)	1.00 (0.99-1.01)	0.98** (0.97-0.99)	0.98** (0.96-0.99)	0.98** (0.97-0.99)	1.16** (1.15-1.17)	1.13** (1.12-1.14)	1.02** (1.01-1.04)	1.00 (0.99-1.01)
Benefits (proportion of dwellings)								
On benefits (units=10%)	1.00 (0.98-1.03)	0.96** (0.94-0.98)	1.05** (1.02-1.07)	0.99 (0.97-1.02)	1.63** (1.59-1.67)	1.50** (1.47-1.54)	1.33** (1.29-1.36)	1.13** (1.1-1.16)
Household age (proportion of dwelling occupants)								
Adults ≥60 years (units=-5%)	1.03** (1.01-1.04)	1.02** (1.01-1.04)	1.02** (1-1.04)	1.01 (0.99-1.02)	1.03** (1.01-1.05)	1.01 (1-1.03)	1.03** (1.01-1.05)	1.01 (0.99-1.03)
Children ≥14 years (units=-5%)	1.01 (0.99-1.03)	1.01 (0.99-1.03)	1.00 (0.98-1.02)	1.01 (0.99-1.03)	1.04** (1.02-1.06)	1.04** (1.02-1.06)	1.02* (1.00-1.04)	1.02** (1.00-1.04)

Note: a) Benefits includes: disability, incapacity, income support, job seekers, pension; b) Core programmes include: EEC, Installers, Fuel Poverty and Home Energy Survey. N.B. In this paper HEED is treated as a 'census' level dataset (i.e. a survey of all households). As such, the reporting of inferential statistical tests such as p-values and confidence intervals is incorrect in this context. Confidence intervals significant at ** $p<0.05$, *$p<0.01$

6.4 Discussion and conclusions

This section offers a discussion and conclusions of the method study findings. The purpose is to consider the implication of the findings in terms of the association of the uptake rate and neighbourhood characteristics and what further hypotheses might be postulated and conclusions drawn.

6.4.1 Energy efficiency uptake

Over the period of interest there has been a large number of reported energy efficiency retrofits in the English housing stock, as shown in HEED. Most retrofits have followed a varying but upward trajectory of installations since 2000. The database shows that reported uptake incidence for loft insulation and cavity insulation are steady post-2004, with a minor decline in cavity filling in 2007. Reported condensing-boiler installations have grown since 2004, this growth likely reflecting both the natural replacement rate of boilers, estimated at 1 million/year (UK CCC, 2012), and also the change to building regulations that required all new and replacement boilers to have a minimum efficiency of 86% (i.e. likely to be condensing) from 2006 (CLG, 2010b). Reported numbers of draught proofing retrofits declined from 2004. Reported double glazing installations increased dramatically post-2002, likely reflecting the source of the glazing data (i.e. FENSA). Those pre-2002 glazing installations reported come from other data suppliers, most notably from Fuel Poverty Schemes (e.g. Warm Front).

The uptake rates of fabric retrofits have been concentrated in the midland and northern regions of England, around the major metropolitan areas. This concentration may be related to the colder climate (shown in the analysis) but is also likely related to the activities of several government schemes in those areas. For example, EAGA and CES, the Warm Front providers, were most active in these areas (EAGA, 2004), which may reflect the generally higher occurrence of fuel poverty in Northern regions (DECC, 2013b). Heating retrofits were also concentrated around major metropolitan areas in England, but not specifically northern cities. Since 1998, the government has required that all landlords have their gas installations inspected yearly (HMSO, 1998); as such the concentration of boiler upgrades will reflect the replacement of defective units and householders' decisions to upgrade, but also the number of rental properties for which regular inspections take place. Although there is a visible concentration of heating measures in areas marked with flats, it is unclear whether regulatory controls around heating system maintenance is truly driving replacement of old boilers in the landlord market or whether the combined trend of high uptake for owner-occupiers and lower uptake in areas of high flats is due to targeting of fuel poverty schemes (which did not apply to socially rented dwellings).

6.4.2 Representativeness of national levels of energy efficiency uptake

HEED has acted as a *de facto* repository for data on many of the energy efficiency installations that have taken place in England over the study period. The comparison between the uptake of a selection of energy efficiency installations in HEED and estimates in the EHS over the 2000 to 2007 period shows some striking similarities and differences. The number of condensing boiler installations and loft insulation to ≥150mm installations are very similar, especially post-2003. By comparison, however, there is a difference between the reported uptake of cavity wall filling over the period. The shape of the curve for cavity walls uptake is similar between the EHS and HEED, but HEED contains approximately 250,000/year more cases. The shape for reported double glazing installation in HEED shows a strong increase post-2002 following the introduction of accredited installation requirements. EHS, in comparison, shows a decline in the additional number of dwellings with a coverage of double glazing in more than 80% of all windows. In trying to reconcile these two seemingly divergent trends, it may be that the double glazing installations in HEED are being carried out in homes that already have a high proportion of double glazing.

6.4.3 Efficiency uptake and neighbourhood characteristics

Across all sources of data in HEED, there is a strong relationship between having a high uptake rate of energy efficiency retrofits and neighbourhood income levels. This effect is particularly evident in areas with low median incomes, which see a high likelihood of being in the highest quintile of energy efficiency uptake incidence rate under Fuel Poverty related schemes. It is also found that, as the proportion of dwellings that are owner-occupied increases so too does the likelihood of having a higher uptake rate. This suggest that home owners, who have more control over the maintenance and operation of the dwelling, are either seeking or accepting more energy efficiency retrofits. Under the fuel poverty schemes, the trend showing that a higher proportion of dwellings with occupants on benefits increases the likelihood of being in a higher uptake group also reflects the relationship with lower income households. There was a small positive effect on uptake related to the number of older persons under the energy efficiency commitment, and a small positive effect on the number of children under the fuel poverty schemes. These relationships support the notion that over the eight-year period from 2000 to 2007, many energy efficiency retrofits have been focused on households that cannot necessarily afford the measure but who have the autonomy to decide on its installation.

Going forward, the results of this study have shown that the uptake rate of energy efficiency retrofits is lower in neighbourhoods with middle and high incomes and also in the rental market. While it may be that the higher-income neighbourhoods are possibly more able to afford investment in energy efficiency retrofits, it would appear that their participation is lower than low-income neighbourhoods who have had government support. The *Green Deal* policy has sought to target this middle-income group by allowing households to avoid having to raise the upfront capital investment required for energy efficiency retrofits by offering a scheme through which the retrofits are paid for through energy bills (DECC, 2012f). The concept is that the fuel savings will go to pay back the cost of the installation. This programme is hoping to cater to those middle and higher-income households that may have interest in improving the energy efficiency of their home but who may not want (or have access) to raise the capital to invest in their home. In terms of the rental market, the Green Deal's focus on the household paying the energy bill being responsible for the long-term cost of the measure means that landlords who may otherwise avoid investing in retrofits that have no direct benefit to their income from rent may be amenable to consent to improvements taking place in their properties - although consent for a retrofit cannot be reasonably withheld.

Understanding the level of uptake of energy efficiency retrofits and how levels vary with associated household characteristics is important for the delivery of future retrofits. Further studies on past programmes will help to understand household responses to particular policies and initiatives, which may improve the evidence base for future policies.

6.4.4 Study limitations

There are retrofits that likely fall outside of these two reporting activities, such as DIY or grey-market installations (e.g. cash-only installations), which may not be included in HEED, unless reported under a home survey. DIY is most likely to have an effect on the number of reported loft insulation and draught-proofing retrofits, as these are easy for the homeowner to carry out. DECC estimates that under the EEC Schemes (i.e. 2002 to 2008) approximately 47 million m^2 of loft insulation was installed as DIY. Assuming an average loft space of $50m^2$ and a 10% wastage factor[10] this could mean that approximately 0.12 million installations per year were carried out during the period. The other retrofits, i.e. boilers, glazing, cavity filling, hot water cylinder insulation would very likely not be undertaken by dwelling owners, but rather by builders and technicians. The reporting of these retrofits may be affected by grey-market activities that attempt to avoid paying taxes or are not certified. These potential sources of selection bias in the reporting will have differing effects on the results. If retrofits are consistently under-reported for certain types of house or household and/or in certain geographic areas then the ratio of the odds would be affected by shifting the balance and changing the relationship with those neighbourhood factors found to be associated with the uptake of retrofits. Another form of bias could be from householders providing incorrect information during surveys, i.e. recall bias. However, given the population size there would need to be a large change in the reported retrofits to have an effect on the odds ratios. For the purposes of this study these biases are treated as being randomly distributed throughout the housing stock population.

There are a number of reasons for differences between the reported national energy efficiency database and national estimates of energy efficiency retrofits from surveys. First, the EHS is a constructed so as to be representative of house features and, although randomly selected from postal addresses, it is possible that the survey and its weightings may not be representative of energy efficiency retrofits installed in England over the period of interest. Also, many of the government-schemes have been focused on lower income areas and as such these types of houses and households will be more fully represented in HEED. There could also be differences between the compared EHS and HEED variables for the selected efficiency retrofits, however a great deal of care was taken to avoid this. The comparisons

[10] This methodology described follows that of DECC, set out in Table 3.21 in the UK Energy Consumptions Statistics 2012 update.

are strongest for those retrofits with only a single item per dwelling, for example boiler replacement or cavity wall filling.

HEED is not a continuous registry and as such differences may also occur due to the periodic nature with which the database is updated. This may mean that the data used in this study is not complete and there may be retrofits not yet accounted for, although every care was taken to ensure that the data in HEED was complete by avoiding later years (i.e. post-2007) for which reporting may be lagging.

6.4.5 Study Conclusions

Over the 2000 to 2007 period, there is evidence that those government programmes targeting vulnerable households or those living on benefits have succeeded, as areas with lower incomes show higher rates of uptake. In addition, the uptake rate is higher in areas with higher proportions of owner-occupied dwellings, suggesting that decision-making autonomy is an important factor in the seeking or acceptance of efficiency retrofits.

The research highlights the hypothesis that there exists a misalignment of landlord and tenant benefits related to energy efficiency retrofits, showing that lower rates of uptake are observed for heating retrofits in areas with flats. A similar trend is seen for fabric retrofits. The findings suggest that the combination of flats, which would likely have a higher proportion of rented tenure (social and private) may play a role in lower uptake rates due to the way retrofit programmes target households that can make decisions on accepting retrofits. Although there is little difference in the trend between heating and fabric uptake among flats, it is likely that many retrofits aimed at individual home owners would be inappropriate for these dwelling types, which may have central heating systems and difficult to insulate construction forms. Although requirements exist for landlords to have their gas heating system checked annually, which does not apply to owner-occupiers, the uptake rate is lower than predominantly owner-occupied areas. It may be that with annual checks, the boilers are being repaired rather than replaced. It is unclear whether the regulatory mechanisms around boiler inspections for landlords is the cause of a lower uptake rate due to repair, or whether other causes are behind the low rate of replacement in areas with high proportion of flats, such as landlord resistance or heating system types. Although regulation may be effective, it would be difficult to roll such a mechanism out into the owner-occupied dwellings unless the ownership structure of boiler units changes over time. For example, a structure based on service contracts, where owner occupiers rented boilers or hot water heaters, would require some form of maintenance standard over the life of the contract.

6.5 Critical discussion of the ecological study approach

This section critically discusses the use of the ecological study as a method of examining the uptake of energy efficiency. First, it examines the limitations of the study method in terms of different forms of bias, the measure of association, the variables used and the available information. Concluding thoughts on the method then follow.

6.5.1 Limitations of the study method

Ecological studies suffer from various forms of bias including: ecological bias, confounding within group classification problems and others such as lack of data, colinearity, and migration across groups (Morgenstern, 1995).

Ecological bias occurs when heterogeneity is present in both the outcome variable and the putative influencing variables and the individual level that are otherwise hidden at the group level. For example, the above findings suggest that the northern regions have a higher level of uptake. This could have been related to the fact that installers took part in programmes more than in other regions regardless of the characteristics of the house. It could also happen that there was a higher level of owner occupiers living in those areas, and because there was more activity and therefore more retrofits being offered, there was a higher level of uptake. Thus, it may not have been that owner-occupiers sought more interventions; rather they may have simply been offered more. In this method study, it was not possible to determine the level of such bias in the data, however because the study is meant to develop hypotheses to test, it is felt that such issues can be tackled subsequently and by more detailed studies. Morgenstern (1995) suggests that strategies to reduce ecological bias include the use of smaller units of analysis, for example postcode sector level (i.e. 25 homes) rather than LSOA level (i.e. ~500 homes).

Misclassification within groups at the individual level is another potential problem. For example, if in the northern regions houses might have been more regularly classified as owner-occupied as a default rather than rental. More likely, misclassification may occur in terms of dwelling features at the individual level (such as age type or size), which could result in associations that are not real. It is not believed that this is a problem for household details as these are drawn from census level statistics and did not rely on the individual record level details of the dwelling or household. However, such effects could be more easily present in householder reported retrofits (compared to those made by surveyors or even installers). Confounding is another possible source of error in the analysis. A confounding variable is one that associates with the outcome as well as the putative influencing variables of interest. Putative confounding variables were controlled for in this study, for example income is known to have an effect on the uptake of energy efficiency retrofits and also the targeting of public assistance programmes. In the method study, income levels of the neighbourhood and programme type were stratified, showing differences in the effects.

Lack of relevant information can cause error in associations. For example, there is no common measure of energy efficiency levels across the UK at the LSOA scale. It may be that some of the areas found to have a low level of uptake were in fact already at a higher base level of energy performance. Until such information is available it is difficult to characterize or control for this effect. A further form of bias over the period of study could be changes at the individual level that occurred between the groups being analysed. Because the information being used is a snapshot, the effect this might have is to change some of the features that were associated at the neighbourhood level over time. As with misclassification, this form of bias is difficult to control for. As a result, judgment is required with respect to the study findings.

Finally, the outcome measure used was the uptake as a rate of interventions over the period, divided by the number of houses at a given point in time. The rate is a continuous variable across all LSOAs. In some circumstances it may be appropriate to select different thresholds above or below the population groupings. In this study the interest was in examining household and dwelling features associated with high uptake rates versus low across England in order to understand what neighbourhood characteristics future policies might need to consider for targeting. For this reason it was felt the grouping was appropriate in this circumstance.

6.5.2 Study method conclusions

Although there are various forms of bias that ecological studies suffer from, overall the ecological approach to studying the uptake of energy efficiency interventions was both plausible and appropriate using this technique. The availability of neighbourhood level information on household and dwelling characteristics made this study approach particularly attractive. The study was able to meet its main objective, which was to investigate differences between populations in terms of uptake and identify those variables that may be influencing the uptake rate, at the neighbourhood level. It also provided the means to generate hypotheses relating the uptake of energy efficiency retrofits that could be the subject of further investigations.

6.6 Summary

This chapter investigated the application of the first proposed method study, a longitudinal ecological study examining the rates of uptake of energy efficiency retrofits in England from 2000 to 2007. It sought to examine whether there were differences in uptake between different dwelling and household characteristics and uptake routes.

In addition to the findings of the applied method study, the chapter also included a discussion of the study design in terms of its appropriateness to the research question and its limitations. The following chapter will describe a method study using a cross-sectional study design to explore the variation in energy demand among the British dwelling population.

Chapter 7 Cross-Sectional Study

"Home energy efficiency and energy demand in UK dwellings"

Chapter Introduction

Using a longitudinal ecological study, Chapter 6 provided an investigation into the uptake of energy efficiency interventions in England for the period 2000 to 2007. That study provided an example of an observational style study using more commonly available aggregated information for the purpose of generating hypotheses.

In support of the UK Government's ambitious retrofit programme, comprehensive and high-quality data on the energy efficiency of buildings and their related energy demand is critical to supporting and targeting investment in energy efficiency. The aim of this chapter is to extend the epidemiological approach through the application of a study design that uses more detailed population data on individuals. In the previous Chapter, HEED was used to examine the uptake of energy efficiency over a selected period. In this chapter, HEED is first assessed for its representativeness in terms of Great Britain's dwelling characteristics. Then, HEED is used to examine the variation in gas and electricity demand among the UK population at the dwelling level for different dwelling characteristics and energy performance levels. The study is designed as a descriptive cross-sectional sample of dwellings in Great Britain in 2006. The objective of this observational study is to identify putative dwelling and household factors that influence energy demand.

The chapter comprises five parts: a) an introduction to the research problem along with relevant background and a description of the key concepts and features of the selected cross-sectional study design; b) a description of the methods specific to the study design and the research problem being investigated; c) a presentation of the study results; d) a discussion and conclusions of the findings specific to the study; and e) a critical appraisal of the study approach, its strengths, weaknesses and limitations.

7.1 Research problem

The UK government has identified the residential building stock as potentially being one of the most cost-effective and technology-ready sectors to substantially reduce greenhouse gas (GHG) emissions over the next decade (DECC, 2012b). Achieving these reductions in practice requires detailed information on the state of the existing stock and the ability to measure and track the energy demand of dwellings, particularly those that have been the subject of energy efficiency retrofits. Basic information on the energy demand of the housing stock, its distribution, variation and influencing factors, is a prerequisite requirement for identifying patterns and problems. Without this information it is difficult to develop practices to control or manage energy demand in line with those pathways set out for achieving GHG reduction commitments. Further, the successful delivery and uptake of efficiency measures in order to achieve key policy objectives for energy demand requires an empirical foundation built on high quality data. In particular, continuous collection of such data is essential for the evaluation of past programmes and the development of future evidence-based policies.

7.1.1 Research question and aims

The datasets in HEED represent 'in action' data, i.e. the product (and by-product) of a range of activities that are centred on dwelling energy efficiency. Its continual collection over the past 15 years has created a large individual-level database, detailing and tracking retrofit activity in the housing stock. Linked to data on energy demand practices, these population databases offer a rich resource from which to draw together evidence on energy performance and the uptake of energy efficiency measures, along with changes in energy demand associated with such measures.

In Chapter 6, it was hypothesised that HEED contains many (if not most) of the energy efficiency measures carried out under government programmes or through certified installers and therefore presents an opportunity from which to develop an energy efficiency evidence base for policy development and evaluation. HEED, connected to energy demand meter data, also provides an opportunity to examine what patterns and variation exists among the housing stock to give more suitable baselines for energy demand.

The research questions being tackled in this study are: does HEED provide a suitably representatives sample of the British housing stock, and thus is it a suitable resource to draw samples for analysis; and, what are the patterns in energy demand associated with different dwelling characteristics and energy performance levels; and, do these patterns persist over time.

The aims of this method study are to: a) describe the HEED data, in particular to assess its overall representativeness as compared to other housing data for Great Britain (GB); b) to describe the differences in energy demand (gas and electricity) of the HEED housing stock, segmented by built form characteristics and level of energy efficiency; and c) to determine the change in energy demand associated with the presence of energy efficiency interventions as they relate to changes in energy demand for a selected period (i.e. 2005 to 2007).

7.1.2 Context and background

Between 1970 and 2008 estimated per capita energy demand for lighting and appliances increased by 88%, meanwhile space heating is estimated to have peaked in the 1980's and has declined by approximately 8% per capita (DECC, 2012c). Total delivered energy demand in dwellings has grown by 30% during the same period, peaking around 2004. In 2010, domestic (i.e. residential) delivered energy accounted for approximately 33% (490 TWh) of total GB energy demand by final consumption, of which gas and electricity accounted for approximately 70% (344 TWh) and 23% (113 TWh) respectively (DECC, 2013c). Figure 19 shows an estimate of the total residential demand by service type and fuel (DECC, 2010b). The majority of residential energy demand is for space and hot water heating (78%) with the remainder for appliances (16%) and cooking (3%)[11].

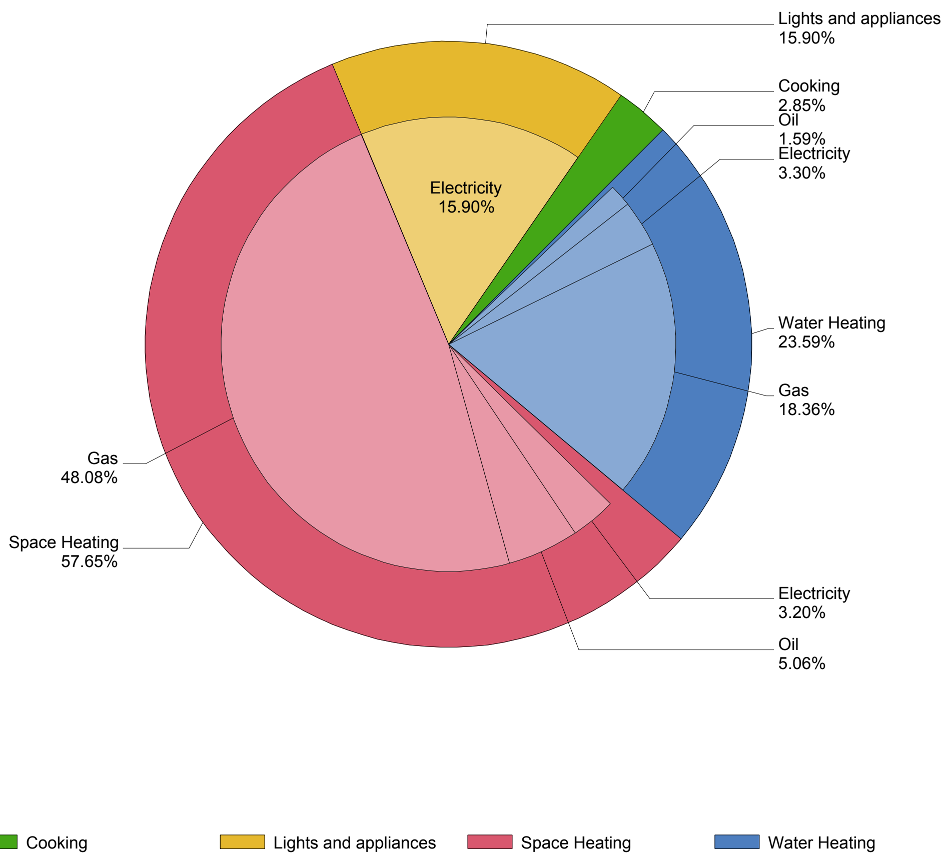


Figure 19 - UK residential fuel by service demand for 2010

Since 1970, estimates of the average UK home energy efficiency, as defined by the Standard Assessment Procedure (SAP) 2005, have risen from 17.6 SAP points in 1970 to 54.7 SAP points in 2010 and the mean heat loss coefficient of dwellings is estimated to have fallen from 376 W/K to below 286 W/K (Palmer and Cooper, 2013). This increase in efficiency has

[11] Residential energy demand by service type is estimated from DUKES data, national totals, and Domestic Energy Fact File data, service fractions. Renewable energy is not included. Services of Fuels <1% of total are not shown but are accounted for in the total.

largely been attributed to the increased uptake in whole house heating systems, more efficient boilers, improved glazing, loft and cavity insulation, and fuel switching to electricity. Per dwelling gas demand has fallen by 13% between 2005 and 2008 (and 25% in 2011); outdoor temperature, price and improvements in efficiency are cited as reasons for this decline (DECC, 2010a).

7.2 Method

This method section describes the study design used in the method study, in term of its features and the concepts that it employs. This is followed by a brief description of the data and the approach used to examine the stated research question.

7.2.1 Cross-sectional study design and key concepts

A descriptive cross-sectional study was selected to investigate the relationship between dwelling characteristics, energy efficiency and energy demand. Cross-sectional studies typically describe outcomes of interest for a population at a single point in time and are often used to examine the prevalence of a particular outcome or condition (Bailey et al., 2005). Information on the frequency and distribution of an outcome within a defined population is the hallmark of a descriptive cross-sectional study. As with other forms of study, missing information or non-response can be a problem when looking at comparing differences between groups. Like ecological studies, cross-sectional studies are typically used for initial investigations and for generating hypotheses for further tests. One of the key features of cross-sectional studies is the representativeness of the sample population against the target population for which inference is to be made. Cross-sectional studies may make use of many different sources of information and data collection, for example interviews and questionnaires, or routine data. When using routine data, such as HEED, issues related to the initial reasons for collection and therefore suitability for study are particularly important. Routine data are typically designed for other purposes and may not contain all the desired information. The main advantage of cross-sectional studies are that they are relatively straightforward in terms of identifying prevalence of putative influencing factors and frequency of outcomes within a population. Successive cross-sectional studies can show changes in these factors over time and their frequency, for example the changes in energy efficiency levels and household features shown by Tovar (2012). Typically cross-sectional designs are not used to study causal associations. Rather, they are used to investigate patterns of outcomes and population characteristics.

This method study is looking at energy demand levels and differences across dwelling characteristics and energy efficiency levels, and is in two main parts. The first is a comparison of HEED for Great Britain against other national cross-sectional surveys, including the EHS, the Scottish Household Conditions Survey, and the England and Wales VOA council tax property data. This comparison evaluates the representativeness of HEED. The second component examines the differences between dwellings that received intervention measures and those that did not, and goes on to describe energy demand by selected dwelling features and levels of energy efficiency.

7.2.2 Data sources and variables

The two main sources of data used in the analysis were annualised meter point gas and electricity data from energy suppliers and the Homes Energy Efficiency Database (HEED). The gas and electricity meter point values were derived from individual meter readings, via aggregators of the data from gas and electricity suppliers (described in Chapter 5).

Gas (non-daily) meters and electricity (not half-hourly) meters for 2004 to 2007 were used in this method study. Both the gas and electricity data underwent a cleaning process to remove or identify potentially erroneous data points, such as negatives and dummy values (e.g. '1' values). In this study, the dataset that removed erroneous data points was used in all energy analysis. Further information on DECC energy data is provided in Chapter 5 and Appendix B.

The sample of HEED used in this method study contains information on the presence of energy efficiency retrofits and a variety of dwelling characteristics for dwellings in England, Wales, Scotland. The extract of the database in February of 2009 used in this study contained approximately 11.5 million distinct home identifiers. Approximately 2.7 million homes appear in at least two programmes (i.e. source datasets) and 1 million in three programmes, while the majority (7.2 million) are present in only one programme.

7.2.2.1 HEED and energy demand

For this study, a dataset containing all matched HEED dwellings and related annualised gas and electricity values for the period 2004 to 2007 was used; Table 15 shows the number of records contained within the source data sets. Note that the numbers of records in electricity and gas represent all meters in Great Britain, both domestic (i.e. residential) and non-domestic, and that the number of records for electricity meters includes those on a time-tariff (i.e. these meters have two records each for on and off-peak time). The two time tariffs are subsequently summed together for a single annual value. Also, the 2007 gas demand is for homes in HEED only and not the whole UK – the remaining data was not made available for use in this research. For those comparisons between HEED and non-HEED energy demand, 2006 data was used. Comparisons of energy use and for installed efficiency measures were based on 2007 data in order to capture a longer time period and more interventions.

Table 15 – Count of records in data sources used in HEED and energy analysis

Data	Records
HEED – Unique Homes in database	11,440,132
HEED – Homes matched with Electricity†	11,685,235
HEED – Homes matched with Gas	9,785,503
Electricity 2004	34,449,299
Electricity 2005	34,660,002
Electricity 2006	35,054,514
Electricity 2007	35,047,989
Gas 2004	21,243,433
Gas 2005	21,994,051
Gas 2006	22,265,312
*Gas 2007**	9,785,500

† Note the number of matched electricity records exceed HEED records due to multiple meter matches; * 2007 gas demand is present for those meters connected to HEED only

7.2.3 Analysis approach

First, the analysis sought to determine how representative of the British (i.e. England, Wales, Scotland) housing stock the meter-matched HEED sample was for a selection of key variables, i.e. age, type, tenure, size and location. This was done by comparing HEED with three other databases: the 2008 English Housing Survey (EHS), the 2007-09 Scottish House Conditions Survey (SHCS), and the 2010 Valuation Office Agency (VOA) Council Tax Property Attributes database for England and Wales. Together these data sources provide more or less complete coverage of the housing stock of Great Britain. Chi-square tests for goodness-of-fit at a 95% confidence interval were used to determine whether the sample data (i.e. dwellings in HEED) was consistent with the hypothesized distribution (i.e. British dwellings). For computational purposes, a 10% randomly selected sample of approximately 1.2 million dwelling records representative of HEED was used for the population comparison, rather than the full HEED database (i.e. 11.5M) (see Appendix C for a Chi-square test for the HEED sample and full HEED database).

The 2008 EHS was used because the collection period aligned with the last year of HEED data, which is also the case for the 2007 to 2009 SHCS. The VOA holds data on both England and Wales and is revised every year; therefore the latest extract was used. Both the EHS and SHCS provide a factor with which to weight variables in order to represent houses or households in England or Scotland; the houses weighting was used for comparison. No weighting was required for the VOA data. With respect to the potential changes in the stock since 2008, approximately 268,000 dwellings were built in 2009 and 2010 (approximately 0.1% of the total GB stock) (CLG, 2010c). Further details of the housing surveys are provided in Appendix D.

The EHS, SHCS, VOA and HEED were all developed for different purposes and were not collected using a common format. As a result only some variables can be compared and in some cases variable classes were banded together to create broadly comparable data categories (e.g. dwelling type and number of bedrooms). Dwelling age is collected using a different age band for each survey and was too complex to band, as dwelling completion rates fluctuate from year to year. Therefore, for dwelling age, a chi-square goodness-of-fit test was not performed and instead the data was visually compared.

Dwellings in HEED with metered energy demand were compared to dwellings not present in HEED (or non-HEED) for the period covering 2004 to 2007 for gas and electricity. Using the date of when the details were collected, it was possible to compare those groups of dwellings across the gas demand period based on when they entered HEED, and therefore were likely to have received an efficiency intervention, to the non-HEED dwellings. For example, a dwelling could enter HEED due to an intervention taking place in 2006 but would also have been connected to the preceding two years of demand (i.e. 2004 and 2005) and the subsequent gas year (i.e. 2007). Changes in gas and electricity demand within the two groups would be broadly affected by a number of exogenous and endogenous drivers; such as fuel price and demand, energy efficiency, income and the ability to pay, behaviour and others; but effects outside of energy efficiency were not investigated.

Gas and electricity demand was analysed for dwellings in HEED by their physical characteristics (i.e. age, size, type) and levels of intervention (i.e. loft insulation level, cavity insulation, glazing type). The 10% randomly selected HEED sample was used for the analysis. Gas and electricity demand were normalised by number of bedrooms as a proxy for dwelling size in an attempt to explore a size effect. For dwellings with 5+ bedrooms, an arbitrary value of 5.5 was used for normalisation.

7.3 Results

The results present the findings from the two analysis strands: first, HEED dwelling characteristics; and second, HEED energy demand by dwelling and energy efficiency characteristics.

7.3.1 Comparison of HEED dwelling characteristics

The characteristics of dwellings in the selected 10% HEED sample are compared against representative samples for England, England and Wales, and Scotland. Table 16 and Table 17 provide overview statistics for the selected compared variables. The results show that the HEED data is not statistically representative of the English and Welsh stock for the selected variables. In all cases of comparison the hypothesis that the compared variables of the HEED data set are the same as those of the English Housing Survey and VOA Council Tax (i.e. all p-values <0.0001 at a 95% confidence limit) is rejected. However, the tests do show that HEED contains a representative sample of the Scottish housing stock, based on the tested variables.

While the analysis of the populations represented in the HEED data does not support the hypothesis that the sample is the same as the other datasets that represent the housing stock of England, and England and Wales, it is not necessarily the case that HEED cannot be used to describe housing energy efficiency demand for those groups. Also, it is known that small divergences are shown to be significant for Chi square goodness-of-fit tests for large samples and that comparisons are often made through visual inspection. A visual comparison of the data suggests that there are small differences for most categories, but many are within 1%. As such, caution should be applied where findings from HEED are interpreted and generalised for the housing stock as a whole.

Overall, in the English and Welsh component of HEED, 'dwelling type' shows fewer flats and more semi-detached houses. There are fewer privately rented dwellings and more socially rented dwellings, likely reflecting the emphasis of the government and energy supplier programmes to target areas of high-deprivation and low-income groups. In terms of geographic coverage, there are fewer homes in the southern regions of England. Despite the targeting of the programmes, given the number of dwellings represented in HEED (approximately 50% of all GB dwellings), HEED does compare *relatively* well to the representative housing stocks of Great Britain. The HEED data can be said to represent the Scottish housing stock, which likely reflects the collection process and inclusion of a proportion of building performance rating data (i.e. Energy Performance Certificates).

Age is compared graphically rather than statistically, due to the difference in category bands. Figure 20 shows that there are more homes in the 1967-82 period and fewer 1990+ homes than in the English and Welsh stocks. The decreased number of post-1975 dwellings may reflect the concentration of cavity wall insulation activity in pre-1976 dwellings.

Table 16 - HEED (England) dwelling characteristics compared to EHS

England[a]	HEED 10% (n)	HEED 10% (%)	EHS 2008 (%)
Dwelling Type			
Flat-Maisonette	96,975	17.2%	18.6%
Bungalow	54,837	9.7%	9.4%
Terrace	141,109	25.1%	28.6%
Semi-detached	183,309	32.6%	26.0%
Detached	86,434	15.4%	17.4%
Chi square (X^2)			12961.22
Degrees of freedom (d.f.)			4
Significance level at 95%			
p-value			<0.0001
Dwelling Tenure			
Social rental	156,195	21.8%	14.8%
Private rental	67,499	9.4%	17.7%
Owner-occupied	493,481	68.8%	67.5%
X^2			51585.46
d.f.			2
p			<0.0001
Dwelling Size (Bedrooms)			
1	71,315	12.6%	9.1%
2	142,619	25.3%	27.1%
3	267,307	47.4%	44.2%
4	58,600	10.4%	15.5%
5+	24,333	4.3%	4.0%
X^2			19219.87
d.f.			4
p-value			<0.0001
Dwelling Region			
North East	70,049	6.2%	5.1%
North West	159,820	14.2%	13.6%
Yorkshire and The Humber	120,624	10.7%	10.6%
East Midlands	91,541	8.1%	8.8%
West Midlands	116,000	10.3%	10.5%
East of England	109,080	9.7%	10.9%
London	132,433	11.8%	14.2%
South East	161,845	14.4%	15.8%
South West	107,767	9.6%	10.3%
X^2			9810.57
d.f.			8
p-value			<0.0001

Notes: [a]10% HEED Sample, England only

Table 17 - HEED (England and Wales) dwelling characteristics compared to VOA

England & Wales[a]	HEED 10% (n)	HEED 10% (%)	VOA 2010 (%)
Dwelling Type			
Flat-Maisonette	96,975	17.2%	21.9%
Bungalow	54,837	9.7%	10.2%
Terrace	141,109	25.1%	27.3%
Semi-detached	183,309	32.6%	24.8%
Detached	86,434	15.4%	15.8%
X^2			20518.77
d.f.			4
p-value			<0.0001
Dwelling Size (Bedrooms)			
1	71,315	12.6%	11.6%
2	142,619	25.3%	28.4%
3	267,307	47.4%	45.4%
4	58,600	10.4%	11.5%
5+	24,333	4.3%	3.0%
X^2			6798.72
d.f.			4
p-value			<0.0001
Dwelling Region			
North East	70,049	6.2%	4.8%
North West	159,820	14.2%	12.9%
Yorkshire and The Humber	120,624	10.7%	9.5%
East Midlands	91,541	8.1%	8.1%
West Midlands	116,000	10.3%	9.7%
East of England	109,080	9.7%	10.4%
London	132,433	11.8%	13.8%
South East	161,845	14.4%	15.1%
South West	107,767	9.6%	9.8%
Wales	55,073	4.9%	5.7%
X^2			14076.56
d.f.			9
p			<0.0001

Notes: [a]10% HEED Sample, England and Wales only

Table 18 shows a comparison of the Scottish dwellings in HEED and accepts the hypothesis that the HEED sample is statistically similar to the Scottish House Conditions Survey.

Table 18 - HEED (Scotland) dwelling demographics comparison to SHCS

Scotland	HEED 10% (n)	HEED 10% (%)	SHCS 2009 (%)
Dwelling Type			
Flat-Maisonette	29,008	36.6%	36.7%
Bungalow	0	0.0%	0.0%
Terrace	20,334	25.6%	25.5%
Semi-detached	15,905	20.1%	20.1%
Detached	14,062	17.7%	17.8%
X^2			1.2293
d.f.			3
p			0.746
Dwelling Tenure			
Social rental	25,334	27.9%	27.7%
Private rental	9,562	10.5%	10.6%
Owner-occupied	56,017	61.6%	61.7%
X^2			1.5907
d.f.			2
p			0.4514
Dwelling Size (Bedrooms)			
1	11,274	19.3%	19.2%
2	22,321	38.1%	38.1%
3	19,314	33.0%	33.1%
4	3,735	6.4%	6.4%
5+	1,867	3.2%	3.1%
X^2			1.9065
d.f.			4
p			0.753
Notes: a10% HEED Sample, Scotland only			

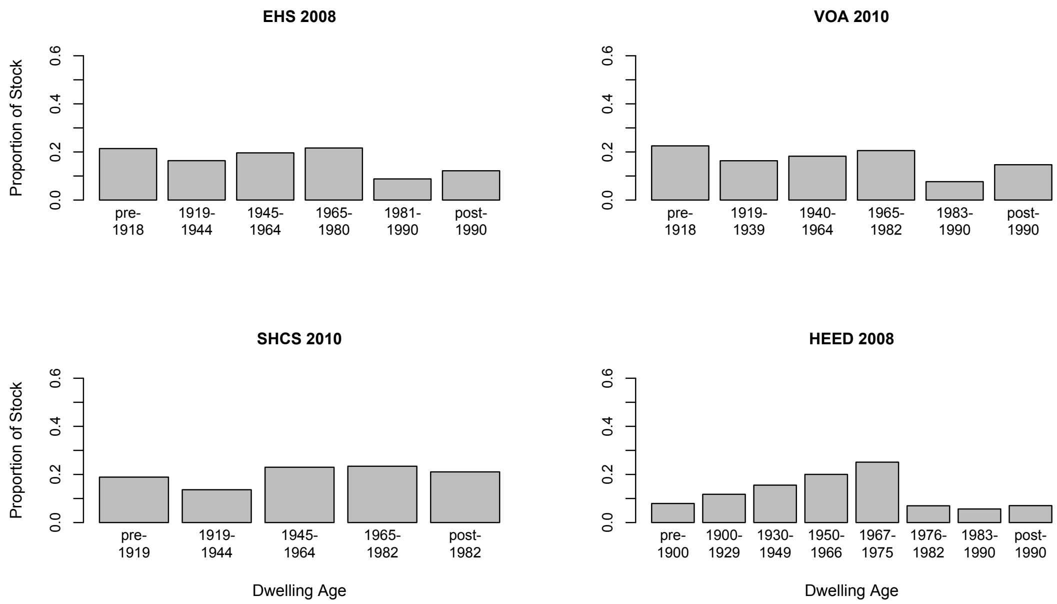


Figure 20 - Housing stock age band comparison

7.3.2 HEED energy efficiency characteristics

Table 19 shows the distribution of a selection of energy efficiency features by dwelling characteristics, as compared to the HEED GB sample. This gives an indication of the coverage for walls, lofts, glazing and heat systems within the selected population and whether there would be any significant population bias expected in any differences found. The differences in coverage by dwelling characteristic appear to be relatively small, although with less coverage of measures in 1967-75 dwellings, and of heat systems in 3-bedroom dwellings.

Table 19 - HEED Stock: Comparison of energy efficiency groups by dwelling characteristic

Dwelling Characteristic	HEED Stock *All*	HEED Stock Energy Efficiency Groups *Wall Type Group*	*Loft Group*	*Glazing Group*	*Heating System Group*
Dwelling Type					
Bungalow	11%	11%	10%	9%	9%
Detached house	19%	20%	19%	20%	22%
Semi-detached house	39%	38%	37%	37%	37%
Terrace house	31%	31%	33%	33%	31%
Dwelling Age					
pre-1900	8%	9%	9%	9%	10%
1900-29	12%	12%	13%	13%	12%
1930-49	16%	16%	17%	17%	17%
1950-66	20%	20%	19%	19%	20%
1967-75	25%	23%	22%	22%	19%
1976-82	7%	7%	7%	7%	7%
1983-90	6%	6%	6%	6%	6%
post-1990	7%	7%	7%	7%	8%
Number of Bedrooms					
1	13%	12%	12%	12%	14%
2	26%	25%	25%	25%	25%
3	46%	47%	47%	47%	43%
4	10%	10%	10%	10%	11%
5+	4%	5%	5%	5%	7%

7.3.3 Energy demand: HEED and non-HEED

In this section, the annualised gas and electricity meter data for the Great Britain (i.e. England, Wales and Scotland) HEED sample is compared against the non-HEED meters. Following this, the gas and electricity use for the HEED stock is described.

7.3.3.1 Gas demand

Table 20 shows that the change in median gas demand in non-HEED meters between 2004 and 2006 is approximately -6.1%. For meters in HEED, the change in median gas demand between 2004 and 2006 is approximately -8.1%. Residential gas demand data is influenced by a long right tail, as can be seen in the <73.2MWh/yr meters gas demand

(Figure 21)[12]. This is an inevitable consequence of the fact that energy demand data cannot be negative but is subject to no well-defined upper limit (other than the very high 73.2 MWh artificial limit). Note also the upward flick in the distribution close to zero demand; dwellings that are unoccupied for part or all of a year may cause this.

Table 20 - Residential gas demand for HEED and non-HEED meters

Profile	Flag HEED	Year	Mean	Median	Std Dev (σ)
Gas[a] - Residential (<73.2MWh/yr)	*Non-HEED*	2004	19,734	18,214	11,137
	n= 8,410,189	2005	19,433	17,877	11,008
		2006	18,625	17,107	10,836
		2007[b]	.	.	.
	HEED	2004	19,623	18,452	9,725
	n= 7,450,540	2005	19,141	17,926	9,511
		2006	18,153	16,958	9,252
		2007	17,468	16,226	9,086

Note: [a]Excludes erroneous data points; [b]Non-HEED 2007 – Gas meter values were only provided for those homes matched in HEED, therefore no statistics are available for this year from the processed data.

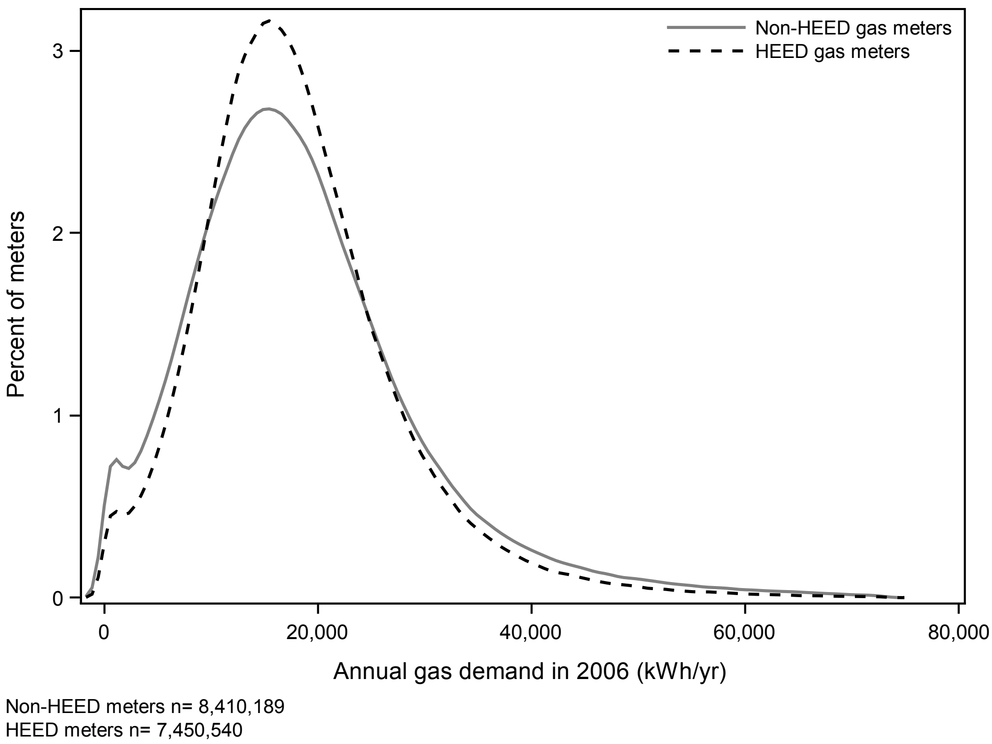


Figure 21 - Distribution of residential gas demand (<73MWh/yr) in 2006 for HEED and Non-HEED meters

[12] For clarity of comparison, the gas and electricity meter data has been plotted using a density function, which is based on the observed data and represents an underlying probability density function to which the population is distributed. For this reason, the line data is smoothed and passes 0, which is artificial. See Appendix B for histograms.

7.3.3.2 Change in gas demand for HEED

HEED contains a time stamp for when a measure was introduced or a survey was carried out for each dwelling. Figure 22 shows meters classified by the home details date, thus entering HEED. It shows that energy demand for homes in HEED with a high likelihood of an intervention in 2005 begin to diverge (i.e. downward slope) from the demands of their non-intervention counterparts in the following year. This is also true for dwellings with interventions in 2006. The change in demand is higher for those dwellings with an intervention within the gas period, with the exception of those entering in 2007, where it is unlikely the gas data would pick up in the change, depending on the reading frequency. Note that this is the bulk trend for all homes in HEED, regardless of the type of measure – more details are provided below on this.

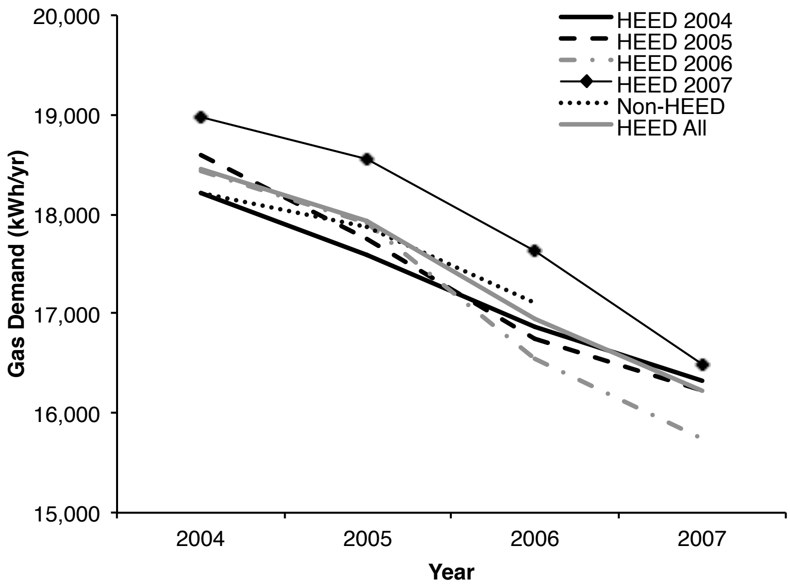


Figure 22 - Gas demand by HEED entry year. N.B. Non-HEED gas data is only available to 2006.

7.3.3.3 Electricity demand

Table 21 shows that the change in the median unrestricted electricity demand in non-HEED meters between 2004 and 2006 is approximately -0.8%. The change between 2004 and 2007 for the same meters is 1.2%. For meters in HEED, the change in median unrestricted electricity demand between 2004 and 2006 is approximately -1.5% and the change between 2004 and 2007 is -0.9%. Non-HEED Economy 7 meters saw a change in median of -5.6% between 2004-2006, compared to -6.2% for HEED Economy 7 meters for the same period (change in medians for 2004 to 2007 is -3.5 and -5.6 for HEED and non-HEED meters respectively).

Table 21 - Residential electricity demand for HEED and non-HEED meters

Profile	Flag HEED	Year	Mean	Median	Std Dev (σ)
Unrestricted[a]	Non-HEED	2004	4,272	3,548	3,304
	n= 9,212,105	2005	4,311	3,551	3,359
		2006	4,231	3,519	3,233
		2007	4,163	3,447	3,230
	HEED	2004	4,023	3,410	2,865
	n= 7,362,544	2005	4,027	3,391	2,894
		2006	3,957	3,359	2,790
		2007	3,888	3,288	2,770
Economy7[a]	Non-HEED	2004	6,960	5,587	5,392
	n= 2,685,662	2005	6,750	5,427	5,225
		2006	6,543	5,275	5,066
		2007	6,675	5,339	5,237
	HEED	2004	6,472	5,069	4,981
	n= 1,735,592	2005	6,199	4,874	4,769
		2006	6,001	4,756	4,593
		2007	6,067	4,749	4,728

[a]Erroneous data points are excluded.

The electricity data (unrestricted and Economy7 meters) is influenced by a long right tail, as can be seen in the distribution of electricity demand (Figure 23). Note that when considering this tail against the gas demand data, electricity meters are classed based on a user type and tariff, whereas the gas data is classified according to consumption. While the long right tail in gas may hold a number of non-domestic users, electricity demand distribution is reflecting actual large domestic users.

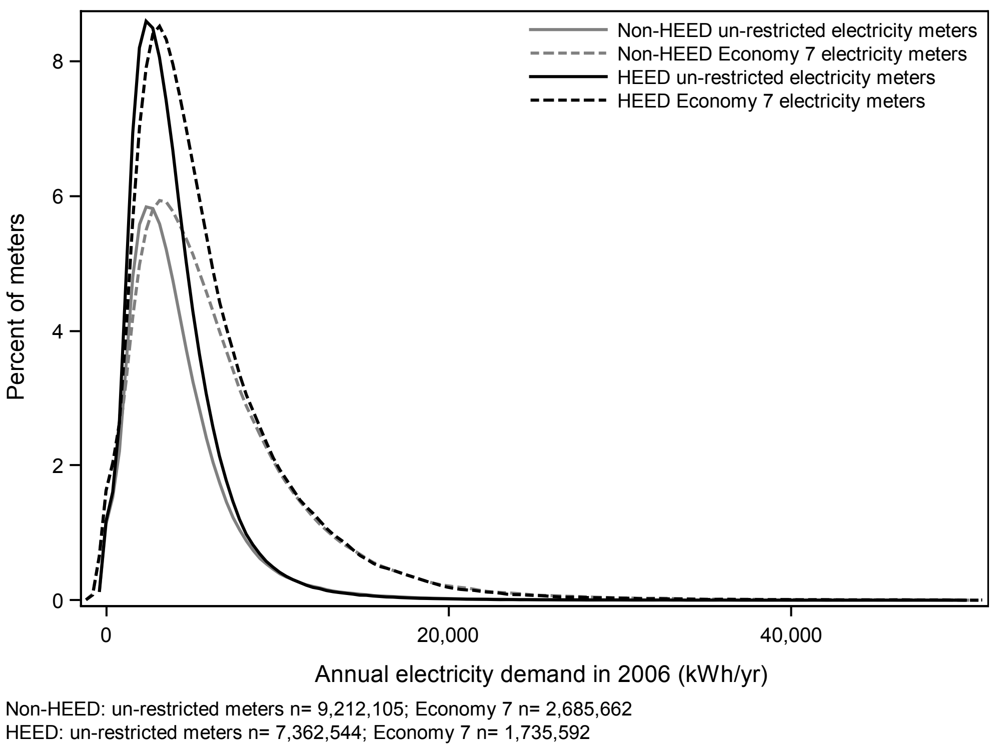


Figure 23 – Distribution of residential unrestricted (ordinary) and Economy 7 electricity demand

7.3.3.4 Change in electricity demand for HEED meters

Figure 24 shows that the year-on-year change for all non-HEED and HEED groups is broadly similar, with non-HEED meters reducing by 0.8% from 2004 to 2007 and HEED meters reducing by 1.2%. Figure 25 shows that change in Economy 7 meters varies more across the period and groups. Note that the Economy 7 demand, which is associated with heating, is not weather-corrected and therefore will be affected by changes in temperature. Note also that the trend change is similar across the groups. The group average change in unrestricted electricity for meters in HEED is a reduction of 3.5% as compared to a reduction of 2.5% for non-HEED meters. Economy 7 meters in HEED broadly show a reduction of around 9.5% from 2004 to 2007 and non-HEED meters show a reduction of 4.1%. Again, note that the Economy 7 is not weather-corrected and this change will reflect weather trends. Note that although the gas weather correction method is known, the values needed to use the method remain unpublished, therefore a weather correction of Economy 7 meters was not attempted.

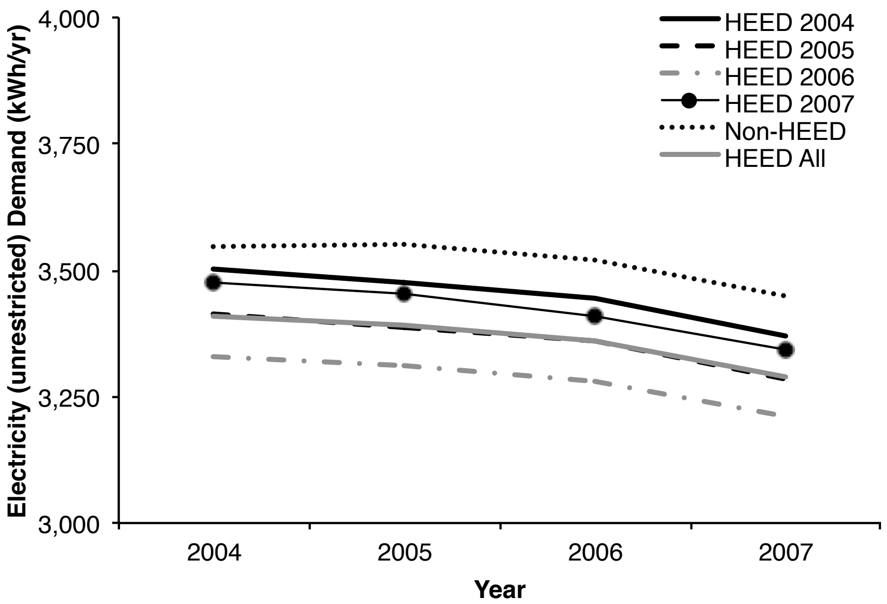


Figure 24 – Unrestricted electricity demand by HEED entry year

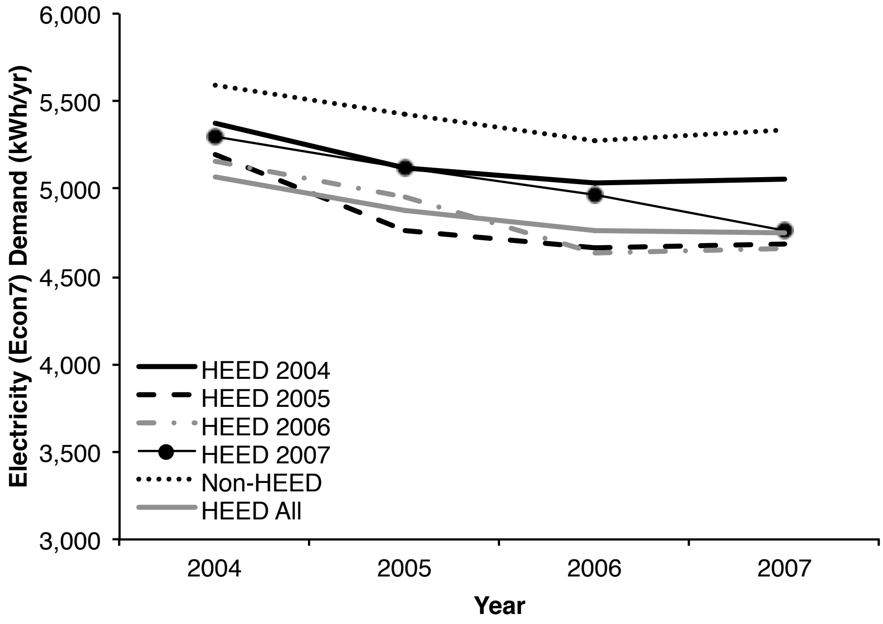


Figure 25 – Economy 7 electricity demand by HEED entry year

7.3.4 Gas and electricity statistics for HEED dwellings

The linked datasets provided an opportunity to tabulate gas and electricity demand by dwelling characteristics. Table 22 and Table 23 provides overview statistics for gas and electricity use in 2006 by a selection of dependent variables. The table shows that older dwellings typically demand more gas and Economy 7 electricity, but that unrestricted electricity demand is very similar in old and new dwellings, with a slight increase in newer dwellings. Detached houses and bungalows record the highest gas demand, with a decline in demand by the extent of detachment; this trend is also true in unrestricted electricity – although terraces seem to use more Economy 7 electricity than semi-detached dwellings, perhaps reflecting Economy 7 as secondary heating for larger dwellings. Median and mean gas and unrestricted electricity demand in private rental dwellings are very similar to demand in social rentals, and owner-occupied dwellings use a third more gas and ~25% more unrestricted electricity. However, median Economy 7 electricity demand in social rental properties is approximately 33% higher than private rentals. Median gas demand increases on average by 22% for every additional bedroom over 1 bedroom. The difference per bedroom is lowest when moving from 4 to 5+ bedrooms (14%) but this is likely due to the banding together of properties above 5 bedrooms as an arbitrary figure of 5.5. Median unrestricted electricity demand increases monotonically from 1 to 4 bedrooms. The increase from 4 to 5+ bedrooms is 12% but is subject to the same caveat as for gas.

Table 22 - HEED Stock: Residential gas demand in 2006 by dependent variables

	Gas[a] (kWh/yr)			
	N	**Mean**	**σ**	**Med-ian**
Dwelling Age				
Missing	575,785	17,333	9,606	16,087
Pre-1900	30,360	18,950	12,121	17,063
1900-29	57,969	18,723	10,763	17,267
1930-49	77,944	17,930	9,169	16,918
1950-66	90,841	16,703	8,567	15,780
1967-75	117,502	16,939	8,661	15,982
1976-82	29,510	15,534	8,408	14,536
1983-90	21,334	15,678	8,754	14,452
post-1990	28,156	16,234	8,677	15,005
Dwelling Type				
Missing	525,816	17,557	9,700	16,300
Flat[1]	70,660	11,557	8,341	10,242
Terrace	140,100	16,004	8,487	14,983
Semi-det.	175,690	17,533	8,276	16,571
Detached	71,521	22,823	10,592	20,992
Bungalow	45,614	17,379	8,527	16,129
Dwelling Tenure				
Missing	379,225	17,538	9,533	16,357
Social[2]	120,802	13,637	7,784	12,964
Private[3]	64,594	13,863	8,485	12,796
Owner	464,780	18,507	9,633	17,186
Number of Bedrooms				
Missing	535,083	17,461	9,714	16,217
1	50,004	12,457	8,541	11,137
2	127,067	14,397	7,737	13,541
3	248,788	17,526	8,261	16,590
4	50,471	23,129	10,503	21,560
5+	17,988	26,292	12,726	24,246

Note: [a]Excluded gas meters = 8,069 due to erroneous values; HEED Sample size is 1,286,372, approximately 20% had no matched gas meter and 7% no matched electricity meter. [1]Flats include purpose built, maisonette and converted; [2]Social includes registered social landlords (RSL) and local authority; [3]Private rental

Table 23 - HEED Stock: Residential electricity demand in 2006 by dependent variables

	Unrestricted[b] (kWh/yr)				Economy 7[b] (kWh/yr)			
	N	Mean	σ	Med-ian	N	Mea n	σ	Med-ian
Dwelling Age								
Missing	512,664	3,796	3,303	3,223	135,471	5,825	5,679	4,502
Pre-1900	31,399	3,881	3,881	3,111	9,703	7,298	6,870	5,561
1900-29	53,366	3,687	3,654	3,098	10,507	6,340	6,377	4,689
1930-49	71,396	3,706	3,084	3,178	14,980	6,186	5,731	4,732
1950-66	83,885	3,484	2,999	2,978	24,317	6,338	6,352	4,906
1967-75	109,336	3,569	3,017	3,086	31,198	6,226	5,710	4,849
1976-82	28,982	3,393	2,804	2,890	8,340	6,159	5,109	4,929
1983-90	19,455	3,474	2,892	2,930	10,128	6,182	4,875	5,082
post-1990	27,808	3,740	3,250	3,235	7,448	6,207	5,479	4,995
Dwelling Type								
Missing	469,962	3,863	3,354	3,284	125,464	5,924	5,664	4,605
Flat[1]	80,964	2,440	3,000	1,967	31,701	5,313	5,857	4,309
Terrace	123,555	3,494	3,033	3,038	29,913	6,364	5,702	4,845
Semi-det.	156,505	3,788	2,911	3,310	32,973	6,341	5,832	4,765
Detached	65,407	4,663	3,683	4,023	20,092	6,898	6,258	5,135
Bungalow	41,898	3,318	2,798	2,784	11,949	6,539	5,913	4,828
Dwelling Tenure								
Missing	344,089	3,801	3,267	3,239	102,485	6,221	5,789	4,819
Social[2]	126,730	3,019	3,136	2,506	35,136	5,960	6,252	4,791
Private[3]	56,245	3,207	3,576	2,651	14,906	4,900	5,779	3,569
Owner	411,227	3,917	3,186	3,347	99,565	6,105	5,601	4,658
Number of Bedrooms								
Missing	476,968	3,806	3,336	3,231	130,776	5,906	5,696	4,558
1	54,634	2,459	2,884	1,934	20,850	5,658	5,903	4,685
2	120,175	2,988	2,791	2,554	35,701	6,102	5,910	4,662
3	223,153	3,807	2,905	3,357	48,354	6,217	5,667	4,637
4	45,824	4,912	3,848	4,358	11,471	6,866	5,778	5,390
5+	17,537	5,589	4,540	4,890	4,940	8,148	7,315	6,171

Note: [b]Excluded electricity meters = 18.190 due to erroneous values; HEED Sample size is 1,286,372, approximately 20% had no matched gas meter and 7% no matched electricity meter. [1]Flats include purpose built, maisonette and converted; [2]Social includes registered social landlords (RSL) and local authority; [3]Private rental

Figure 26 to Figure 28 compare HEED dwelling characteristics (i.e. age, type and tenure) and gas and unrestricted electricity demand per bedroom; the figures give the *mean* gas or electricity use, rather than the preferred median. In order to control the effect that large energy-using meters may have on the results, Tukey's method of determining outliers is used. This method treats any value as an outlier that is greater than the 75th percentile plus 1.5 times the inter-quartile distance, or less than the 25th percentile minus 1.5 times the inter-quartile distance. No data with missing classes is used in these figures. The figures show there is a size effect for electricity (i.e. size and electricity are positively related) but no relationship with dwelling type, age or tenure. Gas demand variation across different dwelling types (excluding bungalows and flats) shows that dwellings with more exposed surface area (i.e. detached houses and bungalows) use slightly more per bedroom. Gas demand by age also shows that older dwellings use more gas, which may be related to their overall level of energy efficiency and/or also reflect large bedrooms. There appears to be only a slight difference between tenure types, with owner-occupied properties consuming more gas per bedroom.

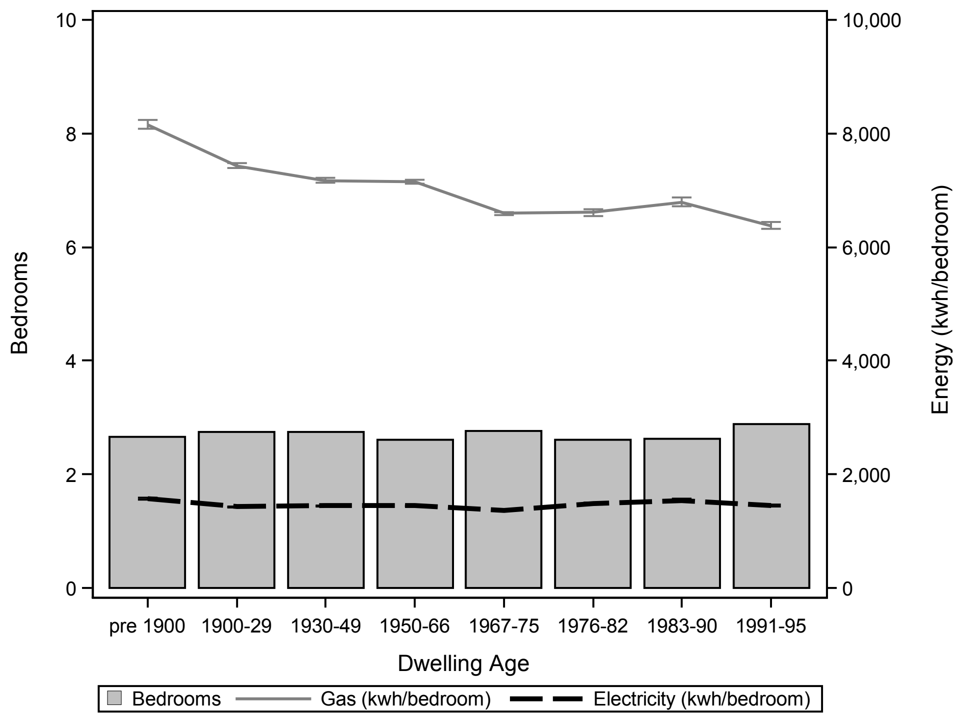


Figure 26 - Mean gas and electricity demand per bedroom by dwelling age

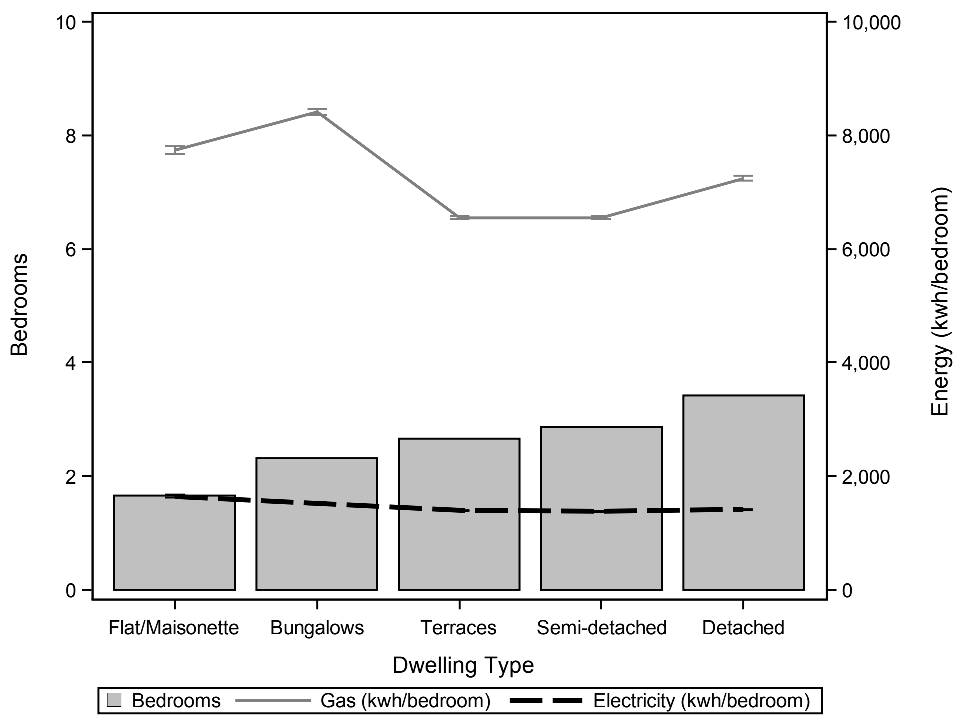


Figure 27 - Mean gas and electricity demand per bedroom by dwelling type

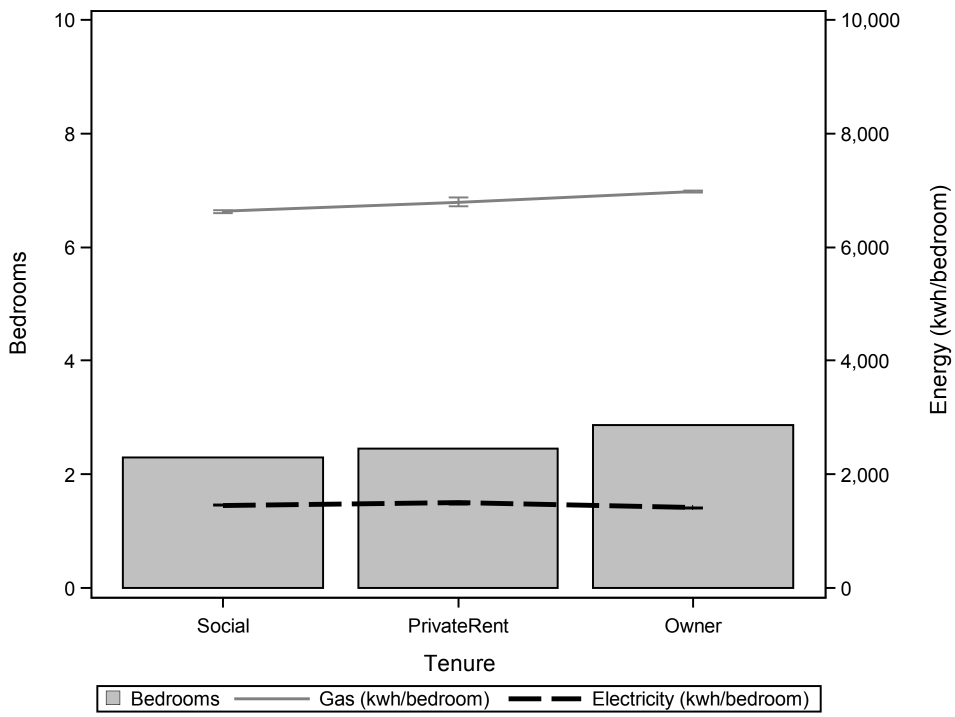


Figure 28 - Mean gas and electricity demand per bedroom by tenure

7.3.5 Energy efficiency characteristics of HEED dwellings

The following section shows the difference in energy demand for varying levels of energy efficiency characteristics (i.e. lofts, wall type, glazing, boiler type) within the HEED data set. Table 24 shows median gas demand by age and dwelling type for loft insulation levels (<50mm, 50-200mm, >200mm) and cavity wall insulation (filled vs unfilled). The average difference across all age bands for dwellings with >200mm of loft insulation is 1.6% less than those with <100mm. Across dwelling types, the average difference between >200mm loft insulation is 6.7% less than for <100mm. The average difference for cavity fillings by age group is 7.9% less than those with cavity unfilled and for dwelling type is 9.4% less than cavity unfilled.

Table 24 – Median annual gas demand by dwelling age and type by loft insulation level and cavity filling

	Stock[a]		Lofts				Cavity Walls[b]		
			Miss-ing	<50 mm	50-200 mm	>200 mm	Miss-ing	Cavity filled	Cavity as built
Gas use 2007	**N**	**Median (kWh/yr)**	**Median (kWh/yr)**				**Median (kWh/yr)**		
Dwelling Age									
Missing	575,785	16,235	16,097	15,712	16,060	16,008	-	-	-
pre-1900	30,360	17,430	16,034	17,311	18,824	18,085	-	-	-
1900-29	57,969	17,593	16,502	17,742	18,256	17,388	-	-	-
1930-49	77,944	17,010	16,153	17,392	17,409	16,960	16,134	18,134	16,185
1950-66	90,841	15,904	14,910	16,473	16,492	15,874	14,387	17,344	15,754
1967-75	117,502	16,011	15,555	16,761	16,881	15,936	15,081	17,228	16,048
1976-82	29,510	14,484	13,324	15,214	15,967	14,516	12,902	15,877	14,723
1983-90	21,334	14,486	13,336	15,137	15,933	14,375	13,000	15,782	14,683
post-1990	28,156	14,950	13,732	15,058	16,024	15,612	13,137	16,060	15,343
Stock	1,029,401	16,201	15,890	16,700	16,887	16,095	16,125	17,227	15,537
Dwelling Type									
Missing	525,816	16,423	16,316	16,523	16,533	14,396	16,631	16,601	14,635
Flat	70,660	10,318	10,072	10,925	10,992	10,402	9,960	10,750	10,006
Bungalow	45,614	15,955	15,858	17,958	17,234	15,608	14,602	18,115	16,268
Terraces	140,100	15,033	14,858	15,362	15,027	14,948	14,964	14,983	14,339
Semi-det.	175,690	16,557	16,163	17,316	16,726	16,517	16,212	17,709	15,904
Detached	71,521	21,012	20,627	21,719	21,132	20,600	21,912	22,531	19,709
Stock	1,029,401	16,201	15,890	16,700	16,887	16,095	16,125	17,227	15,537
		N=	*700,875*	*41,031*	*140,300*	*147,195*	*565,101*	*136,878*	*205,318*

Note: [a]Sample excludes dwellings with no gas meters and erroneous values (256,971); [b]Other wall types have been removed from this sample for the purposes of comparison (solid=116,811 and timber=5,293) and pre-1930's

Table 25 shows median gas demand by age and dwelling type for glazing type (pre-2002 vs post-2002 double) and boiler type (condensing vs non-condensing). The average difference across all age bands for dwellings with post-2002 double glazing is 3% less than those with pre-2002 double glazing. Across dwelling types, the difference between post-2002 double glazing is 4.5% less than pre-2002 glazing. The distinction between pre- and post-2002 double glazing refers to a requirement introduced in the British Building Regulations of 2002 requiring all windows (and replacement windows) to conform to lower U-values. The average difference for condensing boiler upgrades by age group is 8.8% less than those for non-condensing boilers and for dwelling type is 9.2% less.

Table 25 – Median gas demand by dwelling age and type by glazing and boiler type

Gas Use 2007	Stock*		Glazing: Miss-ing	Glazing: Double pre-2002	Glazing: Double post-2002	Boilers: Missing	Boilers: Non-conden-sing	Boilers: Conden-sing
	N		Median (kWh/yr)			Median (kWh/yr)		
Dwelling Age								
Missing	575,785	16,235	15,697	15,896	16,646	16,542	16,010	15,323
pre-1900	30,360	17,430	16,781	17,609	16,818	14,522	18,580	17,633
1900-29	57,969	17,593	16,598	18,008	17,332	16,435	18,790	17,103
1930-49	77,944	17,010	16,158	17,503	16,957	16,231	18,014	16,049
1950-66	90,841	15,904	14,985	16,471	15,853	15,252	16,946	15,039
1967-75	117,502	16,011	15,752	16,541	15,860	16,006	16,836	14,922
1976-82	29,510	14,484	13,510	15,597	14,393	13,514	15,451	13,861
1983-90	21,334	14,486	13,144	15,370	14,765	13,224	15,583	13,970
post-1990	28,156	14,950	13,534	15,544	15,145	14,569	15,367	14,446
Stock	1,029,401	16,201	15,606	16,681	16,482	16,241	17,017	15,330
Dwelling Type								
Missing	525,816	16,423	15,982	16,047	16,738	16,556	16,344	15,521
Flat	70,660	10,318	9,952	10,993	9,594	9,822	11,242	9,800
Bungalow	45,614	15,955	15,656	17,007	15,830	15,933	17,132	15,121
Terraces	140,100	15,033	14,906	15,212	14,519	14,841	15,675	14,429
Semi-det.	175,690	16,557	16,102	17,140	16,293	16,191	17,478	15,767
Detached	71,521	21,012	20,457	21,473	20,345	20,509	21,649	20,226
Stock	1,029,401	16,201	15,606	16,681	16,482	16,150	16,890	15,162
		N=	*462,775*	*201,258*	*315,221*	*576,299*	*232,362*	*200,593*

Note: *Sample excludes dwellings with no gas meters and type1 flags (256,971)

7.4 Discussion and conclusions

This section discusses the method study findings on the variation in the levels of energy demand from 2004 to 2007 and housing characteristics. It also discusses the opportunities an energy and buildings data-framework offers for research, along with the study limitations and the representativeness of HEED against other British housing databases.

7.4.1 Energy demand, energy efficiency and building characteristics

HEED data, when linked to individual annualised gas and electricity meter values, allowed for the description of energy demand between dwelling characteristics, such as age, size, type and tenure and different levels of energy efficiency. From the analysis it is clear that gas demand is influenced by the level of detachment of a property, whereby dwelling forms with a greater exposed surface area have higher gas demand compared to those that are smaller and have less surface area. There is a strong size effect, with large dwellings using both more gas and electricity. It would be expected that electricity and heating demand would be influenced by size and also by occupancy.

The difference in gas demand between similar dwellings with different levels of energy efficiency is very clear. Those dwellings with improved levels of efficiency (i.e. loft insulation, cavity filling, double glazing and boiler replacement) - regardless of form or age - use less than their non-improved counterparts. This comparison suggests that there are long-term savings associated with efficiency measures. This is particularly important for the justification of continued roll out of energy efficiency retrofits, i.e. that higher efficiency levels can indeed maintain a lower demand, and improve financial payback estimates.

While the energy savings for any given dwelling will be influenced by the occupants, the change in gas demand associated with the presence of an energy efficiency measure suggests that real savings do on average occur following an intervention (i.e. a drop in the subsequent years).

From a physical point of view, cavity wall filling reduces the heat loss through the largest exposed area of a house (i.e. the external walls) and is thus associated with a larger change in demand. By comparison, lofts and windows are a much smaller proportion of the exposed area and show a smaller change in demand. Also, in the UK many lofts will already have had some level of insulation and the change between 100mm and 200mm will be smaller as a result. In theory, a boiler upgraded from a non-condensing to a condensing boiler should save gas by the change in efficiency alone; the average efficiency of a non-condensing gas boiler is approximately 70% (Palmer and Cooper, 2013) and industry rating schemes suggest approximately 86% for condensing. A boiler upgrade may also reflect other changes to the heat system, such as thermostatic valves or thermostats, which could also have an effect.

These outcomes are particularly important for the government's flagship energy efficiency policies, in particular the Green Deal that will rely on consumers retrofitting their property voluntarily and paying back the deferred upfront cost of the measure through savings from the energy bill.

7.4.2 Supporting evidence-based policy and research

Creating a data framework that is based on well-structured and consistent data of a high quality begins to lay the foundations for a stronger connection between evidence and policy. While HEED is not a 'gold standard', it does offer a useful resource from which to build such a data foundation, which is reflected in the intention of the government to continue to develop the National Energy Efficiency Data-framework (DECC, 2011a). However, the move towards quantifying the impact of energy efficiency investment in the UK's housing stock requires greater attention to how data is collected and also an acknowledgment of the type of questions that it can attempt to answer.

For policy development that seeks to target certain areas and housing types, the dwellings that HEED represents is of interest as it speaks to those dwellings that have not had efficiency measures through programmes captured under HEED, which have been the bulk of efficiency measures delivered in the UK (as suggested in Chapter 6). Those dwellings not in HEED must be the targets of future energy efficiency programmes, which will need to draw in more households living in semi-detached houses and flats, larger properties (i.e. >3 bedrooms), and social and private rental tenures with a focus on Southern regions.

7.4.3 Study Limitations

HEED contains information on over 50% of dwellings in the UK. The results of the housing stock population comparisons for the English and Welsh sample of HEED and

England and Wales housing stock datasets suggest that the dwellings in HEED are not strictly statistically representative, though major differences are unlikely given the sample size. The English and Welsh sample of HEED has fewer flats and more semi-detached houses, more 1 and 3 bedroom dwellings, more socially rented dwellings, and less coverage in the Southern English regions. However, many of the selected variables in HEED do seem to be similarly distributed (i.e. within 1% point) and can offer a degree of representative descriptiveness. The Scottish sample of HEED was shown to be representative of the Scottish housing datasets. HEED has been expanding by roughly 8% per year in recent years and discrepancies between HEED and the dwelling stock as a whole may reduce in the future.

In terms of potential biases, the majority (~80%) of HEED homes will have had some sort of energy efficiency measure. Also, it is not possible to be exact on the number of homes outside of HEED that have had some level of retrofit. Further, several of the programmes in HEED will have been developed to target certain household types (e.g. fuel poor[13]) who may live in dwellings with certain characteristics that may tend to bias the representativeness of the data.

There will also be limitations to the HEED and energy dataset that have to do with collection methods (i.e. different surveys using different forms), issues of self-selection for surveys and misclassification or assessor bias. Also, a dwelling will enter HEED as a ‘snap shot’, which means that the energy efficiency characteristics recorded for the dwelling will be more or less correct at a particular date. However, these features may not persist over time and changes would only be picked up if dwellings were revisited at a later date. This may occur in the long run through Energy Performance Certificates (currently covers 4.5 million properties in Great Britain), which rate the energy performance of the dwelling and collect characteristics at the time of sale or rental (at some point, nearly every home in Britain will be rented or sold and thus subject to an EPC).

It is unlikely that HEED will offer the same insight as a well-structured research design on the impact of energy efficiency or an omnibus survey in terms of representativeness, but what is clear is that it offers usefulness as a framework within which to collect and link data sources together. Due to the nature and range of its coverage (i.e. containing information on approximately 50% of UK dwellings) it could reasonably be used as a source to describe the broad energy performance characteristics of the UK housing stock. When linked to energy, HEED is capable of offering insight into the differences in demand due to dwelling characteristics and levels of energy efficiency, and the change in demand associated with an energy efficiency retrofit.

[13] Fuel poverty in the UK is the condition whereby a household spends more than 10% of their income on fuel to maintain an adequate level of warmth (DECC, 2010c). This has been more recently revised in 2013 to be defined by low-income levels and notional high heating costs, known as the ‘low-income, high-cost’ definition of fuel poverty (DECC, 2013a)

7.4.4 Study Conclusions

The study found that both gas and electricity are highly related to characteristics of the dwelling and energy efficiency levels, which is important for understanding the 'baseline' of different dwelling types. However, of similar importance are the distributions in energy demand for those dwellings features. This suggests that even for dwellings with seemingly similar characteristics, the energy demand will vary widely. The reasons behind this sub-group variation will be related to both 'hidden' or unknown features of the dwelling that are not accounted for here, such as features of the heating system, and appliance ownership levels. They will also be related to household features and occupant practices and habits, such as occupancy level, working status and location, and temperature set points amongst many others. With further analysis, it may be possible to investigate the contribution of these more detailed factors on energy demand among sub-groups of dwellings and households.

7.5 Critical discussion of the cross-sectional study approach

This section critically discusses the use of the descriptive cross-sectional study as a method of examining energy demand variation. First, it examines the limitations of the study method in terms of different forms of bias, the variables used and the available information. Concluding thoughts on the method then follow.

7.5.1 Limitations of the study method

Cross-sectional studies can suffer from various forms of bias, which include: selection bias, information bias, and confounding. Selection bias occurs when over- or under-sampling from the study population occurs, leaving the sample group unrepresentative of the target population. In this method study the target population are dwellings that have at some point received an energy efficiency intervention. The study sample is a randomly-selected population drawn from an available study population, i.e. HEED. In Chapter 6, HEED was described as a collection of the majority of energy efficiency measures in England over the period 2000 to 2007. This study used all dwellings and HEED to better understand the relationship between energy demand and housing characteristics and energy efficiency levels. Measures were taken to control for interventions that occurred after 2006; however, there is still a chance that the sample did not represent the target population at that time. Information bias occurs when variables are incorrectly measured or mis-reported. Because HEED is an in-action data set compiled from many sources using various different data collection approaches, misclassification has a greater likelihood of occurrence. There are various levels of trust that may be attached to certain forms of data collection, for example trained surveyors are more trusted then self-completed surveys. Installers for a particular measure, say a boiler, may have expertise in that feature but they are not necessarily experts in other features of the dwelling (i.e. age). To examine the occurrence of information bias across that HEED sample, follow-up studies that undertake random checks of dwellings using expert surveyors could look at differences between reported and measured variables. This however was outside the scope of the present research.

Cross-sectional studies typically examine prevalence of conditions; as such they are not ideal for examining causal associations but instead typically describe difference in outcomes between sub-groups. This means it is important to ensure that both the information used to differentiate subgroups is as accurate as possible and that the correct information is used. Numerous studies have identified dwelling characteristics and energy efficiency levels as having a strong association with differing levels of energy demand. The analysis of HEED in this method study is not only limited by the amount of available information on dwellings and households, but also the level of information from the database. Because HEED is effectively a registry of energy efficiency measures it is not a comprehensive source of information on dwelling and household characteristics. Features not collected are also likely to be important influencing factors for the differences in demand seen in this study. For this reason comparisons of subgroups with small sample sizes were omitted.

An advantage of the cross-sectional study design is the ability to examine a range of influencing factors and differing levels of an outcome. A cross-sectional study should capture information from the population with varying levels of the outcome of interest and differing influencing characteristics. For example, not all older dwellings have the same level of energy efficiency, although there may be similarities, and as such these are likely to have different levels of energy demand. In this study, it was found that dwellings with given energy efficiency feature had similar levels of difference in gas demand compared to those without. With more information on the dwelling and households, it may be that a 'true' impact related to energy efficiency interventions could be identified. This will be explored in the subsequent Chapter 8.

Depending on the size of the cross-sectional survey, rare events or occurrences may not appear in the study sample that might be present in the target population. As such, for very rare events or conditions, other study designs are more appropriate (i.e. case-control or cohort).

7.5.2 Study method conclusions

Overall, the application of a descriptive cross-sectional study to examine differences in energy demand and dwelling characteristics was appropriate and offered plausible findings. The study used information from HEED, which was similar to that of a registry. It was able to meet its main objective, which was to identify whether there were differences in energy demand related to dwelling characteristics and energy efficiency levels.

7.6 Summary

This chapter investigated the application of the second proposed method study, a descriptive cross-sectional study of dwellings characteristics and energy efficiency levels and gas and electricity demand. The target group were those dwellings that could have taken part in energy efficiency programmes. It sought to examine whether there were differences in energy demand between different dwelling and household characteristics and energy efficiency levels.

In addition, the chapter also included a discussion of the study design in terms of its appropriateness to the research question and its limitations. In the following chapter, a method study using a cohort study design will be used to explore the uptake of energy efficiency measures and their impact on gas demand.

Chapter 8 Cohort Study

“Energy efficiency uptake and energy savings in English houses”

Chapter Introduction

In Chapter 6, an ecological study was used to examine the association between the levels of uptake rates of energy efficiency interventions over a 7-year period and by neighbourhood level characteristics. The aim of that study was to generate hypotheses around energy efficiency retrofit uptake for further study. The hypotheses were: that privately rented dwellings have low levels of uptake stemming from the landlord-tenant tensions and also differences in construction that affect the type of retrofit being offered; and that low-income neighbourhoods receive and accept more measures due to government targeting. In Chapter 7, a cross-sectional study explored the variation in energy demand associated with dwelling characteristics over a 4-year period with the aim of identifying potential influencing factors. The study found that factors such as dwelling size and age and level of detachedness and energy efficiency levels were all associated with variation in energy demand, and that the presence of energy efficiency retrofits was associated with reduced demand.

The aim of this chapter is to draw together the ecological and cross-sectional study findings and explore the relationships between energy demand, dwelling features and household characteristics and energy efficiency in further detail at an individual dwelling level. Energy savings play a central role in meeting UK climate change mitigation targets, and therefore understanding the take-up of energy efficiency retrofits and their impact on energy demand and variations in these measures across the population is vital to understanding their potential. The research problem being addressed is whether a selection of dwelling features and household characteristics affect the uptake of energy efficiency retrofits and what impact retrofits have on energy demand, adjusting for influencing factors. The research uses a cohort study method, which uses a study sample selected to be representative of the English housing stock based on dwelling age, size (i.e. number of bedrooms), type, household tenure, region and neighbourhood income levels. The sample is used to examine the historic uptake rate and to estimate probability rates, energy demand levels (focusing on gas) and change in gas demand following a selection of efficiency retrofits.

Like the previous two chapters, this chapter comprises five parts: a) an introduction to the research problem along with relevant background and a description of the key concepts and features of the selected cohort study design; b) a description of the methods specific to the cohort design and the research problem being investigated; c) a presentation of the results; d) a discussion and conclusions of the findings specific to the study; and e) a critical appraisal of the study approach, its strengths, weaknesses and limitations.

8.1 Research problem

Government estimates suggest that through increased efficiency, an energy savings potential of 54 TWh is possible by 2020, a reduction of ~10% from 2012 demand levels of 500 TWh (DECC, 2014b), delivered through a range of energy efficiency measures that focus on dwelling fabric and heating systems. A significant potential for energy efficiency retrofit exists in the UK housing stock, including: insulating 7.3 million solid walled homes, 5.1 cavity walled homes, 7.4 million lofts, 19.2 million double glazing installations, 17.6 million boiler upgrades, along with millions of dwellings needing heating controls, draught-proofing, and heat recovery (UK CCC, 2012).

Chapter 6 showed that approximately 12.2 million UK dwellings have received some form of energy efficiency retrofit since 2000. However, the rate of retrofit uptake across UK dwellings has been lower than is required to meet UK targets (UK CCC, 2012). Further, the impact that these retrofits have on energy demand has been less than predicted (Sunikka-Blank and Galvin, 2012). Together, the limited uptake and impact on energy demand pose a clear threat to meeting UK emission reduction targets.

8.1.1 Research question and aims

A pressing question that emerges relates to *who* have (and have not) taken up retrofits and whether household factors affect this uptake over time? A second question is *what* impact have these measures had on demand and how does it differ amongst households? Chapter 6 has shown that uptake has varied amongst English neighbourhoods by income groups, vulnerability, region and age of housing stock. While several cross-sectional studies have shown how dwelling typologies influence retrofit uptake, with older dwellings generally needing more insulation and others requiring specific types of retrofit (i.e. cavity filling insulation) and the influence of household characteristics on retrofit presence with lower income, privately renting households having the lowest levels of efficiency (Brechling and Smith, 1994; Tovar, 2012). However, to date there has been little work to understand a) how individual level household or dwelling characteristics modify uptake over time and the type and combination of retrofits, and b) whether having a retrofit modifies the probability of installing subsequent measures. Further, while studies have attempted to quantify the impact that retrofits have had on energy demand in UK dwellings (Bell and Lowe, 2000; Milne and Boardman, 2000; Wyatt, 2013), there has been little work to understand a) the extent to which dwelling and household characteristics modify changes in energy demand; and b) whether cumulative retrofits result in more savings.

To date, there have been no studies that have examined the uptake of energy efficiency measures in British houses over time by following a group of dwellings and examining what factors might be associated with differences in uptake or differences in the impact of the retrofits installed. The purpose of this study is to provide a better understanding of the uptake of energy efficiency retrofits and the resulting change in energy demand that accounts for individual dwelling and household characteristics, adjusting for potentially confounding and interacting factors. The research questions asked were:

a) What is the rate of uptake of energy efficiency measures in the English housing stock, what dwelling, household and local area features affect this rate, and what differences exist between those dwellings that installed/received efficiency measures;

b) What is the rate of change in energy demand in the English housing stock and what dwelling, household and local area features affect this rate; and,

c) What is the effect (individually and in combination) of heating system and fabric insulation energy efficiency measures on change in energy demand, and what factors affect these changes.

Factors that may be associated with energy efficiency uptake and the impact that retrofit measures have on energy savings include: household practices and their socio-economic characteristics, beliefs and social norms, upfront cost of measures, perception of risks and challenges, perception of institutions such as governments or energy suppliers, ownership, and dwelling characteristics (Mills and Schleich, 2012; Tovar, 2012). Higher-income households may also be more able to reduce their energy demand than lower-income dwellings (Jamasb and Meier, 2010). This chapter tests several hypotheses generated in the literature that relate to dwelling, household and neighbourhood factors that affect uptake of energy efficiency measures (Mills and Schleich, 2012; Tovar, 2012) and changes in energy demand (Galvin, 2010; Wyatt, 2013). The hypotheses tested are:

1. Households with lower incomes accept/receive more measures than higher income levels

2. Households that own their homes accept/receive more measures than other tenures

3. Older dwellings are more likely to take up energy efficiency measures

4. Older dwellings are less likely to achieve energy savings compared to newer dwellings

5. Lower-incomes households are less likely to realise energy savings compared to higher incomes.

A population-based cohort study of English dwellings selected from the Homes Energy Efficiency Database (HEED) was used to investigate the association between household and dwelling characteristics, the uptake of energy efficiency retrofits and changes in energy demand. The study used a sample that was drawn to be representative of English houses, using the English Housing Survey (EHS) as a sample frame. The uptake of energy efficiency measures was examined from 2002 to 2007. The change in gas and electricity demand was followed from 2004 to 2007.

8.1.2 Context and background

The findings of the Chapter 6 ecological study suggests that the rate of uptake of efficiency measures will be most influenced by decision-making autonomy (i.e. dwelling ownership), income levels, existing energy performance, and regulatory requirements. The rate of uptake for all types of energy efficiency measures in England between 2000 and 2007 was lower in neighbourhoods with middle and high incomes and also in the rental market, and the highest rates were among neighbourhoods with lower incomes, more benefits and higher levels of owner-occupied dwellings. An obvious challenge to achieving a high rate of uptake of energy efficiency among UK homes is that there are real differences in terms of the dwellings' physical construction, design and size, energy performance, existing heating and ventilation systems and appliances, access to fuels and their location. Also, there may be even greater differences among households that live in these homes in terms of their energy demand practices, socio-economic circumstances, and interest or ability to undertake retrofits.

Historic and recent research has shown that seemingly similar houses can have very different levels of energy demand (Socolow, 1978; Summerfield et al., 2007), reflecting real differences in the practices around energy demand (Wilhite et al., 2000b). Several studies have examined the uptake of energy efficiency measures across the UK housing stock, including: the investment in efficiency measures in English houses (Brechling and Smith, 1994; Tovar, 2012), and the attitudes and barriers to adopting energy efficiency (Mills and Schleich, 2012). A recent study examined the determinants of energy expenditure using a panel survey of British households from 1997 to 2005; however the data used did not include information on retrofits, or on dwelling features, such as age or energy performance (Meier and Rehdanz, 2010). However, it is not clear how individual dwelling and household characteristics affect uptake rates over time and how they differ across the housing stock.

Introducing efficiency measures with the aim of improving the energy performance of the fabric, heating system or reduce uncontrolled ventilation affects energy demand by reducing heat loss or improving internal temperatures, or (more likely) a combination of the two. The variation in the change in demand has been shown to dependent on the level of efficiency improvement sought (e.g. deep retrofits versus single component improvements) (Milne and Boardman, 2000), the quality of the installation, and the response of household occupants (e.g. upfront cost and savings recuperation, comfort taking) (Oreszczyn et al., 2006a). These interacting factors result in a change in the demand for heating energy that may last for varying periods. Of interest is the *actual* change in energy demand following the introduction of efficiency measures and whether this change can be attributed to the retrofit. Work by Wyatt (2013) showed that installing efficiency measures resulted in changes in gas demand over a three-year period that were greater than dwellings with no efficiency measures, including reductions of: 10% for cavity wall insulation, 3% for loft insulation, 8% for condensing boiler installations, and 2% for double glazing installation (Wyatt, 2013).

To better understand what 'savings' should be expected following a retrofit measure, it is important to know what effect physical dwelling and household characteristics might have on energy demand. Several of the above studies provide evidence of real changes in energy demand following the introduction of fabric and heating efficiency measures, but do not fully examine the potential variation in changes in energy demand due to dwelling or household features.

8.2 Method

Cohort studies are observational studies of a selection of individuals over time. They are well suited for detecting changes in patterns over the period of study. Cohorts are typically used to determine the differences in outcomes for those exposed to a factor or event and those who are not, with the aim of determining aetiological links. The advantage of using a cohort study is that a number of factors and levels can be explored simultaneously.

8.2.1 Datasets

In this study, a cohort study sample of English dwellings was selected to be representative of the English housing stock using the Homes Energy Efficiency Database (HEED), connected to gas and electricity meter data. The study focused on the uptake of energy efficiency measures dwellings with a gas connection.

Homes Energy Efficiency Database: As described in detail in Chapter 5, the Homes Energy Efficiency Database (HEED) comprises information on the energy performance and installation of energy efficiency measures in England, covering a period from 1993 to 2013. Although the database is not statistically a representative sample of dwellings in the UK, it was shown to provide a reasonable breadth of geo-spatial coverage and to account for a majority of the energy efficiency interventions that took place between 2002 and 2007.

Energy supplier meter point data: The energy supplier database of annualised gas and electricity meter point data for the years 2004 to 2007 was used to examine the impact of energy efficiency measures on energy demand. This study uses gas demand as the primary outcome for the energy demand analysis. Gas connection covers over 86% of English dwellings and the majority of gas is used for space heating and hot water (DECC, 2010b). The electricity demand data was not used in the energy demand analysis, as the focus was on space heating, but was used to determine accurate readings for the cohort sampling. However, to determine the influence electricity has on the potential impact of retrofits on gas demand, as a sensitivity test of the impacts are tested using a variable of 'energy demand' combined from the sum of annual electricity and gas demand for the dwelling.

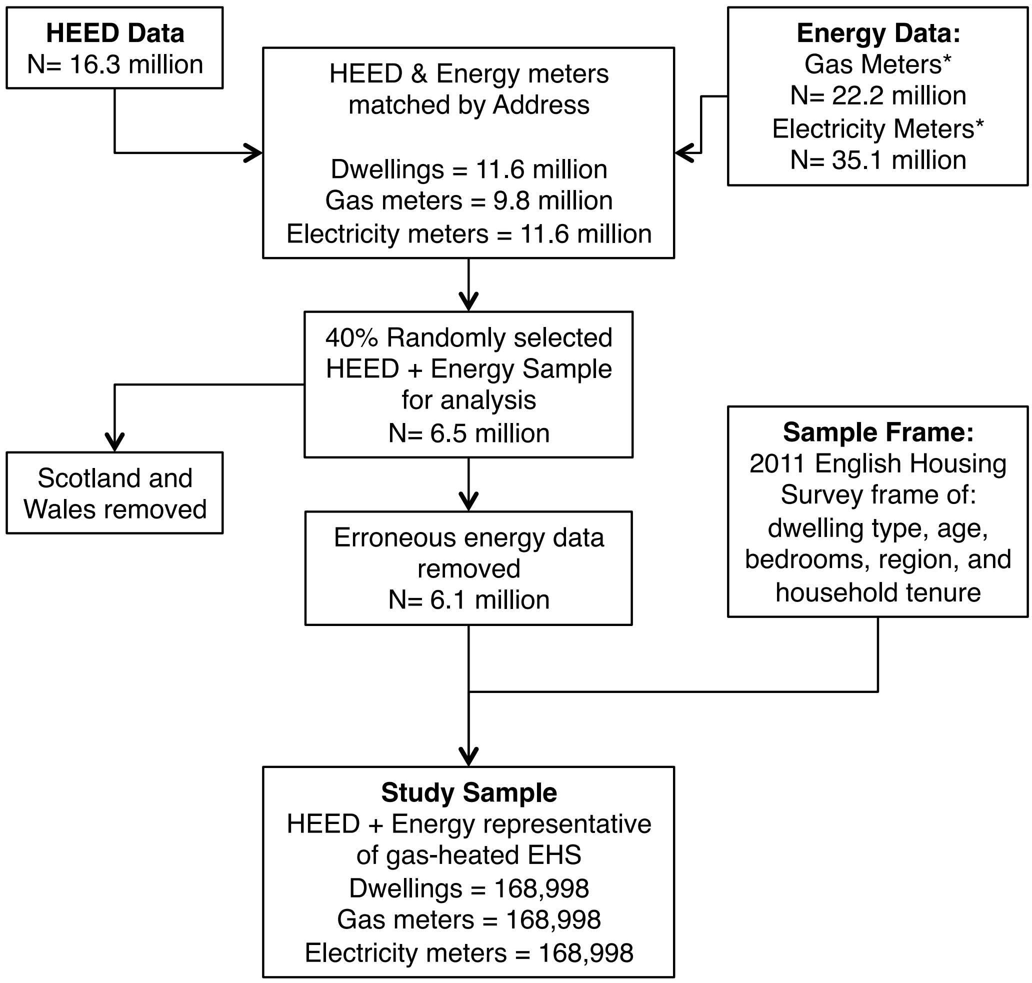
In addition to the cleaning undertaken for Chapter 7 (i.e. removing missing, repeat, zero, negative, and very large values (i.e. above 73.2 MWh/year for gas and 50 MWh/year for electricity), a further cleaning was applied to inter-annual changes in demand. Meters with large changes in demand were also removed, i.e. > ±80% of the preceding year.

Neighbourhood level household characteristics: To examine neighbourhood level effects, data at the lower super output area (LSOA) level was used. Experian Mosaic Public Sector data on median income and household type (based on Mosaic classification) was used (Experian, 2012). Data on age of population, number of benefit claims, and council tax bands were drawn from the Neighbourhood Statistics service (ONS, 2012b). The neighbourhood level data was not collected for every year in the study; therefore data from the nearest year to 2007 was used wherever available. The LSOA level data was connected subsequently using the LSOA codes provided in the anonymised HEED+Energy data.

8.2.2 Study Population – English gas-connected dwelling cohort

To examine the uptake of energy efficiency measures in England's housing stock and its impact on gas demand, a sample from the HEED + Energy Meter data was constructed that was statistically representative for a selection of English dwelling characteristics. The combined HEED + Energy data, relating to approximately 11.6 million unique dwellings along with electricity and gas meters, was used as the basis for selecting the study sample. Although HEED contained 16.3 million dwellings, the matched data made available from DECC used in this study comprised a match for 11.6 million dwellings. For computational purposes, a 40% randomly selected HEED + Energy dataset was drawn from the full dataset for detailed analysis.

The sample frame was constructed using the 2011 EHS, which is a cross-sectional survey that is representative of English dwellings and households (CLG, 2013a). The 2011 EHS comprises survey from 2010 and 2011 and was used because it was the latest data to align with HEED at the time of analysis. The sample frame was constructed to be representative of gas-heated English dwellings and comprised: dwelling age, dwelling type, number of bedrooms, government region, and household tenure. To align with HEED variables, EHS dwelling age, type and tenure were recoded to construct the sample frame (see Table 26). Figure 29 shows the selection process for the study sample. Table 27 shows a comparison between the constructed study sample, the EHS and HEED.



*All meters in the UK were used in matching

Figure 29 - Study sample selection process

Table 26 - EHS sample frame variables recoding for HEED selection

EHS Variable	EHS categories	HEED category
Dwelling type (dwtypenx)	end terrace, mid terrace	terrace (end and mid)
	semi-detached	semi-detached
	detached, bungalow	detached (inc. bungalows)
	converted flat, purpose built flat (low and high rise)	flats (all types)
Dwelling Age (dwage9x)	pre-1850, 1850-1899	pre-1900
	1900-1918, 1919-1944	1900-1944
	1945-1964	1950-1966
	1965-1974	1967-1975
	1975-1980	1976-1982
	1981-1990	1983-1990
	post-1990	post-1990
Number of bedrooms (nbedsx)	1 bedroom	1 bedroom
	2 bedroom	2 bedroom
	3 bedroom	3 bedroom
	4 bedroom	4 bedroom
	≥ 5 bedroom	5+ bedroom
Government office regions (gorehs)	North East	North East
	North West	North West
	Yorkshire and the Humber	Yorkshire and the Humber
	East Midlands	East Midlands
	West Midlands	West Midlands
	East	East
	London	London
	South East	South East
	South West	South West
Household tenure (tenure4x)	owner occupied	owner occupied
	private rented	private rented
	local authority, registered social landlord	social or local authority rented

Table 27 - Comparison between source data (HEED+Energy), EHS data and HEED study sample

	HEED+Energy		2011 EHS (full weighted dataset)		HEED Study Sample	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Dwelling type						
terrace	424,079	21.47	5,872,437	31.32	51,264	30.33
semi detached	660,698	33.46	5,194,761	27.71	53,565	31.70
detached	757,739	38.37	4,907,771	26.18	48,107	28.47
flat, all	132,284	6.70	2,771,945	14.79	16,062	9.50
Frequency Missing = 1,243,037						
Dwelling Age						
pre-1900	65,358	4.30	2,088,451	11.14	11,237	6.65
1900-1949	401,360	26.41	5,137,574	27.40	49,489	29.28
1950-66	267,347	17.59	4,002,610	21.35	38,450	22.75
1967-75	346,591	22.81	2,647,092	14.12	26,355	15.59
1976-82	103,696	6.82	1,171,748	6.25	10,671	6.31
1983-90	77,799	5.12	1,507,953	8.04	13,497	7.99
post-1990	257,528	16.95	2,191,486	11.69	19,299	11.42
Frequency Missing = 1,698,158						
Household tenure						
owner occupied	1,370,498	78.01	12,983,750	69.26	130,403	77.16
private rented	128,607	7.32	2,717,408	14.50	10,096	5.97
social rented	257,717	14.67	3,045,756	16.25	28,499	16.86
Frequency Missing = 1,461,015						
Number of bedrooms						
1 bedroom	116,051	6.95	1,947,798	10.39	10,586	6.26
2 bedrooms	359,325	21.52	4,757,929	25.38	38,886	23.01
3 bedrooms	840,557	50.33	8,372,237	44.66	84,243	49.85
4 bedrooms	194,227	11.63	2,911,931	15.53	27,886	16.5
5+ bedrooms	159,884	9.57	757,019	4.04	7,397	4.38
Frequency Missing = 1,547,793						
Government Office Region						
North East	278,986	8.67	1,045,135	5.57	9,710	5.75
North West	422,740	13.14	2,724,241	14.53	25,633	15.17
Yorkshire and the Humber	441,648	13.72	2,032,134	10.84	18,962	11.22
East Midlands	310,359	9.64	1,644,482	8.77	14,424	8.54
West Midlands	288,355	8.96	1,976,435	10.54	16,757	9.92
East England	352,260	10.95	1,894,819	10.11	17,257	10.21
London	325,643	10.12	2,711,469	14.46	22,352	13.23
South East	509,342	15.83	2,998,696	16.00	28,139	16.65
South West	288,504	8.97	1,719,503	9.17	15,764	9.33

The sample was drawn using SAS 9.3 Proc Surveyselect (SAS Institute Inc., 2011). A sample size of 200,000 dwellings was requested using a *simple random sampling* design, which is selection with equal probability and without replacement. The resulting study sample comprised 168,998 dwellings with gas electricity meters. A comparison is provided of the original HEED + Energy dataset and the 2011 EHS and the study sample.

8.2.3 Energy efficiency interventions

The study period spans 2002 to 2012, which includes a number of government programmes (Warm Front, 2000 to 2013), energy company obligations (Energy Efficiency Commitment (EEC) 1 & 2, 2002 to 2008; Community Energy Savings Programme (CESP), 2008-2012; and Carbon Emission Savings Programme, 2008-2012), retrofit building regulations requirements for double glazing (Fenestration Self-Assessment Scheme (FENSA) from 2002), and gas system safety checks for private and social let properties (Gas Safety Regulations, 1998) (Rosenow, 2012, 2011).

Table 28 details the retrofit interventions examined from HEED, which broadly consist of: insulating cavity walls (with markers for pre and post 1976 dwellings), solid walls, and lofts; installing double glazing units (new and replacement pre and post-2002 double glazing); draught proofing; and heating system and controls upgrades. The heating controls comprised a range of timers and thermostat controls. Heating system upgrades consisted of replacement boilers, both condensing and non-condensing, room heaters and electric systems. A date (including month and year) of survey or retrofit installation was provided for each energy efficiency measure for every dwelling in HEED.

Table 28 - Energy efficiency retrofit details collected in HEED

Component	Energy efficiency interventions
Heating controls	Standby Saver Central Heating Controls Upgrade Delayed start thermostat Thermostatic Radiator Valves Load or Weather Compensation
Heating system	Community Heating Ground Source Heat Pump Replacement: Biomass Boiler, Electric Boiler, Gas Condensing Boiler (Standard and Combi), Gas Boiler (Standard and Combi), Oil Condensing Boiler (Standard and Combi), Oil Boiler (Standard and Combi) Room Heater: Electric, Gas, Solid Fuel Solid Fuel Fire Cassette Storage Heaters Electric and Gas Warm Air System
Cavity walls	Cavity Wall Insulation (pre and post-1976, and Unknown Property Age
Solid walls	External Wall Insulation to U-value of 0.37 W/m2K, U-value of 0.45 W/m2K Internal Wall Insulation to U-value of 0.37 W/m2K Unknown Solid Wall Insulation
Lofts	Loft Insulation: 0 to 250mm, 25 to 250mm, 50 to 250mm, 75 to 250mm, 100 to 250mm, 150 to 250mm
Domestic Hot Water	Installed Modern DHW Cylinder
Ventilation	Draught Proofing (General)
Glazing	Replacement Double Glazing
Smart systems	RTD Long Lifetime Visual Display Unit

8.2.4 Outcome

For the analysis focused on the *uptake of energy efficiency retrofit interventions* the outcome of interest was the presence of an energy efficiency measure installed from 2002 to 2007. The analysis grouped all 'major' measures together, which included: cavity wall insulation, loft insulation to 250mm, double glazing installation, heating system upgrades (including condensing boiler installation), and draught-proofing. Two further subgroups were derived that included 'fabric' measures (wall and loft insulation, glazing, and draught-proofing) and 'heating' measures (all heating system upgrades, including: heating controls, boiler upgrades). In addition to the presence of any intervention, the presence of additional interventions (i.e. any retrofit taking place following an initial retrofit) and the total number of retrofits (i.e. a package of fabric and heating retrofits) were examined. This was in order to determine whether, say, having a fabric intervention (e.g. cavity wall insulation) made a dwelling more or less likely to have subsequent retrofit (e.g. boiler replacement or loft insulation). Also, whether uptake over the period unfolded as packages of energy efficiency retrofits or single interventions. For the analysis, three outcome measures were examined: a) the presence of retrofit intervention(s) any time during the period 2002 to 2007; b) the presence of subsequent intervention measures within the period; and, c) the total number of any retrofits over the period.

For the analysis focused on the *impact of efficiency retrofit interventions on energy demand* the outcome of interest was the change in annualised gas demand between gas years. The available gas data covered only 2004 to 2007; therefore, the impact analysis only examines interventions within that period. The measures of change in annual gas consumption used for the analysis were the absolute change in demand (measured in kWh/year) and the proportion change in annual demand from one year to the next (measured as a proportional change in demand, unitless). The impact on gas demand analysis examines the change in demand across three years (2005-2007) in order to account for uncertainty in the annualisation process and the potential for this to result in delayed signals of change in demand (described in Chapter 5 and Appendix A). Thus, the intervention year is the middle year two (2006). This approach provides a greater certainty that the signal will be identified and also will allow for measures installed any time throughout the intervention year. For the purposes of analysis, all energy efficiency retrofit interventions are allocated to the gas year (i.e. 1 October to 30 September).

8.2.5 Influencing and confounding factors

In this study, putative influencing and potentially confounding factors are identified from the literature and accounted for in the analysis. These factors were classed into two types: physical dwelling characteristics and socio-cultural practices. Physical dwelling characteristics are related to those features of the dwelling that may have an effect on whether a dwelling was eligible for an efficiency retrofit. Dwelling age is likely an important influencing factor on the uptake of energy efficiency retrofits. The type of dwelling will also affect the retrofit take-up. Flats are unlikely to have lofts (unless in converted dwellings) and present more difficulties for wall insulation due to the impractical nature of insulating a single unit (if with external insulation) and more complex ownership structures. These physical factors were also considered to have an effect on changes in gas demand. Older dwellings have been shown to be colder dwellings (Oreszczyn et al., 2006a), and may therefore have a higher potential for temperature take back. Dwelling type will also be a proxy for the number and area of detached walls available for heat loss, which would affect the potential savings from insulation.

Socio-cultural practices are related to the characteristics and preferences of the household occupying the dwellings that could affect energy efficiency uptake and changes in energy demand. Household income or benefit receipt have been shown to affect the ability to afford energy efficiency retrofits (Tovar, 2012), but also eligibility for government assistance (Rosenow, 2012). Household tenure may also affect efficiency uptake due to the decision-making autonomy of a household. Households living in social and private let dwellings are subject to the agreement of landlords to accept retrofits. These issues are also known to affect energy savings that might derive from installed energy efficiency retrofits. Low-income and households on benefits are known to have a higher exposure to poor-quality housing (Healy and Clinch, 2004; Oreszczyn et al., 2006a) and may have a higher temperature take back potential to achieve thermal comfort (Hong et al., 2009).

8.2.6 Statistical analysis

The analysis was carried out using SAS v9.3. Analysis of the uptake of efficiency interventions used logistic regression to examine the presence (0,1) of energy efficiency retrofits during the 2002 to 2012 period for all interventions, fabric and heating system, and a selection of individual interventions (i.e. cavity wall insulation, boiler replacement). The probability of having had an energy efficiency retrofit was modelled for all dwellings (a crude estimate), adjusted to control for putative influencing factors.

General linear models (GLMs) were used to analyse change in energy demand. All categorical variables were entered as classes and a reference class was used against which to determine parameter estimates. Estimates of change in gas demand were made for all dwellings (i.e. crude change) and then adjusted to control for physical and socio-cultural factors. GLM was also selected because the change in energy demand had a Gaussian function distribution.

8.3 Results

In 2007, 168,998 English dwellings were examined as part of the cohort study analysis. From 2002 to 2007, 39% received a major measure, 36% a fabric measure, and 9% a heating measure. The annual average change in energy demand across the stock was approximately -810 kWh/year (-740 kWh/year in 2004/05, -860 kWh/year in 2005/06, and -830 kWh/year in 2006/07). This amounted to an annual average proportion change of -3.6% for 2004/05, -4.4% for 2005/06, and -4.4% for 2006/07. The following sections concentrate on the uptake of energy efficiency measures within the cohort and then the impact of the energy efficiency retrofit interventions on changes in gas demand.

8.3.1 Uptake of energy efficiency retrofits among study sample of English dwellings

The uptake of fabric measures in the study sample over the period 2002 to 2007 was highest for cavity wall insulation and loft insulation, and heat systems (the majority of which were boiler installations) – see Figure 30. In 2009, the annual uptake rates of reported cavity and loft insulation were around their peak of 50 per 1,000 dwellings. Reported condensing boiler installations had a peak uptake rate of 21 per 1,000 dwellings in 2007. Cavity and loft insulation and condensing boiler installations held a relatively constant uptake trajectory from 2002 to 2007 (see Figure 31), though there was a change in the number of added installations in cavity and loft insulation in 2008, coinciding with CERT.

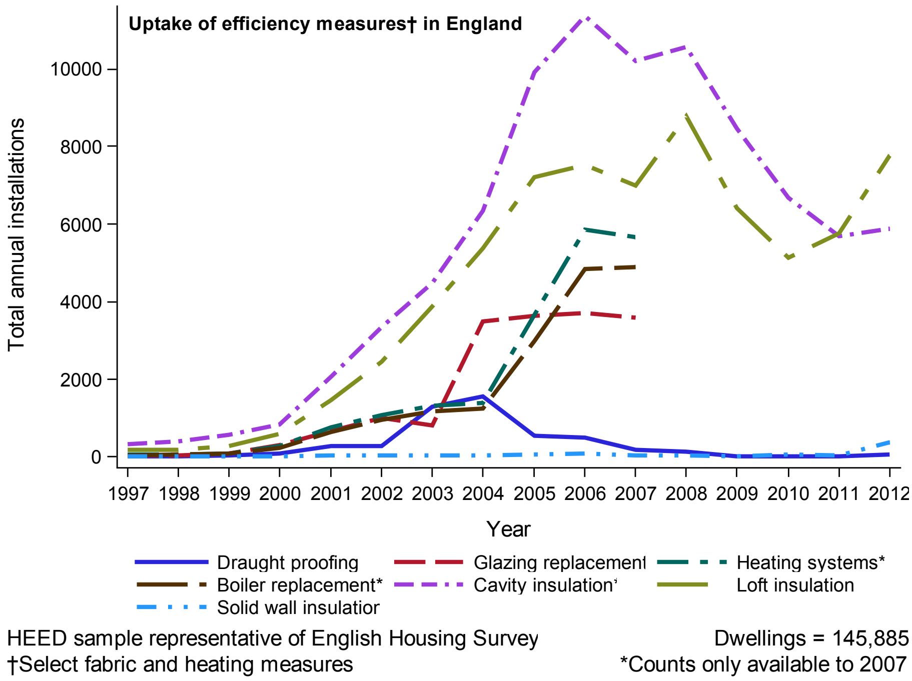


Figure 30 - Uptake of energy efficiency retrofits in sample English stock 1996 to 2012

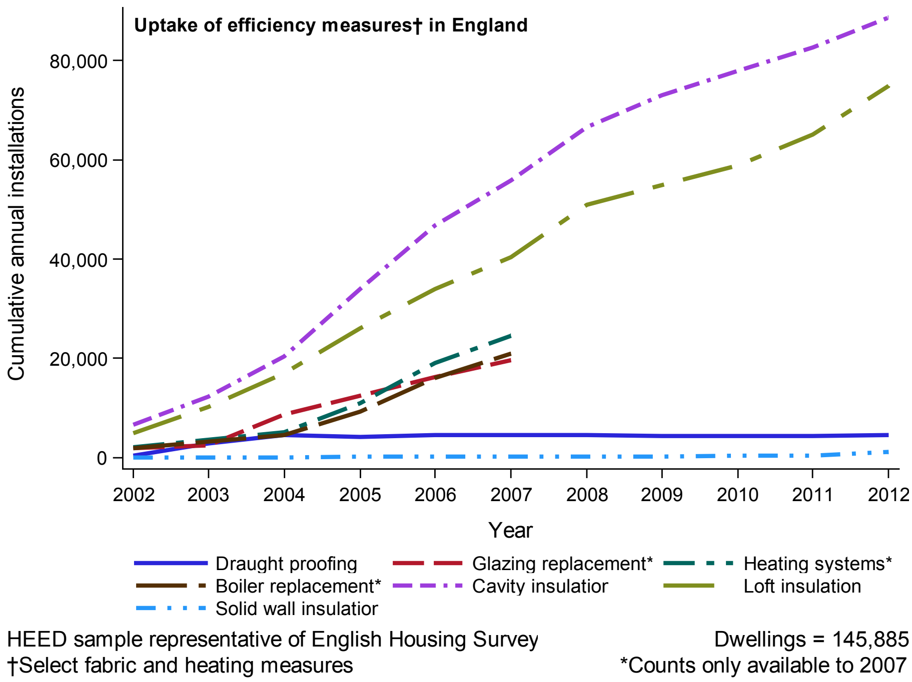


Figure 31 - Cumulative uptake of energy efficiency retrofits in sample English stock 2002 to 2012

The incidence rate (i.e. dwellings with measures installed over all dwellings) of uptake over the study period differed considerably between dwelling characteristics. There was a higher uptake of fabric interventions compared with heating measures within the cohort over the study period (Table 29). By dwelling type, the incidence of all major measures over the period was highest amongst detached dwellings (480 per 1000 dwellings) and lowest amongst flats (280 per 1000 dwellings). Older dwellings had lower rates of fabric measure uptake than newer dwellings (i.e. 190 per 1000 dwellings in pre-1919 dwellings and 460 per 1000 dwellings in pre-1950 dwellings). The majority of the fabric measures are cavity wall filling and therefore these dwellings are more likely to have brick or stone solid walls. The incidence of heating measures shows little difference by dwelling age bands, ranging from 90 to 120 per 1000 dwellings. Also, there is a higher incidence of heating system installation by dwelling type, particularly detached dwellings and for privately let dwelling tenures.

Whilst the incidence rate provided a measure of the uptake over the study period, the likelihood (i.e. dwellings with measures installed over dwellings with no installation) provided a measure of the probability that a dwelling might have a measure installed, and accounts for the size of the population. Using the 'crude' probability (i.e. unadjusted for potentially influencing factors), the average dwelling was 34% less likely to have a major measure installed over the study period than having no measure installed. Table 29 shows that for fabric and heating measures the probability of having a measure was 45% and 90% less likely than having had no measure installed. There was greater variation in the probability of installation from the stock average between different dwelling features. Compared to the 'crude' stock average, dwellings were more likely to have had a major measure installed if they were: detached (22%), constructed between 1967-75 (40%), privately rented (12%), are with 3 bedrooms 6%), and located in the North East (32%), North West (15%) or the West Midlands (21%).

The impact of dwellings features on the probability of uptake among the study sample over the study period was examined using a logistic regression model. Table 30 shows regression coefficients for the association between dwelling features and the probability of having had a major measure, fabric measure or heating measure installed in the study sample from 2002 to 2007.

Table 29 - Number (1000's) of major measures installed in English houses 2002 to 2007, derived from HEED

	Major measure (1000's) 2002-2007			Inci-dence	Likeli-		Fabric measure 2002-2007		Inci-dence	Likeli-		Heating measure 2002-2007		Inci-dence	Likeli-	
Dwelling Characteristic	**Yes**	**No**	**All**	**Rate**[a]	**hood**[b]	**Odds**[c]	**Yes**	**No**	**Rate**[a]	**hood**[b]	**Odds**[c]	**Yes**	**No**	**Rate**[a]	**hood**[b]	**Odds**[c]
Dwelling Type																
flat, all	4.5	11.5	16.1	0.28	0.72	0.61	3.8	12.3	0.24	0.66	0.56	1.2	14.9	0.07	0.80	0.79
terrace	17.4	33.8	51.3	0.34	0.87	0.81	16.0	35.3	0.31	0.88	0.82	2.7	48.5	0.05	0.59	0.57
semi detached	21.0	32.5	53.6	0.39	1.01	1.01	19.9	33.7	0.37	1.04	1.07	2.5	51.0	0.05	0.52	0.50
detached	23.0	25.1	48.1	0.48	1.22	1.43	20.5	27.6	0.43	1.20	1.34	8.8	39.3	0.18	2.03	2.26
Dwelling age																
1900-pre	2.3	9.0	11.2	0.20	0.51	0.39	1.6	9.7	0.14	0.39	0.29	0.9	10.3	0.08	0.91	0.90
1900-1949	17.2	32.3	49.5	0.35	0.89	0.83	15.3	34.2	0.31	0.87	0.81	4.2	45.3	0.09	0.94	0.94
1950-66	14.6	23.8	38.5	0.38	0.98	0.96	13.1	25.3	0.34	0.96	0.94	3.4	35.0	0.09	0.99	0.99
1967-75	14.3	12.1	26.4	0.54	1.39	1.84	13.7	12.6	0.52	1.47	1.97	2.4	23.9	0.09	1.02	1.03
1976-82	4.2	6.5	10.7	0.39	1.01	1.02	3.9	6.8	0.37	1.03	1.05	0.9	9.7	0.09	0.97	0.97
1983-90	4.8	8.7	13.5	0.36	0.92	0.87	4.4	9.1	0.32	0.91	0.87	1.2	12.3	0.09	0.99	0.99
1990-post	8.6	10.7	19.3	0.44	1.14	1.24	8.2	11.1	0.42	1.19	1.33	2.1	17.2	0.11	1.20	1.23
Tenure type																
owner occupied	51.6	78.8	130.4	0.40	1.01	1.02	47.6	82.8	0.37	1.03	1.04	11.5	118.9	0.09	0.98	0.98
private rented	4.4	5.7	10.1	0.44	1.12	1.22	4.0	6.1	0.40	1.12	1.21	1.2	8.9	0.12	1.36	1.42
social rented	9.9	18.6	28.5	0.35	0.89	0.83	8.5	20.0	0.30	0.84	0.77	2.5	26.0	0.09	0.96	0.96
No. Bedrooms																
1 bedroom	2.6	8.0	10.6	0.25	0.63	0.51	2.0	8.6	0.19	0.52	0.41	0.9	9.7	0.08	0.91	0.91
2 bedrooms	14.8	24.1	38.9	0.38	0.98	0.96	13.5	25.4	0.35	0.97	0.96	3.8	35.1	0.10	1.09	1.10
3 bedrooms	34.7	49.5	84.2	0.41	1.06	1.10	32.7	51.6	0.39	1.09	1.15	6.6	77.7	0.08	0.87	0.85
4 bedrooms	10.6	17.3	27.9	0.38	0.97	0.95	9.9	18.0	0.35	1.00	0.99	2.3	25.6	0.08	0.91	0.90
5+ bedrooms	3.2	4.2	7.4	0.44	1.12	1.21	2.2	5.2	0.29	0.82	0.75	1.7	5.7	0.23	2.52	2.97
Region																
North East	5.0	4.7	9.7	0.52	1.32	1.66	4.7	5.0	0.48	1.36	1.70	0.9	8.8	0.10	1.07	1.08
North West	11.5	14.1	25.6	0.45	1.15	1.27	10.7	14.9	0.42	1.17	1.29	2.5	23.1	0.10	1.08	1.09
Yorkshire and the Humber	7.3	11.7	19.0	0.38	0.98	0.97	6.6	12.4	0.35	0.97	0.96	2.0	17.0	0.10	1.15	1.17
East Midlands	6.4	8.1	14.4	0.44	1.13	1.23	5.8	8.6	0.40	1.13	1.21	1.7	12.7	0.12	1.31	1.35
West Midlands	7.9	8.8	16.8	0.47	1.21	1.40	7.3	9.4	0.44	1.23	1.40	1.8	15.0	0.10	1.16	1.18
East England	7.0	10.3	17.3	0.40	1.03	1.05	6.3	11.0	0.37	1.03	1.04	1.7	15.6	0.10	1.09	1.10
London	5.2	17.1	22.4	0.23	0.60	0.48	4.4	18.0	0.20	0.55	0.44	1.3	21.1	0.06	0.64	0.62
South East	9.9	18.2	28.1	0.35	0.90	0.85	9.0	19.2	0.32	0.90	0.85	2.2	25.9	0.08	0.86	0.85
South West	5.8	9.9	15.8	0.37	0.95	0.92	5.4	10.4	0.34	0.96	0.95	1.2	14.6	0.08	0.84	0.83
All	**66.0**	**103.0**	**169.0**	**0.39**	**1.00**	**1.00**	**60.1**	**108.9**	**0.36**	**1.00**	**1.00**	**15.3**	**153.7**	**0.09**	**1.00**	**1.00**

Note: a) Incidence as a rate over study period; b) Likelihood compared to stock average ('All'); c) Odds ratio of dwellings characteristic over stock average ('All')

Table 30 – Logistic (probit) regression coefficients (as probabilities) representing the association between dwelling characteristics and installation of measures (major, fabric and heating) from 2002 to 2007

	Retrofits installed 2002 to 2007		
	Major retrofit	**Fabric† retrofit**	**Heating retrofit¥**
	N=168,988	**N=168,988**	**N=168,988**
Factors	**Coefficient estimate* (confidence limits at 95%)**		
Intercept	-1.14 (-1.2, -1.07)	-1.49 (-1.56, -1.43)	-1.4 (-1.49, -1.31)
Dwelling Type			
terrace			
semi detached	0.1* (0.08, 0.11)	0.28 (0.26, 0.3)	0.96 (0.93, 0.99)
detached	0.38* (0.36, 0.4)	-0.06 (-0.09, -0.03)	0.14 (0.09, 0.18)
flat, all	-0.05 (-0.08, -0.02)	0.1 (0.08, 0.12)	0 (-0.02, 0.03)
Dwelling Age			
Pre-1900			
1900-49	0.4* (0.37, 0.43)	0.53 (0.5, 0.56)	-0.02 (-0.06, 0.02)
1950-66	0.43* (0.4, 0.46)	0.57 (0.54, 0.6)	-0.09 (-0.14, -0.05)
1967-75	0.87* (0.84, 0.9)	1.07 (1.04, 1.11)	-0.15 (-0.2, -0.11)
1976-82	0.5* (0.47, 0.54)	0.69 (0.65, 0.73)	-0.19 (-0.25, -0.14)
1983-90	0.36* (0.33, 0.4)	0.54 (0.5, 0.58)	-0.29 (-0.34, -0.24)
1990-post	0.57* (0.54, 0.6)	0.8 (0.76, 0.83)	-0.19 (-0.23, -0.14)
Household Tenure			
private rented			
owner occupied	-0.13* (-0.15, -0.1)	-0.09 (-0.11, -0.06)	-0.31 (-0.35, -0.28)
social rented	-0.19* (-0.22, -0.15)	-0.22 (-0.25, -0.19)	-0.12 (-0.16, -0.08)
Number of bedrooms			
1 bedroom			
2 bedrooms	0.31* (0.28, 0.34)	0.42 (0.39, 0.45)	0.04 (0, 0.08)
3 bedrooms	0.37* (0.33, 0.4)	0.49 (0.45, 0.52)	0.03 (-0.01, 0.07)
4 bedrooms	0.22* (0.19, 0.26)	0.35 (0.31, 0.39)	-0.15 (-0.2, -0.1)
5+ bedrooms	0.44* (0.4, 0.49)	0.26 (0.21, 0.31)	0.43 (0.37, 0.48)
Government Region			
South East			
East England	0.1* (0.08, 0.12)	0.09 (0.06, 0.11)	0.11 (0.08, 0.15)
East Midlands	0.13* (0.11, 0.16)	0.12 (0.1, 0.15)	0.15 (0.12, 0.19)
London	-0.13* (-0.16, -0.11)	-0.17 (-0.2, -0.14)	0.09 (0.05, 0.13)
North East	0.38* (0.35, 0.41)	0.39 (0.36, 0.42)	0.14 (0.1, 0.19)
North West	0.23* (0.21, 0.26)	0.25 (0.23, 0.27)	0.14 (0.1, 0.17)
South West	0 (-0.03, 0.03)	0.02 (-0.01, 0.04)	-0.06 (-0.1, -0.02)
West Midlands	0.24* (0.22, 0.27)	0.24 (0.21, 0.26)	0.15 (0.11, 0.18)
Yorkshire and the Humber	0.05 (0.02, 0.07)	0.04 (0.01, 0.06)	0.16 (0.12, 0.19)
Median Neighbourhood Income Quintile			
Quintile 1			
Quintile 2	-0.03 (-0.05, 0)	-0.01 (-0.04, 0.01)	-0.07 (-0.1, -0.03)
Quintile 3	-0.1* (-0.12, -0.07)	-0.06 (-0.09, -0.04)	-0.18 (-0.22, -0.14)
Quintile 4	-0.15* (-0.18, -0.13)	-0.11 (-0.14, -0.08)	-0.27 (-0.31, -0.23)
Quintile 5	-0.24* (-0.27, -0.21)	-0.18 (-0.22, -0.15)	-0.37 (-0.41, -0.32)
Neighbourhood Rurality			
Rural or village hamlet			
Town and Fringe	0.04 (0.01, 0.08)	0.03 (-0.01, 0.07)	0.05 (0, 0.1)
Urban > 10K	0.05 (0.01, 0.08)	0.03 (-0.01, 0.06)	0.09 (0.05, 0.14)
Proportion of neighbourhood in receipt of benefit			
<33%			
33-66%	0.02 (0, 0.04)	0.02 (0, 0.04)	0.02 (-0.01, 0.06)
>66%	0.04 (-0.01, 0.08)	0.04 (-0.01, 0.08)	0.08 (0.02, 0.14)
Proportion of neighbourhood in receipt of pension			
≤10%			
>10%	0.03 (0.02, 0.05)	0.05 (0.03, 0.07)	-0.02 (-0.05, 0.01)

Notes: †Fabric retrofits include: cavity wall insulation, loft insulation, or double glazing installation; ¥ Heating retrofits include: boiler replacement, heating controls; *Significant at the 95% confidence level

8.3.2 Change in gas demand in study sample of English dwellings

The mean annual change in gas demand for the sample of English dwellings over the study period 2004 to 2007 was -810 kWh/year, or -4.2% per year. Figure 32 shows the shift in the distribution in annual gas demand.

Using a GLM regression model, the presence of fabric or heating energy efficiency retrofit is shown to be significantly associated with a reduced demand for gas (see Table 31). Adjusting for dwelling type, age, tenure, size, region and median neighbourhood income, the presence of an installed fabric energy efficiency retrofit in English dwellings is on average -790 kWh/year, or a 3.9% reduction from the stock mean gas demand in 2006, and the presence of a heating energy efficiency retrofit is on average -1950 kWh/year, or a 10.4% reduction from the stock mean gas demand in 2006. The fabric and heating measures were not additive, and when installed in the same year represented an average reduction in demand of 2,290 kWh/dwelling/year, or a 11.7% reduction from the 2006 stock mean gas demand. The presence of energy efficiency retrofits does appear to have a significant impact on gas demand even after adjusting for differences in dwelling and household characteristics, such as number of exposed walls (i.e. dwelling type) and proxies of energy performance (i.e. dwelling age). These results suggest there is an impact on gas demand attributable to the retrofit alone.

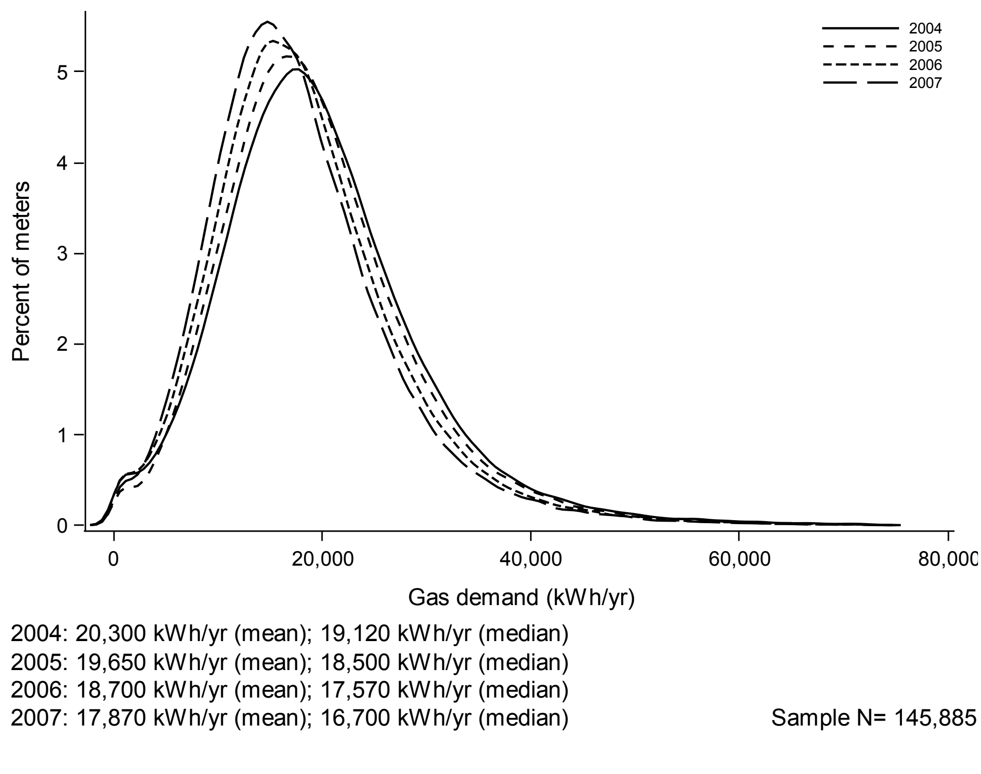


Figure 32 - Distribution of annual gas demand (kWh/year) per dwelling in study sample, 2004 to 2007

The association of the change in energy demand from 2005 to 2007 and specific energy efficiency retrofits installed in 2006 are shown in Table 32, both unadjusted (model 1) and adjusted for dwelling type, age, tenure, size, region and neighbourhood income (model 2). In the following results, only the adjusted values are described, though there was little difference in the resulting values between the two models. The mean change in gas demand associated with the installation of cavity wall insulation for dwellings was -1050 kWh/year, or 5.6% of the stock mean gas demand in 2006. For dwellings with loft insulation installed in 2006, this was associated with a 150 kWh/year increase (~1% of the 2006 stock mean gas demand), though this was not statistically significant at the 95% level. The installation of double-glazed windows was also not statistically significant and was associated with a mean change in demand of -12 kWh/year. Condensing boilers were associated with a mean change in demand of 1060 kWh/dwelling, or 5.7% of the stock mean gas demand in 2006, significant at the 95% level.

The trends described above compare closely to the proportional change in gas demand from 2005 to 2007 compared to 2005, shown in Table 33. The results, using a GLM model, are adjusted for dwelling type, age, tenure, size, region and neighbourhood income, and show that cavity wall insulation and condensing boiler installations had a -4.9% and -5.5% change in demand from 2005 to 2007, respectively. Note that loft insulation and double-glazing installation showed almost no associated change in demand. The combined effect of additional measures showed greater reductions in the change in demand, with combinations that included condensing boiler installations and cavity insulations being associated with the largest changes. The adjusted added effect of cavity wall and loft insulation and a condensing boiler installed in 2006 was associated with an 11.2% change in demand. There is weak evidence to suggest that the retrofits are additive, i.e. combined measures achieving the reductions for single measures added together. The table also contains the sensitivity analysis that includes electricity in the dependant variable. Whilst the trend is the same, the magnitude of change is less when including electricity. Unrestricted electricity demand is approximately a one fifth the demand for gas, which means that it should have little overall effect. However, due to its low annual rate of change (i.e. ~1%) it slightly reduces the magnitude.

When stratified by household tenure, the presence of an energy efficiency retrofit is significantly associated with changes in gas demand for owner-occupiers (Table 34). The change in demand for both socially and privately rented dwellings with the presence of a measure is not statistically significant. An analysis of variance using least squared means (due to the unbalanced nature of the design, i.e. uneven group sizes) showed no significant difference in the change in gas demand between tenure types (test not shown). However, the associated proportional change in demand for owner-occupiers was higher than the stock averages shown in Table 33. Focusing on cavity wall insulation, socially rented dwellings show the lowest change in demand (~ -3%), while privately rented dwellings change the most (~ -8%). Changes in energy demand associated with condensing boiler installations are greatest for owner occupiers (~6%), while socially and privately rented dwellings show changes around -4%. The stratified change in gas demand associated with retrofits by dwelling age shows larger changes in gas demand for the 1967-75 group (see Table 35).

Stratifying by neighbourhood income shows a more consistent trend with neighbourhoods in the lower-income quintile associated with on average lower changes in gas demand and higher incomes having greater changes (see Table 36). The results for cavity wall insulation and condensing boiler installations for the most part appear to be significant.

Table 31 – Regression coefficients (standard deviation) of gas demand per dwelling adjusting for selected dwelling, household and neighbourhood characteristics with and without fabric and heating measures

	Gas Demand in 2006			
	Energy Demand	Fabric† retrofit	Heating retrofit¥	Heating and Fabric retrofit¥
	N=168,988	N=168,988	N=168,988	N=168,988
Factors	Coefficient Estimate* (Standard Error)			
Sample mean	18400	18400	18400	18400
Intercept	22030* (216)	21800* (149)	21820* (149)	21850* (149)
Energy Efficiency Retrofit in 2005				
Fabric retrofit		-800* (60)		-460* (63)
No Fabric Retrofit		0		0
Heating Retrofit			-1990* (96)	-1760* (101)
No Heating Retrofit			0	0
Dwelling Type				
Detached	3950* (59)	4000* (59)	4120* (59)	4130* (59)
Semi-detached	1440* (51)	1450* (51)	1440* (51)	1450* (51)
Flat, all	-1990* (88)	-1990* (88)	-1960* (88)	-1970* (88)
Terrace	0	0	0	0
Dwelling Age				
Pre-1900	3930* (95)	3900* (95)	3980* (95)	3950* (95)
1900-49	2630* (69)	2630* (69)	2680* (69)	2660* (69)
1950-66	1410* (71)	1410* (71)	1450* (71)	1440* (71)
1967-75	1100* (75)	1170* (75)	1120* (75)	1160* (75)
1976-82	-210 (94)	-200 (94)	-200 (94)	-190 (94)
1983-90	-1160* (88)	-1190* (88)	-1180* (87)	-1200* (87)
Post-1990	0	0	0	0
Dwelling Tenure				
Owner occupied	1290* (62)	1270* (62)	1240* (62)	1240* (62)
Private rented	500* (96)	460* (96)	500* (96)	480* (96)
Social rented	0	0	0	0
Number of Bedrooms				
1 bedroom	-7320* (129)	-7360* (129)	-7400* (129)	-7420* (129)
2 bedrooms	-6610* (102)	-6620* (102)	-6700* (102)	-6700* (102)
3 bedrooms	-3960* (98)	-3960* (98)	-4050* (98)	-4040* (98)
4 bedrooms	260 (102)	240 (102)	120 (103)	130 (103)
5+ bedrooms	0	0	0	0
Government Region				
East England	-2080* (83)	-2050* (83)	-2060* (83)	-2050* (83)
East Midlands	-1560* (86)	-1540* (86)	-1560* (86)	-1540* (86)
London	-1060* (89)	-1000* (86)	-980* (86)	-990* (86)
North East	800* (98)	840* (97)	800* (97)	820* (97)
North West	-440* (75)	-400* (75)	-420* (74)	-410* (74)
South East	-2170* (76)	-2160* (76)	-2170* (76)	-2170* (76)
South West	-3340* (84)	-3350* (84)	-3370* (84)	-3370* (84)
West Midlands	-1100* (83)	-1030* (83)	-1050* (83)	-1030* (83)
Yorkshire	0	0	0	0
Median Neighbourhood Income				
Quintile 1	-3210* (96)	-3010* (73)	-2970* (73)	-2960* (73)
Quintile 2	-2830* (74)	-2710* (67)	-2680* (67)	-2670* (67)
Quintile 3	-2390* (65)	-2350* (64)	-2330* (64)	-2330* (64)
Quintile 4	-1750* (62)	-1740* (61)	-1730* (61)	-1730* (61)
Quintile 5	0	0	0	0
Model R-Square	0.242	0.242	0.243	0.243

Notes: †Fabric retrofits include: cavity wall insulation, loft insulation, or double glazing installation; ¥ Heating retrofits include: boiler replacement, heating controls; *Significant at the 95% confidence level

Table 32 - Regression coefficients (standard error) of change in gas demand (2005 to 2007) per dwelling adjusting for selected dwelling, household and neighbourhood characteristics with and without fabric and heating measures

	Change in gas demand from 2005 to 2007			
	Cavity Insulation	Loft Insulation	Double Glazing Installation	Condensing Boiler Replacement
Factors	Coefficient Estimate* (Standard Error)			
Model 1 - Unadjusted				
Intercept	-1456* (15)	-1456* (15)	-1456* (15)	-1456* (15)
Measure in 2006	-1107* (76)	99 (88)	40 (184)	-1055* (137)
No Measure 2005 to 2007				
Model 2 - Fully Adjusted				
Intercept	-1497* (126)	-1525* (127)	-1524* (128)	-1496* (128)
Measure in 2006	-1047* (77)	153 (88)	-12 (183)	-1059* (138)
No Measure 2005 to 2007				

Notes: *Significant at the 95% confidence level; †Adjusted for dwelling type, age, tenure, number of bedrooms, region and neighbourhood income quintile.

Table 33 - Regression coefficients (standard error) of proportional change in gas demand (2005 to 2007) per dwelling adjusting for selected dwelling, household and neighbourhood characteristics with and without fabric and heating measures

		Proportional change in demand from 2005 to 2007 with measure in 2006	
		Gas	Gas + Electricity
		Adjusted†	
Interventions¥	**N**	Coefficient Estimate* (Standard Error)	
Cavity insulation	104,623	-0.049* (0.003)	-0.042* (0.003)
Loft insulation	103,615	0.009 (0.004)	0.008 (0.004)
Double glazing installation	101,391	0 (0.008)	0.003 (0.008)
Boiler installation	101,897	-0.055* (0.006)	-0.045* (0.006)
Cavity and loft insulation	102,661	-0.057* (0.005)	-0.052* (0.005)
Boiler, cavity and loft Insulation	101,061	-0.112* (0.012)	-0.1* (0.011)
Glazing, boiler, cavity and loft insulation	100,771	-0.1 (0.033)	-0.131 (0.004)
Glazing, cavity and loft Insulation	101,160	-0.034 (0.01)	-0.034 (0.01)
Glazing, boiler and loft insulation	100,778	-0.099 (0.014)	-0.104 (0.007)
Glazing and cavity wall insulation	101,474	-0.031* (0.008)	-0.019 (0.007)

Notes: *Significant at the 95% confidence level; †Adjusted for dwelling type, age, tenure, number of bedrooms, region and neighbourhood income decile; †Intercept not shown

Table 34 - Regression coefficients (standard error) of proportional change in gas demand (2005 to 2007) per dwelling by household tenure, adjusting for selected dwelling, household and neighbourhood characteristics with and without fabric and heating measures

	Proportional change in gas demand from 2005 to 2007		
	Household Tenure		
	Owner Occupied	Private Rented	Social Rented
Interventions† in 2006	Coefficient¥ Estimate* (Standard Error)		
Cavity insulation	-0.053* (0.004)	-0.076 (0.026)	-0.03 (0.009)
Sample size n=	83,122	6,280	15,221
Loft insulation	0.003 (0.004)	0.008 (0.023)	0.034 (0.011)
Sample size n=	82,271	6,302	15,042
Double glazing installation	0.008 (0.009)	-0.048 (0.047)	-0.016 (0.02)
Sample size n=	80,547	6,216	14,628
Boiler installation	-0.063* (0.008)	-0.043 (0.028)	-0.041 (0.013)
Sample size n=	80,759	6,271	14,867

Notes: *Significant at the 95% confidence level; †Adjusted for dwelling type, age, number of bedrooms, region and neighbourhood income quintile; †Intercept not shown

Table 35 - Regression coefficients (standard error) of proportional change in gas demand (2005 to 2007) per dwelling by dwelling age, adjusting for selected dwelling, household and neighbourhood characteristics with and without fabric and heating measures

		Proportional change in gas demand from 2005 to 2007						
		Dwelling Age						
		Pre-1900	1900-1949	1950-1966	1967-1975	1976-1982	1983-1990	Post-1990
Interventions† in 2006		Coefficient¥ Estimate* (Standard Error)						
Cavity insulation		-0.06 (0.053)	-0.035* (0.008)	-0.029* (0.007)	-0.072* (0.006)	-0.04 (0.011)	-0.015 (0.013)	-0.054* (0.011)
	Sample size n=	7,258	31,762	23,857	15,056	6,800	8,563	11,327
Loft insulation		-0.012 (0.019)	0.015 (0.008)	0.013 (0.009)	0.007 (0.007)	0.019 (0.017)	0.004 (0.016)	-0.002 (0.012)
	Sample size n=	7,374	31,907	23,527	14,446	6,580	8,491	11,290
Double glazing installation		-0.018 (0.035)	0.018 (0.014)	-0.013 (0.016)	-0.05 (0.029)	0.002 (0.033)	0.025 (0.025)	-0.008 (0.025)
	Sample size n=	7,281	31,341	23,213	13,602	6,476	8,404	11,074
Boiler installation		-0.058 (0.022)	-0.037 (0.011)	-0.055* (0.011)	-0.092* (0.019)	-0.112* (0.025)	-0.054 (0.022)	-0.035 (0.025)
	Sample size n=	7,349	31,491	23,390	13,668	6,502	8,423	11,074

Notes: *Significant at the 95% confidence level; †Adjusted for dwelling type, tenure, number of bedrooms, region and neighbourhood income quintile; ‡Intercept not shown

Table 36 - Regression coefficients (standard error) of proportional change in gas demand (2005 to 2007) per dwelling by neighbourhood income quintile, adjusting for selected dwelling, household and neighbourhood characteristics with and without fabric and heating measures

		Proportional change in gas demand from 2005 to 2007				
		Quintile ranking of neighbourhood income				
		Rank 0	Rank 1	Rank 2	Rank 3	Rank 4
Interventions† in 2006		Coefficient¥ Estimate* (Standard Error)				
Cavity insulation		-0.039* (0.008)	-0.051* (0.008)	-0.049* (0.007)	-0.055* (0.008)	-0.056* (0.008)
	Sample size n=	19,708	20,586	21,105	21,373	21,851
Loft insulation		0.018 (0.009)	0.001 (0.009)	0.008 (0.009)	0.012 (0.009)	0.001 (0.009)
	Sample size n=	19,541	20,412	20,792	21,135	21,735
Double glazing installation		0.011 (0.021)	0.024 (0.018)	-0.039 (0.018)	-0.013 (0.019)	0.013 (0.015)
	Sample size n=	18892	19,930	20,383	20,757	21,429
Boiler installation		-0.01 (0.013)	-0.053* (0.013)	-0.057 (0.016)	-0.081* (0.015)	-0.105* (0.013)
	Sample size n=	19,087	20,075	20,429	20,822	21,484

Notes: *Significant at the 95% confidence level; †Adjusted for dwelling type, tenure, number of bedrooms, and region; ‡Intercept not shown

8.4 Discussion

This section discusses the findings related to the uptake of energy efficiency retrofits and their variation among the sample of English dwellings over the study period, along with the association between change in energy and energy efficiency retrofits and the influence of dwelling and household characteristics. Finally, study limitations are discussed.

8.4.1 Determinants of energy efficiency retrofits uptake

Owner-occupied, 3 bedroom detached dwellings built in the mid-20th century in areas of lower neighbourhood income in the northern English regions were associated with a higher probability of having measures. This reflects both the nature of the measures (i.e. older more inefficient homes), ability to accept or undertake measures, and being the target of government programmes. Further, the presence of fabric and heating energy efficiency retrofits was shown to have a significant effect (a reduction) on energy demand within the same gas year with ~4% for fabric measures and 10% for heating measures, after adjusting for dwelling and household characteristics. These results suggest that retrofits have an attributable impact on energy demand. Retrofits were significantly associated with changes in energy demand (after adjusting for dwelling and household characteristics) across a three-year period, and the change in demand increases as more retrofits are installed - suggesting a dose-response like relationship (non-linear). Some retrofits (i.e. loft insulation and double glazing) do not appear to have a significant impact on energy demand over the period.

The uptake model found that as neighbourhood incomes increased, the probability of having a major measure installed over the study period decreased, offering further support that households living in areas marked by higher incomes are not investing in their property compared to low-income areas that are the focus of government policy, therefore supporting the hypothesis (H1) that low-income households are more likely to receive and accept energy efficiency retrofits. Broadly speaking, ownership and income remain important determinants of having energy efficiency retrofits.

The findings supports the notion that there is a lack of investment by owner occupiers but ultimately rejects the hypothesis that people who own their home receive and accept more measures than other tenure types (H2). However, the finding is not necessarily suggesting that this household type is a driver of uptake, but rather reflects the investment in energy efficiency for vulnerable customers through supplier obligation (which comprise the bulk of the interventions in England) and government schemes (such as Warm Front) over the study period (Rosenow, 2012).

Older dwellings were less likely than the stock average to have reported retrofits during the study period. The model showed an inverted 'U-shape' curve for the uptake of fabric retrofits, with both older (and in theory less efficient) and newer dwellings not having insulation installed, and with dwellings built in the 1967-75 and 1976-82 age bands having the highest probability of retrofits over the period. This was particularly the case for fabric measures and the high uptake rates of cavity wall insulation. This finding rejects the notion that older dwellings are more likely to have energy efficiency measures (H3). While older dwellings are likely to be relatively more inefficient and are therefore in greater 'need' of retrofits that improve their fabric, the finding points to the nature of the insulation needed for older dwellings. Older English dwellings are more likely to be constructed of solid brick or stone and that means 'cheap' insulation techniques such as blown insulation in cavity walls is not a viable option. Both the supplier and government retrofit programmes excluded insulation for solid-walled dwellings, instead opting for the 'low-hanging' fruit of cavity wall and loft insulation. This leaves solid-wall dwellings as the single largest retrofit action to be tackled.

8.4.2 Determinants of energy savings

Introducing energy efficiency retrofits resulted in attributable energy savings, after controlling for the effects of dwelling type, size, age, tenure, region and neighbourhood income. The retrofit associated with the largest change in (adjusted) energy demand over the three-year period was the installation of a condensing gas boiler, -5.2%, with the second largest being cavity wall insulation at around -3.8%. The effect of the combined installation of a condensing boiler, and cavity and loft insulation was around -11%. These findings are very similar in scale to results from a previous study of British houses under the Warm Front scheme, which found that loft and full cavity wall insulation reduced demand within a one-year period by 10-17% (Hong et al., 2006). While the attributable change in energy demand over the period associated with the reported installation of a retrofit may seem lower compared to notional 'savings' or the Hong et al. study, it is important to bear in mind that these changes control for physical, household and area-based factors. The effect of controlling for physical factors on energy 'savings' means that the effect of number and area of exposed walls is removed as is any effect related to the age of the dwelling, while household effects could reflect ability to afford larger areas to heat and greater comfort conditions. By controlling for these factors, the effect of the retrofit can be isolated, which is important for determining a 'baseline' of expected change in demand on which future estimates could rely.

After adjustment, neither loft insulation or double glazing were associated with significant energy savings over the three-year period. It is not necessarily that these measures do not save energy, but that the effect is on average fairly small and/or cannot easily be detected using annualised energy data. Although glazing is one of the thermally weakest elements of the building fabric, the area of double glazing replaced will have an effect on the potential energy savings. However, because of the way the data was reported, it was not possible to adjust account for glazing area replaced. The majority of loft insulations were top-ups of around 5-75mm and therefore the change in gas demand would be minimal.

The findings also show that there are differences in the savings associated with certain household/ dwelling groups. Dwelling age appears to have an inconsistent effect on changes in gas demand, with the 1967-75 group having much greater reduction in demand for all single retrofit measures compared to other age bands. It is not necessarily the case, therefore, that older dwellings are less likely to have greater energy savings than newer dwellings (H4). The variation may in part be explained by the eligibility and type of retrofits installed. Cavity wall filling is most applicable to mid-century and onward dwelling age bands, with few being applicable to pre-1950 or post-1990 dwellings. The impact of boilers amongst this mid-century group was also greater (after controlling for size), which could reflect a number of building design features, such as the nature of the installed heating systems which according to the 2010 EHS have a higher prevalence of gas central heating (CLG, 2013b),

Changes in energy demand were lower among households with low socio-economic levels, such as renting or living in areas of lower income, therefore supporting the hypothesis that lower incomes are less likely to realise energy savings compared to higher-income households (H5). This trend may be attributed to these households have higher levels of energy utility (i.e. greater need for the amount used) (Meier and Rehdanz, 2010). Both social renting and living in lower income neighbourhoods is associated with reduced energy demand, even after controlling for type and age of dwelling (Table 31), which may also suggest that these households have a greater potential for increasing demand that energy efficiency retrofits enable. The differences could also be construed as 'comfort taking', whereby these households in areas of lower income reduce the potential 'energy savings' by taking the savings in the form of temperature increases, an effect that has been shown in a study of vulnerable households in England (Hong et al., 2009).

8.4.3 Strengths and weaknesses

The study relies on reported measures drawn from a number of programmes over the study period collected by the Energy Saving Trust into the Home Energy Efficiency Database. Whilst EST undertook precautions to check data for erroneous entries and applied 'trust' flags to data from different suppliers (i.e. accredited installers and surveyors were more trusted than web-based surveys), using this data means that it is not possible to verify the accuracy of the reported data. This could mean that some homes may have reported some measures when none were installed (or vice versa). However, it was assumed in this study that such events would likely occur randomly (i.e. without systematic bias) because of the number of data providers and the low theoretical probability of installers, assessors and homeowners consistently mis-reporting the same class of measure. Also, such occurrences were assumed unlikely to be very widespread, as many retrofits require specialist installers (cavity wall insulation and double glazing installation) and are regulated (i.e. condensing boiler and double glazing installation). Despite this, it is acknowledged here that there was a particular problem with the way double-glazing installations were reported in HEED. The reporting did not include the number of glazing units replaced (or the percentage) in the dwelling and therefore it was not possible to control for this effect, possibly underestimating the potential impact.

In terms of representativeness, the study sample was drawn to be representative of six key English dwelling and household variables, including: dwelling age, type, and number of bedrooms, the region and household tenure. This means that sample cannot necessarily represent other non-sampled variables, particularly as they relate to the household (e.g. occupants or income levels). The study should only be used for the purposes of describing the housing stock and not the households therein.

8.4.4 Study conclusions

This study has shown that it is possible to construct a reasonably well-defined cohort sample from pre-existing datasets that is broadly representative of the English housing stock and use population-level analysis techniques to assess the uptake of energy efficiency retrofits and the impact of these interventions on energy demand over a defined study period. This study provides a step towards an empirically-based population-level approach to studying energy demand that accounts for variation amongst different dwelling and household groups. The method emphasises associations, rather than causation, as a means for generating hypotheses that can be further explored in more detailed studies.

Energy efficiency retrofits *do* have an attributable impact on reducing space-heating related energy demand, and combining retrofits displays a dose-response like effect on energy demand, after controlling for household and dwelling factors. In order to meet the intended energy efficiency targets, the retrofit uptake rate will need to significantly increase. Meeting these uptake and energy savings targets can be broadly achieved using existing technologies and deployment processes.

8.5 Critical discussion of cohort study approach

This section critically discusses the use of the cohort study as a method of examining the uptake of energy efficiency and the impact of those measures on change in energy demand. First, it examines the limitations of the study method in terms of different forms of bias, the variables used and the available information. Concluding thoughts on the method then follow.

8.5.1 Limitations of the study method

A cohort study is primarily designed to answer the question: what are the effects of a given exposure? The cohort study method used here allowed for the comparison of a number of potentially interacting factors that influence the uptake of efficiency measures and also changes in demand. The strength of a cohort study is that these factors can be analysed simultaneously. In the case of energy demand, the exposure could be a socio-economic, occupant or physical factor that is expected to influence energy demand. In this method study, the 'exposures' of interest were the presence of an energy efficiency retrofit. Within this structure, it was possible to examine what influence household and dwelling characteristics had on the energy efficiency retrofit 'exposure'.

Sampling of the population is very important to ensure that external validity and internal validity are achieved. Sampling requires deciding on 'who' the target population is, into whom the study will attempt to provide insight. The study needs to draw from a source population, for which not all will be eligible or available for study. In this method study, the target population was the English housing stock (not households). The HEED+Energy database provided a group of eligible dwellings from which to construct a study sample, using the EHS as the frame from which to draw the sample. The study sample comprised those dwellings with non-missing data for a selection of variables and valid energy meter data. External validity is achieved if the sample is applicable to the wider target population. In this method study the key variables and the location were used to try and ensure that a representative sample of English houses was drawn.

Internal study validity is where bias exists due to bias in the observations, differences in the population compared, and confounding. Due to the nature of the collection process of HEED and its use in this study, it is not possible to determine the extent of misclassification in the sample, which is a potential limitation. However, as stated earlier, there is very likely a random misclassification of variables. Further, in comparison studies there must be a true measure of actual exposure, requiring good evidence to describe the exposure levels in those classified as 'not exposed'. As with misclassification, it was determined that HEED likely covered most substantial energy efficiency retrofits by way of regulations and accredited installers. Unlike a case-control study for looking at the change in energy demand, this cohort study made comparisons against all other dwellings that did not have a reported efficiency measure within the study period. It was possible that they had a measure outside of the period (i.e. pre 2004). Regardless, the focus was a comparison of the change in gas demand for those who did and did not report a measure over the study period and not a comparison between those with and without an efficiency retrofit. The difference is that a case-control study would be more suitable in compare similar dwellings that did and did not have, say, condensing boilers, to determine what impact these might have on gas demand.

A limitation of retrospective cohort studies is that observational studies (particularly retrospective) do not provide any protection against confounding variables that have not been measured (Elwood, 2007). It may be that outside the sample frame and the putative influencing variables selected there are other factors that influence gas demand that were not measured and therefore could confound or add bias to the study. For example, the study did not have a measure of energy prices at the household level (though data is available at the regional level), which is known to have an effect on energy demand, though these impacts tend to be fairly moderate over the long term (Meier and Rehdanz, 2010). The method study only had a limited selection of variables available and it is very likely that other factors would further influence the change in demand, in particular household features such as number of persons and employment status.

This cohort study sought to represent the English housing stock. Some sub-group analysis was undertaken to determine if there were significant differences between tenure, age of dwelling and neighbourhood income level. However, because the study sample was not constructed for all the potential variations in energy efficiency retrofits it is possible that some subgroups were too small for detecting differences. Ideally subgroup analysis should be defined from the outset to ensure that the samples are constructed to be internally valid.

It was for this reason that only major retrofits were examine in sub-group analysis where there was sufficient group size.

Finally, cohort studies provide an effective method of examining causal concepts through the use of longitudinal data. Using Hill's guidelines, a causal link could exist if the timing is correct, there is a strong relationship, there is a dose-response, and the impact is consistent and specific. In this study, there is certainly some potential that several major energy efficiency retrofits *cause* a reduction in energy demand. Both cavity wall insulation and condensing boilers appear to display many causal signs, such as: timing (i.e. the reduction in demand greater than the averages occurs following the retrofit), consistency (i.e. the reduction happens across household groups), physical processes, and dose-response (i.e. the impact is greater as more retrofits are combined).

Cohort studies, if properly designed and carried out, "should have less observation bias, give clearer evidence of the time relationships of association, and have a comparison group whose results are more easily interpreted" (Elwood, 2007). In this study, the cohort study provides evidence to show what dwellings and household features are associated with the uptake of a retrofit, and also what impact the installation of a retrofit has on gas demand.

8.5.2 Study method conclusions

Overall, the application of a cohort study to examine the uptake of energy efficiency retrofits and their impact on energy demand at the individual dwelling level was coherent and offered plausible findings. The study met its main aim, which was to determine the association between dwelling and household features and variation in retrofit uptake, the impact of retrofits on energy demand.

This study was the culmination of a series of investigations on energy demand using an epidemiological approach that began with an ecological study to generate hypotheses around what dwelling and household features affected uptake rates at a neighbourhood level. A cross-sectional study was then used to explore the variation in energy demand amongst the population and identify potentially important drivers of energy demand, including energy efficiency levels. Finally, this cohort study was used to examine the uptake of retrofits and also the impact that those retrofits had on gas demand, while controlling for known influencing factors.

8.6 Summary

This chapter investigated the application of the third proposed method study, i.e. a cohort study of the uptake of energy efficiency retrofits and their impact on energy demand. The association was examined between the uptake of energy efficiency measures (insulation, heating and draught proofing) and changes in gas and electricity demand for a cohort of English dwellings was examined. In the sample, the effects after adjustment for dwelling size, age, tenure, demonstrated a dose-response effect whereby higher levels of energy efficiency were associated with large changes in energy demand. There were also strong effects on change in energy demand for individual measures.

In the following chapter, an overarching discussion is provided of the studies and their contribution to a population-level empirically-based method for studying energy demand.

Chapter 9 Discussion

"Supporting the study of population-level end-use energy demand in dwellings through energy epidemiology"

Chapter Introduction

The method studies illustrated the application of the epidemiological approach, study designs, analysis techniques and key concepts to the study of relevant practical problems of energy demand. The ecological study examined the uptake of energy efficiency in English households from 2000 to 2007 and investigated whether there were differences in uptake levels associated with neighbourhood features. The descriptive cross-sectional study examined the distribution of energy demand among the UK population and putative influences from dwelling characteristics and energy efficiency levels. The cohort study examined the association that dwelling and household characteristics had on the uptake energy efficiency retrofits and the impact those retrofits had on changes in energy demand.

Up to this point the thesis has drawn together information, applied research methods and undertaken analyses but has not discussed the potential implications this approach has for research. This chapter considers the implications of applying an epidemiological approach to the study of energy demand.

The chapter first considers whether the approach is able to support the challenge posed to the energy demand and buildings research field in the face of the pressing challenges of shifting to a low-carbon society. It then considers whether an epidemiological approach is able to overcome the limitations of the current approach previously identified. It then considers the benefits and challenges of applying the approach within the research field. It goes on to highlight the findings from the methods studies and the application of the study designs in addressing the research questions posed. Finally, it concludes with a discussion of how the approach could be extended in support the development of a strong evidence base.

9.1 Energy epidemiology and the study of population-level energy demand

The shift to a low carbon economy and the need to address energy demand related priorities, such as fuel access and affordability, are drivers of change in the way energy is used and demanded in buildings at the national and international scale. In the UK, such change will involve the retrofit of millions of buildings and changes in energy demand services. A number of authors have pointed out that studying energy demand in buildings at a population level is different than in individual or small samples of buildings or users because of the heterogeneity present among populations for which studies need to account (Oreszczyn and Lowe, 2010; Skea, 2012; Summerfield and Lowe, 2012). Essential to the development of a strong evidence base is the use of empirically derived data from large populations that can represent the real-world conditions of a complex building stock and population (Skea, 2012; Summerfield and Lowe, 2012). Evaluating policies and determining the effect of technologies *in situ* in millions of buildings means using techniques that support that level of analysis. However, the literature shows that the current approach to studying populations of energy demand in buildings is lacking a methodological framework for the study of populations of energy use and buildings, meaning there is limited evidence available to respond to and address pressing issues.

The argument being made here is that in order to address the challenge of shifting to a low-carbon society through the modification of millions of buildings, changing energy using practices of populations, and the application of technologies across the building stock, a more sophisticated approach to studying population-level energy demand is needed. It is proposed here that an epidemiological approach provides a framework through which to do this.

This thesis tested the assertion of whether the epidemiological approach provided the concepts, methods and tools for undertaking research on outcomes among populations. It asked if (RQ1) an epidemiological approach could support the population-level study of energy demand (focused on the residential sector)? And, if so, (RQ2) could the aims and objectives of an epidemiological approach as applied to energy demand be identified and would they support of the research paradigm?

The following sections discuss the implications of using the epidemiological approach to study energy demand and how it supports the call for action under the dominant low-carbon society paradigm. The implications of an empirically-driven approach are discussed in terms of how it supports drawing together the various disciplines involved in energy demand research. Finally, there is discussion as to whether the epidemiological approach can offer improvements to studying problems of energy demand in buildings, which are complex and subject to many interacting components. The aims and objectives of an energy epidemiological approach are discussed in terms of addressing these complexities.

9.1.1 Responding to the call for action

The call for action to transform the built environment and respond to the threat of climate change has been clearly made and has major implications for energy demand in buildings. For example, in the UK the plans set out under the *Energy Efficiency Strategy* include a number of polices to reduce energy demand through improvements in the energy performance of buildings and shifting occupant behaviour, along with an aggressive decarbonization of the energy supply system. Under these plans the UK government has set out pathways that could see millions of energy efficiency retrofits installed in houses leading up to 2050, resulting in £ billion's invested. However, current research and available evidence leaves the ability to achieve these reductions uncertain. A recent all-party Parliamentary Group (APPG) on Excellence in the Built Environment (EBE) highlighted the scale of the intervention required: 180,000 installations per year by 2050 - or 500 per day. In the EBE report, the APPG emphasized that the major challenges to delivering these changes were: limited installer capacity, homeowner apathy towards energy efficiency, and minimal market pressure.

The evidence needs have been set out by a number of authors and have focused on: more empirically-based analysis of good-quality population level datasets, more research that is of a high quality, the use of methods and analysis techniques that are appropriate for understanding variation in energy demand, and improving inputs into models (Skea, 2012; Summerfield and Lowe, 2012). They point out that the existing approach to the study of energy demand, which is largely undertaken through engineering/physics-based methods or social science economic approaches, have not provided the breadth of evidence to effectively address gaps in knowledge on energy demand and the impact of energy efficiency retrofits.

It is proposed here that epidemiology offers a relevant framework for undertaking complex, interdisciplinary energy demand research at the population-level through an empirically-based approach that reflects variation among the population. It is argued here that the proposed central paradigm for energy epidemiology supports the shift to a low-carbon society. This paradigm is:

> That the shift to a low-carbon society along with the alleviation of energy-related social and environmental phenomena, such as fuel poverty and comfort, can be improved through population-based methods that analyse patterns and systems of energy demand services in order to better understand the practices, drivers, causes and differences of energy demand outcomes.

It is further argued that the epidemiological approach supports this paradigm through its conceptual and methodological framework from which to develop evidence for building models, and to support policy to manage energy demand. This thesis offers a selection of study examples that illustrate the research approach in practice.

9.1.2 An empirically driven approach

The intent of the adaptation of the epidemiological approach to the study of energy demand is to illustrate a means to observe and describe the trends and patterns of energy demand at a population level, to undertake and contextualize interventional studies, and to identify associations between factors that lead to an energy demand-related outcome or event. For research, the thesis aims to show at the energy epidemiological approach draws together tools and methods that can serve the study of the complex interactions between social, economic, technological, physical and environmental factors that lead to an energy demand outcome among a population.

This thesis therefore intended to determine the extend to which the epidemiological method supports a process through which to develop a body of evidence that is representative of a given population and to establish causal associations. This process first relies on the empirical nature of observed outcomes in the first instance in order to a) identify a problem, b) define and characterize the condition, and c) collect information from a study group. The second part of the process is exploratory in nature and seeks to d) examine influencing factors, e) explore whether patterns and trends exist. Leading from these initial studies of association, the third part of the process looks to f) develop hypotheses, g) test hypotheses through appropriate study designs. The development of hypotheses will build on the available detailed understanding of underlying mechanisms that may be acting within a population, and also on existing theory. Repeating the processes goes to h) building a body of evidence around a problem, i) developing policies for control or management, and j) evaluating efficacy and effectiveness of implemented programmes. As stated in Chapter 4, the proposed basic aims of the energy epidemiological approach are to:

- Describe and measure the distributions of variable(s) of interest, e.g. energy demand per unit
- Explain the distribution by its determinant factors: physical, environmental, social, behavioural
- Support models that predict the changes expected in the distribution due to interventions, particularly energy efficiency and behavioural control measures
- Provide an evidence base for informing policy related to the management of energy demand.

Through the review of research literature study methods and the applied method studies, this thesis makes a case that the energy epidemiology approach supports the proposed population-level research paradigm through its conceptual and methodological framework. Figure 33 illustrates the interaction of the aims of the approach, which are underpinned by the identified key concepts, with their framework. For example, the description and measurement of distributions is supported by concepts such as population, frequency (i.e. prevalence and incidence), variation, and patterns, which can be examined using ecological, cross-sectional or cohort study designs. Understanding the determinant factors of the distribution and its variation is supported by concepts such as patterns, association, and risk factors and can be again examined under study designs such as ecological, cross-sectional and cohort. Understanding what interventions or controls can be applied to manage the occurrence of an outcome within the population is supported by concepts such as cause and effect and can be supported by study designs such as case-control, cohort or intervention and field trials. At various stages, the outcomes and evidence provided through these studies can be used to inform policy or to evaluate the past implementation of policy. It is through this conceptual and methodological framework that energy epidemiology provides an over-arching approach for the many disciplines involved in energy demand research.

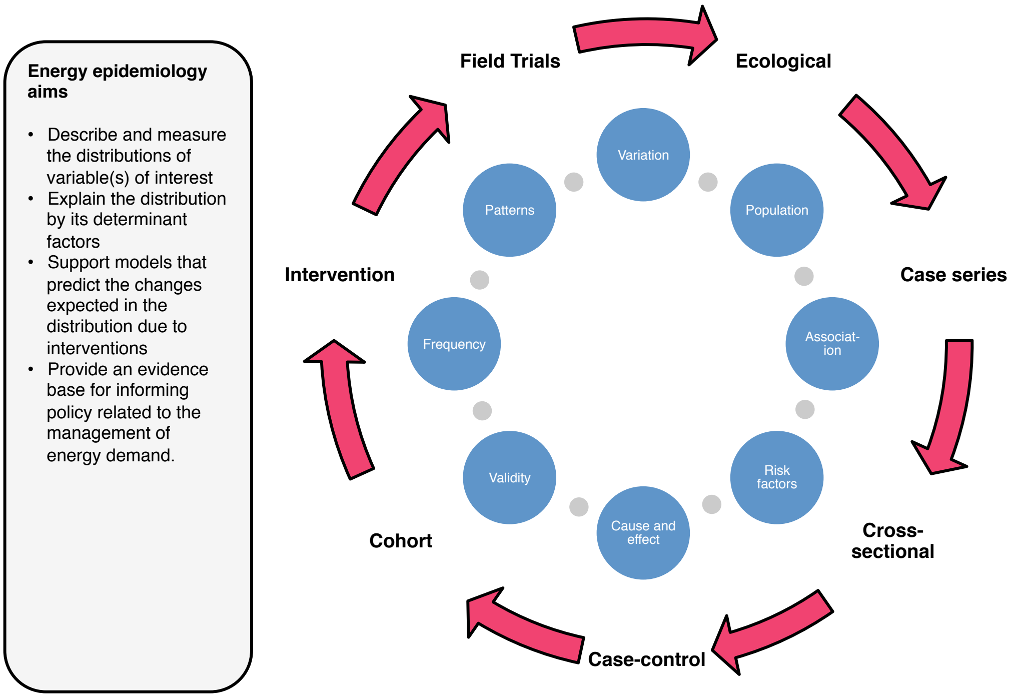


Figure 33 - Energy epidemiology conceptual and methodological framework

9.1.3 Interdisciplinary approach

Like health epidemiology, energy epidemiology is proposed to be inter-disciplinary in nature. By drawing on evidence from numerous disciplines and by providing a conceptual and methodological framework within which different disciplines can study complex socio-technical problems the approach aims to be interdisciplinary and therefore study relationships among populations from multiple viewpoints. An individual’s demand for energy is related to an array of specific characteristics and circumstances, such as the availability and cost of fuels, technologies and services, physical surroundings, environmental conditions, as well as social norms and behavioural preferences. In aggregate, these individuals may exhibit tendencies or similarities in characteristics and attributes that may only become apparent at the population level. Further, because populations are made up of heterogenous individuals, accounting for an array of characteristics and factors may help in clarifying or determining new relationships or refuting others. Summerfield and Lowe argued that the current approach to energy and building research has often been preoccupied by a singular discipline's set of methods of studying energy use in buildings from either a technical/building physics or social/behavioural approach. They make the point that achieving a closer degree of collaboration requires multi-disciplinary working environments and interdisciplinary approaches.

Epidemiology uses a principle of interdisciplinarity (to varying degrees of success) in order to gain deeper insights into population level outcomes. Interdisciplinarity is the interaction and collaboration of multiple disciplines working jointly on a problem, with the aim of integrating techniques and synthesizing theories (Choi and Pak, 2007; Cooper, 2002). Epidemiology encompasses a tradition of multidisciplinary collaboration. Epidemiological studies frequently involve various clinical or medical researchers whose interest primarily may lie in biological or genetic mechanisms for the disease under investigation, but their expertise can ensure that appropriate data collection and subsequent analysis can inform the research on potential mechanisms.

The interdisciplinary approach is not a new one for those working in the built environment. Cooper (2002) highlighted that disciplinary boundaries among practitioners are not so clear. However, he put forward that interdisciplinary working practices are time-intensive, yet working methods that transcend discipline boundaries are essential to identifying and resolving problems (Cooper, 2002). However, to date, with regard to *population level energy demand research* of the built environment, there has been little experience of integrating the engineering and physical sciences-oriented research with the insights provided by social sciences, nor is there an environment of empirically collected *in situ* data. Thus, contextualizing first principles models and laboratory testing to derive realistic assessments of real-world performance of engineered systems remains a challenge

Energy epidemiology could provide a 'neutral' territory to facilitate interaction and understanding. For instance, while many of the concepts of survey methodology are likely to be familiar to social sciences, building physicists or engineers may operate like clinicians by investigating underlying mechanisms. By themselves the approaches are insufficient to deal with the complex nature of energy demand among a population. Under an epidemiology approach there is the potential to be complimentary by allowing detailed findings to be used to determine prevalence within a population.

9.1.4 Addressing a diversity of problems

Epidemiology attempts to provide a conceptual and methodological framework relevant for the study of energy demand at the population level. The review and methods studies illustrated this process, i.e. identifying appropriate methods to design and carry out studies, drawing representative samples, determine strength of evidence, and evaluate past practices.

As illustrated through the method studies, the epidemiological conceptual and methodological framework may provide the tools needed to address the unique challenges of studying energy demand at the population level. Some key characteristics of the proposed approach are that:

- Energy epidemiology provides a framework for multidisciplinary research for diverse perspectives ranging from engineering and building physics, to sociology and economics;
- Research methods and protocols should be formally established, for instance in dealing with sampling, measurement error, and missing values.
- Using the results of comprehensive population level studies provides a means for contextualising laboratory tests and technical analysis and small scale and field trial studies.
- Findings are focussed on reducing risk of adverse outcomes over the long term.

9.2 Adapting epidemiological methods to energy demand: a critical discussion

This thesis examines the question of whether an epidemiological approach to studying energy demand among populations can help to address the challenges set out by the low-carbon paradigm. Its central premise is that the epidemiological approach is well placed to address the challenges of studying energy demand among populations by providing appropriate methods and concepts.

In examining whether epidemiology can be applied to population-level energy demand research, two further research questions were asked, which were: whether (RQ3) the key epidemiological concepts are appropriate and applicable to the population-level study of energy demand, and whether (RQ4) epidemiology study designs and methods can be applied to population-level energy demand research using currently available people, energy and buildings data, and strengths and weaknesses of available data.

Whilst the discussion provided in Chapter 4 and Chapter 5 outlined the case for the adaption of the conceptual and methodological frameworks and the subsequent method studies appraised the application. Here the implications of these concepts and methods are discussed more broadly in terms of the limitations, benefits and challenges that the approach offers to the energy demand research field.

9.2.1 Critique of the approach

There are many challenges being faced throughout the field of health epidemiology and it is not being suggested that an epidemiology approach is a panacea for the many issues related to population-level energy demand research. Applying the approach is likely to excite debate and will no doubt lead towards identifying concepts and methods that are not covered here that may be further explored. In the following sections, a number of aspects of the applied epidemiological approach are criticised.

9.2.1.1 To adapt, not replace

Epidemiology is unlikely to replace how building science or energy demand research is currently being done, and nor should it as this would suggest that current practices have no value. Rather, the examination of epidemiology here has sought to adapt (i.e. make suitable for a new use or purpose) those concepts and methods that are appropriate and relevant to the study of population-level energy demand. The emphasis of adaption is to take a feature or system and modify and adjust it in order to capture its benefits and improve the current situation.

In the thesis (Chapter 4), the epidemiological concepts were examined and assessed for whether they could be applied to energy demand. They were judged on the basis that i) the concepts were maintained, ii) the concepts were appropriate to energy demand, and iii) the concepts would advance the research field. For the most part, it was found that most of the identified concepts had already been applied to the study of energy demand and were shown to be broadly consistent with the epidemiological basis they embodied.

The population-level approach advocated by epidemiology is likely appropriate for many avenues of study, whether it is energy demand or some other population-level phenomena. There is likely to be some challenge to the use of the word epidemiology because health researchers most commonly use it. However, its literal translation and its broader objectives offer an opportunity for it to be used by others. As has been noted, even those within health epidemiology suggest that the approach may be of benefit to other areas of study (Porta et al., 2008). Further, epidemiology is best applied alongside other sciences and research techniques, drawing on the evidence of other studies to examine an issue.

9.2.1.2 Suitability

Epidemiology, at its core, is an empirically-based population-level approach and is therefore not necessarily suitable to non-population types of study. Studies of detailed mechanisms for a deeper understanding of the features of a particular system are needed to examine specific issues, events or mechanisms. For example, the third method study considered the impact of reported condensing boiler installations on the change in annual gas demand. However, whilst informative for examining the effect of thousands of installed boilers across many homes, the research provided very little information on detailed mechanisms that would affect the change in gas use. Turn-down rates (i.e. modulation of the thermal flow), for example, of gas condensing boilers and its implication for the overall seasonal coefficient of performance is an important issue to understand at the individual heating system level and would be the focus of in-depth study. Detailed studies are complimentary to the population approach examined here and provide deeper insight to issues seen amongst the population. For example, it may be that condensing boilers are systematically underperforming due to issues of under/over-sizing boilers and heating systems, which impacts on overall efficiency levels and therefore the amount of gas used.

An epidemiological approach could help tackle this problem by identifying the prevalence of boiler sizes among the population, with further investigation of the impact that turn-down rates have on gas demand.

9.2.1.3 Contextualisation

A key feature of the epidemiological approach is the concept of using population-level data to contextualize detailed studies. In energy and buildings research, many population-level studies use a case series design, which can means their findings are difficult to generalize to the broader population. Contextualizing case findings, in terms of the key physical and social features, provides clues for understanding how common (i.e. prevalent) the identified problem is among the target population (i.e. the population under study). The key to contextualization is the ability to draw data and information on the wider population. The first and second method studies (Chapter 7 and 8) both provided an examination of the broader conditions of the uptake of energy efficiency retrofits and energy demand among the British housing stock. While illustrating several concepts, these studies also provided a means for contextualising the uptake or energy demand in other studies that sampled from British houses. Doing this requires good quality population level data that can support these detailed studies.

For health research, there are a number of population level surveys and samples that are undertaken that can be used including Census, annual mortality and population estimates, and a number of registries provide these opportunities. In the UK, there is a shift in government thinking around data and its use by researchers that has meant that some sources of information on energy demand, people and buildings is becoming available, for example the National Energy Efficiency Database (NEED) and the 2011 Energy Follow-Up Survey (EFUS). These cross-sectional surveys are essential although they do not supplant the need for longitudinal data sets that allow one to look at trends over time.

9.2.1.4 Organisation – standards and review

In health epidemiology, there are a number of bodies and organizations that set out standards for undertaking epidemiological studies. Many focus on providing guidance on designing, conducting, interpreting, and presenting epidemiological research.

For example, the UK Medical Research Council (MRC) funds seven epidemiology-related institutes, units and centres that are mandated with:

> adopting broad multidisciplinary approaches to address major challenges in health-related research often requiring ground breaking methodology and technology development. (MRC, 2013)

These institutions are governed by the MRC's ethics and research guidance, including detailed plans on undertaking trials, data sharing, ethical standards and public participation, open access publishing, and to protocols for emergency situations. Another international health research organization is the Cochrane Collaboration, a network of researchers along with 17 global centres that adheres to a strict set of protocols when undertaking evaluations for the purpose of providing evidence-based healthcare (Cochrane Collaboration, 2013; Higgins and Green, 2008). Two essential features of these organizations is the clarity of the standards and regulations that have been established to govern and guide the research taking place while identifying the need for multiple disciplines to address the complex issues related to health. Whilst the nature of the standards is debated amongst the health community, they have had the effect prioritising consistency and transparency.

The methods studies presented in this thesis did not attempt to comply with health epidemiology study standards. Therefore, it would not be possible to say that these studies set a standard. However, the use of study guidance and standards that would be applicable to energy demand and buildings studies could be helpful for others undertaking research using an epidemiological approach. A further area of research would be to begin developing study guidance. The newly funded RCUK Centre for Energy Epidemiology (CEE) does offer an opportunity to begin to set down these methodological approaches for energy demand and gather together and advocate the approach.

9.2.2 Challenges

Using an epidemiological approach to study energy demand in buildings is faced with a number challenges. These include: accessing and collecting data of a necessary quality and required scale, and the difficulty of establishing methods among the wider research community. Other challenges are likely to exist, but in order to apply and epidemiological approach to the study of population-level energy demand in buildings, data and the take-up of methods are the most important.

The absence of or limited access to high-quality people and buildings data and high-resolution energy data of the statistical and methodological quality is a major challenge. In this thesis, the different datasets accessed (under very strict controls), such as HEED, gas and electricity energy supplier data, and VOA data vary in their quality, transparency and accessibility. HEED was derived from 'in-action' data, collected from a number of programmes and data providers without any common template or classification process, which meant the data was subject to limitations, such as missing data or misclassification. Access to the energy supplier data was tightly controlled and the process of annualisation largely complicated by a lack of detailed data to unpick the effect of the modelling. At present, neither dataset are widely available to researchers, although a version of NEED has been published by DECC subsequently. To address the challenge of data availability and quality, research data generated from current and future studies should be logged in a suitably accessible repository for future analysis and connection. As this thesis shows, access to detailed and comprehensive data collection offers an opportunity to undertake reviews of past programmes that can help feed into the policy process.

Alongside the problems of data, the review of current population-level energy demand studies (Chapter 4) suggests that coherent analytical methods have yet to be refined and applied widely within the field. Building an interdisciplinary culture capable of illuminating the complex co-evolution of practices and infrastructure that ultimately drive energy demand will be a challenge for the research community.

A number of researchers in the energy field have highlighted that barriers relating to different world views, language and methodological practices that make interdisciplinary working difficult can be alleviated through careful and consistent attention to research practices (Cooper, 2002; Shipworth, 2005; Wilhite et al., 2000a). One of the key challenges will be the integration of or at least collaboration between disciplines with very differing 'world views'. Shipworth (2005) highlights the problem of integrating exploratory social sciences with current empirical approaches to energy policy includes differences in language and conceptual frameworks (Shipworth, 2005).

Perceptions and practices of disciplines can be barriers to interdisciplinary research. However, the strong conceptual framework of epidemiology, along with its focus on definitions, structured methods, and empirical data, may provide a platform on which a wide range of disciplines related to the study of energy and buildings can interact.

9.3 Extending the energy epidemiology approach

The following section considers how the epidemiological approach might be extended beyond this thesis and how the approach might be taken up more widely by practicing researchers, supported by funders and used by policymakers.

9.3.1 A shift towards interdisciplinary studies

Shifting the practices around energy demand and seeking to reduce energy use and improve energy performance of buildings through the application of millions of energy efficiency retrofits across a heterogeneous building stock and population requires techniques and methods that can support population-level research. Through the method studies this thesis provides evidence that the epidemiological approach can be applied to the study of energy demand in buildings and generate interesting insights.

For the approach to have wider appeal and to have wider impact on policymaking, it will need to fit with and extend the current approach being used by researchers. Energy epidemiology is not intended to displace conventional small scale or qualitative research studies on energy and building use; rather it can complement and add to their value by contextualising findings among the wider population.

A potential advantage of setting out a methodological framework for energy research based on epidemiology is that it provides a foundation for consistent study development, enabling a legacy to build on past research. Through the use of consistent research design and study methods it provides a mechanism to formalise the peer-review process. In an international context, it can provide the ability to undertake comparison studies and to learn and replicate findings elsewhere.

9.3.2 Establishing levels of evidence

The epidemiological approach provides a way of looking at evidence with different levels of trust and accuracy, and moving from simple association towards causal associations. Using epidemiological techniques, a problem may be looked at from various different approaches. This may include single studies and case-series, associations through ecological or cross-sectional studies of association or case-control studies that look at the impact of an intervention on a given outcome.

Formalised results through summaries of evidence would allow for the evidence from energy demand studies results to be more effectively judged and assessed. It could also include using a hierarchical framework for quantitative and qualitative studies for energy demand, with the lowest order being single-case studies, through descriptive, conceptual and, the most important, generalizable studies. Through this and with the use of more formalised evidence reviews, greater project-by-project learning could occur.

Approaching an issue using several different study designs and techniques offers the greatest ability to clarify and determine the strength of evidence. How might this be achieved? The conceptual framework of epidemiology focuses on the influences, drivers and pathways that act on a given set of identified outcomes. For example, in the case of solid walls this could be to consider how a range of factors influences the effectiveness of insulation in reducing energy demand. Then, working from this conceptual framework of potential mechanisms, it would be necessary to begin to build a description of energy use and its drivers (e.g. fuels, service demands, timing, occupant patterns and preferences, etc.) in a range of building types in order to provide a baseline of energy use. Interviews with residents also could help to identify and clarifying factors that could affect their decision to accept or invest in insulating their home and the influence that social institutions may have. Beyond the specific study, extending detailed observational surveys of building and occupant characteristics such as the English Housing Survey, for example, would help to identify issues as they arise (as was the case for damp in 2009 – 2010). The broader population-level studies would provide a route through which more detailed mechanisms and factors associated with the intervention and energy demand outcomes (i.e. changes in gas demand, or 'savings') could be investigated at a detailed level. For example, using field trials that investigate a set of well-defined problems to establish how observed practices or systems (i.e. mechanisms) and interventions affect, for example, energy demand. The purpose of such trials would be to determine whether statistical evidence for particular mechanisms can be established, and whether interventions are sufficiently effective and well delivered to form the basis for practical policies.

9.3.3 Collaboration and communication

The tools and methods described in this thesis are not solely the domain of health epidemiologists. The types of studies mentioned above have been (and are still) commonly used in energy demand studies. For example, in the early 1980's there were 'case-control' studies of sorts in Milton Keynes, known as the Pennyland Project (Chapman et al., 1985), and Twin Rivers, New Jersey by Sonderegger (1978) and Socolow (1978). The latter used cohort-style analysis to look at the effect that differences in occupant practices had on energy demand in similarly constructed dwellings (Socolow, 1978; Sonderegger, 1978). This suggests that an epidemiology concept would not be foreign to many practicing researchers. However, although these methods have been employed to study energy demand, they have not necessarily had the occurrence, consistency or context typical of epidemiological studies. However, given the experience of many researchers with the concepts, it is suspected that many researchers may become comfortable with the concepts of the energy epidemiology framework.

9.3.4 Developing policy

The practical implication of adapting and using the epidemiological approach is that it is interdisciplinary and inclusive and can support feedback between researchers, practitioners and policymakers. Figure 34 offers a proposed practical structure through which the energy epidemiological approach is applied with interdisciplinary interaction and policy implications. In this practical approach, findings from large-scale studies both inform energy policy while providing a context for conventional small-scale studies. This type of structure exists already in some form, but with a limited focus on population-level research.

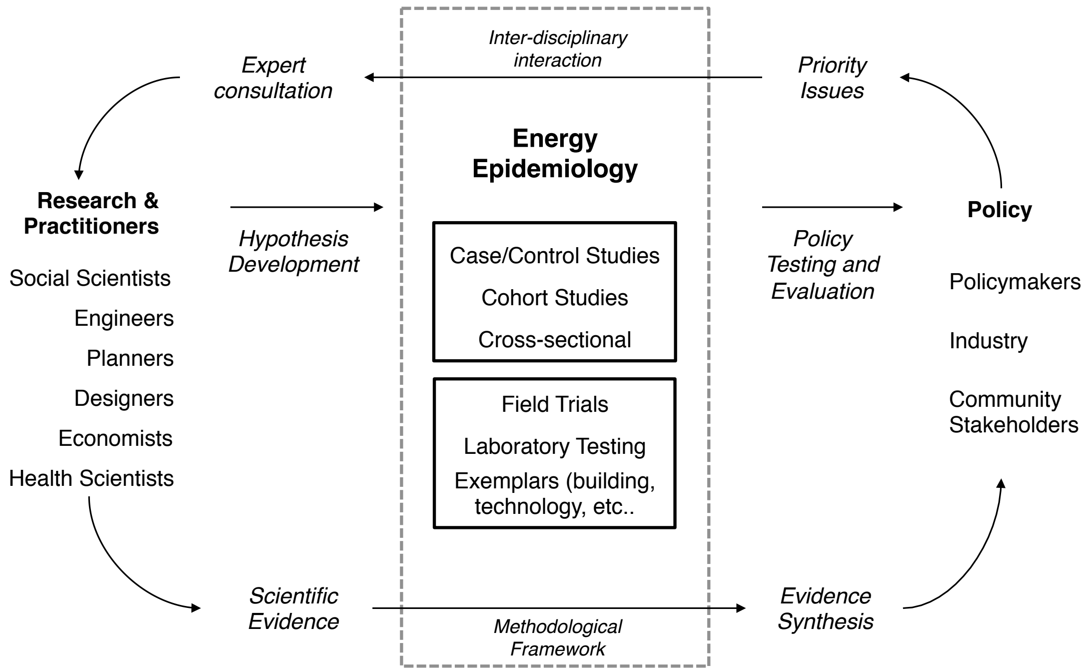


Figure 34 – Energy epidemiology in practice and interaction with policy development and evaluation

9.4 Summary

This thesis proposed that the research structure of health epidemiology can be reinterpreted and adapted in order to support an analysis framework from which to study and address end-use energy demand at the population level.

Whilst every method of study has limitations, there is considerable evidence from the successes and on-going challenges of public health over recent decades that an epidemiological approach can address complex issues. Adapting the epidemiological approach to population-level energy demand research could provide the means to describe the trends and patterns of energy demand and begin establishing causal factors that lead to outcomes of interest. It could also provide the means to undertake and contextualize intervention studies. The benefits of such an approach would be to strengthen the empirical foundation from which evidence is drawn to inform policy decisions and evaluate past intervention programmes or regulatory actions while also accounting for the complexity of the system within which the studies occur.

Chapter 10 Summary of Findings

"Findings from an energy epidemiological approach"

Chapter Introduction

The main thesis has focused on the methodological value of undertaking epidemiological studies applied to energy demand in buildings. In the process of demonstrating the methods applied to problems in the UK residential sector, several interesting insights have been generated and this chapter discusses aspects of these findings in further detail.

10.1 Whole-house retrofit packages

In Chapter 8, the findings of the research showed that the impact of energy efficiency retrofits, after adjustment for physical and household factors, demonstrated a dose-response effect whereby combined packages of retrofits was associated with increasing changes in energy demand (Table 33). Larger increases in reductions in gas demand were associated with boilers and cavity wall insulation, with only minor additional effects from lofts and glazing. The largest proportional change in gas demand was associated with the combined installation of a condensing boiler, and cavity and loft insulation at -11.2%. Although not always statistically significant, when combined the changes in energy demand were not additive, e.g. the individual change attributable to cavity insulation (-4.9%) and condensing boiler installation (-5.5%) and loft insulation (~1.0%) resulted in a change of -11.4%, which was instead greater (though in some other combinations less). The findings point to the potential impact that undertaking deep retrofits could have on energy demand. Combining retrofits into single package may have benefits in achieving energy demand reduction and potential cost-savings of installation (e.g. wall scaffolding is only set up once).

It may also be that there is a 'take-back' threshold after which rebound related to thermal comfort is lessened. That is, the potential for rebound has been met through the first (or second) retrofit and that subsequent interventions achieve a greater proportion of the potential energy savings. For example, a dwelling with unfilled cavity walls and a non-condensing boiler may not satisfy the demand temperature of the household. Filling the cavity walls with insulation would reduce the heat losses and allow the boiler to operate more effectively in meeting otherwise unmet demand. The subsequent installation of a condensing boiler with a similar power rating could achieve near expected energy use reductions through performance but also by operating less.

The impact and the additive nature of retrofits could have significant implications for meeting carbon reduction targets through energy efficiency retrofits in British households. A 'whole-house' retrofit package delivered to all homes in England is needed in order realise the potential energy savings set out in the DECC energy efficiency strategy. If an average energy savings of 10% (e.g. ~2300 kWh reduction) were achieved from the average UK dwelling, it would take approximately 9,565,000 'whole-house' retrofits to achieve the estimated 22 TWh of energy savings by 2020, which is equivalent to retrofitting 40% of UK dwellings. To achieve a 10% reduction in 2006 levels by 2020 (i.e. 54 TWh) through energy efficiency alone would take the equivalent of every home in the UK being refurbished (i.e. 23,500,000). Although further efficiencies may be gained from water heating and appliances, space-heating related energy comprises the bulk of residential demand and therefore should remain a high priority under government policy. Achieving these savings is an enormous task that will require a significant increase in historical rate of retrofit uptake. However, this research shows that these savings are achievable using widely available technologies and insulating techniques that rely on an existing deployment system and skill base.

10.2 Empirical compared to modelled energy demand and savings

For the buildings sector, the CCC has budgeted a 75% decline in carbon emissions from current levels by 2022, which includes ~40% reduction for residential emissions. Allowing for decarbonising of electricity generation and the target of constructing 'zero carbon' new buildings by 2019, this equates to ~25% reduction in energy demand from the existing stock over the next decade (UK CCC, 2012). Equivalent savings are specified across the entire building stock, including commercial and public sector buildings. For the residential stock, the CCC's 'Extended Ambition' scenario proposes to insulate 2 million solid wall houses, replace 13 million boilers and that 90% of lofts and cavities are insulated from 2008, with an expected savings of 17 $MtCO_2$ (UK CCC, 2010).

These estimates from the CCC (and DECC) are made using building-physics and engineering-based models. These models are both limited by a lack of sufficient information on the housing stock and are shown to be sensitive to a select number of inputs, including indoor thermostat set points and solar orientation (Stone et al., 2014). These models provide estimates of notional energy demand, but they have not been designed to model 'actual' energy use and therefore make for unreliable estimates when attempting to characterise real demand reductions.

Consider, for example the UK building regulations in the UK to improve the energy performance of dwellings. Since 1972 the Building Regulations in England and Wales have required energy efficiency measures in new buildings to reduce energy use, especially for space heating.(CLG, 2010b) Following that date, changes in building fabric performance should have resulted in reduced heat loss with age. However, analysis in Chapter 7 of annualised gas consumption data and dwelling age shows only an average of 2.0% change in metered energy per decade of construction date. How does this compare to 'actual' annualised energy use? Figure 35 below uses model results from a BREDEM-based stock model, the Cambridge Housing Model (CHM) version 3.0, to estimate energy consumption of the English housing stock. The stock model, described in detail by Hughes *et al.* (2013a), uses a modified version of SAP2009 to calculate energy demand (along with other outputs, such as CO_2 emissions). The figure shows modelled heating demand against mean annualized gas demand by age from Chapter 7. The energy values are centred on the mid period dwelling vintage in order to illustrate the proportional difference (rather than the absolute slope). When compared to the 'actual' annualized energy demand the rate of reduction is less than half the rate of decline expected due to building improvements from standard normative models. Older properties use less energy than expected and newer dwellings using more. The findings suggest that older dwellings lose significantly less heat and/or are heated to lower temperatures than previously assumed in these models.

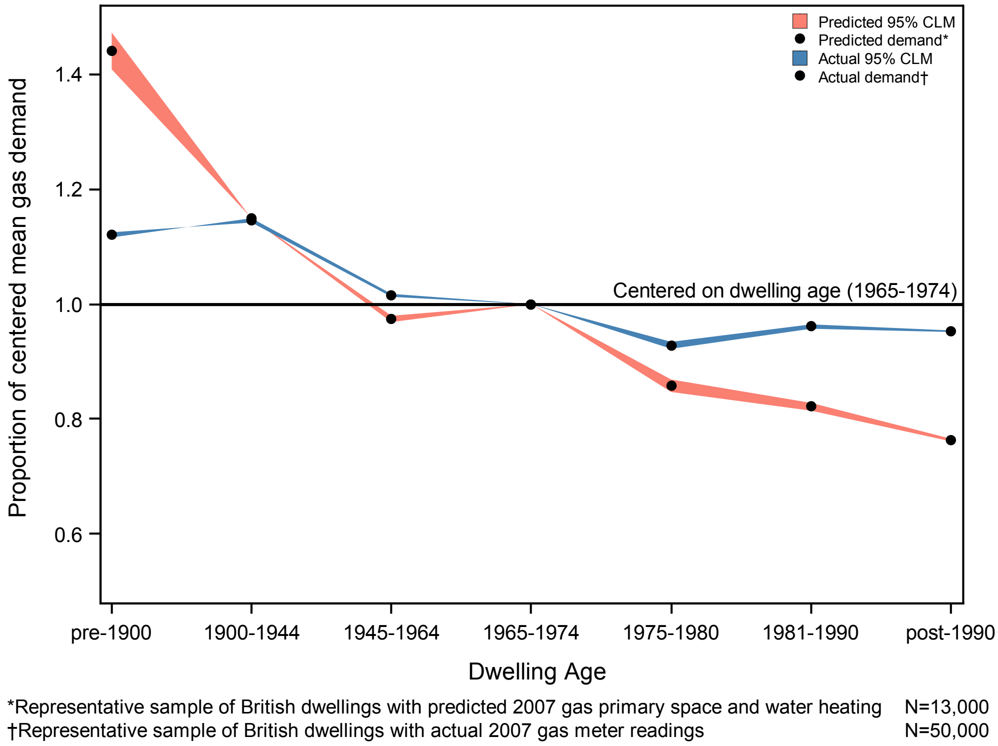


Figure 35 - Comparison of annual modelled heating demand against annualised gas demand by dwelling age bands

The implication of this clear discrepancy between modelled and 'actual' annualised energy demand poses a challenge for estimating energy savings. Using the CHM model, it is possible to compare the estimated savings derived from the analysis of HEED with gas demand values from Chapter 8 against modelled energy savings. Table 37 shows that the modelled savings are between two to three and half times greater than what the analysis above suggests. Note that the actual savings are different than those presented in Chapter 8 above as these values are not adjusted for dwelling characteristics. The modelled savings are derived from a Reduced SAP implementation applied to the English Housing stock, described in further detail elsewhere (Hughes et al., 2013b; Stone et al., 2014). In the model, the retrofits were applied to all eligible dwellings, i.e. those marked in the English Housing Survey as not having the measure. The savings were the difference in the modelling total heating demand (including hot water but excluding cooking). DECC have recently issued a report that shows actual savings which are slightly higher than estimated here, but the description of their methodology is limited, i.e. selection and population comparison, and the savings appear to be unadjusted for factors that may influence their impact within the selected.

Table 37 - Comparison of annual estimated change in gas use compared to modelled savings using CHM

Interventions	Average¥	Actual Change from 2005	Actual Savings (from trend)	Modelled Savings+
All 2005	17,570	-	-	-
No efficiency† 2007	16,240	-7.5%	-	-
Boiler only 2007*	14,500	-17.4%	-9.9%	-20.0%
Loft & Boiler only 2007*	14,490	-17.6%	-10.0%	-25.2%
Cavity & Boiler only 2007*	14,170	-19.4%	-11.8%	-41.1%

Notes:
†No efficiency measures from 2005 to 2007;
*Intervention in 2006;
¥median;
+ using EHS 2009 with no rebound factors

These findings are not just confined to the UK. A recent large scale analysis of dwellings in Germany found a similar pattern, whereby the difference between the predicted energy used for heating and actual energy consumed increased as the energy rating increased (in this case indicating a less energy efficient home) (Sunikka-Blank and Galvin, 2012). Such discrepancies in the predicted and measured performance of older 'energy inefficient' dwellings compared with newer 'energy efficient' dwellings have serious implications in terms of their potential impact on energy efficient refurbishment policies.

10.3 Implications for energy efficiency policy

There is a strong historic track record in the UK of policy helping to improve the energy efficiency of dwellings occupied by vulnerable and low-income households. However, future policies will need to address the gap in the uptake of retrofits amongst older, owner occupied dwellings in areas of higher incomes. This middle-income household group also use more energy on average, after controlling for home size, age and type, and thus the potential impact on absolute energy savings is greater. However, the shift in government policy towards encouraging middle-income households to self-invest in their dwelling's energy efficiency has faced an uphill struggle. The Green Deal prioritised self-investment by providing access to upfront capital and a payback process that assured the cost of the retrofit would be equal to the notional savings, known as the 'golden rule' (DECC, 2012f). The policy, however, has been slow to be taken up. Initial government estimates sought to have

20,000 dwellings sign up to Green Deal by the end of 2014, the most recent statistics show that over an 18 month period only 1,587 dwellings installed a retrofit through Green Deal (DECC, 2014c). The Green Deal cashback scheme has proven to be more popular, with 14,452 retrofits installed over the same period.

In comparison, the energy company obligation (ECO) is a continuation of government-driven supplier obligations focused on reducing energy consumption and supporting vulnerable costumers. Its target for 2013-2015 was an expenditure of £1.3 billion and the latest statistics show that £0.4 billion has been invested through ECO by energy suppliers or 31% of the total commitment (DECC, 2014c). The research findings provided in this thesis suggest that the Green Deal is broadly targeting the right household groups and dwelling types (i.e. those with historically lower uptake rates). However, the fact that retrofit take up remains low suggests that the Green Deal is not addressing actual barriers to uptake or exploiting the motivations of the targeted household groups. A government survey in 2011 (prior to the launch of the Green Deal) found that households felt that the proposed interest rates were either too high (compared to existing available finance interest rates) or that the payback periods were too long (DECC, 2011b). Regardless, the current annual percentage rate (APR) for Green Deal amounts ranges from 10.3% for £1,5000 over 15 years to 7.9% for £5,000 over 25 years. These rates have been criticised as being a cause for low uptake, particularly when many middle income households would be eligible for market interest rates of nearer to 2.3% (Hough and White, 2014). Further, a recent survey looking at motivating factors found that most households considering energy efficient retrofits were doing so for reasons that related to amenity renovations and not for energy savings alone; that energy efficiency retrofitters were a select group more likely to be older, owner-occupier homes with few dependants; and that emergency repairs was a strong trigger for retrofitting existing systems (Wilson et al., 2013). Given the size and potential of this group, barriers around the cost of financing and trigger points could be addressed to potentially improve uptake.

At present, UK energy policy includes a measure of variation in energy savings for vulnerable households, known as a ‘comfort’ factor. The comfort factor is measured at 15% and is applied to predicted energy savings as an amount of potential energy savings that is taken back in the form of thermal comfort improvements (DECC, 2012a). In addition to the comfort factor, several UK energy efficiency policies also include an ‘in use’ factor that is applied to notional energy savings mean to account for ‘comfort’ and ‘in use’ factors (DECC, 2012a). The ‘in use’ factor is meant to account for physical variations between the actual installed system and the model (e.g. limited measured information or differences in exact system specification). In the Green Deal, the factors are used to adjust the notional savings to reflect actual savings and therefore what retrofits would meet the ‘golden rule’. This introduces a number of complications: first, using modelled estimates of energy demand will inevitably fail to be representative of actual demand and therefore savings (Sunikka-Blank and Galvin, 2012); second, the ‘in use’ factors should not penalize potential savers or undermine the potential payback of the retrofit. Instead, it would be preferable to directly use empirical data to estimate energy savings that reflects a number of socio-technical factors, such as those in this study. In doing so, it would be possible to provide more accurate estimates of energy savings for individual households (with appropriate uncertainty bands) and more widely for the housing stock.

10.4 The state of energy and buildings data

The Homes Energy Efficiency Database is an example of what can be characterised as 'in action' data. HEED is not the product of a large omnibus survey or a concerted monitoring and reporting exercise; instead HEED is the product (and by-product) of a range of disparate activities that are centred on home energy efficiency. HEED offers a repository and framework for these sources, one that is clearly flexible to a range of data types and quality.

It is unlikely that HEED will offer the same insight as a well structured research design on the impact of energy efficiency or an omnibus survey in terms of representativeness, but what is clear is that is has an extraordinary usefulness as a framework within which to collect and link data sources together. Due to the nature and range of its coverage (i.e. containing information on over 50% of UK dwellings) it can reasonably be used as a source to describe the broad characteristics of the UK housing stock. When linked to energy, HEED is capable of offering insight into the differences in demand due to dwellings characteristics and levels of energy efficiency and the change in demand associated with an energy efficiency measure.

The use of HEED linked to Energy Supplier data in this thesis also highlights another aspect of the state of energy and buildings data for the UK population, and that is the lack of other appropriate data sources that are available to support this type of study. The EHS is an obvious candidate for supporting energy and buildings analysis in the residential sector. However, there is not yet a mechanism for linking the energy data with the EHS variables. Although the EHS was not constructed to be representative of energy use in English houses and households, it would be plausible that an appropriate weighting could be constructed to allow for this. The EHS, however, would not support longitudinal study of energy demand in the residential sector due to its cross-sectional nature. A possible candidate for this is the British Household Panel Survey (BHPS), which is now known as 'Understanding Society'. In this dataset, British households provide reported expenditure on their energy and energy efficiency retrofits but there is no actual energy demand or information on the energy tariffs paid by the residents.

The non-domestic building sector, not addressed in this thesis, has even less structured data available on energy demand and buildings. Steadman et al (2009) points out that the non-domestic sector is characterised by a heterogeneous building stock and that the concept of a 'building' itself is problematic in this sector due to limitations in their classification by mapping agents, government administrative agents (e.g. VOA) and commercial sensitivities (Kohler et al., 2009; Steadman et al., 2009, 2000).

The development and analysis of data frameworks that link energy and dwelling characteristics together open up extraordinary and unprecedented opportunities for improving the precision and reliability with which low carbon strategies are framed, guided and evaluated. These databases offer the possibility of a step change in our understanding of current patterns of consumption at the level of individual dwellings, premises and buildings; and in the years to come, will make it possible to follow changes and trends in different sectors, building types, end uses and fuels. Further, these large data-framework approaches, if carried out consistently and transparently, will allow researchers and industry to continue to carry out field trials and detailed studies with the knowledge that they will be able to contextualise their work within the population.

Chapter 11 Conclusions

“An epidemiological approach to end-use energy demand”

A case for energy epidemiology

This thesis has outlined a case for an epidemiological approach to studying population-level energy demand. Energy epidemiology aims to investigate the causes and effects of key factors on energy demand within a population or subpopulations at various scales (e.g. from individuals and buildings to communities or building complexes). It considers the complex interactions between the physical and built environment, socio-economic features, and individual interactions and practices through an inter-related conceptual framework. It also provides a methodological framework within which to identify and describe the interacting factors acting on the complex energy demand system.

The central paradigm of an energy epidemiological approach supports the shift to a low-carbon society along with the alleviation of energy-related social and environmental phenomena, such as fuel poverty and comfort. This can be achieved through population-based methods that analyse patterns and systems of end-use energy demand for services in order to better understand the practices, drivers, causes and differences of energy demand outcomes.

Adapting the epidemiological approach is not a panacea to dealing with the challenges facing the field of research in energy demand in buildings. However, it can provide a set of concepts, methods and analysis tools that are capable of supporting an empirically-based population-level research approach, identified as a necessary step towards to developing a deeper foundation of evidence.

The epidemiological approach should only be applied when appropriate to the research question. This means developing research questions that can be examined through appropriate study designs and, where possible, draw on detailed studies that examine suspected underlying mechanisms. Contextualising detailed studies within the wider population requires suitable datasets against which to compare and identify differences of a sufficient quality and comprehensive nature. While some of the existing cross-sectional datasets provide some details on energy demand, there is a general lack of such information over a longitudinal horizon. In time, with further funding, datasets such as the National Energy Efficiency Data-framework and subsequent cross-sectional Energy Follow-up

Surveys may fill this gap. However, there remains a need for a high-quality longitudinal panel survey, with detailed monitoring of energy and environmental conditions. The Dept. of Energy and Climate Change have recently begun to look at a longitudinal study of UK energy, known as the Longitudinal UK Energy Survey (LUKES) (Cooper et al., 2014).

The development of guidance, protocols and a process by which to develop definitions is essential to the epidemiological approach. It is not necessary that these definitions are wholly agreed on, but the process should be transparent and openly debated with input from the many researchers actively taking part in energy demand research.

The method studies illustrated several population level epidemiological study designs and concepts and how evidence from studies using empirical data from large populations can help to provide insight into the energy performance of buildings. Not without their limitations, the selected method studies offer some evidence that the epidemiological approach could support analysis of population-level energy demand research questions. The findings from the ecological study provided evidence on the uptake of energy efficiency retrofits over the study period (i.e. 2000 to 2007) and illustrated the concepts of incidence and prevalence. The findings from the descriptive cross-sectional study of energy demand and dwellings characteristics illustrated the variation seen within the British housing stock. The findings from the cohort study identified that energy efficiency interventions are associated with changes in gas demand over time and that energy efficiency retrofits have are additive when combined.

Extending the approach beyond this thesis has a number of potential benefits to the energy demand in buildings research field. Energy epidemiology represents an innovative multidisciplinary approach to change the culture and practice of energy demand research in a way that is commensurate with the scale and scope of the policy agenda to reduce energy demand from buildings. The approach complements and adds value to small-scale and exploratory studies by allowing their findings to be placed into a wider context and therefore to increase their use within the field. By relying on a well-defined conceptual and methodological framework and using strong study designs, energy epidemiology provides a way of looking at evidence with different levels of trust and accuracy, moving from simple association towards causal associations. Energy epidemiology can help drive an evidence-based policy cycle by identifying and describing problems, undertaking studies around their differences and causes, develop policy interventions and evaluate past practices.

Evaluation is an important aspect of policy development because it lets policymakers understand and adapt policies to increase the likelihood of their success, to identify policies or elements of policies that achieved expected outcomes, and to provide evidence of how policies worked in practice and their impact as part of a drive towards accountability or related requirements. Innovative evaluation methods are needed to address the advancing nature of government policies, the complexity of delivery chain, the number of interacting actors and the competing policy priorities and objectives.

Developing policies in order for them to be evaluated means putting in place a prospective conceptual framework that describes the events leading to a desired long-term outcome (e.g. ‘theory of change’ or ‘realistic evaluation’) (Blamey and Mackenzie, 2007). Setting out the drivers, actions and events that seek to bring about the outcome the policy intends for provides the basis for identifying testable evaluation questions. With a more explicit statement of the questions and understanding of the expected processes that will delivery the policy objectives an evaluation framework can be put in place. Such a framework would comprise identifying the evaluation needs, determining aims and objectives, study designs, data collection, methods and expected outputs. At present, however, evaluations of energy and buildings policies face several challenges, including: a scarcity of strategic energy policy management data, and an absence of evaluation theory and practice that can effectively deal with the socio-technical nature of energy policy. The epidemiological approach is capable of responding to challenges of evaluating policies that seek to control and change end-use energy demand in face of a host of complex issues. Embedded in the epidemiological approach is a focus on evaluating policies and their programmes, along with unintended impacts of ‘natural’ experiments.

Meeting the policy evaluation needs requires an evidence base that can addresses the complex nature of the socio-technical interactions between people, energy and the built environment. For policymakers, this means putting in place mechanisms that can provide for the collection of data throughout the policy development process. Evidence must be based on a strong theoretical foundation, be representative of the population being assessed, account for differences, and be undertaken at scale commensurate with the problem at hand. To achieve this, the evidence should come from consistent studies that are properly designed, conducted, interpreted and presented with the necessary details to make findings transferable and able to withstand scrutiny. This effectively means policymakers would need to integrate the research design and data collection into the policy development process for the purposes of evaluation. For example, the UK government has put in place a number of Evaluation Frameworks that seeks to identify evaluation priorities and objectives to assist in guiding government towards the right type/level of evaluation that would best meet ‘their’ objectives and also the wider known and possible or prospective objectives. Doing so will provide policymakers with conclusive evidence for how policies have delivered intended outcomes, the reasons behind those that have not, and importantly a view towards understanding the unintended consequences of their actions.

Related to policy evaluation is the use of systematic approaches (e.g. reviews, or meta-analyses) to evaluate and judge the strengths and weakness of the existing available evidence. Systematic review in health sciences provides a method of assessing the validity of studies pertaining to an issue. To support systematic reviews and meta-analyses requires that quantitative results be presented with sufficient statistical and sample information to judge the strength and precision of the findings and assess study soundness. The benefits of well-defined systematic reviews are to inform policymakers on the strength of evidence and to identify evidence gaps.

Although at present end-use energy demand research does not offer the same degree of formalised result presentation or study design, applying more rigorous study design and reporting standards would assist. The review and evidence assessment approach would assist policymakers and researchers understand the strength of evidence on an energy issues. For example, the effectiveness of energy efficiency intervention programmes in achieving persistent change in energy demand (or energy 'savings'). The concept of the systematic review has been extended to energy demand and is described as being able to improve the evidence based for developing energy policy by improving the definition of key issues, increasing the number of studies considered, identifying gaps and minimising uncorroborated results (Sorrell, 2006). Embedding evidence reviews using systematics processes as the first step to undertaking energy research will help to shape the evidence landscape. This understanding of the evidence can be used in developing comprehensive studies that limit bias and clearly address the given issue. It is unlikely that the current evidence base would support an approach to systematic reviews undertaken by groups such as the Cochrane Collaboration who strongly emphasis the use of randomised control trials (RCT). However, using an epidemiological approach will ultimately enable the wider use of systematic reviews in examining evidence from energy research.

To support epidemiological study approaches and systematic reviews, large-scale studies and putting in place mechanisms for routine data collection that can be accessed for research and evaluation are required. Large-scale studies representative of the population form the backbone of epidemiological research approaches. With few exceptions, empirical collection of data on building characteristics and energy demand has historically been ad-hoc or inconsistent, with little tradition of reporting data and findings in a formal sense. There is significant potential of linking together existing routine datasets such as those collected for tax and energy metering purposes (as has been done in DECC's NEED). In doing so, however, issues and perceptions around privacy and commercial sensitivity need to be addressed by offering a secure process with appropriate assurances to data providers (and their subjects) on privacy. In the health sciences, the need to link data (argued as a public good) to address urgent or critical health problems or events has overcome the concerns of individuals' privacy by putting in place the National Information Governance Board (NIGB) for Health and Social Care, whose task it is to oversee how an individual's data are used, stored and shared. They have a legal obligation to review research requests that would make use of individual's data outside its original collection purpose and judge whether the research is significant enough to allow access to anonymised data to proceed (UK Parliament, 2006). While the field of energy demand and information on the built environment does not have the same legal or governance background as health data, there is a slowly growing momentum to collect and make accessible such data. Under the *Energy Act of 2011*, for example, the Secretary of State for DECC has the ability to collect information on energy efficiency measures installed in UK houses. This, along with the licensing and management requirements of storing and accessing high-frequency energy data output from UK 'smart' meters will further add to the necessary legal framework (DECC, 2012g). Also, the UK government has committed to making data available under its open data strategy

(HM Government, 2012). For energy and the built environment, departments such as DECC are making a commitment to 'big' data availability, under appropriate privacy controls, for use by research ((DECC, 2012h) p. 9). Putting in place a national energy research advisory board who could act as a review panel for determining the value of proposed studies to research and government and who could act in a similar manner as the NIGB for granting access to data would help energy research extend its current study design limitations.

Energy epidemiology places an emphasis on the crucial role of research translation for policymakers and stakeholders. In general, findings are of little value to policymakers unless they are expressed or formulated in a way that is relevant to policy development or its implementation. Their implications need to be expressed in terms of the benefits of optimal type, timing, and targeting of energy related interventions. In the energy demand field, policymakers are faced with challenges from the political process, stakeholder interests and the wider society. It may be that policymakers respond to numerous forms of evidence that can provide both an understanding of influencing factors on the energy demand outcome, along with a wider body of evidence that can offer co-benefits, such as health or economic growth. This serves to further underline the role of multidisciplinary approaches, from engineering, social sciences and economics to planning, health and finance that is likely needed for policy evaluation.

11.1 Next Steps

In terms of next steps, there is considerable work to be done around extending the epidemiological approach to energy demand. This includes continuing to work on guidance for researchers, developing definitions and metrics for measurement, and undertaking reviews of current research evidence. It also involves extending the study methods and designs to other energy demand problems being faced.

Finally, policymakers and research funders will need to be convinced that an empirically-based population-level approach to studying energy demand is able to address the challenge of shifting to a low-carbon society. The epidemiological approach will prove its value when applied to pressing issues, such as fuel affordability, energy savings, shifting occupant practices, evaluating programmes. This application will serve to develop trust around the approach and work towards establishing a robust evidence base.

Chapter 12 Referenced works

Alberini, A., Banfi, S., Ramseier, C., 2013. Energy Efficiency Investments in the Home: Swiss Homeowners and Expectations about Future Energy Prices. Energy J. 34. doi:10.5547/01956574.34.1.3

Andersson, K., 1998. Epidemiological Approach to Indoor Air Problems*. Indoor Air 8, 32–39.

Aroonruengsawat, A., Auffhammer, M., Sanstad, A.H., 2012. The Impact of State Level Building Codes on Residential Electricity Consumption. Energy J. 33, 31–52. doi:10.5547/ISSN0195-6574-EJ-Vol33-No1-2

Attari, S.Z., DeKay, M.L., Davidson, C.I., Bruine de Bruin, W., 2010. Public perceptions of energy consumption and savings. Proc. Natl. Acad. Sci. U. S. A. 107, 16054–9. doi:10.1073/pnas.1001509107

Bailey, L., Vardulaki, K., Langham, J., Chandramohan, D., 2005. Introduction to epidemiology, Understanding Public Health. McGraw-Hill International.

Baker, K.J., Rylatt, R.M., 2008. Improving the prediction of UK domestic energy-demand using annual consumption-data. Appl. Energy 85, 475–482. doi:10.1016/j.apenergy.2007.09.004

Bell, M., Lowe, R.J., 2000. Energy efficient modernisation of housing: a UK case study. Energy Build. 32, 267–280. doi:10.1016/S0378-7788(00)00053-0

Bhopal, R., 2008. Concepts of Epidemiology: An integrated introduction to the ideas, theories, principles and methods of epidemiology, 2nd ed. Oxford University Press, USA.

Biddulph, P., Gori, V., Elwell, C.A., Scott, C., Rye, C., Lowe, R., Oreszczyn, T., 2014. Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements. Energy Build. 78, 10–16. doi:10.1016/j.enbuild.2014.04.004

Bielecki, J., 2002. Energy security: is the wolf at the door? Q. Rev. Econ. Financ. 42, 235–250.

Biggart, N.W., Lutzenhiser, L., 2007. Economic Sociology and the Social Problem of Energy Inefficiency. Am. Behav. Sci. 50, 1070 –1087. doi:10.1177/0002764207299355

Bircher, J., 2005. Towards a dynamic definition of health and disease. Med. Health Care. Philos. 8, 335–41. doi:10.1007/s11019-005-0538-y

BIS, 2010. Low Carbon Construction. Department for Business Innovation and Skills, London, UK.

Blamey, A., Mackenzie, M., 2007. Theories of change and realistic evaluation: peas in a pod or apples and oranges? Evaluation 13, 439–455. doi:10.1177/1356389007082129

Boardman, B., 1991. Fuel Poverty: From Cold Homes to Affordable Warmth. Belhaven Press, London, UK.

Boardman, B., 2004. New directions for household energy efficiency: evidence from the UK. Energy Policy 32, 1921–1933. doi:10.1016/j.enpol.2004.03.021

Boden, M., 1996. Paradigm Shift and Building Services. Serv. Ind. J. 16, 491–510. doi:10.1080/02642069600000043

Booth, A.T., Choudhary, R., 2013. Decision making under uncertainty in the retrofit analysis of the UK housing stock: Implications for the Green Deal. Energy Build. 64, 292–308. doi:10.1016/j.enbuild.2013.05.014

BRE, DECC, 2009. The Government's Standard Assessment Procedure (SAP) for Energy Rating of Dwellings. Department of Energy and Climate Change, Garston, Watford, UK.

Brechling, V., Smith, S., 1994. Household Energy Efficiency in the UK. Fisc. Stud. 15, 44–56. doi:10.1111/j.1475-5890.1994.tb00196.x

Brownson, R.C., Chriqui, J.F., Stamatakis, K. a, 2009. Understanding evidence-based public health policy. Am. J. Public Health 99, 1576–83. doi:10.2105/AJPH.2008.156224

Caballero, B., 2007. The Global Epidemic of Obesity: An Overview. Epidemiol. Rev. 29, 1–5. doi:10.1093/epirev/mxm012

California Air Resources Board, 2008. Climate Change Scoping Plan: a framework for change.

Cameron, D., Jones, I.G., 1983. John Snow, the Broad Street Pump and Modern Epidemiology. Int. J. Epidemiol. 12, 393–396. doi:10.1093/ije/12.4.393

Canoy, D., Buchan, I., 2007. Challenges in obesity epidemiology. Obes. Rev. 8, 1–11. doi:10.1111/j.1467-789X.2007.00310.x

Cayla, J., Maizi, N., Marchand, C., 2011. The role of income in energy consumption behaviour : Evidence from French households data. Energy Policy 39, 7874–7883. doi:10.1016/j.enpol.2011.09.036

Cayre, E., Allibe, B., Laurent, M.H., Osso, D., 2011. There are people in the house! How the results of purely technical analysis of residential energy consumption are misleading for energy policies'. Proc. ECEEE 2011 Summer Study, Pap. 7, 277.

Chapman, J., Lowe, R.J., Everett, R., 1985. The Pennyland Project. Open University, Milton Keynes, UK.

Chiu, L.F., Lowe, R., Raslan, R., Altamirano-Medina, H., Wingfield, J., 2014. A socio-technical approach to post-occupancy evaluation: interactive adaptability in domestic retrofit. Build. Res. Inf. 42, 574–590. doi:10.1080/09613218.2014.912539

Choi, B.C.K., Pak, A.W.P., 2007. Multidisciplinarity, interdisciplinarity, transdisciplinarity in health research, services, education and policy: 2. Promotors , barriers , and strategies of enhancement. Clin. Investig. Med. 30, 224–233.

Clery, D., 2007. A sustainable future, if we pay up front. Science (80-.). 315, 782–3. doi:10.1126/science.315.5813.782

CLG, 2005. Dwelling Stock by Council Tax Band - LSOA level statistics. Office of National Statistics, London, UK.

CLG, 2010a. English Housing Survey: Household report 2008–09. Communities and Local Government Publications, London, UK.

CLG, 2010b. Approved Document L: Conservation of fuel and power (2010 Ed.). The Stationery Office, London, UK.

CLG, 2010c. Housing Stock, England - Dwelling Stock estimates, England, 2011. Department for Communities and Local Government, London, UK.

CLG, 2013a. English Housing Survey Technical Advice Note - Survey Overview and Methodology: 2011-12 Update. Department for Communities and Local Government, London, UK.

CLG, 2013b. English Housing Survey, 2010: Housing Stock Data. HMSO, London, UK. doi:10.5255/UKDA-SN-7039-4

Cochrane Collaboration, 2013. The Cochrane Collaboration: Cochrane Reviews [WWW Document]. Cochrane Rev. URL http://www.cochrane.org/cochrane-reviews

Cooper, A., Shipworth, D., Elam, S., Hamilton, I., Huebner, G.M., Liddiard, R., Morrell, G., Oskala, A., Park, A., Shipworth, M., Steadman, P.J., 2014. UK Energy Lab : A feasibility study for a longitudinal , nationally representative sociotechnical survey of energy use Final Report October 2014. London, UK.

Cooper, I., 2002. Transgressing discipline boundaries : is BEQUEST an example of " the new production of knowledge "? Build. Res. Inf. 30, 116–129. doi:10.1080/09613210110097406

CPUC, 2008. California long term energy efficiency strategic plan: achieving maximum energy savings in California for 2009 and beyond. California Public Utilities Commission, Sacramento, CA.

Daly, J., Willis, K., Small, R., Green, J., Welch, N., Kealy, M., Hughes, E., 2007. A hierarchy of evidence for assessing qualitative health research. J. Clin. Epidemiol. 60, 43–49. doi:10.1016/j.jclinepi.2006.03.014

Davies, M., Oreszczyn, T., 2012. The unintended consequences of decarbonising the built environment: A UK case study. Energy Build. 46, 80–85. doi:10.1016/j.enbuild.2011.10.043

DECC, 2008a. Energy consumption in the United Kingdom. Department of Energy and Climate Change, London, UK.

DECC, 2008b. Regional and local electricity consumption statistics for 2007, Energy Trends, Special feature – Regional and local electricity consumption.

DECC, 2008c. Regional and local gas consumption statistics for 2007.

DECC, 2009. Guidance note for the DECC MLSOA/IGZ and LLSOA electricity and gas consumption data. Department of Energy and Climate Change, London, UK.

DECC, 2010a. Digest of United Kingdom Energy Statistics 2010, 2010th ed. The Stationery Office, London, UK.

DECC, 2010b. Energy Consumption in the United Kingdom - Domestic Data Tables. Department of Energy and Climate Change, London.

DECC, 2010c. Annual report on fuel poverty statistics 2010. London, UK.

DECC, 2011a. National Energy Efficiency Data-Framework: Report on the development of the data-framework and initial analysis. Department of Energy and Climate Change, London, UK.

DECC, 2011b. Consumer Needs and Wants for the Green Deal. Department of Energy and Climate Change, London, UK.

DECC, 2012a. Final Stage Impact Assessment for the Green Deal and Energy Company Obligation. Department of Energy and Climate Change, London, UK.

DECC, 2012b. The Energy Efficiency Strategy: The Energy Efficiency Opportunity in the UK. Department of Energy and Climate Change, London, UK.

DECC, 2012c. Energy consumption in the United Kingdom 2012: Domestic energy consumption. Department of Energy and Climate Change, London, UK.

DECC, 2012d. National Energy Efficiency Database: Summary of analysis. Department for Environment and Climate Change.

DECC, 2012e. Great Britain's housing energy factfile Report 2011. Department of Energy and Climate Change, London, UK.

DECC, 2012f. The Green Deal: Which energy efficiency improvements qualify for Green Deal Finance. Department of Energy and Climate Change, London, UK.

DECC, 2012g. Smart Metering Implementation Programme Smart Energy Code: Consultation. Department of Energy and Climate Change, London, UK.

DECC, 2012h. DECC's Open Data Strategy. Department of Energy and Climate Change, London, UK.

DECC, 2013a. Fuel Poverty: a Framework for Future Action. The Stationery Office Ltd., London, UK.

DECC, 2013b. Trends in Fuel Poverty - England 2003 to 2010. Department of Energy and Climate Change, London, UK.

DECC, 2013c. Total final energy consumption at regional and local authority level: 2005 to 2010. Department of Energy and Climate Change, London, UK.

DECC, 2013d. National Energy Efficiency Data-Framework: domestic energy consumption tables. Department of Energy and Climate Change, London, UK.

DECC, 2014a. Developing DECC's Evidence Base. Department of Energy and Climate Change, London, UK.

DECC, 2014b. UK National Energy Efficiency Action Plan. Department of Energy and Climate Change, London, UK.

DECC, 2014c. Domestic Green Deal and Energy Company Obligation in Great Britain, Monthly report. Department of Energy and Climate Change.

Dietz, T., 2010. Narrowing the US energy efficiency gap. Proc. Natl. Acad. Sci. 107, 16007 – 16008. doi:10.1073/pnas.1010651107

Dietz, T., Gardner, G.T., Gilligan, J., Stern, P.C., Vandenbergh, M.P., 2009. Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. Proc. Natl. Acad. Sci. U. S. A. 106, 18452–6. doi:10.1073/pnas.0908738106

Diez Roux, A. V, 2007. Integrating social and biologic factors in health research: a systems view. Ann. Epidemiol. 17, 569–74. doi:10.1016/j.annepidem.2007.03.001

Dixon, R.K., McGowan, E., Onysko, G., Scheer, R.M., 2010. US energy conservation and efficiency policies: Challenges and opportunities. Energy Policy 38, 6398–6408. doi:10.1016/j.enpol.2010.01.038

Doll, R., Hill, A., 1950. Smoking and carcinoma of the lung. BMJ 2, 739–748.

Doll, R., Peto, R., Boreham, J., Sutherland, I., 2004. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 328, 1519. doi:10.1136/bmj.38142.554479.AE

Dooley, J., 1998. Unintended consequences: energy R&D in a deregulated energy market. Energy Policy 26, 547–555.

Dos Santos Silva, I., 1999. Cancer Epidemiology: Principles and Methods, 1st ed. International Agency for Research on Cancer (IARC) Press, Lyon, France.

Dowson, M., Poole, A., Harrison, D., Susman, G., 2012. Domestic UK retrofit challenge: Barriers, incentives and current performance leading into the Green Deal. Energy Policy 50, 294–305. doi:10.1016/j.enpol.2012.07.019

Druckman, A., Jackson, T., 2008. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. Energy Policy 36, 3177–3192. doi:10.1016/j.enpol.2008.03.021

DTI, DEFRA, 2001. UK Fuel Poverty Strategy. DTI, London, UK.

Du Plessis, C., Cole, R.J., 2011. Motivating change: shifting the paradigm. Build. Res. Inf. 39, 436–449. doi:10.1080/09613218.2011.582697

EAGA, 2004. Delivering Warmth Through Partnership: The Warm Front Scheme Annual Report 2002/03. EAGA Partnership, London, UK.

Egger, G., Swinburn, B., 1997. An "ecological" approach to the obesity pandemic. BMJ 315, 477–480. doi:10.1136/bmj.315.7106.477

EIA, 2013. International Energy Outlook 2013. U.S. Energy Information Administration, Washington, DC.

Elexon, 2009. Load profiles and their use in electricity settlement.

Elwood, M., 2007. Critical Appraisal of Epidemiological Studies and Clinical Trials, 3rd ed. Oxford University Press, Oxford, UK.

EPSRC, 2013. £39 million for UK energy efficiency research to cut carbon use [WWW Document]. EPSRC News. URL http://www.epsrc.ac.uk/newsevents/news/2012/Pages/energyefficiencycutcarbonuse.aspx (accessed 3.28.13).

EPSRC, 2014. Energy efficiency (end use energy demand) [WWW Document]. EPSRC Res. Portf. URL https://www.epsrc.ac.uk/research/ourportfolio/researchareas/energyefficiency/ (accessed 5.7.15).

EST, 2009. Homes Energy Efficiency Database (HEED) – DECC Data Frameworks Residential Pilot: Final Report to Department of Energy & Climate Change. Energy Saving Trust, London, UK.

EST, 2010. HEED - Data Interpretation Guide 2.0.

EST, 2011. HEED Online 2 - Full Technical Guide 2.0. Energy Saving Trust, London, UK.

European Commission, 2010. Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast). Off. J. Eur. Union 53, 13–35.

European Commission, 2011a. A Roadmap for moving to a competitive low carbon economy in 2050. European Commission, Brussels, Belgium.

European Commission, 2011b. Energy Efficiency Plan 2011. European Commission, Brussels, Belgium.

Experian, 2009. Median Household Income 2004 [WWW Document]. Experian Data. URL http://cdu.mimas.ac.uk/experian/index.htm

Experian, 2012. Household Income 2011. Experian Ltd., Nottingham, UK.

Firth, S.K., Lomas, K.J., Wright, A.J., 2010. Targeting household energy-efficiency measures using sensitivity analysis. Build. Res. Inf. 38, 25. doi:10.1080/09613210903236706

Formiguera, X., Cantón, A., 2004. Obesity: epidemiology and clinical aspects. Best Pract. Res. Clin. Gastroenterol. 18, 1125–1146. doi:10.1016/j.bpg.2004.06.030

Galvin, R., 2010. Thermal upgrades of existing homes in Germany: The building code, subsidies, and economic efficiency. Energy Build. 42, 834–844. doi:10.1016/j.enbuild.2009.12.004

Geller, H., Harrington, P., Rosenfeld, A.H., Tanishima, S., Unander, F., 2006. Polices for increasing energy efficiency: Thirty years of experience in OECD countries. Energy Policy 34, 556–573. doi:10.1016/j.enpol.2005.11.010

Gemserv, 2007. iGT UNC AQ Procedures Document v1.0. UK.

Geweke, J., Horowitz, J., Pesaran, H., 2008. Econometrics, in: The New Palgrave Dictionary of Economics. Palgrave Macmillan.

Gillingham, K., Newell, R., Palmer, K., 2006. Energy Efficiency Policies: A Retrospective Examination. Annu. Rev. Environ. Resour. 31, 161–192. doi:10.1146/annurev.energy.31.020105.100157

Gillingham, K., Newell, R.G., Palmer, K., 2009. Energy Efficiency Economics and Policy (No. 15031), NBER Working Paper Series. Cambridge, MA.

Goldman, C.A., 1985. Measured energy savings from residential retrofits: Updated results from the BECA-B project. Energy Build. 8, 137–155. doi:10.1016/0378-7788(85)90022-2

Green, G., Gilbertson, J., 2008. Warm Front, Better Health: Health Impact Evaluation of the Warm Front Scheme. Sheffield, UK.

Greening, L.A., Greene, D.L., Difiglio, C., 2000. Energy effciency and consumption - the rebound effect - a survey. Energy Policy 28, 389–401. doi:10.1016/S0301-4215(00)00021-5

Greenland, S., 1998. Induction versus Popper: substance versus semantics. Int. J. Epidemiol. 27, 543–548. doi:10.1093/ije/27.4.543

Greenland, S., Gago-Dominguez, M., Castelao, J.E., 2004. The Value of Risk-Factor ("Black-Box") Epidemiology. Epidemiology 15, 529–535. doi:10.1097/01.ede.0000134867.12896.23

Guerra Santin, O., Itard, L., Visscher, H., 2009. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy Build. 41, 1223–1232. doi:10.1016/j.enbuild.2009.07.002

Gupta, R., Gregg, M., 2012. Appraisal of UK funding frameworks for energy research in housing. Build. Res. Inf. 40, 446–460. doi:10.1080/09613218.2012.683240

Guy, S., Shove, E., 2000. The sociology of energy, buildings and the environment: constructing knowledge, designing practice. Routledge, London, UK.

Haines, A., McMichael, A.J., Smith, K.R., Roberts, I., Woodcock, J., Markandya, A., Armstrong, B.G., Campbell-Lendrum, D., Dangour, A.D., Davies, M., Bruce, N., Tonne, C., Barrett, M., Wilkinson, P., 2009. Public health benefits of strategies to reduce greenhouse-gas emissions: overview and implications for policy makers. Lancet 374, 2104–2114. doi:10.1016/S0140-6736(09)61759-1

Halliday, S., 2001. Death and miasma in Victorian London: an obstinate belief. BMJ 323, 1469–71.

Harrington, B.E., Heyman, B., Merleau-Ponty, N., Stockton, H., Ritchie, N., Heyman, A., 2005. Keeping warm and staying well: findings from the qualitative arm of the Warm Homes Project. Health Soc. Care Community 13, 259–67. doi:10.1111/j.1365-2524.2005.00558.x

Hartenberger, U., Lorenz, D., Lützkendorf, T., 2013. A shared built environment professional identity through education and training. Build. Res. Inf. 41, 60–76. doi:10.1080/09613218.2013.736202

Hassett, K. a., Metcalf, G.E., 1995. Energy tax credits and residential conservation investment: Evidence from panel data. J. Public Econ. 57, 201–217. doi:10.1016/0047-2727(94)01452-T

Hassett, K.A., Metcalf, G.E., 1993. Energy conservation investment. Energy Policy 21, 710–716.

Healy, J.D., 2003. Excess winter mortality in Europe: a cross country analysis identifying key risk factors. J. Epidemiol. Community Heal. 57, 784–789. doi:10.1136/jech.57.10.784

Healy, J.D., Clinch, J.P., 2002. Fuel poverty, thermal comfort and occupancy: results of a national household-survey in Ireland. Appl. Energy 73, 329–343.

Healy, J.D., Clinch, J.P.P., 2004. Quantifying the severity of fuel poverty, its relationship with poor housing and reasons for non-investment in energy-saving measures in Ireland. Energy Policy 32, 207–220. doi:10.1016/S0301-4215(02)00265-3

Helm, D., 2003. Energy, the State, and the Market: British Energy Policy since 1979. OUP USA, Oxford, UK.

Helm, D., 2005. The Assessment: The New Energy Paradigm. Oxford Rev. Econ. Policy 21, 1–18. doi:10.1093/oxrep/gri001

Herring, H., 2006. Energy efficiency—a critical view. Energy 31, 10–20.

Herring, H., Roy, R., 2007. Technological innovation, energy efficient design and the rebound effect. Technovation 27, 194–203. doi:10.1016/j.technovation.2006.11.004

Heyman, B., Harrington, B., Heyman, A., 2011. A Randomised Controlled Trial of an Energy Efficiency Intervention for Families Living in Fuel Poverty. Hous. Stud. 26, 117–132. doi:10.1080/02673037.2010.512787

Higgins, J., Green, S., 2008. Cochrane handbook for systematic reviews of interventions, 5.1.0 ed. Wiley-Blackwell, Chichester, UK.

Hills, J., 2012. Getting the measure on fuel poverty - final report of the fuel poverty review. Department of Energy and Climate Change, London, UK.

HM Government, 2011. The Carbon Plan: Delivering our low carbon future. DECC, London, UK.

HM Government, 2012. Open Data White Paper: Unleashing the Potential. The Stationery Office, London, UK.

HM Government, 2013. Maintaining UK energy security [WWW Document]. Gov.uk. URL https://www.gov.uk/government/policies/maintaining-uk-energy-security--2

HMSO, 1998. The Gas Safety (Installations and Use) Regulations 1998. Her Majesty's Stationery Office, United Kingdom.

HMSO, 2006. Building and Approved Inspectors (Amendment) Regulations 2006. United Kingdom.

Hong, S.H., Gilbertson, J., Oreszczyn, T., Green, G., Ridley, I., 2009. A field study of thermal comfort in low-income dwellings in England before and after energy efficient refurbishment. Build. Environ. 44, 1228–1236. doi:10.1016/j.buildenv.2008.09.003

Hong, S.H., Oreszczyn, T., Ridley, I., 2006. The impact of energy efficient refurbishment on the space heating fuel consumption in English dwellings. Energy Build. 38, 1171–1181.

Horton, R., 2012. GBD 2010: understanding disease, injury, and risk. Lancet 380, 2053–4. doi:10.1016/S0140-6736(12)62133-3

Horton, R., 2013. Non-communicable diseases: 2015 to 2025. Lancet 381, 509–10. doi:10.1016/S0140-6736(13)60100-2

Hough, D., White, E., 2014. The Green Deal (No. SN/SC/05763), Science and Environment Section.

House of Commons, 2011. House of Commons Energy and Climate Change Committee UK Energy Supply: Security or Independence ? (No. HC 1065). London, UK.

House of Commons Library, 1999. A Century of Change: Trends in UK statistics since 1900, Social and General Statistics Section. London, UK.

Howden-Chapman, P., Viggers, H., Chapman, R., O'Dea, D., Free, S., O'Sullivan, K., 2009. Warm homes: Drivers of the demand for heating in the residential sector in New Zealand. Energy Policy 37, 3387–3399. doi:10.1016/j.enpol.2008.12.023

Hu, F., 2008. Obesity Epidemiology, in: Hu, F. (Ed.), Obesity Epidemiology. Oxford University Press, New York, USA, pp. 53–83. doi:10.1093/acprof:oso/9780195312911.003.0005

Huebner, G.M., McMichael, M., Shipworth, D., Shipworth, M., Durand-Daubin, M., Summerfield, A., 2013. Heating patterns in English homes: Comparing results from a national survey against common model assumptions. Build. Environ. 70, 298–305. doi:10.1016/j.buildenv.2013.08.028

Hughes, L., 2009. The four "R"s of energy security. Energy Policy 37, 2459–2461. doi:10.1016/j.enpol.2009.02.038

Hughes, M., Palmer, J., Cheng, V., Shipworth, D., 2013a. housing energy model Sensitivity and uncertainty analysis of England's housing energy model. Build. Res. Inf. 41, 156–167. doi:10.1080/09613218.2013.769146

Hughes, M., Palmer, J., Cheng, V., Shipworth, D., 2013b. Sensitivity and uncertainty analysis of England's housing energy model. Build. Res. Inf. 41, 156–167. doi:10.1080/09613218.2013.769146

Hutton, S., 1984. Domestic fuel expenditure: An analysis of three national surveys. Energy Econ. 6, 52–58.

IEA, 2010. World Energy Outlook 2010, 2010th ed, World Energy Outlook. OECD Publishing, Paris, France. doi:10.1787/weo-2010-en

IEA, 2012. World Energy Outlook 2012, 2012th ed, World Energy Outlook. OECD Publishing, Paris, France. doi:10.1787/weo-2012-en

IEA, 2013. Energy Efficiency Market Report 2013. OECD Publishing, Paris, France.

International Energy Agency (IEA), 2012. Glossary of terms [WWW Document]. URL http://www.iea.org/glossary/glossary_A.asp

IPCC, 2014. Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

IPCC Secretariat, WMO, UNEP, 2007a. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

IPCC Secretariat, WMO, UNEP, 2007b. Climate Change 2007: Mitigation of Climate Change. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Jabareen, Y., 2009. Building a Conceptual Framework: Philosophy, Definitions, and Procedure. Int. J. Qual. Methods 8, 49–62.

Jamasb, T., Meier, H., 2010. Household Energy Expenditure and Income Groups: Evidence from Great Britain, EPRG Working Paper. Cambridge, UK.

Jenicek, M., 1997. Epidemiology, evidenced-based medicine, and evidence-based public health. J. Epidemiol. 7, 187–197.

Joffe, M., Gambhir, M., Chadeau-Hyam, M., Vineis, P., 2012. Causal diagrams in systems epidemiology. Emerg. Themes Epidemiol. 9, 1. doi:10.1186/1742-7622-9-1

Jollands, N., Waide, P., Ellis, M., Onoda, T., Laustsen, J., Tanaka, K., de T'Serclaes, P., Barnsley, I., Bradley, R., Meier, A., 2010. The 25 IEA energy efficiency policy recommendations to the G8 Gleneagles Plan of Action. Energy Policy 38, 6409–6418. doi:10.1016/j.enpol.2009.11.090

Karhausen, L.R., 1995. The Poverty of Popperian Epidemiology. Int. J. Epidemiol. 24, 869–874. doi:10.1093/ije/24.5.869

Kaza, N., 2010. Understanding the spectrum of residential energy consumption: A quantile regression approach. Energy Policy 38, 6574–6585. doi:10.1016/j.enpol.2010.06.028

Kelly, M.J., 2009. Retrofitting the existing UK building stock. Build. Res. Inf. 37, 196–200.

Kelly, S., 2011. Do homes that are more energy efficient consume less energy?: A structural equation model of the English residential sector. Energy 36, 5610–5620. doi:10.1016/j.energy.2011.07.009

Kelly, S., Shipworth, M., Shipworth, D., Gentry, M., Wright, A., Pollitt, M., Crawford-Brown, D., Lomas, K., 2013. Predicting the diversity of internal temperatures from the English residential sector using panel methods. Appl. Energy 102, 601–621. doi:10.1016/j.apenergy.2012.08.015

Kohler, N., Steadman, P., Hassler, U., 2009. Research on the building stock and its applications. Build. Res. Inf. 37, 449–454. doi:10.1080/09613210903189384

Krieger, N., 1994. Epidemiology and the web of causation: has anyone seen the spider? Soc. Sci. Med. 39, 887–903.

Kuhn, T.S., 1996. The structure of scientific revolutions, 3rd ed. The University of Chicago Press, Chicago, USA.

Laitner, J. a. “Skip,” 2009. Understanding the size of the energy efficiency resource: Ten policy recommendations to accelerate more productive investments. Policy Soc. 27, 351–363. doi:10.1016/j.polsoc.2009.01.009

Laurent, M., Allibe, B., Tigchelaar, C., Oreszczyn, T., Hamilton, I., 2013. Back to reality: How domestic energy efficiency policies in four European countries can be improved by using empirical data instead of normative calculations, in: ECEEE Summer Study: Rethink, Renew, Restart. pp. 2057–2070.

Leahy, E., Lyons, S., 2010. Energy use and appliance ownership in Ireland. Energy Policy 38, 4265–4279. doi:10.1016/j.enpol.2010.03.056

Li, F.G.N., Smith, A.Z.P., Biddulph, P., Hamilton, I.G., Lowe, R., Mavrogianni, A., Oikonomou, E., Raslan, R., Stamp, S., Stone, A., Summerfield, A.J., Veitch, D., Gori, V., Oreszczyn, T., 2014. Solid-wall U -values: heat flux measurements compared with standard assumptions. Build. Res. Inf. 43, 238–252. doi:10.1080/09613218.2014.967977

Liddell, C., Morris, C., 2010. Fuel poverty and human health: A review of recent evidence. Energy Policy 38, 2987–2997. doi:10.1016/j.enpol.2010.01.037

Lomas, K.J., 2009. Decarbonizing national housing stocks: strategies, barriers and measurement. Build. Res. Inf. 37, 187–191.

Lomas, K.J., 2010. Carbon reduction in existing buildings: a transdisciplinary approach. Build. Res. Inf. 38, 1. doi:10.1080/09613210903350937

Lowe, R.J., Oreszczyn, T., 2008. Regulatory standards and barriers to improved performance for housing. Energy Policy 36, 4475–4481. doi:10.1016/j.enpol.2008.09.024

Lowe, R.J., Wingfield, J., Bell, M., Bell, J.M., 2007. Evidence for heat losses via party wall cavities in masonry construction. Build. Serv. Eng. Res. Technol. 28, 161–181. doi:10.1177/0143624407077196

Lutsey, N., Sperling, D., 2008. America's bottom-up climate change mitigation policy. Energy Policy 36, 673–685. doi:10.1016/j.enpol.2007.10.018

Lutzenhiser, L., Shove, E., 1999. Contracting knowledge: the organizational limits to interdisciplinary energy efficiency research and development in the US and the UK. Energy Policy 27, 217–227.

Mallaburn, P.S., Eyre, N., 2013. Lessons from energy efficiency policy and programmesin the UK from 1973 to 2013. Energy Effic. 7, 23–41. doi:10.1007/s12053-013-9197-7

Margolis, R.M., Kammen, D.M., 1999. Underinvestment: the energy technology and R&D policy challenge. Science (80-.). 285, 690.

Margolis, R.M., Kammen, D.M., 1999. Evidence of under-investment in energy R&D in the United States and the impact of Federal policy. Energy Policy 27, 575–584.

McLeod, K.S., 2000. Our sense of Snow: the myth of John Snow in medical geography. Soc. Sci. Med. 50, 923–935. doi:10.1016/S0277-9536(99)00345-7

Meier, H., Rehdanz, K., 2010. Determinants of residential space heating expenditures in Great Britain. Energy Econ. 32, 949–959. doi:10.1016/j.eneco.2009.11.008

Mills, B., Schleich, J., 2012. Residential energy-efficient technology adoption, energy conservation, knowledge, and attitudes: An analysis of European countries. Energy Policy 49, 616–628. doi:10.1016/j.enpol.2012.07.008

Milne, G., Boardman, B., 2000. Making cold homes warmer: the effect of energy efficiency improvements in low-income homes. Energy Policy 28, 411–424. doi:10.1016/S0301-4215(00)00019-7

Mooney, S.J., Baecker, A., Rundle, A.G., 2013. Comparison of anthropometric and body composition measures as predictors of components of the metabolic syndrome in a clinical setting. Obes. Res. Clin. Pract. 7, e55–e66.

Moore, R., 2012. Definitions of fuel poverty: Implications for policy. Energy Policy 49, 19–26. doi:10.1016/j.enpol.2012.01.057

Moran, F., Natarajan, S., Nikolopoulou, M., 2012. Developing a database of energy use for historic dwellings in Bath, UK. Energy Build. 55, 218–226. doi:10.1016/j.enbuild.2012.09.016

Morgenstern, H., 1995. Ecologic studies in epidemiology: concepts, principles, and methods. Annu. Rev. Public Health 16, 61–81. doi:10.1146/annurev.pu.16.050195.000425

Nair, G., Gustavsson, L., Mahapatra, K., 2010. Factors influencing energy efficiency investments in existing Swedish residential buildings. Energy Policy 38, 2956–2963. doi:10.1016/j.enpol.2010.01.033

Nemet, G.F., Kammen, D.M., 2007. U.S. energy research and development: Declining investment, increasing need, and the feasibility of expansion. Energy Policy 35, 746–755.

Noailly, J., 2012. Improving the energy efficiency of buildings: The impact of environmental policy on technological innovation. Energy Econ. 34, 795–806. doi:10.1016/j.eneco.2011.07.015

NRS, 2013. 2011 Census : Key Results on Population , Ethnicity , Identity , Language , Religion , Health , Housing and Accommodation in Scotland - Release 2A. National Records of Scotland, Edinburgh, UK.

OECD/IEA, 2010. Energy Poverty - How to make modern energy access universal? OECD/IEA, Paris, France.

OECD/IEA, 2012. World Energy Statistics - Energy use per capita 1960 to 2010 [WWW Document]. URL http://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE

Ofgem, 2002. Energy Efficiency Commitment 2002-2005 Technical Guidance Manual Issue 3. Office of the Gas and Electricity Markets, London, UK.

OFGEM, 2009. Uniform Network Code - Generic LDZ CSEP NExA: Annex A. London, UK.

Ofgem, Energy Saving Trust, 2003. A review of the Energy Efficiency Standards of Performance 1994 - 2002. Office of the Gas and Electricity Markets, London, UK.

ONS, 2010. Mid-2005 Population estimates for UK, England and Wales, Scotland, Northern Ireland [WWW Document]. Natl. Stat. Popul. URL http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2002-to-mid-2010-revised--national-/index.html

ONS, 2011. Statistical Bulletin National Population Projections , 2010-Based Statistical Bulletin. Office of National Statistics, London, UK.

ONS, 2012a. 2011 Census, Population and Household Estimates for England and Wales [WWW Document]. 2011 Census. URL http://www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-england-and-wales/index.html

ONS, 2012b. ONS Neighbourhood Statistics [WWW Document]. Neighb. Stat. URL http://www.neighbourhood.statistics.gov.uk/

ONS, n.d. A Beginner's Guide to Geography.

OPSI, 2008. Climate Change Act 2008. HMSO, London, United Kingdom.

Oreszczyn, T., 1993. The energy duality of conservatory use: a survey of conservatory use. Florence, pp. 522–525.

Oreszczyn, T., Hong, S.H., Ridley, I., Wilkinson, P., 2006a. Determinants of winter indoor temperatures in low income households in England. Energy Build. 38, 245–252. doi:10.1016/j.enbuild.2005.06.006

Oreszczyn, T., Lowe, R.J., 2010. Challenges for energy and buildings research: objectives, methods and funding mechanisms. Build. Res. Inf. 38, 107–122. doi:10.1080/09613210903265432

Oreszczyn, T., Ridley, I., Hong, S.H., Wilkinson, P., 2006b. Mould and winter indoor relative humidity in low income households in England. Indoor built Environ. 15, 125–35. doi:10.1177/1420326X06063051

Oxford English Dictionary Online, 2010. Oxford English Dictionary Online, Oxford English Dictionary, OED Online. Oxford University Press, Oxford, UK.

Oxman, A.D., 1994. Systematic Reviews: Checklists for review articles. BMJ 309, 648–651. doi:10.1136/bmj.309.6955.648

Pachauri, S., 2004. An analysis of cross-sectional variations in total household energy requirements in India using micro survey data. Energy Policy 32, 1723–1735. doi:10.1016/S0301-4215(03)00162-9

Palmer, J., Cooper, I., 2013. United Kingdom housing energy fact file 2012. Department of Energy and Climate Change, London, UK.

Palmer, J., Cooper, I., Hughes, M., 2013. Openness in household energy use: the new housing energy fact file, in: Swan, W., Brown, P. (Eds.), Retrofitting the Built Environment. John Wiley & Sons, London, UK, pp. 20–35. doi:10.1002/9781118273463.ch3

Palmer, J., Terry, N., Pope, P., 2012. How much energy could be saved by making small changes to everyday household behaviours ? Department of Energy and Climate Change, London, UK.

Parnell, R., 2005. Informing the Development of Domestic Energy Efficiency Initiatives: An Everyday Householder-Centered Framework. Environ. Behav. 37, 787–807. doi:10.1177/0013916504274008

Pearce, D., 2006. The political economy of an energy tax: The United Kingdom's Climate Change Levy. Energy Econ. 28, 149–158. doi:10.1016/j.eneco.2005.10.001

Pearce, N., 1996. Traditional epidemiology, modern epidemiology, and public health. Am. J. Public Health 86, 678–683. doi:10.2105/AJPH.86.5.678

Pearce, N., 2011. Epidemiology in a changing world: variation, causation and ubiquitous risk factors. Int. J. Epidemiol. 40, 503–12. doi:10.1093/ije/dyq257

Pelenur, M.J., Cruickshank, H.J., 2012. Closing the Energy Efficiency Gap: A study linking demographics with barriers to adopting energy efficiency measures in the home. Energy 47, 348–357. doi:10.1016/j.energy.2012.09.058

Perdiguero, E., 2001. History of health, a valuable tool in public health. J. Epidemiol. Community Heal. 55, 667–673. doi:10.1136/jech.55.9.667

Perrels, A., Ostertag, K., Henderson, G., 2006. Reshaping markets for the benefit of energy saving. Energy Policy 34, 121–128. doi:10.1016/j.enpol.2004.08.023

Petticrew, M., Whitehead, M., Macintyre, S.J., Graham, H., Egan, M., 2004. Evidence for public health policy on inequalities: 1: the reality according to policymakers. J. Epidemiol. Community Health 58, 811–6. doi:10.1136/jech.2003.015289

Porta, M., Greenland, S., Last, J.M., 2008. A Dictionary of Epidemiology, 5th ed. Oxford University Press, New York, USA.

Poyer, D.A., Henderson, L., Teotia, A.P.S., 1997. Residential energy consumption across different population groups: comparative analysis for Latino and non-Latino households in USA. Energy Econ. 19, 445–463. doi:10.1016/S0140-9883(97)01024-4

Poyer, D.A., Williams, M., 1993. Residential energy demand: additional empirical evidence by minority household type. Energy Econ. 15, 93–100. doi:10.1016/0140-9883(93)90027-O

Price, C.W., 2005. The Effect of Liberalizing UK Retail Energy Markets on Consumers. Oxford Rev. Econ. Policy 21, 128–144. doi:10.1093/oxrep/gri007

Reddy, A., 2002. The new paradigm for energy. Curr. Sci. 82, 258–259.

Rose, G., 2001. Sick individuals and sick populations. Int. J. Epidemiol. 30, 427–32; discussion 433–4.

Rose, G., Barker, D.J., 1978. Epidemiology for the uninitiated. Rates. Br. Med. J. 2, 941–2.

Rose, G., Khaw, K.-T., Marmot, M., 2008. Rose's Strategy of Preventive Medicine. Oxford University Press.

Rosenow, J., 2011. Different paths of change: home energy efficiency policy in britain and germany. ECEEE Summer Study 261–272.

Rosenow, J., 2012. Energy savings obligations in the UK—A history of change. Energy Policy 49, 373–382. doi:10.1016/j.enpol.2012.06.052

Rothman, K.J., Greenland, S., 2005. Causation and causal inference in epidemiology. Am. J. Public Health 95 Suppl 1, S144–50. doi:10.2105/AJPH.2004.059204

Rothman, K.J., Lash, T.L., Greenland, S., 2008. Modern Epidemiology, 3rd ed. Lippincott Williams & Wilkins, London, UK.

Rowland, D.T., 2003. Demographic Methods and Concepts. Oxford University Press, Oxford, UK.

Rychetnik, L., Frommer, M., Hawe, P., Shiell, A., 2002. Criteria for evaluating evidence on public health interventions. J Epidemiol Community Heal. 56, 119–127.

Saidel, M. a, Alves, S.S., 2003. Energy efficiency policies in the OECD countries. Appl. Energy 76, 123–134. doi:10.1016/S0306-2619(03)00053-9

SAS Institute Inc., 2011. SAS Base Version 9.3.

Schipper, L., 1987. Energy conservation policies in the OECD did they make a difference? Energy Policy 15, 538–548.

Schweber, L., Leiringer, R., 2012. Beyond the technical: a snapshot of energy and buildings research. Build. Res. Inf. 40, 481–492. doi:10.1080/09613218.2012.675713

Scottish Government, 2009. Scottish House Condition Survey - Key Findings 2008, Scottish House Condition Survey. Edinburgh, UK.

Sefton, T., 2002. Targeting fuel poverty in England: is the government getting warm? Fisc. Stud. 23, 369–399. doi:10.1111/j.1475-5890.2002.tb00065.x

Severinsen, M., 2001. Principles Behind Definitions of Diseases – a Criticism of the Principle of Disease Mechanism and the Development of a Pragmatic Alternative. Theor. Med. Bioeth. 22, 319–336. doi:10.1023/A:1011830602137

Seyfang, G., 2010. Community action for sustainable housing: Building a low-carbon future. Energy Policy 38, 7624–7633. doi:10.1016/j.enpol.2009.10.027

Shipworth, M., 2005. Transdisciplinarity in the sociology of energy use, in: 37th World Congress of the International Institute of Sociology. WCIIS, Stockholm, Sweden.

Shipworth, M., 2011. Thermostat settings in English houses: No evidence of change between 1984 and 2007. Build. Environ. 46, 635–642. doi:10.1016/j.buildenv.2010.09.009

Shipworth, M., Firth, S.K., Gentry, M.I., Wright, A.J., Shipworth, D.T., Lomas, K.J., 2010. Central heating thermostat settings and timing: building demographics. Build. Res. Inf. 38, 50–69. doi:10.1080/09613210903263007

Shorrock, L., Dunster, J., 1997. The physically-based model BREHOMES and its use in deriving scenarios for the energy use and carbon dioxide emissions of the UK housing stock. Energy Policy 25, 1027–1037. doi:10.1016/S0301-4215(97)00130-4

Shortt, N., Rugkåsa, J., Rugkasa, J., 2007. "The walls were so damp and cold" fuel poverty and ill health in Northern Ireland: results from a housing intervention. Heal. Place 13, 99–110. doi:10.1016/j.healthplace.2005.10.004

Shove, E., 2003. Converging Conventions of Comfort, Cleanliness and Convenience. J. Consum. Policy 26, 395–418. doi:10.1023/A:1026362829781

Skea, J., 2012. Research and evidence needs for decarbonisation in the built environment: a UK case study. Build. Res. Inf. 40, 432–445. doi:10.1080/09613218.2012.670395

Skea, J., Hannon, M., Rhodes, A., 2013. Investing in a brighter energy future : Energy research and training prospectus 2013. London, UK.

Skea, J., Nishioka, S., 2008. Policies and practices for a low-carbon society. Clim. Policy 8, S5–S16. doi:10.3763/cpol.2008.0487

Snow, J., 1857. Cholera and the water supply in the southern districts of London. BMJ 2, 864.

Socolow, R.H., 1978. The Twin Rivers program on energy conservation in housing: Highlights and conclusions. Energy Build. 1, 207–242. doi:10.1016/0378-7788(78)90003-8

Sonderegger, R.C., 1978. Movers and stayers: The resident's contribution to variation across houses in energy consumption for space heating. Energy Build. 1, 313–324. doi:10.1016/0378-7788(78)90011-7

Sorrell, S., 2006. Improving the Evidence Base for Energy Policy: The Role of Systematic Reviews (No. 146), Energy Policy, SPRU Electronic Working Paper. Brighton, UK. doi:10.1016/j.enpol.2006.06.008

Sorrell, S., 2007. The Rebound Effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency. UK Energy Research Centre, London, UK.

Sorrell, S., Dimitropoulos, J., Sommerville, M., 2009. Empirical estimates of the direct rebound effect: A review. Energy Policy 37, 1356–1371. doi:10.1016/j.enpol.2008.11.026

Steadman, P., Bruhns, H.R., Holtier, S., Gakovic, B., Rickaby, P. a, Brown, F.E., 2000. A classification of built forms. Environ. Plan. B Plan. Des. 27, 73–91. doi:10.1068/bst7

Steadman, P., Evans, S., Batty, M., 2009. Wall area, volume and plan depth in the building stock. Build. Res. Inf. 37, 455. doi:10.1080/09613210903152531

Steg, L., 2008. Promoting household energy conservation. Energy Policy 36, 4449–4453. doi:10.1016/j.enpol.2008.09.027

Stern, N., 2007. The economics of climate change. Cambridge University Press, Cambridge.

Stone, A., Shipworth, D., Biddulph, P., Oreszczyn, T., 2014. Key factors determining the energy rating of existing English houses. Build. Res. Inf. 42, 725–738. doi:10.1080/09613218.2014.905383

Summerfield, A.J., Bruhns, H.R., Caeiro, J., Steadman, P., Oreszczyn, T., 2005. Life course building epidemiology: An alternative to the collection and analysis of carbon emission data, in: The Queensland University of Technology Research Week International Conference. Brisbane, Australia.

Summerfield, A.J., Lowe, R.J., 2012. Challenges and future directions for energy and buildings research. Build. Res. Inf. 40, 391–400. doi:10.1080/09613218.2012.693839

Summerfield, A.J., Lowe, R.J., Bruhns, H.R., Caeiro, J.A., Steadman, J.P., Oreszczyn, T., 2007. Milton Keynes Energy Park revisited: Changes in internal temperatures and energy usage. Energy Build. 39, 783–791. doi:10.1016/j.enbuild.2007.02.012

Sunikka-Blank, M., Galvin, R., 2012. Introducing the prebound effect: the gap between performance and actual energy consumption. Build. Res. Inf. 40, 260–273. doi:10.1080/09613218.2012.690952

Susser, M., Susser, E., 1996a. Choosing a future for epidemiology: I. Eras and paradigms. Am. J. Public Health 86, 668–73.

Susser, M., Susser, E., 1996b. Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. Am. J. Public Health 86, 674–7.

Swinburn, B., Egger, G., 2002. Preventive strategies against weight gain and obesity. Obes. Rev. 3, 289–301.

Thomas, G., 2011. A Typology for the Case Study in Social Science Following a Review of Definition, Discourse, and Structure. Qual. Inq. 17, 511–521. doi:10.1177/1077800411409884

Tovar, M.A., 2012. The structure of energy efficiency investment in the UK households and its average monetary and environmental savings. Energy Policy 50, 723–735. doi:10.1016/j.enpol.2012.08.019

UK CCC, 2008. Building a low-carbon economy – the UK's contribution to tackling climate change: The first report of the Committee on Climate Change. Norwich, UK.

UK CCC, 2010. Fourth Carbon Budget. UK Committee on Climate Change, London, UK.

UK CCC, 2012. Meeting Carbon Budgets - 2012 Progress Report to Parliament. UK Committee on Climate Change, London, UK.

UK Met Office, 2012. Heating degree days 1961-2006 [WWW Document]. UKCP09 Annu. Data sets. URL http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/available/annual.html

UK Parliament, 2006. National Health Service Act 2006.

UNEP, 2007. Buildings and Climate Change: Status, challenges and opportunities. United Nations Environment Programme, Paris, France.

UNEP, 2011. Towards a Green Economy: Pathways to Sustainable Development and Poverty Eradication. United Nations Environment Programme, Paris, France.

Van den Bergh, J.C.J.M., 2010. Energy Conservation More Effective With Rebound Policy. Environ. Resour. Econ. 48, 43–58. doi:10.1007/s10640-010-9396-z

VOA, 2010. Council Tax Property Attribute Data.

Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C., Vandenbroucke, J.P., 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. Ann. Intern. Med. 147, 573. doi:10.7326/0003-4819-147-8-200710160-00010

Waddams Price, C., Brazier, K., Wang, W., 2012. Objective and subjective measures of fuel poverty. Energy Policy 49, 1–7. doi:10.1016/j.enpol.2011.11.095

Walker, J.J., Mitchell, R., Platt, S.D., Petticrew, M.P., Hopton, J., 2006. Does usage of domestic heating influence internal environmental conditions and health? Eur. J. Public Health 16, 463–9. doi:10.1093/eurpub/cki186

Walker, R., Liddell, C., McKenzie, P., Morris, C., 2013. Evaluating fuel poverty policy in Northern Ireland using a geographic approach. Energy Policy 63, 765–774. doi:10.1016/j.enpol.2013.08.047

Wall, L.W., Goldman, C.A., Rosenfeld, A.H., Dutt, G.S., 1983. Building energy use compilation and analysis (BECA). Part B: Retrofit of existing North American residential buildings. Energy Build. 5, 151–170. doi:10.1016/0378-7788(83)90003-8

Whitehead, M., Petticrew, M., Graham, H., Macintyre, S.J., Bambra, C., Egan, M., 2004. Evidence for public health policy on inequalities: 2: assembling the evidence jigsaw. J. Epidemiol. Community Health 58, 817–821. doi:10.1136/jech.2003.015297

Whitesides, G.M., Crabtree, G.W., 2007. Don't forget long-term fundamental research in energy. Science (80-.). 315, 796–8. doi:10.1126/science.1140362

WHO, 2010. International Classification of Diseases, ICD-10th ed.

Wilhite, H., Shove, E., Lutzenhiser, L., Kempton, W., 2000a. The Legacy of Twenty Years of Energy Demand Management: we know more about Individual Behaviour but next to Nothing about Demand, in: Jochem, E., Sathaye, J., Bouille, D. (Eds.), Society, Behaviour, and Climate Change Mitigation. Kluwer Academic Publishers, Dordrecht, pp. 109–126.

Wilhite, H., Shove, E., Lutzenhiser, L., Kempton, W., 2000b. Twenty Years of Energy Demand Management : We Know More About Individual Behavior But How Much Do We Really Know About Demand?, in: ACEEE. ACEEE, Alimore, California, USA, pp. 435–453.

Wilkinson, P., Landon, M., Armstrong, B., Stevenson, S., McKee, M., Fletcher, T., 2001. Cold comfort: the social and environmental determinants of excess winter death in England, 1986-1996, York: Joseph Rowntree Foundation. Joseph Rowntree Foundation, York, UK.

Wilkinson, P., Smith, K.R., Beevers, S., Tonne, C., Oreszczyn, T., 2007. Energy, energy efficiency, and the built environment. Lancet 370, 1175–87. doi:10.1016/S0140-6736(07)61255-0

Wilson, C., Chryssochoidis, G., Pettifor, H., 2013. Understanding Homeowners ' Renovation Decisions : Findings of the VERD Project. UK Energy Research Centre, London, UK.

Wong, J.K.W., Li, H., Wang, S.W., 2005. Intelligent building research: a review. Autom. Constr. 14, 143–159.

Woodward, M., 2004. Epidemiology: study design and data analysis, 2nd ed, Tropical Medicine, Texts in Statistical Science. CRC Press, Boca Raton, Florida, USA.

Wyatt, P., 2013. A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England. Energy Policy 60, 540–549. doi:10.1016/j.enpol.2013.05.037

Yohanis, Y.G., 2012. Domestic energy use and householders' energy behaviour. Energy Policy 41, 654–665. doi:10.1016/j.enpol.2011.11.028

Chapter 13 Appendix A

Home Energy Efficiency Database

Appendix introduction

This appendix provides further details on the Home Energy Efficiency Database (HEED) used in this thesis. Access to HEED was provided by the Energy Saving Trust (EST) as part of the EPSRC funded *Buildings and Energy Data Framework* (EP/H021957/1) and *New Empirically-Based Models on Energy Demand in Buildings* (EP/I038810/1) projects. Portions of this text are featured in the BEDF project report.

Two versions of HEED were used in this thesis analysis. The first was provided for use in the BEDF project in January of 2010 and comprised approximately 11.4 million unique dwelling identifiers. The second wave of data was provided in July of 2012 for use in the EBBS project and comprised approximately 16.3 million unique dwelling identifiers. The following text describes the 11.4 million HEED. The main difference between the two databases was the additional data draw from the Carbon Emission Reduction Target (CERT) and the Community Energy Savings Programme (CESP).

13.1 Homes Energy Efficiency Database

The Homes Energy Efficiency Database (HEED) was collected and maintained by the Energy Saving Trust (EST) on behalf of the Department of Energy and Climate Change (DECC). HEED is now wholly the responsibility of EST. HEED was subsumed within the National Energy Efficiency Database (NEED) (DECC, 2013d). The following details describe the HEED data used in this thesis in further detail.

HEED contains information on the characteristics and energy efficiency of over 50% of the GB housing stock (EST, 2011). EST collects the data from a variety of suppliers and existing sources of data. Although HEED is collected from this range of sources, EST go through a classification and cleaning process that attempts to rationalise the classes against which variables are collected. The bulk of HEED data has been classified using the Reduced Standard Assessment Procedure (rdSAP) format, which attempts to categorise dwellings into common bands relevant to modelling energy demand. Where other forms were used, the rdSAP variables were expanded to include more information (e.g. dwellings surveyed by CORGI include considerable additional information on heating systems). Otherwise, collected variables may also be allocated to the nearest class using rdSAP type categories.

The data provided in HEED typically draws from energy and other survey data (i.e. Home Energy Check, fuel poverty schemes, etc) and installed measures data (i.e Energy Efficiency Commitment 1 and 2, Carbon Emissions Reduction Targets, CORGI boiler inspections, FENSA, etc). The dataset contains information at a property level, i.e. dwelling level rather than by households or occupants, on the physical features of the dwelling as described in Appendix Table 1 and Appendix Table 2 below. For every record in HEED there is a date for when a detail of the dwelling was collected and when an intervention occurred; the dates range from 1995 to 2009 in the extract used for this study. The majority of HEED data was collected from 2004 onward. Every dwelling in HEED has a unique home identifier code that links together the property data collected from different years. HEED data contains address information for the property, but this was not provided for this study. Instead, the HEED data that was made available to us contains a postcode sector and lower layer super output area (LLSOA), also known as lower super output area (LSOA), code for purposes of approximate geographic identification.

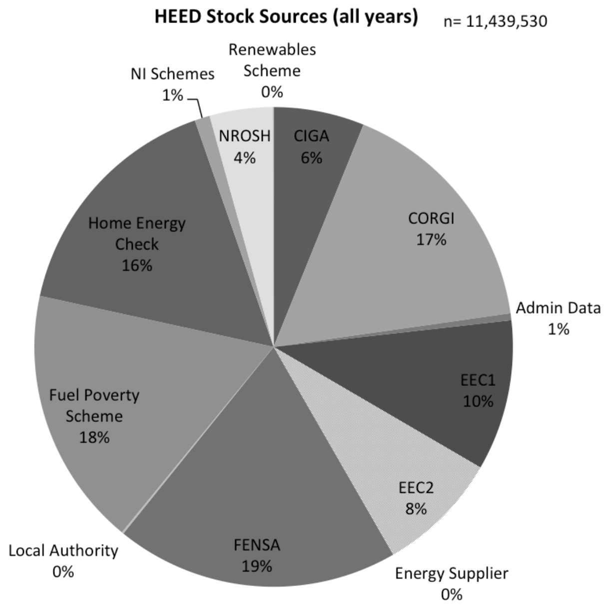
Appendix Table 1 - Homes Energy Efficiency Database (HEED) data suppliers and programmes

Programme	*Provider(s)*	*Survey/ Measures example*
Fuel Poverty Scheme	EAGA Warm Front Warm Homes Scottish Central Heating Programme The Warm Deal	Survey (e.g. dwelling type, tenure, age) Measures (e.g loft insulation)
Home Energy Check (HEC)	DAX SENSE	Survey (as above)
Boiler Replacement	CORGI	Measures (e.g. boiler type, heat system control type)
Window Replacement	FENSA	Measures (e.g. double glazing)
Clear Skies and Microgeneration Certification Scheme	Suppliers and Certified Installers	Measures (e.g. solar photovoltaic panels)
Social Housing Survey	NROSH Local Authorities Northern Ireland	Survey (as above)
Insulation Installers	CIGA NIA	Measures (e.g. cavity wall filling)

Appendix Table 2 - Homes Energy Efficiency Database (HEED) example data

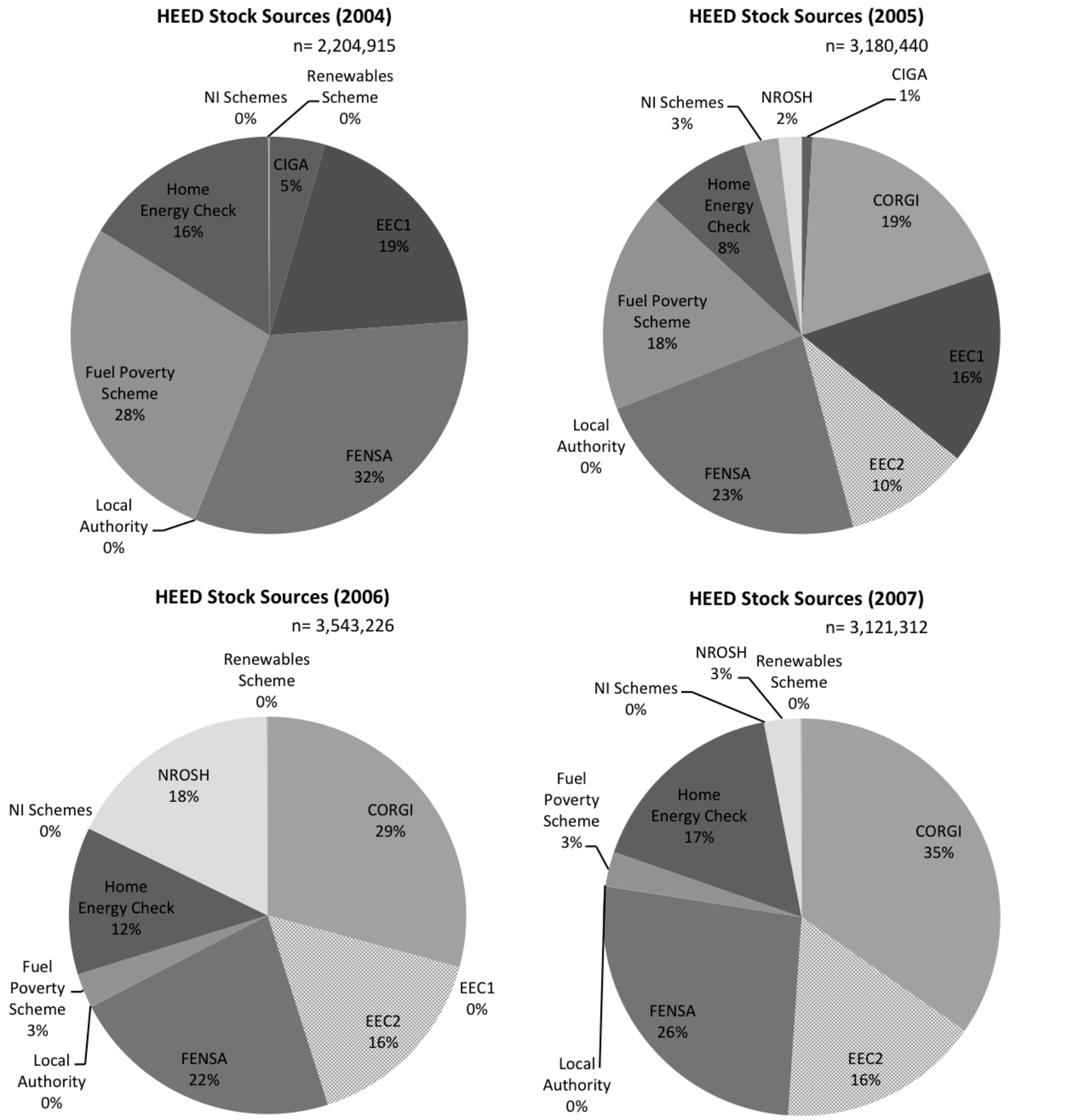
Data Type	*Data Examples*	
Survey Data	• Property type • Tenure • No of Bedrooms • Year of construction • Space heating fuel • Water heating fuel • Loft insulation thickness • External wall type	• Window type • Window frame type • Levels of draught-proofing • Main heating system • Secondary heating system • Hot water system • Heating controls (various types) • Energy rating (SAP/NHER) • Hot water tank insulation
Measures Data	• Loft insulation and depth • Cavity wall insulation • Solid wall insulation/flexible linings • Boiler replacements • Heating control upgrades	• Compact Fluorescent Lamps • Solar Thermal • Solar PV • Biomass heating • Heat Pumps

Appendix Figure 1 shows a breakdown of the sources from which the extract of HEED analysed was drawn. Note that the variables collected under each source will vary and many sources for measures will include survey data.



Appendix Figure 1 – HEED stock data sources

Appendix Figure 2 below shows the sources of data for each year of available energy data (2004 to 2007).



Appendix Figure 2 – HEED stock data sources by year of available energy demand (calendar year)

13.2 HEED: Data features

HEED is combined from approximately 60 datasets from approximately 20 organisations and numerous government-backed or sponsored fuel poverty and carbon reduction schemes. When combined these datasets feature 128 million records in approximately 3000 classes. Overall, there are 11,439,530 distinct dwelling identifiers. Appendix Table 3 below shows that approximately 2.7 million homes appear in two programmes (i.e source datasets), and approximately 1 million in three programmes. The great majority are present in only one home programme.

Appendix Table 3 – HEED data source count

No. Sources	No. dwellings Count HomeID
1	7,252,156
2	2,724,825
3	1,018,161
4	339,040
5	86,561
6	16,413
7	2,164
8	197
9	13
-	11,439,530

Many dwellings have multiple variables for which details of the dwellings are known. Appendix Table 1 below shows that approximately 50% of dwellings present in HEED have between 4 to 10 variables with information.

Appendix Table 4 – HEED detail count

No. Details	Count HomeID
1	576,099
2	15,646
3	100,752
4	2,629,785
5	356,787
6	305,878
7	103,770
8	603,796
9	574,007
10	1,009,987
…	…
35	7
36	3
	11,439,530

13.3 HEED Coverage

HEED consists primarily of information on the physical characteristics of the dwelling, including: dwelling type, age, number of bedrooms, wall type and insulation level, loft insulation, glazing type, heating system, and fuel type. It comprises information at the individual dwelling level rather than by households or occupants. It contains no information on households or dwelling occupants, aside from the tenure; so socio-economic factors cannot be determined directly.

HEED contains details on a range of energy efficiency interventions, including: loft insulation, cavity wall insulation, double-glazing installation, condensing boiler installation, draught-proofing, heating controls (e.g. thermostats), and hot water cylinder upgrades. A home enters HEED as a result of being part of a survey (self-completed by occupants or collected by others) or having an efficiency intervention. Approximately 70% of all dwellings in HEED had at least one efficiency intervention between 2000 and 2007.

Appendix Table 5 shows the number of installations per year in England for a selection of energy efficiency retrofits in HEED.

Appendix Table 5 - Total number of energy efficiency installations per year in England 2000 to 2007

Efficiency installations[a] in England (1000's)	Year								
	2000	2001	2002	2003	2004	2005	2006	2007	Total
Loft insulation	28	96	190	334	591	612	682	644	3,177
Condensing boiler install	18	59	110	240	264	600	969	1,071	3,331
Heating system[b] install	25	89	174	315	401	927	1,395	1,474	4,801
Double glazing install	33	95	157	148	1,810	2,733	2,617	2,707	10,300
Cavity insulation	92	168	251	362	652	853	917	764	4,060
Hot water cylinder install	0	1	2	2	12	46	57	66	186
Draught proofing	13	70	77	224	388	175	158	52	1,158
Solar hot water install	0	0	1	2	14	15	38	41	110
Total[c]	192	519	851	1,388	3,867	5,361	5,863	5,749	23,790

Note: a) Figures are for total number of installations, dwelling may have more than one measure; b) Heating system includes condensing boiler replacement; c) Total does not include condensing boiler install, as this is included in heating system

HEED has several limitations with respect to the interpretation of the data. In particular, the variety of sources from which the data are drawn can mean that the quality is not standardised with respect to collector bias or sample control. Also, there is a large amount of data missing on a number of physical characteristics of the home, which is due to the some data providers only collecting information required by their accreditation body or programme policies. The coverage of any given variable depends on the scheme or survey from which information on the homes was collected under. For example, dwellings in the CORGI data will have a high coverage of boiler-related variables but may have a lower coverage of other variables such as loft insulation levels. Appendix Table 6 gives the percentage covered (i.e. $n_{variable}/N$) for a selection of house characteristics and energy efficiency measures.

Appendix Table 6 - Percent coverage of selected variables in HEED for period from 2000 to 2007

Dwelling characteristic		**% Coverage**
Dwelling type		
	Missing	33%
	Flat/Maisonette	8%
	Bungalow	8%
	Mid-Terrace	12%
	End-Terrace	5%
	Semi-Detached	23%
	Detached	12%
No. Bedrooms		
	Missing	39%
	1	6%
	2	15%
	3	30%
	4	7%
	5+	2%
Dwelling Age		
	missing	54%
	Pre-1900	3%
	1900-1929	5%
	1930-1949	8%
	1950-1966	9%
	1967-1975	11%
	1976-1982	3%
	1983-1990	3%
	1991-1995	2%
	1996-2002	1%
	post-2002	0%
Household Tenure		
	missing	41%
	Owner Occupier	44%
	Privately Rented	4%
	Rented from Local Authority	5%
	Rented from Housing Association	5%
Region		
	North East	7%
	North West	16%
	Yorkshire and The Humber	12%
	East Midlands	8%
	West Midlands	11%
	East of England	10%
	London	11%
	South East	15%
	South West	10%
Energy Efficiency Installation		
Loft insulation to 250mm		
	Missing	81.81%
	0 - 250mm	5.45%
	25 - 250mm	2.16%
	50 - 250mm	4.57%
	75 - 250mm	1.02%
	100 - 250mm	4.64%
	150 - 250mm	0.33%
Cavity insulation		
	Missing	75%
	Cavity Wall Insulation (pre 1976)	15%
	Cavity Wall Insulation (post 1976)	3%
	Cavity Wall Insulation (Unknown Age)	7%
Glazing replacement		
	Missing	66%
	Double glazing replacement	34%
Heating system Replacement		
	Missing	81%
	Other	3%
	Condensing boiler	13%
	Non-condensing boiler	2%

13.4 HEED: comparison of dwelling characteristics

A great deal of the value associated with HEED is its coverage of the UK housing stock. Therefore, in using HEED it is important to determine how representative the characteristics of the dwellings in HEED are of the UK stock, as compared to other UK housing datasets.

The first step is to determine how representative of the housing stock as a whole the meter-matched HEED sample is. This is done by comparing HEED with three other databases: the English Housing Survey, the Scottish House Conditions Survey, and the Valuation Office Agency database of residential properties, which together provide more or less complete coverage of the housing stock of Great Britain.

In the first instance, the dwelling characteristics (i.e. age, type, tenure, size, location) are compared against other representative samples of the respective housing stocks in Great Britain (i.e. England, Wales, Scotland). For the purposes of testing the representativeness of HEED in terms of dwelling characteristics we use Chi-square tests for goodness-of-fit. For computational purposes, a 10% randomly selected sample of approximately 1.2 million dwelling records representative of HEED is used for the population comparison, rather than the full HEED database (i.e. 11.4 million).

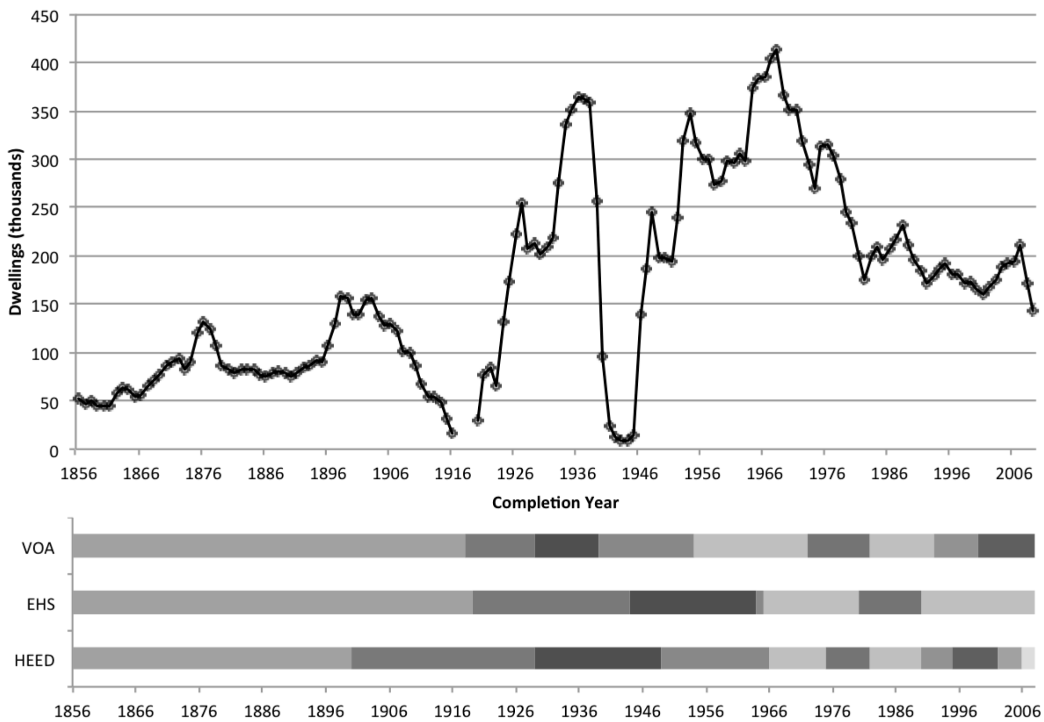
A set of building characteristic features within the HEED population was compared to the 2008 English Housing Survey (EHS), the 2007-09 Scottish House Condition Survey (SHCS) and the Valuation Office Agency's (VOA) Council Tax Property Attributes for 2010. The 2008 EHS was used because the collection was approximately in line with the last year of HEED data provided in the extract (i.e. the 11.4 million dataset); this is also the case for the 2007 to 2009 SCHS. The VOA updates the council tax database every year and therefore the latest extract is used (i.e. 2008). Both the EHS and SHCS provide a factor with which to weight variables in order to represent houses or households in England or Scotland; for the comparisons we use the houses weighting. No weighting is required of the VOA data. With respect to the potential changes in the stock since 2008 and therefore the implication of comparing the data to subsequent years, approximately 268,000 dwellings were built in 2009 and 2010 (approximately 0.1% of the total GB stock) (CLG, 2010c).

The English Housing Survey reports on the overall condition of English dwellings and the households living in them. The survey provides data on housing stock characteristics (including age, type and size) based on survey work undertaken between 2007 and 2009. The surveyed sample of properties where physical inspections were carried out contains 16,150 occupied or vacant dwellings, or 0.7% of the housing stock of 22.7 million dwellings in England (CLG, 2010a). The EHS provides a statistically random sample of the English stock against which HEED can be compared.

The Scottish House Condition Survey reports on the households and physical condition of the Scottish stock. The survey includes an interview and physical survey of approximately 3,000 to 4,000 dwellings, undertaken in a continuous format with the aim of achieving a 15,000 sample over 5 years (Scottish Government, 2009). The survey collects physical characteristics of the dwellings (e.g. age, type, size, energy efficiency, etc.) and household features (e.g. tenure, income, etc.) The survey sample is stratified by area and randomly selected and attempts to be representative of the 2.4 million dwellings in Scotland.

The VOA's Council Tax Property Attributes tables are collected as part of the VOA's responsibility to group properties into the appropriate council tax band (VOA, 2010). As part of the banding process, the VOA collects information on the characteristics of the property that may affect its value (including age, type, area, and number of bedrooms). The VOA data was considered a 'global' listing, as it contains information on all dwellings within an area (i.e. local council) and thus should have all dwellings in the stock. Note that the VOA only collects information for England and Wales. The VOA in maintaining a current valuation list revises the data annually; the data used in the comparison comes from an extract made in September 2010. The VOA dataset provides information at the Local Authority level for approximately 24.7 million residential properties.

The EHS, SHCS, VOA and HEED are not collected using a common format (i.e. rdSAP), since they have all been developed for different purposes; as a result only some variables can be compared. In certain cases this has meant that variable classes have been banded in order to create comparable data categories (e.g. dwelling type and number of bedrooms). Other variables are too complex to do this; age, for example, is collected using a different age band for each survey and therefore 6 age bands have been developed to capture all periods. The dwelling completion rate can fluctuate from year to year and therefore it is difficult to aggregate the bands for comparison. The banding in the different source datasets can group years that exhibited a high build rate, as was seen in 1982-87 (House of Commons Library, 1999), and may therefore be difficult to compare when not exactly matched; see Appendix Figure 3. Care has been taken to check for these occurrences; the bands were developed to limit such errors. However for the comparison of age, the chi-square goodness-of-fit test is not performed and is instead compared for information.



Appendix Figure 3 – Historic house completion rate and survey and database age bands

13.5 Comparison of HEED dwelling characteristics

The characteristics of dwellings in the selected HEED sample (i.e. the randomly selected 10% dataset) are compared against representative samples for England, England and Wales, and Scotland. Appendix Table 7 and Appendix Table 8 provide overview statistics for the selected variables that are compared against the above-mentioned British housing datasets.

The results show that the HEED data is not statistically representative of the English and Welsh stock for the selected variables. In all cases of comparison we reject the hypothesis that the compared variables of the HEED data set are the same as those of the English Housing Survey and VOA Council Tax (i.e. all p-values <0.0001 at a 95% confidence limit).

Appendix Table 7 - HEED (England) dwelling characteristics compared to EHS

England[a]	HEED 10% (n)	HEED 10% (%)	EHS 2008 (%)
Dwelling Type			
Flat-Maisonette	96,975	17.2%	18.6%
Bungalow	54,837	9.7%	9.4%
Terrace	141,109	25.1%	28.6%
Semi-detached	183,309	32.6%	26.0%
Detached	86,434	15.4%	17.4%
X^2			12961.22
d.f.			4
p			<0.0001
Dwelling Tenure			
Social rental	156,195	21.8%	14.8%
Private rental	67,499	9.4%	17.7%
Owner-occupied	493,481	68.8%	67.5%
X^2			51585.46
d.f.			2
p			<0.0001
Dwelling Size (Bedrooms)			
1	71,315	12.6%	9.1%
2	142,619	25.3%	27.1%
3	267,307	47.4%	44.2%
4	58,600	10.4%	15.5%
5+	24,333	4.3%	4.0%
X^2			19219.87
d.f.			4
p			<0.0001
Dwelling Region			
North East	70,049	6.2%	5.1%
North West	159,820	14.2%	13.6%
Yorkshire and the Humber	120,624	10.7%	10.6%
East Midlands	91,541	8.1%	8.8%
West Midlands	116,000	10.3%	10.5%
East of England	109,080	9.7%	10.9%
London	132,433	11.8%	14.2%
South East	161,845	14.4%	15.8%
South West	107,767	9.6%	10.3%
X^2			9810.57
			8
			<0.0001

Notes: [a]10% HEED Sample, England only

Appendix Table 8 - HEED (England and Wales) dwelling characteristics compared to VOA

England & Wales[a]	HEED 10% (n)	HEED 10% (%)	VOA 2010 (%)
Dwelling Type			
Flat-Maisonette	96,975	17.2%	21.9%
Bungalow	54,837	9.7%	10.2%
Terrace	141,109	25.1%	27.3%
Semi-detached	183,309	32.6%	24.8%
Detached	86,434	15.4%	15.8%
X^2			20518.77
d.f.			4
p			<0.0001
Dwelling Size (Bedrooms)			
1	71,315	12.6%	11.6%
2	142,619	25.3%	28.4%
3	267,307	47.4%	45.4%
4	58,600	10.4%	11.5%
5+	24,333	4.3%	3.0%
X^2			6798.72
d.f.			4
p			<0.0001
Dwelling Region			
North East	70,049	6.2%	4.8%
North West	159,820	14.2%	12.9%
Yorkshire and The Humber	120,624	10.7%	9.5%
East Midlands	91,541	8.1%	8.1%
West Midlands	116,000	10.3%	9.7%
East of England	109,080	9.7%	10.4%
London	132,433	11.8%	13.8%
South East	161,845	14.4%	15.1%
South West	107,767	9.6%	9.8%
Wales	55,073	4.9%	5.7%
X^2			14076.56
d.f.			9
p			<0.0001

Notes: [a]10% HEED Sample, England and Wales only

Appendix Table 9 shows a comparison of the Scottish dwellings in HEED and accepts the hypothesis that the HEED sample is statistically similar to the Scottish House Conditions Survey.

Appendix Table 9 - HEED (Scotland) dwelling demographics comparison to SHCS

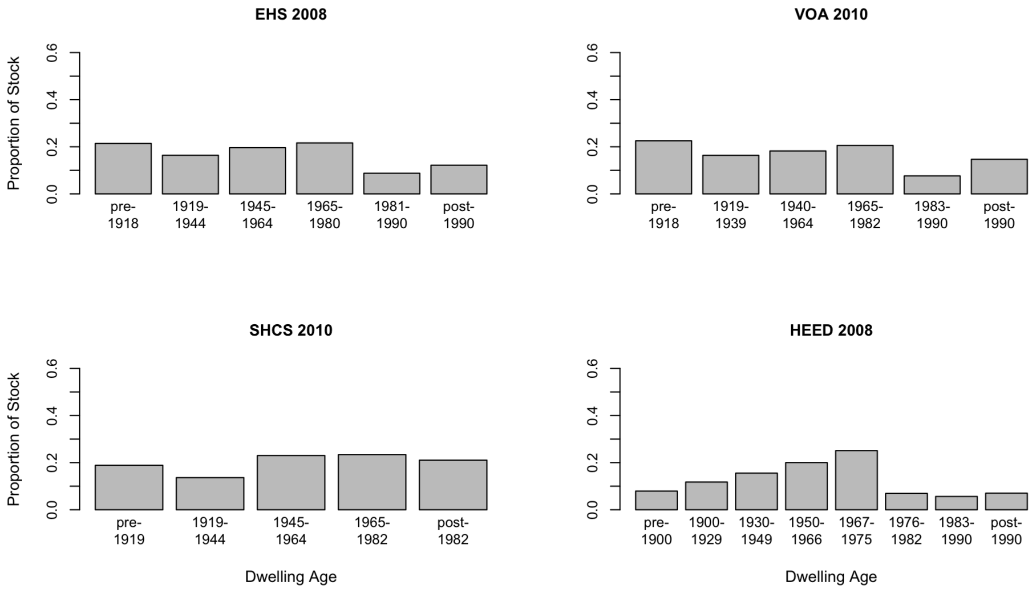
Scotland	HEED 10% (n)	HEED 10% (%)	SHCS 2009 (%)
Dwelling Type			
Flat-Maisonette	29,008	36.6%	36.7%
Bungalow	0	0.0%	0.0%
Terrace	20,334	25.6%	25.5%
Semi-detached	15,905	20.1%	20.1%
Detached	14,062	17.7%	17.8%
X^2			1.2293
d.f.			3
p			0.746
Dwelling Tenure			
Social rental	25,334	27.9%	27.7%
Private rental	9,562	10.5%	10.6%
Owner-occupied	56,017	61.6%	61.7%
X^2			1.5907
d.f.			2
p			0.4514
Dwelling Size (Bedrooms)			
1	11,274	19.3%	19.2%
2	22,321	38.1%	38.1%
3	19,314	33.0%	33.1%
4	3,735	6.4%	6.4%
5+	1,867	3.2%	3.1%
X^2			1.9065
d.f.			4
p			0.753

Notes: a) 10% HEED Sample, Scotland only

The analysis of the populations represented in the HEED data rejects the hypothesis that the sample is the same as the other data sets that represent the housing stock of England, and England and Wales. This does not necessarily mean however that HEED cannot be used to describe housing energy efficiency demand for those groups – merely that caution should be applied where findings from HEED are interpreted and generalised for the UK housing stock as a whole.

Overall, in the English and Welsh component of HEED, 'dwelling type' shows fewer flats and more semi-detached houses. There are fewer private let dwellings and more socially rented dwellings, likely reflecting the emphasis of the Government and Energy Supplier programmes to target areas of high deprivation and low-income groups. In terms of geographic coverage, there are fewer homes in the southern regions of England. Despite the targeting of the programmes, given the number of dwellings represented in HEED (approximately 50% of all dwellings), the HEED dataset does compare *relatively* well to the representative housing stocks of Great Britain. The HEED data can be said to represent the Scottish housing stock, which likely reflects the collection process and inclusion of a proportion of building performance rating data (i.e. Energy Performance Certificates).

Age is compared graphically rather than statistically, due to the difference in collection age bands. Appendix Figure 4 shows that there are more homes in the 1967-82 period and fewer 1990+ homes in the English and Welsh stocks. Scotland again is the same, and for comparison has the added benefit of using the same age bands.



Appendix Figure 4 – Housing stock age band comparison

A further visual comparison of the data using the above groups suggests where the differences lie within the HEED data as compared to the other datasets (see Figure 36). In HEED, there are more homes in the 1967-82 period and fewer 1990+ homes. This is the product of the programmes that provide HEED with data. 'Dwelling type' shows fewer flats and more semi-detached houses. There are fewer private let dwellings and fewer homes in the 'Southern' region. Interestingly, however, the proportion of bedrooms in HEED is comparable to the VOA housing stock. The differences may be due in large part to the way programmes collected into HEED targeted households, with many programmes seeking to assist low-income or vulnerable households. The data would also therefore reflect the types of houses in which these households live. Despite these programmes, however, given the number of houses represented in HEED (i.e. nearly 50% of all houses – this is significantly in excess of estimates of the number of dwellings that could be described as social housing) a good deal of the data will also represent homes that are privately owned and not social housing.

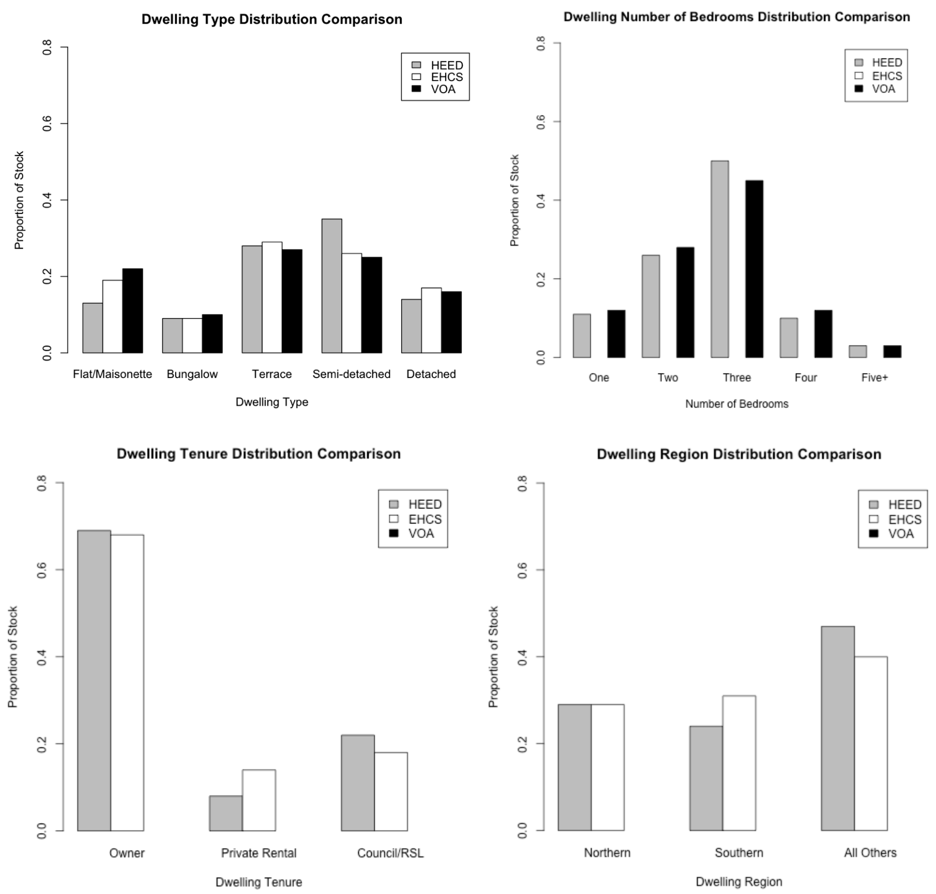


Figure 36 – HEED Stock: Key housing characteristic comparison

13.6 HEED: segmentation of dwelling characteristics

The distribution of key HEED variables is very important for how cross-segmentations of the HEED dataset are interpreted, in terms of the representativeness within the dataset and when undertaking further analysis of energy demand and energy efficiency interventions. The following section provides a basic overview of cross-tabulations of key HEED variables, allowing for a layered evaluation of the data. The coverage for the different dwelling characteristics is different and the tables below will show different total counts. Also, missing classes have been removed from these tables. Note that the coverage of HEED variables range from ~30% to 70%, therefore when assessing any combination of variables the number with non-missing class data changes; however the sample size remains the same. This is highlighted in the following tables with a 'missing' category.

Appendix Table 10 - HEED Stock: Dwelling Age by Number of Bedrooms

	Number of bedrooms							
	Missing	1	2	3	4	5+	All	
Dwelling Age	*count*	*count*	*count*	*count*	*count*	*count*	*% of stock*	*count*
pre-1900	37,776	53,643	80,524	104,257	36,222	43,424	3.1	355,846
1900-29	51,475	46,911	152,175	260,203	54,888	31,474	5.2	597,126
1930-49	78,449	65,371	173,621	424,052	56,570	25,341	7.2	823,404
1950-66	121,780	113,344	260,396	451,030	63,880	25,516	9.1	1,035,946
1967-75	76,864	120,910	297,658	655,402	132,633	24,274	11.4	1,307,741
1976-82	34,659	69,004	85,279	125,659	37,265	14,168	3.2	366,034
1983-90	34,809	61,830	65,815	72,505	31,202	18,431	2.5	284,592
1991-95	23,044	27,293	54,627	68,952	31,176	14,386	1.9	219,478
1996-06	9,960	24,202	27,104	36,848	18,311	23,305	1.2	139,730
Missing	5,427,174	79,294	235,434	461,206	101,784	4,741	55.1	6,309,633
All	5,895,990	661,802	1,432,633	2,660,114	563,931	225,060	100	11,439,530

Appendix Table 11 - HEED Stock: Dwelling Age and Tenure

	Tenure					
	Missing	Owner	Private Rental	Social Rental/ RSL	All	
Dwelling Age	*count*	*count*	*count*	*count*	*% of stock*	*count*
pre-1900	66,746	244,476	23,543	21,081	3.1	355,846
1900-29	82,913	398,600	39,590	76,023	5.2	597,126
1930-49	111,892	491,700	37,449	182,363	7.2	823,404
1950-66	157,252	541,738	30,594	306,362	9.1	1,035,946
1967-75	170,917	803,051	23,813	309,960	11.4	1,307,741
1976-82	46,948	190,867	10,444	117,775	3.2	366,034
1983-90	37,630	166,599	8,745	71,618	2.5	284,592
1991-95	25,783	143,236	10,541	39,918	1.9	219,478
1996-06	25,527	84,154	3,492	26,557	1.2	139,730
Missing	3,580,916	1,854,678	455,464	418,575	55.2	6,309,633
All	4,306,524	4,919,099	643,675	1,570,232	100	11,439,530

Note that a higher proportion of pre-1900 dwellings fall into the extreme categories of either 1 or 5+ bedrooms. Fewer dwellings post-1900 fall into these categories. Approximately 80% of dwellings in the age bands 1900-1975 are 2 or 3 bedroom; with approximately 70% of dwellings between 1976 and 1995 having 2 to 3 bedrooms (see Appendix Table 12 below).

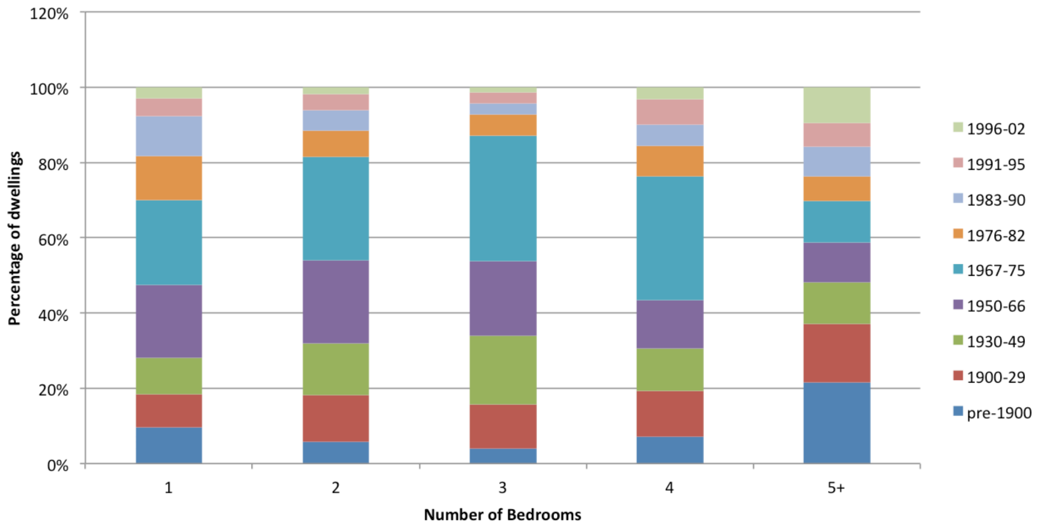
Appendix Table 12 - HEED Stock: Dwelling type and age

	Dwelling Age										All	
	Missing	pre-1900	1900-29	1930-49	1950-66	1967-75	1976-82	1983-90	1991-95	1996-06		
Dwelling Type	*count*	*count*	*count*	*count*	*count*	*count*	*count*	*count*	*count*	*count*	*% of stock*	*count*
Detached bungalow	65,537	1,936	6,748	14,044	40,871	97,717	17,846	14,727	11,755	4,957	2.4	276,138
Semi-detached bungalow	61,018	1,430	6,226	18,177	55,882	83,422	15,850	12,223	6,899	4,209	2.3	265,336
Flat, maisonette	322,984	38,975	44,986	67,226	125,729	144,184	68,982	54,013	27,750	19,832	8	914,661
Detached house	108,890	59,769	50,267	64,795	106,932	203,607	73,651	75,782	62,719	49,776	7.5	856,188
Semi-detached house	290,483	53,912	141,210	372,599	389,328	411,935	76,680	49,709	42,172	26,083	16.2	1,854,111
End-terrace house	58,560	39,371	62,728	60,689	76,677	88,926	27,788	15,928	14,850	7,123	4	452,640
Mid-terrace house	131,118	110,553	225,518	117,325	128,023	174,012	47,576	25,733	34,082	11,151	8.8	1,005,091
Missing	5,271,043	49,900	59,443	108,549	112,504	103,938	37,661	36,477	19,251	16,599	50.8	5,815,365
All	6,309,633	355,846	597,126	823,404	1,035,946	1,307,741	366,034	284,592	219,478	139,730	100	11,439,530

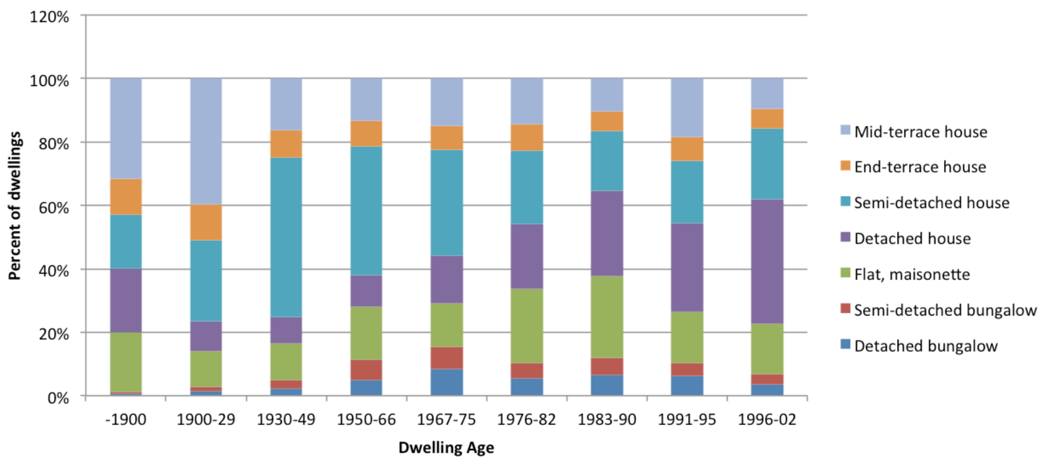
Appendix Table 13 - HEED Stock: Dwelling type and number of bedrooms

	Number of bedrooms						All	
	Missing	1	2	3	4	5+		
Dwelling Type	*count*	*count*	*count*	*count*	*count*	*count*	*% of stock*	*count*
Detached bungalow	223	20,579	94,202	126,422	27,104	7,608	2.4	276,138
Semi-detached bungalow	3,132	60,416	140,001	56,304	2,944	2,539	2.3	265,336
Flat, maisonette	283,166	321,054	243,735	57,708	5,869	3,129	8	914,661
Detached house	94,083	25,971	77,799	287,579	255,779	114,977	7.5	856,188
Semi-detached house	151,701	70,133	312,707	1,138,171	137,045	44,354	16.2	1,854,111
End-terrace house	40,914	28,730	120,671	233,547	19,723	9,055	4	452,640
Mid-terrace house	85,612	55,547	286,091	518,603	41,623	17,615	8.8	1,005,091
Missing	5,237,159	79,372	157,427	241,780	73,844	25,783	50.8	5,815,365
All	5,895,990	661,802	1,432,633	2,660,114	563,931	225,060	100	11,439,530

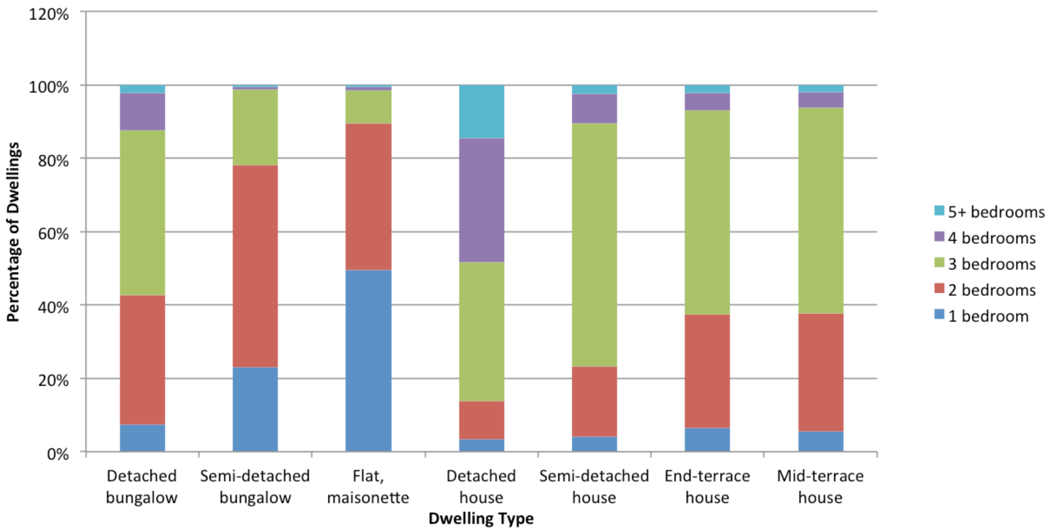
The tables above and the figures below provide a cross-tabulation between HEED dwelling type and age, which shows that pre 1930's dwellings are predominantly mid-terrace, with semi-detached housing becoming more visible between 1930 to 1982 after which time detached and semi-detached forms make up the higher proportion of dwelling types.



Appendix Figure 5 – HEED: Number of bedrooms by dwelling age



Appendix Figure 6 - HEED: Dwelling age by dwelling type



Appendix Figure 7 - HEED: Dwelling type by number of bedrooms

13.7 Summary

HEED contains information on over 50% of dwellings in the UK. The results of the housing stock population comparisons for the English and Welsh sample of HEED and England and Wales housing stock datasets suggest that the dwellings in HEED are not strictly statistically representative; but note that this is unlikely given the large sample size.

There are limitations to the HEED dataset that need to be considered in any analysis, for example: the collection methods (i.e. different surveys using different forms), issues of self-selection for surveys, and misclassification or assessor bias. Also, a dwelling will enter HEED as a 'snapshot', which means that the energy efficiency characteristics recorded for the dwelling will be more or less correct at a particular date.

The Homes Energy Efficiency Database is an example of what can be characterised as 'in action' data. HEED is not the product of a large omnibus survey or a concerted monitoring and reporting exercise; instead HEED is the product (and by-product) of a range of disparate activities that are centred on home energy efficiency. HEED offers a repository and framework for these sources, one that is clearly flexible to a range of data types and quality.

Chapter 14 Appendix B

UK gas and electricity meter supplier data

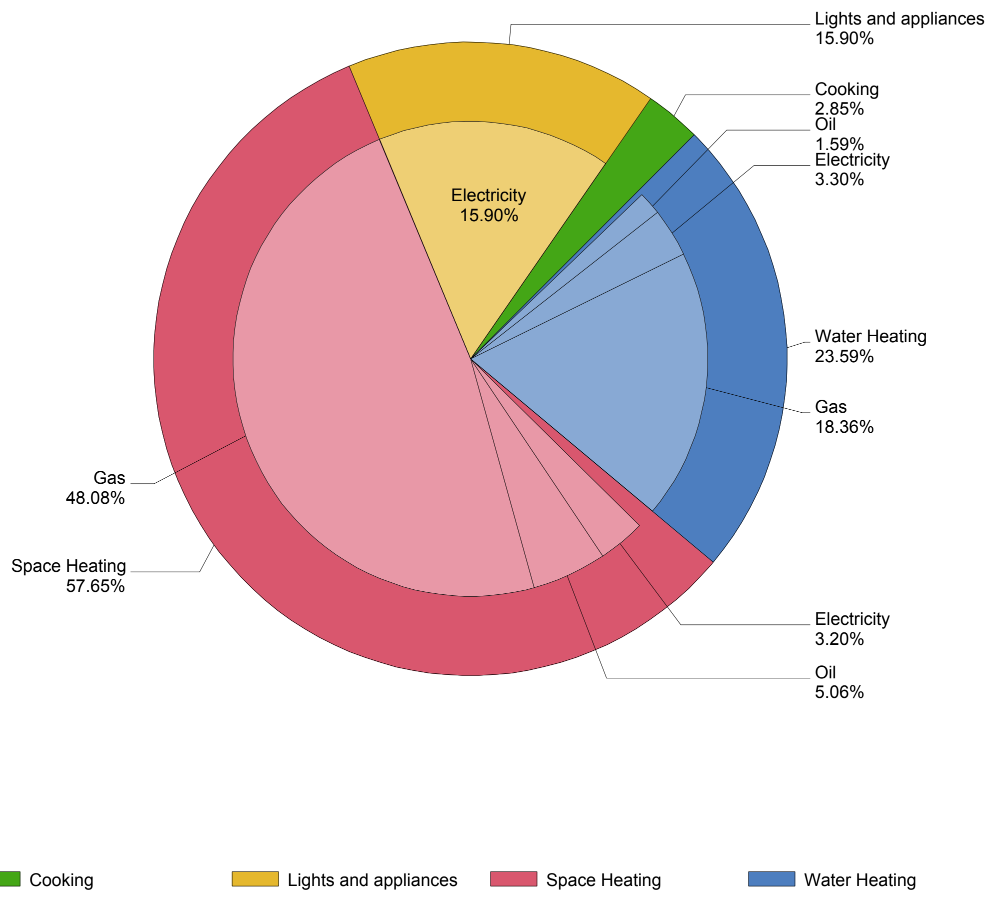
Appendix introduction

This appendix provides further in-depth details on the UK gas and electricity meter supplier data made available by the Department of Energy and Climate Change (DECC) as part of the EPSRC funded *Buildings and Energy Data Framework* (EP/H021957/1) and *New Empirically-Based Models on Energy Demand in Buildings* (EP/I038810/1). Portions of this text are featured in the BEDF project report.

14.1 Meter point energy consumption data

14.1.1 UK level gas and electricity demand

In 2010, domestic (i.e. residential) delivered energy accounted for approximately 33% (490 TWh) of total GB energy demand by final consumption, of which gas and electricity accounted for approximately 70% (344 TWh) and 23% (113 TWh) respectively (DECC, 2013c). Appendix Figure 8 shows an estimate of the total residential demand by service type and fuel[14] (DECC, 2010b). The majority of residential energy demand is for space and hot water heating (78%) with the remainder for appliances (16%) and cooking (3%). Electricity and gas annualised meter data was used in this thesis. These fuels account for over 90% of total UK domestic energy demand by final consumption, for the period 2004 to 2008.



Appendix Figure 8 – UK domestic energy demand by fuel and end-use

[14] Residential energy demand by service type is estimated from DUKES data, national totals, and Domestic Energy Fact File data, service fractions. Renewable energy is not included. Services of Fuels <1% of total are not shown but are accounted for in the total.

14.1.2 Meter point gas and electricity meter classification

The UK government, currently via DECC, collect annualised final consumption gas and electricity data for individual meter points from energy suppliers for the purpose of various statistical outputs; in 2008 there were approximately 22.6 million gas meters (22.3million domestic and 0.3 million non-domestic) and 29.1 million electricity meters (26.7 million domestic meters and 2.4 million non-domestic meters) (DECC, 2009).

UK property-level gas and electricity meters are classified into two types: daily metered (gas) or half-hourly (electricity) metered, and non-daily metered (gas) or non-half hourly (electricity). Daily metered[15] gas meters are typically for industrial users with a demand of >2.2GWh/year and may be interruptible. Half-hourly metered electricity meters are also typically for large industrial users[16] and are required for all users with a maximum demand >100kW. Non-daily and non-half hourly meters represent all other meters and are typically smaller users. Consumption figures are collected from meter readings or estimated from historic customer demand patterns. These users are characterised by annualised energy demand values (kWh's). Annualised data for all gas and electricity meter points were accessed for this thesis, under strict confidentiality agreements with the suppliers.

The above meter types do not account for all gas and electricity use in the UK. Within the data certain meters were not covered. These are: electricity users known as 'Central Volume Allocation' who interact directly with the high voltage mains (i.e. >275kV); electricity users who do not draw directly from the network, i.e. private distribution or 'private wire'; large gas users who draw directly from the national transmission system; and any gas not passing through the National Grid owned network (DECC, 2008b, 2008c). In addition, Northern Ireland (which accounts for about 3% of the UK population) was not covered.

The annualised gas and electricity meter point values were derived from individual meter readings, via gas and electricity suppliers and their collecting agents. The energy suppliers convert the meter readings into annual consumption values using a common methodology that was developed and agreed upon with the Office of the Gas and Electricity Markets (OFGEM). Gas non-daily meters (NDMs) are divided into three categories based on their total annual demand: domestic, small NDMs and large NDMs. Electricity non-half hourly (NHH) meters are grouped into 8 classes, representing probable demand profiles (see Appendix Table 14 below). Gas demand values (NDMs and DMs) are annualised for a standard weather year and normalised to represent the appropriate end-user consumption patterns for each meter for a year. Electricity demand values (NHH and HH meters) are annualised using profile class information. The annualisation process is described for each fuel meter type in further detail below. The electricity data annual period is from 30 January

[15] In 2005 there were approximately 2,500 daily metered users

[16] There were approximately 85,000 half-hourly metered users in 2005

to 29 January. The gas data annual period is 1 October to 30 September and represents a full heating season[17].

The meter data provided from the suppliers contained no features of the households or buildings, only the unique meter number and the annualised gas and electricity demand value. DECC used the address associated with the electricity Meter Point Administration Number (MPAN) and gas Meter Point Reference Number (MPRN) to link the gas and electricity meters to the Homes Energy Efficiency Database (HEED) using Ordnance Survey AddressPoint (OSAPR) (EST, 2009). The gas and electricity demand data for each meter point was provided in an anonymised format (i.e. without an address or an OSAPR) containing only the unique meter point number and the energy quantity used for each energy year, i.e. 2004 to 2008. A linking file was provide that contained a unique home identifier for each HEED dwelling and its MPAN and MPRN. The gas and electricity meter data and HEED were provided in separate files and were subsequently joined for the study using the linking file. Although the gas and electricity data has been provided publicly at an aggregated census level (i.e. MLSOA and LSOA) since 2004, the work by DECC and EST was the first time such data had been linked to built form features and analysed at an anonymised individual meter point level. Since then, DECC have completed their initial NEED analysis report (DECC, 2011a).

14.1.3 Meter point gas and electricity collection methodology

14.1.3.1 Electricity Meters

Annualised consumption data for each Meter Point Administration Number (MPAN), or electricity meter, is provided to DECC by data aggregators, agents of the electricity suppliers who collate/aggregate electricity consumption levels for each meter, and from Gemserv, who provide a central access point for electricity suppliers, distributors and their agents to obtain addresses and postcode information about each meter (DECC, 2008b). DECC receives address point data for each meter in order to match each MPAN to census geographies (e.g. local authority or LLSOA) (ONS, n.d.) using the National Statistics Postcode Directory and the Postal Address File (PAF). Individual meter addresses were not provided for this study. Instead, in order to identify the electricity meter points geographically a postcode sector and lower super output area (LLSOA) code were provided.

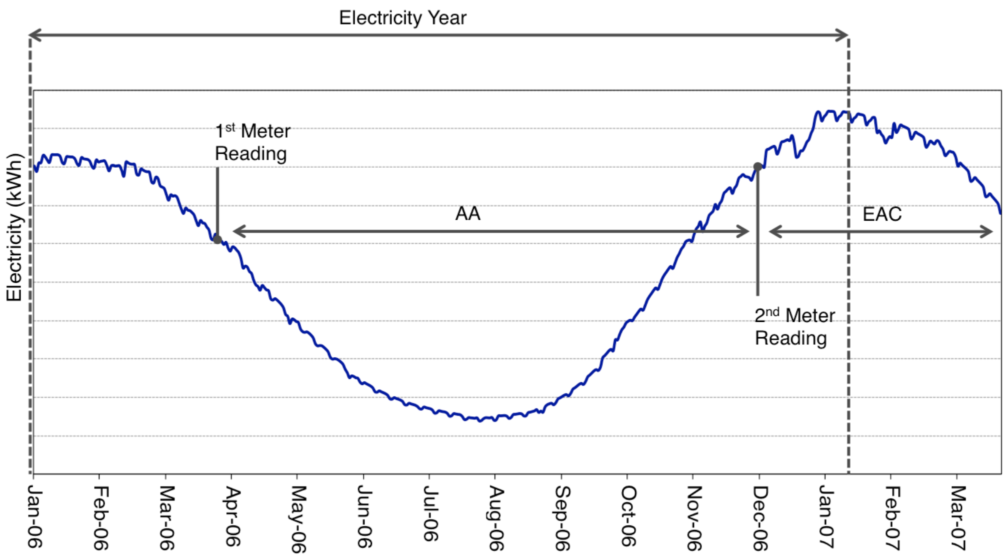
Annualised electricity meter data are estimates based on an annual advance (AA) or an estimated annual consumption (EAC) for all NHH meters. An annual advance is calculated by applying the sum of daily profile coefficients[18] for the meter year. The effect of this is to scale the meter advance up or down to derive an annualised value. An EAC is calculated using the AA and the previous EAC to determine an estimate of consumption for a period not covered within the meter reading. Where no previous EAC exists, a nominal EAC is

---

[17] Gas AQ data is not directly comparable to the national UK statistics (i.e. *Digest of United Kingdom Energy Statistics*) as those represent a calendar year. In addition DUKES represents all gas flow within the gas national transmission system, i.e. gas users outside the distribution network [9].

[18] Daily profile coefficients are presented as a fraction of the yearly consumption and are used to estimate what a meter would have consumed for any given half-hour for a year based on the profile class and grid supply point (GSP) group, i.e. groups segmented by total annual electricity demand. There are 12 GSPs in Profile Class 1 (Domestic Ordinary meters) for England and Wales [12].

provided based on the meter profile's class average. An EAC is used where two meter readings are not available, and an estimate of annualised consumption is produced by the energy company using historical information and the profile information relating to the meter. In both cases the meter profile class determines how the estimates are derived (see Appendix Figure 9). The AA and EAC are summed for the relevant period to create a total annual demand (measured in kWh). The AA and EAC consumption data for each MPAN is not weather-corrected and represents 365 days (DECC, 2008c).



Appendix Figure 9 - Example of Annualised Advance (AA) and Estimated Annual Consumption (EAC)

The meters are allocated to one of eight profile classes (see Appendix Table 14) each representing different user types. Domestic electricity meters are 'ordinary' or 'Economy7', listed as profiles 1 and 2 respectively. Economy7 refers to meters that are on a time charge tariff that offers cheaper electricity during off-peak hours, typically an 8 hour period, and are either time or radio switched; in houses, these meters are most often associated with electric heating, either space heating (e.g. storage heaters) or hot water, offering the customer the advantage of off-peak electricity stored as heat for daytime use, thus lowering fuel bills. 'Ordinary' meters are all other types of meters; these meters may be used for heating but are not on an economy tariff.

Appendix Table 14 - Electricity meter profile types

Electricity Profile Class	*Representative Group*
Profile Class 1	Domestic Unrestricted 'Ordinary' Customers
Profile Class 2	Domestic Economy 7 Customers
Profile Class 3	Non-Domestic Unrestricted Customers
Profile Class 4	Non-Domestic Economy 7 Customers
Profile Class 5	Non-Domestic Maximum Demand (MD) Customers with a Peak Load Factor (LF) of less than 20%
Profile Class 6	Non-Domestic Maximum Demand Customers with a Peak Load Factor between 20% and 30%
Profile Class 7	Non-Domestic Maximum Demand Customers with a Peak Load Factor between 30% and 40%
Profile Class 8	Non-Domestic Maximum Demand Customers with a Peak Load Factor over 40%

Although it is likely that every dwelling will have an electricity meter, the number of MPANs will be greater than the number of households as a result of multiple meters within households from two and three-rate meters (e.g. Economy7) along with second residences and additions to the stock. In England and Wales the number of meters exceed households by approximately 9%. In Scotland this is around 18% due to the high number of multi-rate meters (DECC, 2009).

DECC undertake an annual validation process to check the number of electricity meters against the number of homes within the geographic areas; it is acknowledged that such a comparison will not necessarily be an exact match as there may be households with multiple meters. In addition, DECC advises that domestic electricity meters should not record consumption of 100,000 kWh/year or above, and special attention is paid to meters showing consumption in excess of 50,000 kWh/year. In the aggregated publicly available statistics, all profile 1 and 2 customers with a recorded consumption greater than 100,000kWh/year and those users >50,000kWh/year identified as non-domestic (using information from the linked address data, such as business names) are reclassified as industrial and commercial customers and removed from the housing data.

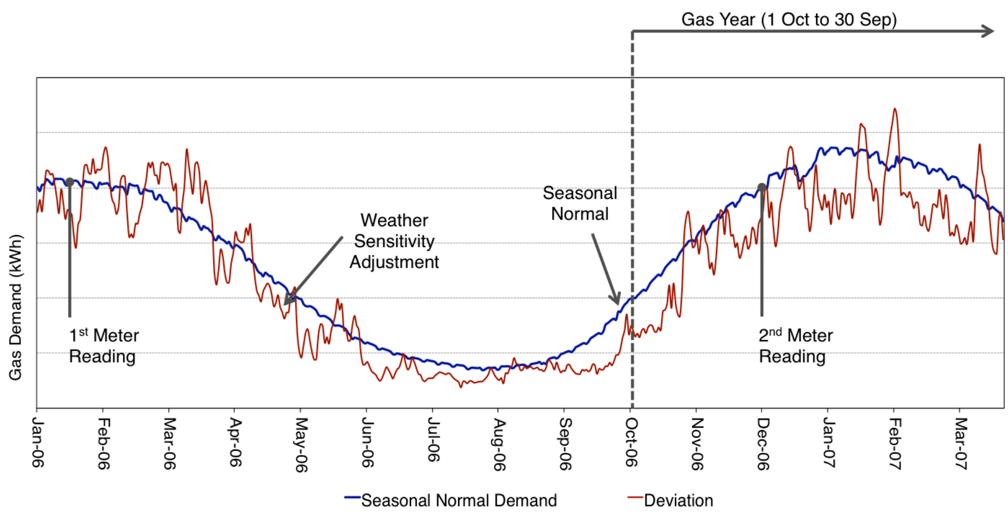
In this study, a similar validation and flagging process was undertaken (described below) in order to identify all 'non-irregular' domestic electricity meters, i.e. those meters that are most likely to correspond to normal or typical customers. The aim of the research was to analyse individual users and populations rather than total aggregate energy demand, therefore we did not attempt to reconcile the total electricity demand by user profiles to the DECC sub-regional statistics.

14.1.3.2 Gas Meters

Annualised quantity (AQ) data for each meter point reference number (MPRN) or gas meter, for all non-daily meters (NDMs) and daily meters (DMs), is collected by DECC from gas suppliers via XoServe and independent gas networks, along with information on the location of the meters (DECC, 2008c). DECC allocate gas meters to various census unit levels using the National Statistics Postcode Directory. As with the electricity data, individual address information was not provided as part of the gas data for this study, but rather LLSOA codes.

The AQ is an estimate of the annualised gas consumption using two meter readings at least six months apart. Gas demand is subject to a seasonal and an end-user climate sensitivity adjustment. The purpose of the corrections is to allow for inter-year comparisons that are independent of weather; in other words, raw metered gas demand from cold or warm years is corrected to a seasonal normal demand to simplify comparison across different years. The methodology, described in some detail below, can be found in the UNC AQ Procedure Document produced by Gemserv (Gemserv, 2007).

An AQ demand model is required for each end user category (EUC) within a gas local distribution zone (LDZ). The demand model consists of an overall seasonal weather correction that represents the past 17-year conditions to adjust the demand to a seasonal normal demand curve. The demand is then further adjusted for each day in the case period according to: an annual load profile, which determines the seasonal normal demand; a daily adjustment factor, which determines the impact of weather sensitivity for the user; and an LDZ weather correction factor. The AQ for each MPRN is an annualised value representing the period between 1 October and 30 September (see Appendix Figure 10).



Appendix Figure 10 - Annualised quantity (AQ) example

The weather and climate corrections have the overall effect of allocating gas demand to the colder parts of the collected period or gas year, accounting for the likely sensitivity a dwelling may have to the weather experienced. Where a reading does not cover the collection period, the annualisation process uses the seasonal normal demand and weather sensitivity factors for the appropriate end-user category to 'fill in the blanks'. Unlike the electricity meters, the AQ does not differentiate between the portions of gas use which are directly metered and estimated. However, for both electricity and gas meters no distinction is made in the data used.

The National Grid apply the weather correction to the data prior to it being supplied to XoServe and DECC. The process of weather correction is therefore impossible to reverse based on information currently in the public domain and introduces unknown effects into the available AQ records for each individual house that cannot be interrogated. The different treatment of gas and electricity use means that only part of the energy used to heat UK dwellings is weather-corrected. Given the likelihood that the UK climate will warm significantly over the coming century (UKCIP), and that the weather sensitivity of dwelling energy use will change systematically over coming decades due to changes in mix of dwelling types, insulation levels, the electrification of heat supply and the amount of energy used for non-heating purposes, the process as it currently stands will need to be revised.

Although each MPRN is assigned to an end-user category, the AQ values provided with the MPRNs by the suppliers do not bring forward this categorisation, unlike electricity meters. There are approximately 594 EUC codes used for calculating the AQs in Great Britain, i.e. 18 local distribution zones (LDZs) of gas by 33 user categories for each zone. As a result of this crucial detail missing from the gas data, user groups cannot be identified (e.g. domestic ordinary); instead, the general approach of the gas industry is to allocate all meters with an annual consumption of less than 73,200 kWh to the domestic sector and those above to the commercial/ industrial sector. This figure, which has been inherited from the days before the privatisation of British Gas (it is the metric equivalent of 2500 Therms) is roughly 5 standard deviations above mean gas consumption. The gas data cannot be exactly aligned to data in the *Digest of UK Energy Statistics* (DUKES), which are for calendar years, and are not weather-corrected (DECC, 2009). It should be noted that efforts were made to contact those responsible for setting and implementing the weather correction process, but no details of the relevant years correction factors were made available. Therefore, the following analysis uses the corrected value, with the basic understanding that all gas meter values have been subject to a weather correction.

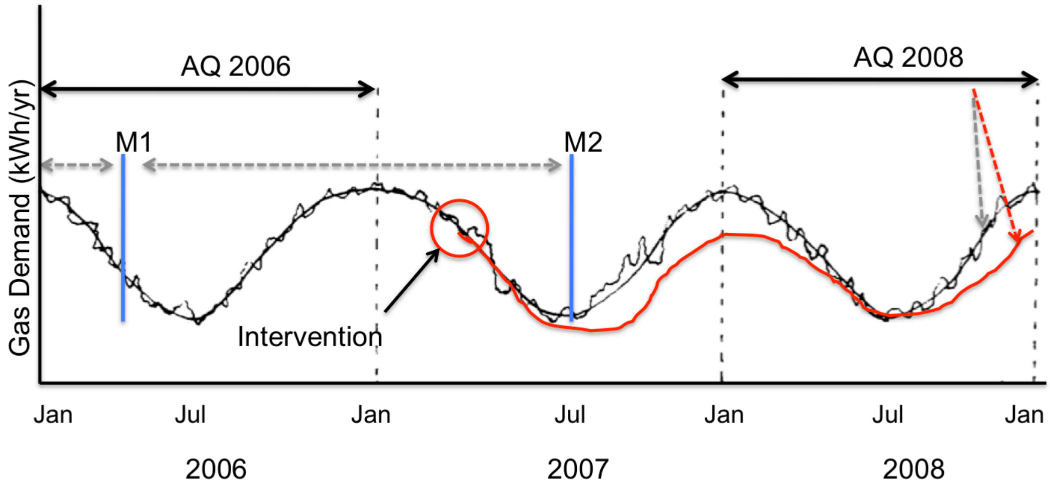
14.1.4 AQ weather correction and energy efficiency interventions

As described above, the gas AQ data is subject to a weather correction but more importantly, may not contain an actual meter reading within any given year. According to the UNC Network Code, a meter must be read at a minimum once within any two-year period. In terms of what this might mean for assessing the impact of energy efficiency interventions through the detection of changes in energy demand between years, it may be that long-term trends are more significant than year-on-year changes. There is a risk that the methodology used to develop the annualised quantity (AQ) may not pick up a significant change in energy use for a given dwelling for two or more years after the change.

In theory, the AQ should have had at least two meter readings within a year that could be used in the AQ calculation and would show any significant changes in demand. However, no information exists within the meter point data to determine how many meter readings occur within the energy year.

Appendix Figure 11 illustrates the potential implications; the two blue actual meter readings occur a year apart in 2006 and 2007 and the red line denotes a change in actual demand due to an intervention. However, the nearest meter reading for AQ 2007 may not occur until 2008. Thus it is possible that an actual reduction in energy consumption may not become apparent for this individual meter point through the AQ data for as long as two or three years.

The analysis subsequently shown in this thesis does not attempt to correct for any effect that the annualisation process may have on gas or electricity meter point data. Therefore it is unknown what variability in gas or electricity demand should be assigned to the annualisation and the weatherization factoring of the meter point data.



Appendix Figure 11 – AQ energy efficiency intervention (black line denotes weather and normal profile; red line denotes change due to intervention; M1 and M2 denote meter reading points)

14.2 Data Flagging

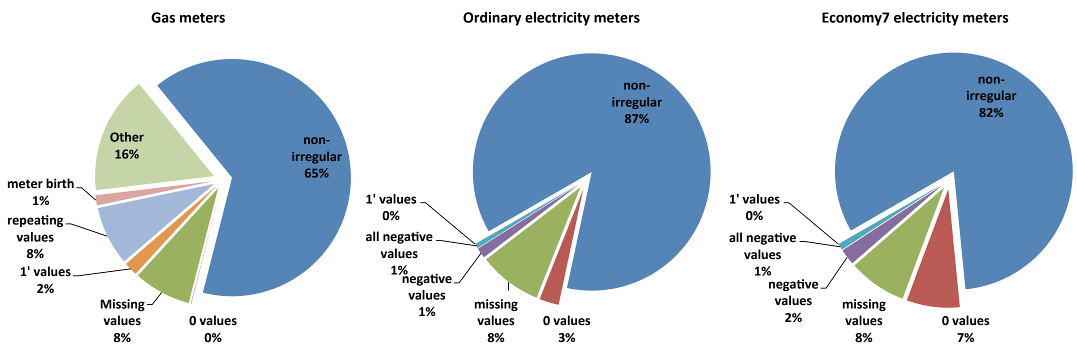
A data flagging exercise was undertaken in order to further make use of the linked electricity and gas data by identifying potentially erroneous, inconsistent, and missing data. The flagging created a consistent dataset that was suitable for analysis by identifying 'non-irregular' gas values. The term 'non-irregular' was employed here as a way of describing data that was not necessarily normal, but rather was not associated with any features that might render the data unreliable when describing gas demand for individual households. Note that the analysis was concerned with meter and dwelling populations and not with reconciling the meter data with DECC statistics.

Two types of flag were attached to the data: intra-year or within year, and inter-year or between years. Within the flags were inclusive and exclusive flagging statements, or Type1 and Type2 flags respectively. Type1 flags were used to define the data values that would not be included in the dataset used in the analysis (subject to Type2 flag), whereas Type2 flags were used to identify *potential* errors or explanatory features that were not necessarily wrong (i.e. a step change in energy demand between two years) and included for subsequent analysis. Appendix Table 15 provides a summary of the various gas and electricity flags. From this flagging process, a processed gas and electricity dataset was prepared for analysis.

Appendix Table 15 - Annualised energy meter value data flags

Domestic Gas and electricity flags	*flag purpose*
Within-year	
'missing' values [type1]	blank or missing values
'0' values [type1]	'0' values or non-active meters
negative values [type1]	meters with negative values (possible export)
'1' values [type1]	meters with '1' values
'dummy values' [type1]	known dummy values used by gas
>3stdev outliers [type1]	aggregators[a]
>50,000kWh (electricity and gas) [type2]	flag larger/extreme users
	flag larger users
Across-Years	
Meter 'birth' [type2]	no preceding meter readings
Meter 'death' [type2]	no following meter readings
all repeated value [type1]	all meters readings are the same
1 repeated value [type2]	Two meter readings are the same
>50% change (upward or downward) [type2]	step change in meter reading value

Notes: [a] Dummy or default values are used by the energy industry where meter readings are not available (OFGEM, 2009)

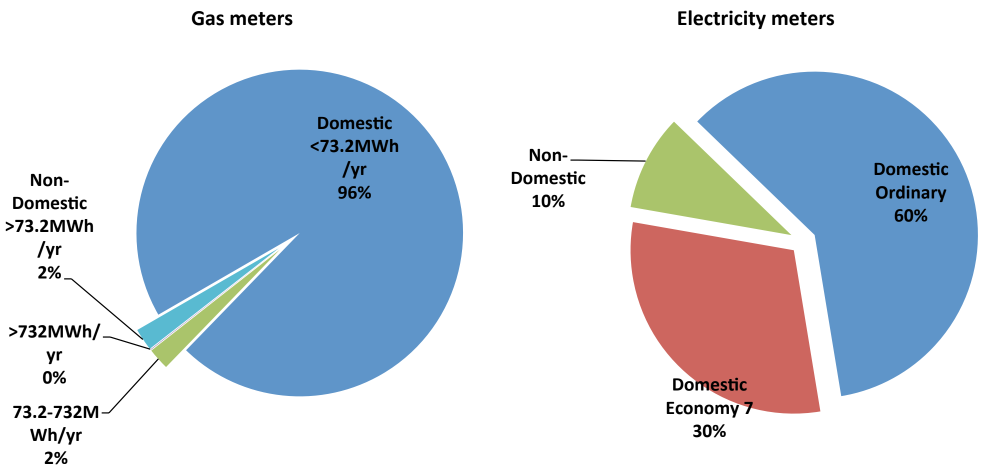


Appendix Figure 12 – Gas and electricity meter point flagging

In addition to the above, since no user information was associated with the gas AQ data, a profile variable was developed in order to identify meters that were likely to be domestic. This involved identifying three user profiles for gas meters (i.e. <73.2MWh/year, between 73.2MWh/year and 732MWh/year, and >732MWh/year). These broad divisions will result in some large domestic users being wrongly allocated to other user categories, and conversely small non-domestic users being included in data for dwellings. In contrast, DECC energy statistics contain address data that allows these users to be allocated to the appropriate user class. When the gas and electricity data are joined to HEED (described below) large domestic users are more readily identified and reallocated to a domestic profile class, which was included as part of the linking of the HEED and energy datasets carried out by EST. Note that >73 MWh/year may still be a domestic user, e.g. a block of flats with a single meter. These large users affect the mean and are flagged and excluded when generating per meter level statistics.

Also, profile class 2 electricity meter data (i.e. economy tariffs) has two records for each meter by time pattern, one being 'off-peak' and the other 'normal tariff' electricity use.

Although a time pattern code was included with the data, no associated key to decode the time period was available for this study. Therefore, the class 2 data were summed together to create a single stream of electricity data per household.



Appendix Figure 13 - Gas and electricity meter point classes

14.3 UK Energy Statistics

Gas and electricity meter data was available for the whole of Great Britain and therefore it was possible to compare consumption for dwellings in HEED with dwellings not in HEED. The gas and electricity supplier data covered the period 2004 to 2007, which coincided with the majority of HEED data collection; approximately 60% of all the dwellings information in HEED was collected over that period (see Appendix Table 16).

Although HEED covers over 50% of the stock, it is reasonable to imagine that biases might exist in parts of HEED as a result of the targeting of specific programmes. However, it is possible to see only small differences in demand between dwellings whose meters are linked to HEED and those whose are not. Although the vast majority of those dwellings in HEED have been the subject of energy interventions, home efficiency improvements are also taking place outside government programmes.

Using the flagged gas and electricity data set, a set of statistics was developed that describe, in a basic manner, the gas and electricity use by meter within the UK stock.

Note that the annualisation process of the energy meter data means that a change in energy demand will depend on the frequency of meter readings within a year. This means that measures installed in later gas years may not be fully reflected for those meters that have not had twice-yearly readings or where the measure occurs after a reading for the given gas year. Therefore this thesis looked at the change in demand between 2005 and 2007 for those dwellings that have received a measure in 2006 or before.

Appendix Table 16 - HEED stock data sources for energy data period

	Period				
HEED data source	All	2004	2005	2006	2007
Energy suppliers	18.5%	19.3%	25.9%	15.9%	16.4%
Government Schemes	17.5%	27.7%	17.9%	2.8%	2.8%
Installers	42.0%	36.9%	43.0%	51.6%	61.2%
Survey	22.1%	16.1%	16.1%	29.7%	19.6%
N	11,439,530	2,204,915	3,178,845	3,543,226	3,121,310

14.4 Gas meter point energy demand

14.4.1 Gas Statistics

Summary statistics for gas meters that use less than 73.2MWh/year from the processed dataset are provided in Appendix Table 17. The meters are further classified into i) all meters in HEED, and ii) those meters that entered HEED in the relevant gas year. The table offers, firstly, a broad comparison between those meters whose house details have been collected into HEED compared to those who have not. By showing the gas demand for these groups, it is possible to compare the direction and magnitude of change across the energy period.

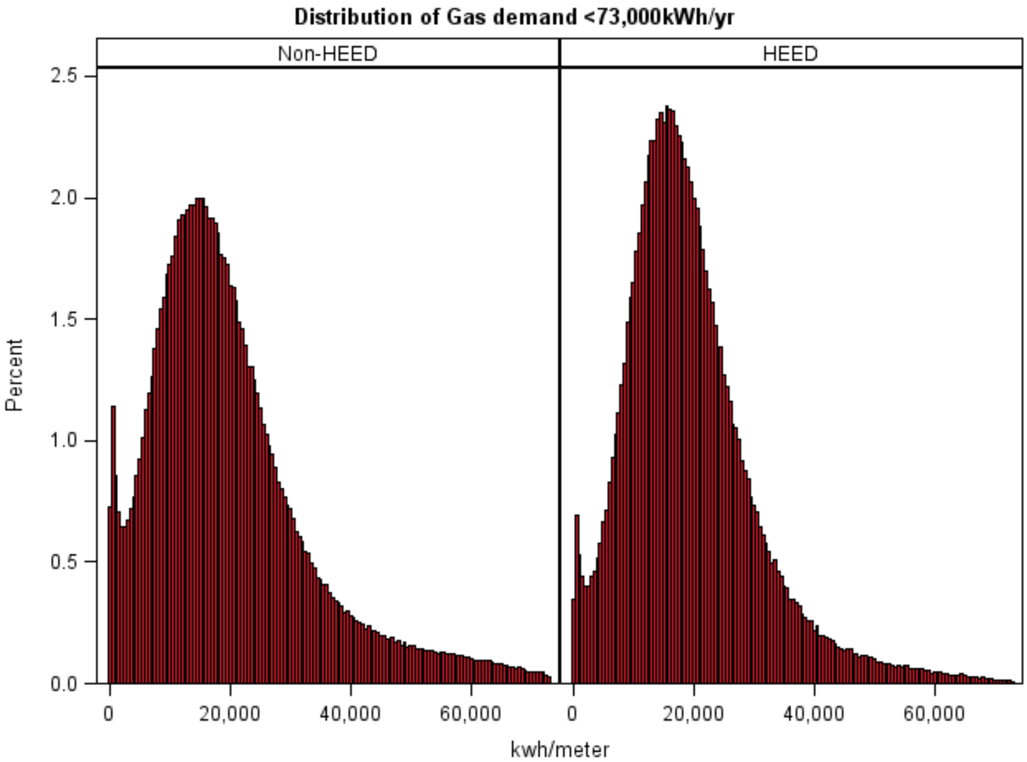
Appendix Table 17 - Domestic gas meter demand by year

Profile	HEED N	Year	Max[a]	Mean	Med-ian	Std Dev	Std Error	Upper 95 % CL for Mean	Lower 95 % CL for Mean
<73.2 MWh/year	Non-HEED	2004	73,200	19,734	18,214	11,137	3.8	19,741	19,726
	8,410,189	2005	73,200	19,433	17,877	11,008	3.8	19,440	19,425
		2006	73,198	18,625	17,107	10,836	3.7	18,633	18,618
		2007[b]	.			.	.	.	.
	All HEED	2004	73,200	19,623	18,452	9,725	3.6	19,630	19,616
	7,450,540	2005	73,198	19,141	17,926	9,511	3.5	19,148	19,135
		2006	73,199	18,153	16,958	9,252	3.4	18,159	18,146
		2007	73,199	17,468	16,226	9,086	3.3	17,475	17,462
	In HEED 2004	2004	73,197	19,351	18,221	9,609	10.3	19,371	19,331
	876,133	2005	73,193	18,805	17,595	9,348	10.0	18,824	18,785
		2006	73,187	18,044	16,875	9,192	9.8	18,063	18,025
		2007	73,194	17,523	16,318	9,136	9.8	17,542	17,504
	In HEED 2005	2004	73,200	19,755	18,597	9,822	8.0	19,770	19,739
	1,495,272	2005	73,196	19,030	17,760	9,442	7.7	19,045	19,015
		2006	73,199	18,040	16,757	9,190	7.5	18,055	18,025
		2007	73,199	17,548	16,229	9,164	7.5	17,563	17,534
	In HEED 2006	2004	73,199	19,577	18,430	9,988	7.6	19,592	19,562
	1,705,025	2005	73,198	19,069	17,909	9,805	7.5	19,084	19,054
		2006	73,192	17,746	16,545	9,256	7.1	17,759	17,732
		2007	73,188	17,005	15,736	9,020	6.9	17,018	16,991
	In HEED 2007	2004	73,200	20,187	18,975	10,145	8.1	20,203	20,171
	1,584,058	2005	73,184	19,811	18,564	9,992	7.9	19,827	19,796
		2006	73,196	18,811	17,632	9,784	7.8	18,826	18,796
		2007	73,195	17,797	16,493	9,378	7.5	17,811	17,782

[a]Excludes all Type 2 (i.e. exclusive) flags. [b]Non-HEED 2007 – Gas meter values were only provided for those homes matched in HEED, therefore no statistics are available for this year from the processed data.

The change in median gas demand in 'non-HEED' meters between 2004 and 2006 is approximately -6.1%, with an average year-on-year change of -3.1% between the medians. For meters in HEED, the change in median gas demand between 2004 and 2006 is approximately -8.1%, with a year-on-year change of -4.2% between the medians. The AQ methodology, described above, may mean that the gas demand meter values for later years are more likely showing the lagged effect of efficiency interventions (as a result of the frequency of meter readings). As subsequent years of energy data are made available the rate of change in metered gas demand may more truly reflect the impact of the interventions that took place in the period 2004-2006.

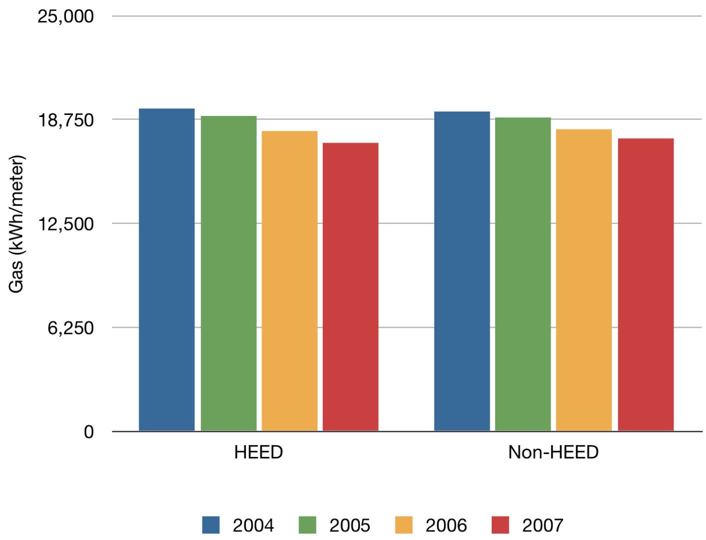
The data shown above suggests that domestic gas demand data is influenced by a long right tail, as can be seen in the <73.2MWh/year meters gas demand (Appendix Figure 14). This is an inevitable consequence of the fact that energy demand data cannot be negative, but is subject to no well-defined upper limit (other than the very high 73.2 MWh limit). Note also the upward flick in the distribution close to zero demand. This may be caused by dwellings that are unoccupied for part or all of a year.



Appendix Figure 14 - Domestic gas distribution

14.4.2 Change in Gas Demand

The reduction in gas demand in HEED dwellings over the period 2004 to 2006 was higher than in those that were not part of the HEED data set (identified via unmatched gas MPRNs). It is not clear however what rate or level of energy efficiency is present in non-HEED homes.



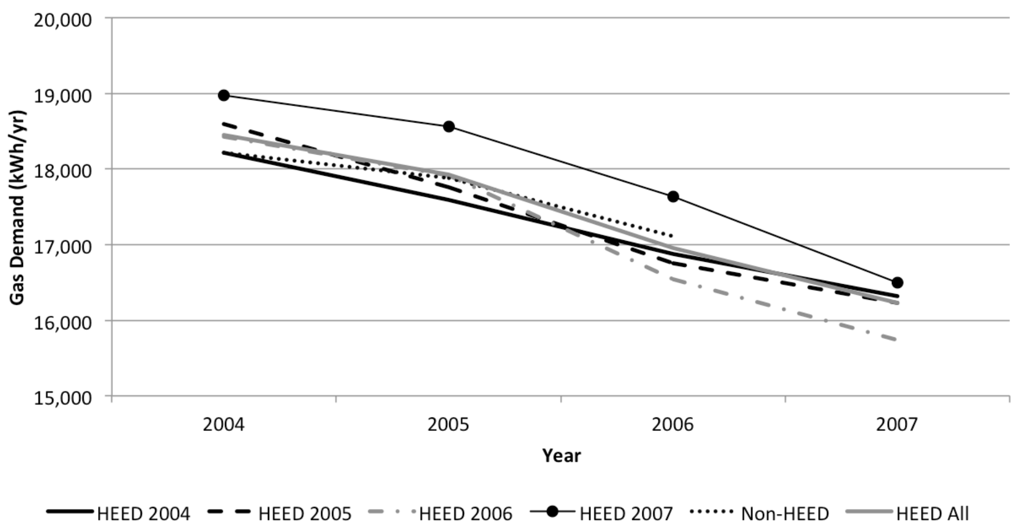
Percent difference between years

Period	HEED	Non-HEED
2004 to 05	-2.3%	-1.9%
2005 to 06	-4.8%	-3.8%
2006 to 07	-3.9%	-3.0%

Appendix Figure 15 - Median domestic gas demand and change in gas demand 2004 to 2007

14.4.3 Change in Gas Demand for HEED

HEED contains a time stamp for when a measure of survey was carried out for each dwelling. This stamp provides a point in time from which a comparison of pre- and post-intervention or survey can be made. In Appendix Table 16 above, meters were classified by the date the home details entered to HEED. For example, a dwelling enters HEED due to an intervention taking place in 2006 but is also then connected to the preceding two years of demand (i.e. 2004 and 2005) and the subsequent gas year (2007). It is thus possible to compare those groups of dwellings across the gas demand period available by when they entered HEED, and therefore were likely to have received an efficiency intervention, to the non-HEED dwellings. Appendix Figure 16 shows this comparison. The figure shows that non-HEED dwellings with an intervention in 2004 used less gas than the non-HEED dwellings during the gas period 2004 to 2005. Homes with an intervention in 2005 begin to use less than their no-intervention counterparts in the following year. This is also true for dwellings with interventions in 2006. It appears that preceding an intervention date the following year's gas demand appears to be lower than both the non-HEED group and the no-interventions group. Appendix Table 18 shows the change from 2004 to 2006 for each of the respective groups. The change in demand is higher for those dwellings with an intervention within the gas period, with the exception of those entering in 2007, where it is unlikely the gas data would pick up in the change, depending on the reading frequency.



Appendix Figure 16 - Gas demand by HEED entry year

Appendix Table 18 - Change in gas demand 2004 to 2006 by HEED group

Group	Period	Mean	Median
Non-HEED	2004 to 2006	-5.6%	-6.1%
All HEED	2004 to 2006	-7.5%	-8.1%
In HEED 2004	2004 to 2006	-6.8%	-7.4%
In HEED 2005	2004 to 2006	-8.7%	-9.9%
In HEED 2006	2004 to 2006	-9.4%	-10.2%
In HEED 2007	2004 to 2006	-6.8%	-7.1%

14.5 Electricity meter point energy demand

Overview statistics for the electricity meters, by profile class, from the processed dataset are provided in Appendix Table 19. As with gas demand, different HEED groups are provided for comparison, i.e. those meters not in HEED, all HEED meters and those meters that entered into HEED by electricity year.

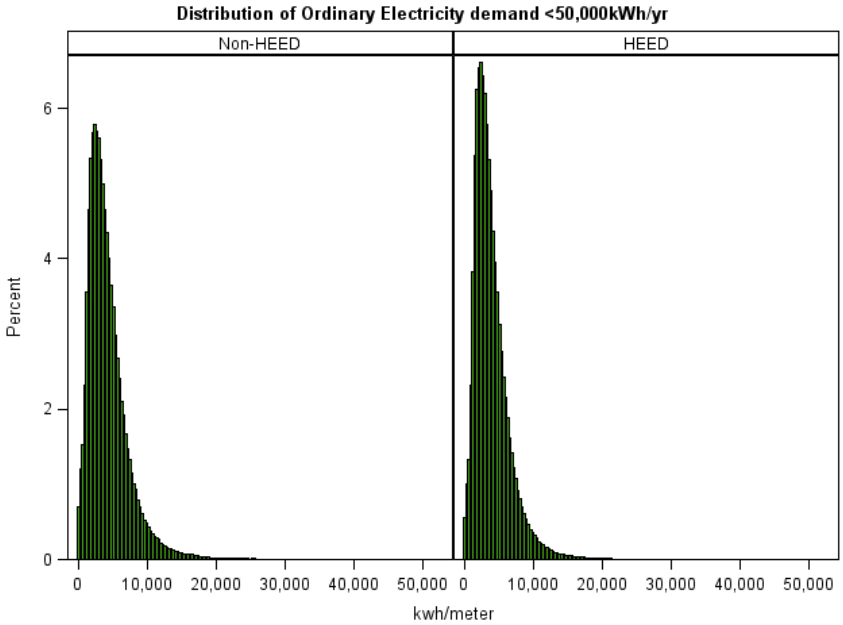
Appendix Table 19 - Domestic electricity meter demand by year

Profile	Flag HEED	Year	Max[a]	Mean	Med-ian	Std Dev	Std Error	Upper 95 % CL for Mean	Lower 95 % CL for Mean
Ordinary	Non-HEED	2004	49,991	4,272	3,548	3,304	1	4,274	4,270
	9,212,105	2005	49,994	4,311	3,551	3,359	1	4,313	4,309
		2006	49,999	4,231	3,519	3,233	1	4,233	4,229
		2007	49,994	4,163	3,447	3,230	1	4,165	4,161
	HEED	2004	49,968	4,023	3,410	2,865	1	4,025	4,021
	7,362,544	2005	50,000	4,027	3,391	2,894	1	4,029	4,025
		2006	49,963	3,957	3,359	2,790	1	3,959	3,955
		2007	49,995	3,888	3,288	2,770	1	3,890	3,886
	HEED 2004	2004	49,761	4,117	3,502	2,891	3	4,123	4,111
	833,854	2005	49,953	4,119	3,477	2,909	3	4,125	4,112
		2006	49,764	4,051	3,446	2,826	3	4,057	4,045
		2007	49,775	3,974	3,370	2,808	3	3,980	3,968
	HEED 2005	2004	49,964	4,041	3,415	2,889	2	4,046	4,036
	1,374,163	2005	49,982	4,037	3,386	2,888	2	4,042	4,032
		2006	49,963	3,973	3,359	2,796	2	3,978	3,968
		2007	49,988	3,897	3,283	2,783	2	3,902	3,892
	HEED 2006	2004	49,947	3,944	3,327	2,874	2	3,949	3,940
	1,650,374	2005	49,989	3,950	3,311	2,912	2	3,954	3,946
		2006	49,963	3,880	3,281	2,776	2	3,885	3,876
		2007	49,995	3,811	3,211	2,751	2	3,815	3,807
	HEED 2007	2004	49,910	4,100	3,474	2,924	2	4,105	4,095
	1,499,238	2005	49,960	4,104	3,453	2,963	2	4,109	4,099
		2006	49,922	4,017	3,411	2,868	2	4,021	4,012
		2007	49,902	3,951	3,343	2,805	2	3,956	3,947
Economy7	Non-HEED	2004	49,998	6,960	5,587	5,392	3	6,967	6,954
	2,685,662	2005	49,995	6,750	5,427	5,225	3	6,757	6,744
		2006	49,996	6,543	5,275	5,066	3	6,549	6,537
		2007	49,996	6,675	5,339	5,237	3	6,682	6,669
	HEED	2004	49,965	6,472	5,069	4,981	4	6,480	6,465
	1,735,592	2005	49,990	6,199	4,874	4,769	4	6,206	6,192
		2006	49,966	6,001	4,756	4,593	3	6,008	5,994
		2007	49,995	6,067	4,749	4,728	4	6,074	6,060
	HEED 2004	2004	49,879	6,788	5,369	5,104	10	6,809	6,768
	241,057	2005	48,715	6,480	5,122	4,886	10	6,499	6,460
		2006	49,924	6,318	5,034	4,726	10	6,337	6,299
		2007	49,743	6,433	5,059	4,903	10	6,453	6,413
	HEED 2005	2004	49,913	6,675	5,191	5,129	9	6,692	6,658
	355,134	2005	49,852	6,104	4,756	4,706	8	6,119	6,088
		2006	49,701	5,944	4,664	4,560	8	5,959	5,929
		2007	49,978	6,040	4,684	4,724	8	6,056	6,025
	HEED 2006	2004	49,953	6,580	5,158	5,081	8	6,596	6,565
	415,312	2005	49,741	6,276	4,948	4,838	8	6,291	6,262
		2006	49,966	5,829	4,629	4,491	7	5,843	5,816
		2007	49,989	5,934	4,665	4,631	7	5,948	5,920
	HEED 2007	2004	49,965	6,819	5,293	5,281	9	6,836	6,802
	380,520	2005	49,946	6,569	5,119	5,093	8	6,585	6,553
		2006	49,768	6,321	4,960	4,923	8	6,336	6,305
		2007	49,995	6,153	4,759	4,892	8	6,168	6,137

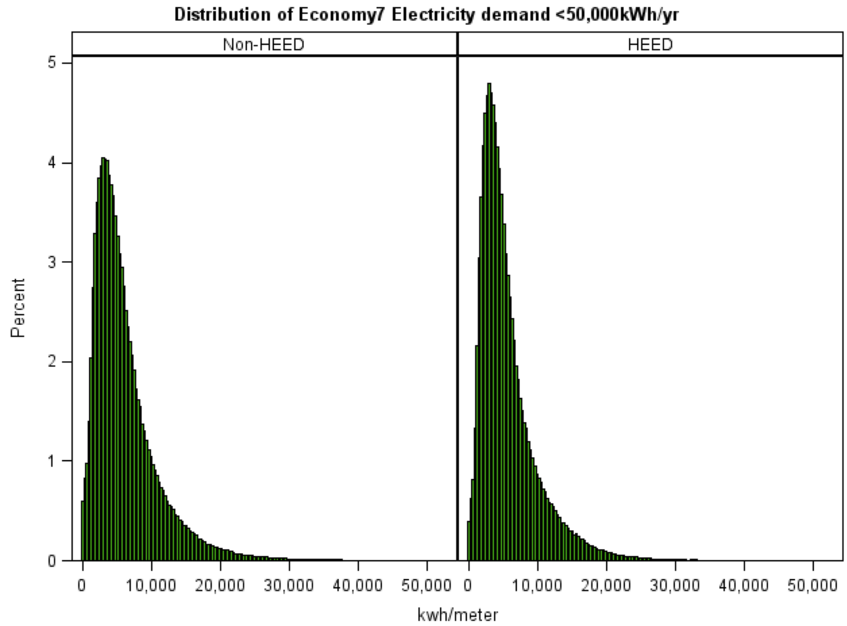
[a]Max – all domestic meters (i.e. profile 1 and 2) greater than 50,000kWh/year were flagged and not included in determining the summary statistics.

Focusing on domestic meters, the change in the median ordinary electricity demand in 'non-HEED' meters between 2004 and 2006 is approximately -0.8%, with an average year-on-year change of -1%. The change between 2004 and 2007 for the same meters is 1.2%. For meters in HEED, the change in median ordinary electricity demand between 2004 and 2006 is approximately -1.5%, with a year-on-year change of -1.2% between the medians. The change between 2004 and 2007 is -0.9%. Non-HEED Economy7 meters saw a change in median of -5.6% between 2004-2006, compared to -6.2% for HEED Economy7 meters for the same period (change in medians for 2004 to 2007 is -3.5 and -5.6 for HEED and non-HEED meters respectively).

The electricity data (ordinary and Economy7 meters) is influenced by a long right tail, as can be seen in the distribution of electricity demand (Appendix Figure 17 and Appendix Figure 18). Note that when considering this tail against the gas demand data, electricity meters are assigned based on their user class, where as the gas data is classified according to the consumption. This means that the long right tail in gas may hold a certain number of non-domestic users, whereas the electricity demand is reflecting high domestic users.



Appendix Figure 17 - Domestic ordinary electricity demand distribution

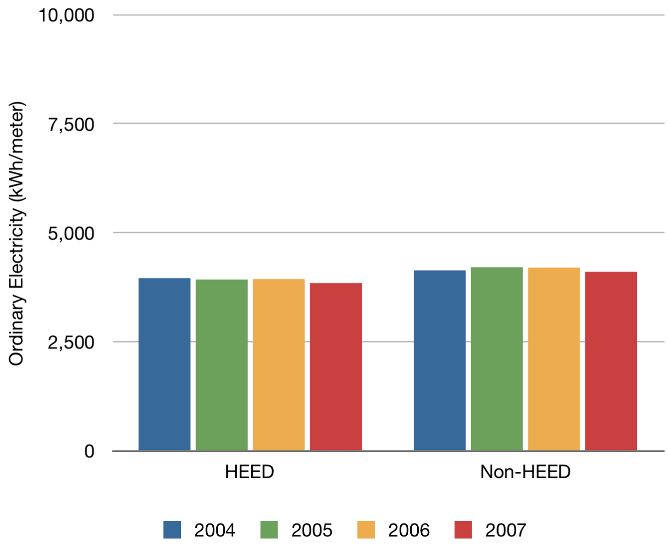


Appendix Figure 18 - Domestic economy7 electricity demand distribution

14.5.1 Change in Electricity Demand

The rate at which dwellings in HEED changed their electricity demand was higher than those that were not part of the HEED data set (identified via unmatched MPANs). The change in demand, however, was not always by the same sign (i.e. 2004 to 2005 for ordinary electricity). The change in ordinary electricity demand is probably less associated with changes in efficiency, although measures such as low energy lighting are present in HEED (and probably also in non-HEED).

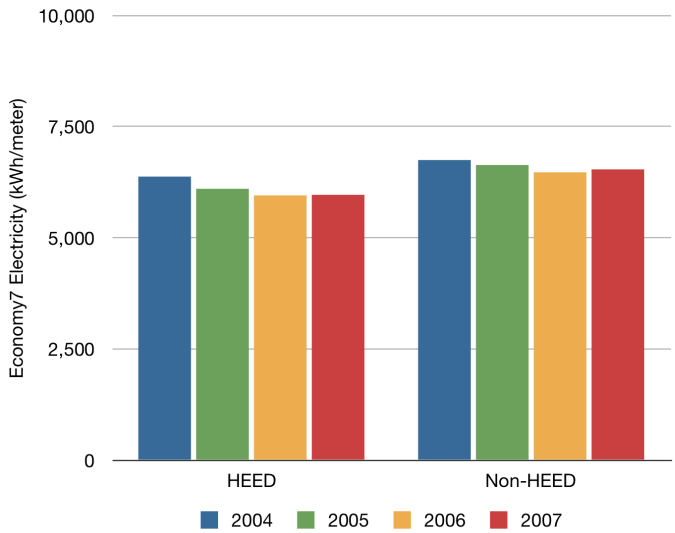
Note also that the change in Economy7 meters was different from their gas-using counterparts, where in fact demand for electricity increased over the same period. This could be due to the fact that electricity data is not corrected for weather, while gas data is.



Percent difference between years

Period	HEED	Non-HEED
2004 to 05	-0.8%	1.7%
2005 to 06	0.3%	-0.2%
2006 to 07	-2.3%	-2.3%

Appendix Figure 19 – Median Ordinary Electricity and change in Ordinary Electricity 2004 to 2007



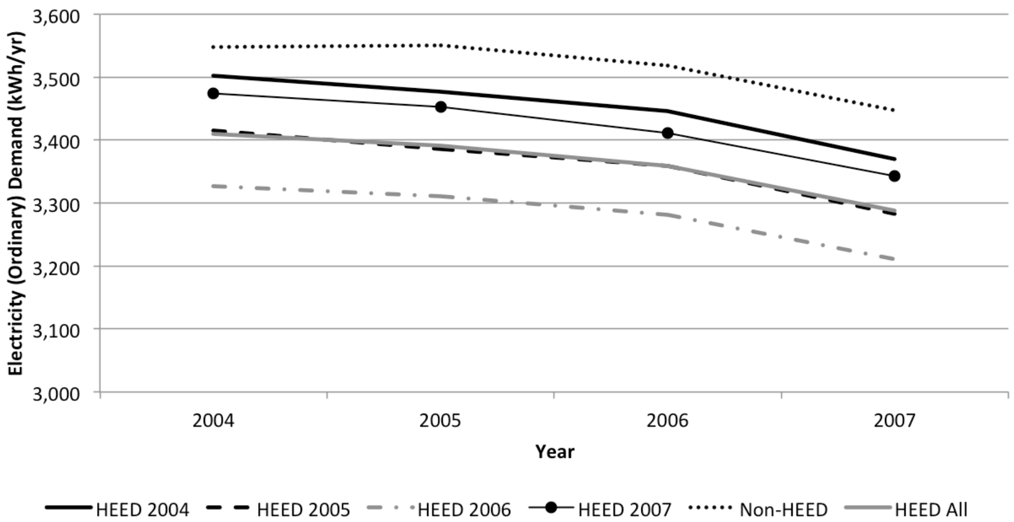
Percent difference between years

Period	HEED	Non-HEED
2004 to 05	-4.3%	-1.7%
2005 to 06	-2.4%	-2.5%
2006 to 07	0.2%	1.0%

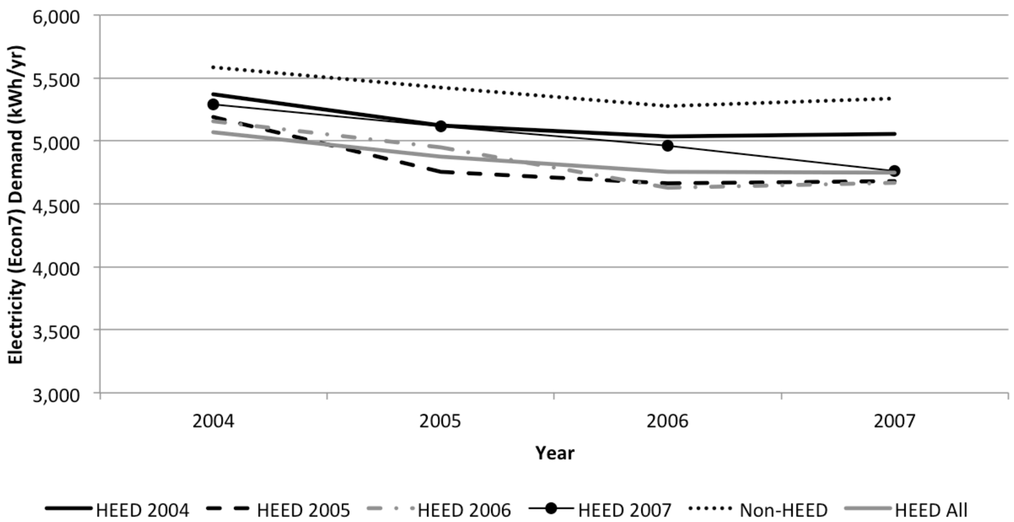
Appendix Figure 20 – Median Economy7 demand and change in Economy7 Electricity 2004 to 2007

14.5.2 Change in electricity demand for HEED meters

Using the date stamp, a comparison of the change in ordinary and economy7 electricity demand can be made. Appendix Figure 21 shows that the year-on-year change for all non-HEED and HEED groups is broadly similar, with non-HEED meters reducing by 0.8% from 2004 to 2007 and HEED meters reducing by 1.2%. Appendix Figure 22 shows the change in Economy7 meters varies more across the period and groups. Note, however, that the change in demand cannot be compared directly as this data has not been weather corrected; rather what is interesting is that the both the trend change is similar across all groups and that those meters in HEED all show a reduced demand as compared to those meters not in HEED.



Appendix Figure 21 - Ordinary electricity demand by HEED entry year



Appendix Figure 22 - Economy7 electricity demand by HEED entry year

Appendix Table 20 provides the change in electricity demand (ordinary and economy7) from 2004 to 2007. The group average change in ordinary electricity for meters in HEED is a reduction of 3.5% as compared to a reduction of 2.5% for non-HEED meters. Economy7 meters in HEED broadly show a reduction of around 9.5% from 2004 to 2007 and non-HEED meters show a reduction of 4.1%. Again, note that the economy7 is not weather corrected and this change may reflect weather trends.

Appendix Table 20 - Change in electricity use 2004 to 2007 by HEED group and meter profile

Profile	Group	Period	Mean	Median
Ordinary	*Non-HEED*	2004 to 2007	-2.5%	-2.8%
	All HEED	2004 to 2007	-3.3%	-3.6%
	In HEED 2004	2004 to 2007	-3.5%	-3.8%
	In HEED 2005	2004 to 2007	-3.6%	-3.9%
	In HEED 2006	2004 to 2007	-3.4%	-3.5%
	In HEED 2007	2004 to 2007	-3.6%	-3.8%
Economy7	*Non-HEED*	2004 to 2007	-4.1%	-4.4%
	All HEED	2004 to 2007	-6.3%	-6.3%
	In HEED 2004	2004 to 2007	-6.9%	-6.2%
	In HEED 2005	2004 to 2007	-9.5%	-9.8%
	In HEED 2006	2004 to 2007	-9.8%	-9.6%
	In HEED 2007	2004 to 2007	-9.8%	-10.1%

14.6 Summary

The energy supplier meter data used in this thesis provides a means of examining both trends and variation in gas and electricity demand amongst the UK dwelling stock. The gas and electricity meter data is subject to a process of annualisation and, in the case of gas demand, a weather correction. Furthermore, the process of the annualisation and weather correction is done in such a manner as making it very difficult to 'uncorrect' the data; and although the methodology is published the specific values used in the correction process are not. This meant that the energy demand data had limitations in terms of how change in demand might be detected.

Chapter 15 Appendix C

UK population data

Appendix introduction

This appendix provides further details on UK dwelling population data used in the method studies.

Sections of this appendix appeared in:

Hamilton IG, Shipworth D, Summerfield AJ, Steadman P, Oreszczyn T, Lowe R. Uptake of energy efficiency interventions in English dwellings. Build Res Inf. 2014 May 4;42(3):255–75.

Hamilton IG, Steadman PJ, Bruhns H, Summerfield AJ, Lowe RJ. Energy efficiency in the British housing stock: Energy demand and the Homes Energy Efficiency Database. Energy Policy. 2013 Sep;60:462–80.

15.1 English Housing Survey (English House Conditions Survey)

The English Housing Survey (EHS) was used as an independent source of comparison of energy efficiency measures over the period of interest at a national level. The EHS uses an un-clustered stratified sample drawn randomly from a list of all addresses in England and has been updated and made available quinquennially since 1967 and yearly since 2008 (CLG, 2010a). Approximately 17,500 households have been interviewed on the details of their home and household, and a further physical survey has been undertaken of approximately 8,000 of the interviewed households. Weighting factors are used for both dwellings and households in order to 'scale up' and represent the full stock; dwelling weighting factors were used for consistency, because HEED is reported by dwelling and not by household.

The EHS contains details of selected efficiency features present in the dwelling (i.e. loft insulation thickness, the predominant type of window, type of wall and insulation, and boiler type). Loft insulation thickness distinguishes none, less than 100mm, 100mm up to 150mm and 150mm or more. Windows are categorised as single and double glazed by casement material (i.e. UPVC, metal or wood). Wall insulation includes cavity with and without insulation and 'other' (e.g. solid wall, wood, pre-fabricated, etc…). Boilers include standard (i.e. non-condensing), back boilers, combination boilers, condensing boilers and condensing-combination boilers.

15.2 Scottish House Conditions Survey

The Scottish House Conditions Survey reports on the households and physical condition of the Scottish stock. The survey includes an interview and physical survey of approximately 3,000 to 4,000 dwellings, undertaken in a continuous format with the aim of achieving a 15,000 sample over 5 years (Scottish Government, 2009). The survey collects physical characteristics of the dwellings (e.g. age, type, size, energy efficiency, etc.) and household features (e.g. tenure, income, etc.) The survey sample is stratified by area and randomly selected and attempts to be representative of the 2.4 million dwellings in Scotland.

15.3 Valuation Office Agency Council Tax Property Attributes

The Valuation Office Agency (VOA) Council Tax Property Attributes tables are collected as part of the VOA's responsibility to group properties into the appropriate council tax band (VOA, 2010). As part of the banding process, the VOA collects information on the characteristics of the property that may affect its value (including, age, type, area, and number of bedrooms). The VOA data is considered a 'global' listing, as it will contain information on all dwellings within an area (i.e. local council) and thus should have all dwellings in the stock. Note that the VOA only collects information for England and Wales. The VOA in maintaining a current valuation list revises the data annually; the data used in the comparison comes from an extract made in September 2010. The VOA dataset provides information at the Local Authority level for approximately 24.7million residential properties.

15.4 LSOA level Neighbourhood statistics

The Neighbourhood statistics provides information for a range of geographic levels across England, including census geography. For the most part, the data is area-based and draws together Census and other administrative datasets. Appendix Table 21 provides a list of LSOA level datasets and variables used in this thesis.

Appendix Table 21 – List of datasets and variables (with geographic levels) used in the energy efficiency uptake analysis in England for the period 2000 to 2007

Dataset	Source	Level	Year	Variables used	Measurement	Description	Reference
Mid-2005 Population Estimates, All Persons	Office for National Statistics	LSOA	2005	0-15', '16-29', '30-44', '45-64 Males &45-59 Females', '65+ Males & 60+ Females'	Estimate of number of persons	Dataset of the number of persons by age bands and sex for England. The estimates are made using the Kannisto-Thatcher method, based on modified survival ratios for the population.	(ONS, 2010)
Dwelling Type	Valuation Office Agency	LSOA	2011	Bungalow' 'Flat/Maisonette', 'Terrace', 'Semi-detached', 'Detached', 'Other', 'Unknown'	Count of domestic properties	Dataset of the number of domestic properties in dwelling type bands from the Valuation Office Agency property attribute tables for council tax	(ONS, 2012b)
Benefits Data: Summary Statistics	Office for National Statistics	LSOA	2005	'Disability Living Allowance', 'Incapacity Benefit/Severe Disablement Allowance', 'Income Support', 'Jobseekers Allowance', 'Pension Credit'	Count of claimants (persons)	Dataset of summary statistics from Department of Work and Pensions covering benefit claims during the period of August 2005	(ONS, 2012b)
Median Household Income	Experian	LSOA	2004	Median income'	Estimate of median LSOA level income	Dataset of median income levels of households in an LSOA estimated by Experian using a multi-stage modelling approach.	(Experian, 2009)
Household Tenure	Office for National Statistics	LSOA	2001	Owned', 'Rented from council', Other social rented', 'Private rented', 'Living rent free'	Count of households in domestic properties	Dataset from 2001 Census describing the household tenure related to the accommodation in question.	(ONS, 2012b)
Heating degree days	Met Office	LSOA	2005	Heat degrees'	Estimate of the annual average degrees below 15.5 in °C	Dataset of annual sum of heating degrees below 15.5 °C over a 5 x 5 km2 grid of England. Data are converted to LSOA by an overlay and averaging of the grid points.	(UK Met Office, 2012)
Dwelling Stock by Council Tax Band	Office for National Statistics	LSOA	2005	Band A' to 'Band H' for England	Count of domestic properties	Dataset of the number of domestic properties in council tax bands provided by the Valuation Office Agency, covering 23,101,020 dwellings in England.	(ONS, 2012b)

Chapter 16 Appendix D

Method studies: additional material

Appendix introduction

This appendix provides additional material for the method studies.

16.1 Interpolation of English Housing Survey data

Appendix Table 22 below shows the results from the linear interpolation between English Housing Survey years (i.e. 1996, 2001, 2003, 2005, 2007, and 2009). The results of these interpolations for a selection of energy efficiency levels are used for comparison against reported efficiency installations in the HEED.

Appendix Table 22 – Results from interpolation of English Housing Survey data 1996 to 2007 for selected dwelling energy efficiency levels

Efficiency Measure (10,000's)	**Survey Year**											
	1996	**1997***	**1998***	**1999***	**2000***	**2001**	**2002***	**2003**	**2004***	**2005**	**2006***	**2007**
Loft insulation thickness												
none	69.0	62.2	56.5	53.0	52.9	57.4	66.6	76.8	84.1	85.7	80.7	73.6
<100mm	926.4	803.0	690.5	599.8	541.8	527.3	556.6	587.5	580.2	547.5	512.1	483.3
100-150mm	584.8	650.5	710.9	760.8	795.0	808.2	798.1	773.9	747.3	729.5	726.1	719.6
150mm+	214.5	265.4	314.4	359.6	399.1	431.1	456.0	483.0	522.5	577.8	647.4	714.2
no loft	233.8	267.4	295.7	313.5	315.6	296.6	257.9	227.2	227.5	237.5	233.3	228.2
Type of wall and insulation												
cavity with insulation	284.3	348.5	408.5	460.3	499.7	522.6	528.7	533.4	553.5	597.4	664.6	726.8
cavity uninsulated	1,033.7	1,014.7	996.3	979.3	964.2	951.8	942.5	935.7	928.6	909.3	869.6	826.0
other	710.5	685.2	663.1	647.2	640.5	646.3	664.0	679.3	679.5	671.4	665.4	666.1
Extent of double glazing												
no double glazing	821.9	760.9	700.1	639.8	580.2	521.5	464.2	409.4	358.7	315.4	281.7	253.2
less than half	288.3	268.6	249.9	233.4	220.0	211.0	205.8	198.1	183.0	168.1	160.8	154.8
more than half	302.2	295.1	290.1	289.5	295.2	309.5	331.6	349.5	352.3	346.0	338.3	326.4
entire house	616.1	723.9	827.9	924.2	1,009.1	1,078.7	1,133.5	1,191.5	1,267.7	1,348.6	1,418.8	1,484.5
Extent of double glazing												
less than 80% double glazed	1,267.9	1,184.1	1,102.0	1,023.6	950.6	884.9	825.8	763.3	690.9	625.0	580.6	542.7
80% or more double glazed	760.5	864.3	965.9	1,063.2	1,153.8	1,235.8	1,309.4	1,385.1	1,470.7	1,553.1	1,619.0	1,676.2
Boilers												
Standard boiler	1,042.1	1,057.0	1,068.0	1,070.9	1,061.9	1,037.0	997.3	964.2	954.1	942.5	907.6	878.2
Back boiler	276.6	278.9	280.7	281.3	280.2	276.8	270.1	258.0	239.1	218.1	201.2	194.4
Combination boiler	280.3	305.8	333.1	364.1	400.6	444.5	496.0	549.2	595.9	625.4	631.5	628.7
Condensing boiler	0.0	4.4	8.5	12.0	14.4	15.5	15.4	15.4	18.2	30.0	52.8	69.8
Condensing (combi) boiler	0.0	7.6	15.0	21.7	27.4	31.9	35.0	37.3	43.4	72.7	134.7	183.7
No boiler	429.4	394.6	362.7	336.9	319.9	315.0	321.4	324.4	311.0	289.4	271.8	264.2
Total	2,028.5	2,048.4	2,068.0	2,086.8	2,104.5	2,120.7	2,135.1	2,148.4	2,161.6	2,178.1	2,199.6	2,218.9

16.2 Analyses of energy efficiency uptake using HEED 2002 to 2007

Appendix Table 23 below shows the results of the ecological analysis for the total uptake of energy efficiency measures by programme from 2002 to 2007 and the association with LSOA-level dwelling and household characteristics.

Appendix Table 23 - Odds ratios of incidence rate of uptake of energy efficiency measures from 2002 to 2007 at LSOA level and energy efficiency programme in England

	Energy efficiency measures uptake incidence rate 2002 to 2007											
	Energy Efficiency Commitment		Installers		Fuel Poverty Schemes		Home Energy Survey		Core Programmes[c]		All sources	
Variable	Fabric	Heat	Fabric	Heat	Fabric	Heat	Fabric	Heat	Fabric	Heat	Fabric	Heat
Quintile of median income in 2005												
Q1 vs Q5	2.71	2.04	5.07	4.04	9.31	6.03	2.30	1.57	11.86	6.98	11.03	6.68
Q2 vs Q5	1.51	1.26	2.68	2.12	5.49	3.70	1.78	1.42	4.86	3.23	4.73	3.16
Q3 vs Q5	1.36	1.07	1.95	1.42	3.76	2.57	1.61	1.24	3.11	1.86	3.05	1.85
Q4 vs Q5	1.11	0.95	1.36	1.06	2.38	1.89	1.31	1.11	1.79	1.29	1.78	1.29
Tenure (proportion of dwellings)												
Owner occupied (units[a]=10%)	1.27	1.11	1.44	1.18	1.88	1.7	1.27	1.07	1.82	1.33	1.74	1.30
Dwelling type (proportion of dwellings)												
Flats (units[a]=10%)	0.84	0.91	0.79	0.87	0.79	0.89	0.84	0.92	0.72	0.81	0.71	0.81
Quintile of climate (heat degree days in 2005)												
Q2 vs Q1	1.56	1.55	1.77	1.51	0.74	0.73	0.98	0.85	1.47	1.28	1.54	1.30
Q3 vs Q1	2.11	1.86	2.25	1.76	0.66	0.66	1.14	0.97	1.92	1.63	1.96	1.65
Q4 vs Q1	2.20	1.89	2.61	1.94	0.6	0.57	1.15	0.95	2.04	1.63	2.14	1.67
Q5 vs Q1	1.99	1.71	2.03	1.69	0.75	0.8	1.32	1.01	1.96	1.79	1.96	1.79
Council Tax Band (proportion of dwellings)												
Band A&B (units[a]=10%)	1	0.98	0.98	0.98	1.14	1.11	1	0.98	1.03	1	1.02	1
Benefits[b] (proportion of dwellings)												
On benefits (units[a]=10%)	1	0.96	1.04	0.99	1.62	1.52	1.11	1.03	1.31	1.13	1.32	1.13
Household age (proportion of dwelling occupants)												
Adults ≥60 years (units[a]=-5%)	1.03	1.02	1.02	1.01	1.04	1.02	1	0.99	1.04	1.01	1.04	1.01
Children ≥14 years (units[a]=-5%)	1.01	1.01	1	1.01	1.04	1.04	1.01	1.02	1.02	1.03	1.02	1.03

Note: a) ORs estimates correspond to each additional unit of change; b) Benefits includes: disability, incapacity, income support, job seekers, pension; b) Core programmes include: EEC, Installers, Fuel Poverty and Home Energy Survey

N.B. In this paper HEED is treated as a 'census' level dataset (i.e. a survey of all households). As such, p-values are not reported.

16.3 HEED and sample comparisons

Appendix Table 24 - HEED full sample and HEED 10% comparison

Great Britain	HEED[a] (n)	HEED (%)	HEED[a] 10% (%)
Dwelling Age			
Pre-1900	462,141	7.9%	7.9%
1900-29	684,956	11.7%	11.7%
1930-49	906,658	15.6%	15.5%
1950-66	1,167,139	20.0%	20.0%
1967-75	1,463,750	25.1%	25.2%
1976-82	405,962	7.0%	6.9%
1983-90	328,952	5.6%	5.6%
post-1990	410,421	7.0%	7.0%
X^2			11.9843
d.f.			7
p			0.1011
Dwelling Type			
Flat-Maisonette	1,256,778	19.6%	19.6%
Bungalow	588,662	9.2%	9.2%
Terrace	1,619,047	25.2%	25.1%
Semi-detached	1,993,108	31.1%	31.0%
Detached	960,679	15.0%	15.0%
X^2			2.1417
d.f.			4
p			0.7097
Dwelling Tenure			
Social rental	1,811,016	22.4%	22.5%
Private rental	772,352	9.6%	9.5%
Owner-occupied	5,494,295	68.0%	68.0%
X^2			4.9031
d.f.			2
p			0.0862
Dwelling Size (Bedrooms)			
1	822,167	13.2%	13.3%
2	1,647,076	26.5%	26.5%
3	2,866,957	46.1%	46.0%
4	621,709	10.0%	10.0%
5+	261,396	4.2%	4.2%
X^2			2.7461
d.f.			4
p			0.6012
Dwelling Region			
North East	702,766	5.7%	5.6%
North West	1,599,855	12.9%	12.9%
Yorkshire and The Humber	1,199,671	9.7%	9.7%
East Midlands	914,464	7.4%	7.4%
West Midlands	1,161,697	9.4%	9.3%
East of England	1,088,751	8.8%	8.8%
London	1,323,260	10.7%	10.7%
South East	1,617,462	13.0%	13.0%
South West	1,076,835	8.7%	8.7%
Wales	1,168,235	9.4%	9.4%
Scotland	553,061	4.5%	4.4%
X^2			27.9954
d.f.			10
p			0.0018

Notes: [a]HEED and HEED 10% Sample are of England and Wales only