Department for
Business, Energy
& Industrial Strategy

EVALUATION OF THE WARM HOME DISCOUNT SCHEME

Analytical Paper 2: Quantitative Research into the impact of the WHD on energy expenditure and the indoor environment.

March 2018

This document is available in large print, audio and braille on request. Please email enquiries@beis.gov.uk with the version you require.

EVALUATION OF THE WARM HOME DISCOUNT SCHEME

Analytical Paper 2: Quantitative Research into the impact of the WHD on energy expenditure and the indoor environment.

Acknowledgements

This independent research report was produced by James Leather and Andrew Jarvis from ICF, and Ian Hamilton, Paulo Agnolucci and Chris Grainger from UCL.

We are grateful to all interview participants for their time and assistance with this research and for the insights and guidance of our (formerly DECC) research managers, Jonathan Smetherham, Sam Jenkins, and Charlotte Dann.

Any enquiries regarding this publication should be sent to us at enquiries@beis.gov.uk

This publication is available for download at www.gov.uk/government/publications

Contents

Executive summary

Background to the evaluation

In March 2015, ICF, in association with UCL Consultants (UCL), was commissioned by the Department of Energy and Climate Change (DECC) to undertake a combined process and impact evaluation of the Warm Home Discount (WHD) scheme. This report (Analytical Paper 2) presents the results of a quantitative analysis of WHD impacts. It is one of three complementary reports that present the results of the evaluation. The companion reports provide: (i)) the results of qualitative research into the WHD customer journey; and (ii an overall synthesis and set of answers to the evaluation questions[1].

The WHD was an energy supplier-funded scheme which ran from 2011 to 2016. It provided an annual rebate via a single annual transfer of £120-£140 to recipients' electricity accounts[2]. There were two groups of customers (beneficiaries): i) a 'Core Group' that consisted of low income pensioners; and ii) a 'Broader Group' that consisted of other vulnerable or low income individuals. The eligibility criteria for the Core Group were set by DECC and were the same across all energy suppliers, whereas the eligibility criteria for the Broader Group were defined by individual energy suppliers and applied to their respective customer bases (albeit using similar benefits-based parameters).

Aims of the WHD Scheme

DECC sought to tackle fuel poverty by acting on three identified drivers of fuel poverty, i.e. thermal efficiency of the dwelling, household income levels, and the cost of energy.

The Warm Home Discount (WHD) scheme was part of DECC's wider strategy for tackling fuel poverty in England. DECC's motivation for tackling fuel poverty (as described in the Consultation Fuel Poverty Framework) is in response to evidence that low-income families often live in dwellings with poor energy performance, face high energy expenditure and may not heat to achieve adequate warmth. Living in cold or damp buildings is known to affect health and wellbeing. Fuel poverty is wrapped up in wider issues of poverty, though fuel poverty is seen as distinct because fuel costs (compared to use) exist outside the control of the household and

[1] These papers are titled 'Evaluation of the Warm Home Discount Scheme – Analytical Paper 1: Qualitative Research into the Delivery and Customer Journey of the Warm Home Discount' and 'Evaluation of the Warm Home Discount Scheme – Synthesis Evaluation Report' respectively. Both are available at: https://www.gov.uk/government/publications/warm-home-discount-evaluation-2010-to-2015

[2] The WHD scheme also involved 'industry initiatives' (which consisted of a range of measures implemented by energy suppliers to support customers in fuel poverty or at risk of fuel poverty), and 'legacy spend' (which consisted of a continuation of the discounted / social tariffs that suppliers had offered to certain customers as part of a voluntary agreement that preceded the WHD scheme. Neither the industry initiatives nor the legacy spend fell within the scope of this evaluation.

because inadequate heating is related to poor health and wellbeing, with the tacit link to excess winter death burden.

Quantitative research objectives

The objectives of the quantitative research were as follows:

- To determine the extent to which the WHD scheme was responsible for moving a significant number of households out of fuel poverty, both in relation to the 10% fuel poverty measure under which the scheme was initially set, and also against the Low-Income High-Costs (LIHC) measure of fuel poverty;
- To determine the extent to which the WHD scheme alleviated the distributional impacts of higher energy bills as a result of energy policies and prices – on low-income and vulnerable households;
- To determine the impact that the WHD scheme had on improving thermal comfort in recipient homes.

Impacts of the WHD on heating expenditure

The effect of the WHD scheme on heating fuel expenditure was estimated using a model based on 'Engel curves', a commonly used economic relationship which explains the relationship between household budget shares for particular goods and household income. The model tested the potential 'labelling effect' which could occur when a payment is made with a specific label (i.e. calling the payment Warm Homes Discount on the energy bill). This labelling effect would cause a household to spend differently relative to an unlabelled cash transfer.

This research addressed whether the WHD payment increased a household's energy consumption relative to an unlabelled cash transfer and analysed how the rebate had been used by customers. This research also considered whether the same / higher impact could be achieved via different and lower / same cost means.

Analysis approach

This research used an approach called 'Regression Discontinuity Design' (RDD) to determine the labelling effect of the WHD in conjunction with survey data from the Living Cost and Food Survey (LCFS) on household weekly expenditure. RDD is a 'quasi-experimental method', which means that it employs statistical techniques to make causal connections between observational data. The modelling in this analysis took advantage of the arbitrary (in statistical terms) eligibility cut-offs for the policy in income and age as a means of building counterfactuals, as households directly above and below these cutoffs are (controlling for other observed factors) directly comparable.

To address the question on different means of payment, RDD was used to examine the impact that cash payments had for elderly households (i.e. the Winter Fuel Payment), replicating and building upon the work done by Beatty et al (2014). Beatty et al's study reported a statistically significant labelling effect for the Winter Fuel Payment. That is, they determined that the Winter Fuel Payment increased expenditure on fuel more than an equal increase in unlabelled income.

Findings from WHD scheme analysis

- Using RDD, some evidence of a labelling effect was found due to the WHDS. While a statistically significant result of a positive labelling effect was found (i.e. a WHD payment resulted in an increase in fuel expenditure) in a minority of the cases, the result was sensitive to the model specification and identifying assumptions of the RDD model, and thus was not always reliable.

- Given the mix of evidence, this study is unable to determine whether households treat the additional payment as a pure income increase or not. If not, the WHDS would still result in an increase in energy expenditure due to increase in income but not any more than an equivalent unlabelled increase in income.

- Because the effect size is consistent across the significant and non-significant findings, this study may also be interpreted to show an underlying labelling effect for which there is simply insufficient data.

Findings from the Winter Fuel Payment analysis

- In order to test whether a similar impact on household fuel expenditure could be achieved using a similar or less costly payment type, a comparison of the WHD and Winter Fuel Payment (WFP) was conducted. The WFP is a direct cash payment of between £100-300 for eligible individuals (those of pension age and living in the UK).

- As a basis for comparison, a recent study (Beatty et al 2014) by the Institute for Fiscal Studies (IFS) found that households in receipt of a WFP changed their fuel expenditure as compared to an equivalent increase in income. The IFS study of the WFP was replicated to test robustness of this study's findings. The analysis found that the results of the IFS paper were also sensitive to aspects of the design of the study (the assumptions for RDD) as well as model structure. The study therefore finds that, in terms of a labelling effect, there is no difference between the direct cash payment used for the WFP and the method used by the WHD.

Implications

- Households may spend an increased proportion of the WHD scheme on fuel relative to an unlabelled cash transfer, indicating a labelling effect, but this finding is not consistent across model specifications and samples. A labelling effect would indicate an economic inefficiency, as it would mean that households are not spending the increased income on those items or services they need or value most.

- If the significant findings are true, this study finds that the WHD would increase spending on fuel by approximately £0.05 per week or £2.60 per year.

Conclusions

This study found some evidence of a labelling effect for the WHD scheme, but it is not robust across all specifications and samples. This indicates that the WHDS may cause households to allocate additional income from what they would have done in the case of an unlabelled cash transfer, but the evidence for this limited. Finally, there appears to be no difference in the impact on energy expenditure when considering different methods of payment, i.e. through a direct cash payment under Winter Fuel Payment.

Impact of the WHD on household wellbeing

Evidence has shown that there is a relationship between the energy performance of a dwelling and the internal temperatures experienced during wintertime conditions (Oreszczyn et al., 2006). An added benefit of an increase in indoor temperature is an improvement in the health and mental wellbeing of the occupants, such as mental wellbeing, cardio-respiratory disease and asthma in children.

The evaluation investigated whether the WHD rebate improved recipients' wellbeing, the net benefit/impact of the WHD scheme for individual recipients and society, and whether the value of the benefits was sufficient to justify the cost of the rebate process. Three analysis models (energy expenditure, indoor temperature and energy performance) were used to estimate the impact of the WHD scheme on indoor environment, energy and wellbeing.

Analysis Approach

A model quantifying the change in indoor environmental conditions and the health impacts of housing energy efficiency and fuel payment measures was used to determine the change in indoor temperature during typical wintertime conditions (i.e. 5 °C outdoors) and the corresponding potential health impact and mental wellbeing. This wintertime temperature was chosen in order to compare all homes and has been found to have a significant relationship with cardio-respiratory health.

The modelling used an empirical relationship between indoor temperature and the dwelling energy performance to estimate the impacts on household health and mental wellbeing due to a change in fuel expenditure.

Dwellings eligible for the receipt of the WHD scheme (i.e. in Pension Credit, Guarantee Credit) were identified and targeted using the English Housing Survey (EHS). The number of households eligible were approximately 1.2 million households as defined in the EHS 2013.

The WHD scheme was applied as an additional £140 directly added to the energy expenditure. For those households that were eligible for the Winter Fuel Payment (WFP) the two payments were combined but only the impact of the WHD was estimated.

Changes in indoor temperature wintertime thermal conditions within the home and the associated change in energy demand alongside the potential change in health impact. The health impacts were quantified as 'quality adjusted life-years' (QALYs), which is a measure of the burden of disease on a year of life and accounts for both the quality and quantity of life. One year of good health is worth 1 QALY.

Findings from the impact of WHD scheme on indoor environment, energy and wellbeing

- The WHD health impact modelling found that more eligible households live in flats than dwellings, most are present or ex-council flat tenants, and most live in post-1945 flats. WHD scheme eligible dwellings tend to be more efficient with much lower ventilation heat losses, reflecting the dwelling typology they live within. Most WHD scheme eligible households tend to already be warmer and use less energy than the non-eligible dwellings, due to the type of dwellings they live in.

- The analysis found that households in receipt of the WHD could on average see an increase in temperature of around 0.25 °C during wintertime conditions (i.e. when 5 °C outdoors). The corresponding change in health for the Core Group households, including benefits to cardio-respiratory health, and common mental disorder, was a benefit of approximately 6,100 QALYS over a 15-year period for the WHD scheme eligible group and a per capita improvement of 50 QALYs per 10,000 persons. The Broader Group health improved (which also include impact on childhood asthma) by approximately 33,000 QALYs over the period, with a per capita improvement of 30 QALYs per 10,000 persons. These are relatively small increases in health and these impacts reflect the low risk of cold-related deaths among the wider population and the treated households compared to other diseases.

- If monetised these benefits could result in a societal benefit of approximately £150 per capita of non-discounted health benefit over the 15-year period for the Core Group and £132 per capita for the Broader Group, assuming a value of £30,000 per QALY. These wider societal benefits are accrued to the wider economy, for example through quality of life or employment gains. These gains do not consider potential impacts on the NHS budget expenditure on cold-related disease treatment.

Implications

WHD scheme eligible households (both Core Group and Broader Group) are predominantly already warmer than non-eligible households, reflecting the type of dwellings they live in. Among the WHD eligible households, increases in income will likely result in some thermal improvement. The largest gains are concentrated in dwellings that are generally less energy efficient and so the increase expenditure may only see a moderate improvement in temperatures due to their dwelling being harder to heat. For the majority of WHD eligible households, there may only be a small change in temperature for the reason that most of the households live in dwellings with higher energy performance. WHD scheme payments may provide some improvements in cardio-respiratory and mental health.

Conclusions

The health impact analysis of the WHD found that there was a potential improvement in health over 15 years of life (average of +65 year olds) for both the Core Group and Broader Group eligible households measured as a positive change in Quality Adjusted Life Years. These improvements were due to the small change in indoor temperature related to an overall increase in fuel expenditure due to the increased income.

The modelling suggested that the WHD would result in generally only a small improvement in temperature (and therefore health) due the type of dwellings that WHD eligible households lived in. Because most of those in receipt of the payment live in dwellings that had a higher average energy performance level, the was a greater probability that the rebate would not be needed to increase indoor temperatures.

Effectiveness of the WHD in identifying vulnerable customers

The aim of the WHD is to address the risk of being in fuel poverty amongst those households that are at highest risk and severity, which is defined under the WHD scheme as those being in receipt of Pension Credit Guarantee Credit, i.e. the Core Group. However, the risk of fuel poverty and of living in cold homes may not fully align with the eligibility criteria of WHD.

The analysis considered how well the definitions of the Core and Broader Groups target the fuel poor, how well the definitions target the neediest part of the fuel poor group, and if a new proxy might be developed to better target fuel poor customers under the low income high cost (LIHC) fuel poverty definition[3].

Analysis approach

A Random Forest classification approach was developed to determine important variables in identifying households that were at risk of living within cold homes (defined as having a wintertime average indoor temperature <18 °C). This temperature was selected because it aligned with a recent review by Public Health England that heating homes to at least 18°C (65F) in winter poses minimal risk to the health of a sedentary person, wearing suitable clothing.

These variables were then used in a logistic regression model and the outcome of this model is compared against WHD eligibility, fuel poverty risk (both LIHC and 10%) for predicting living in a cold home.

Findings from the vulnerable customer identification analysis

- The following important variables were identified by the Random Forest method for identifying households that were at risk of living within cold homes: the E-value (i.e. a measure of the dwelling energy performance), length of residency, household type, dwelling age, presence of a boiler, age of the household reference person, number of people in the home, household income, number of bedrooms, and whether the household reference person is employed.

[3] Under the Hills Low Income High Costs (LIHC) definition a fuel poor household is one in which: 1) a household has required fuel costs that are above the median level; and, 2) where the household to spend that amount, they would be left with a residual income below the official poverty line.

- The logistic regression model built using the variables above was able to correctly identify more homes which were observed as being ‘cold’ (defined as having a wintertime average indoor temperature <18 °C) in the sample used than those identified by the WHD scheme Core Group eligibility criteria, and the LIHC and 10% fuel poverty definitions.

- The chance of incorrectly identifying (i.e. false positives) a home as being a ‘cold home’ using the model variables is 34%, compared to 77% using the WHD scheme Core Group definition. That is, out of the households the logistic regression model predicted would be ‘cold’, 34% were not. This is compared to 77% of the households predicted to be ‘cold’ by the WHDS Core Group definition.

Implications

The WHD scheme is aimed at reducing the cost of fuel expenditure amongst vulnerable households, who may be at risk of living in fuel poverty. An extension of this risk of being fuel poor is that households may be living in a cold home.

The use of the WHD scheme Core Group eligibility criterion is not necessarily a strong indicator that households in receipt of the payment would be living in cold homes (i.e. <18 °C). This reflects the predominant type of home that those households occupy, i.e. mid-century flats that are social rentals, and therefore built to a higher energy performance standard.

The stronger predictors of coldness were a measure of the dwelling energy performance and other measures of household and dwelling age. While the Core Group part of the WHD scheme targets those who are older, it may not necessarily reflect the energy performance of dwellings they reside in.

Conclusions

In addressing the research questions focused on identifying the neediest part of the fuel poor it was found that households in receipt of the WHD may not necessarily be those with the highest risk of living in cold homes. Using a measure of energy performance, which would also reflect dwelling age, and some form of length of residence within the LIHC definition of fuel poverty could provide a more appropriate proxy to better target households vulnerable to living in cold homes.

Background and Objectives

Introduction

In March 2015, ICF, in association with UCL Consultants (UCL), was commissioned by the Department of Energy and Climate Change (DECC) to undertake a combined process and impact evaluation of the Warm Home Discount (WHD) scheme. The aim of the evaluation was to determine the extent to which the WHD scheme was responsible for removing households from fuel poverty, to establish the impact on customers, and to review the process by which the scheme was delivered.

This report (*Analytical Paper 2*) presents the results of quantitative research into the labelling and health impacts of the WHD scheme, as well as the effectiveness of the scheme in identifying vulnerable customers. It should be read in conjunction with two other reports that present the results of the evaluation:

- *Analytical Paper 1*: Presents the results of qualitative research into the WHD customer journey;
- *Synthesis Paper*: Draws together the results of both analytical papers to provide answers to the overarching process and impact evaluation questions.

The Warm Home Discount scheme

The WHD was an energy supplier-funded scheme that operated in England, Wales and Scotland. It came into operation on 1 April 2011 and was originally envisaged to be a four year programme ending on 31 March 2015, but was extended through to 31 March 2016.

The WHD was developed at a time when energy bills were relatively high and were expected to increase still further, thus putting greater numbers of households into fuel poverty. The WHD scheme had two objectives[4]:

- "To remove a significant number of households from fuel poverty and improve the thermal comfort and health of assisted households by providing direct support with energy bills; and
- To help to mitigate the burden of rising energy prices on low-income households, who will be worse affected than higher income households".

[4] DECC (2011) The Warm Home Discount Scheme: Final Stage Impact Assessment

The WHD scheme was thus expected to contribute towards the UK Government's target for reducing fuel poverty. It was intended to complement other Government initiatives, including the Affordable Warmth target within the Energy Companies Obligation – ECO[5] – and the Winter Fuel Payment and Cold Weather Payment. The WHD would directly mitigate the impacts of rising energy prices by providing a rebate on energy bills, whereas ECO would improve the thermal efficiency of homes, and the Winter Fuel Payment would improve general household income.

DECC identified five broad principles that guided the design of the WHD scheme[6]:

- "Delivers a fair and clear benefit for consumers: consumers should have certainty on the absolute level of support that they will receive, allowing them to plan and budget for their energy costs;
- Provides focused support for vulnerable households: support should be targeted at households vulnerable to fuel poverty;
- Delivers good value for money: support should be a cost-effective tool for tackling fuel poverty, without undue administrative costs;
- Is consistent with competitive energy markets: has a minimal impact on the incentives of consumers and suppliers to engage with the domestic energy market; and
- Ensures a smooth transition from the current arrangements[7] for consumers and suppliers".

The WHD scheme provided a £120-£140 one-off annual rebate on the electricity bills of eligible individuals[8]. To be eligible for the rebate, individuals had to fall within one of two groups:

- A 'Core Group', consisting of low income pensioners. To be eligible for the rebate, pensioners had to be in receipt of the Guarantee Credit part of Pension Credit[9], and had to

[5] The Energy Companies Obligation, ECO, was launched in 2013 and replaced two previous schemes: the Carbon Emissions Reduction Target (CERT) and the Community Energy Saving Programme (CESP)

[6] DECC (2011) The Warm Home Discount Scheme: Final Stage Impact Assessment

[7] The WHD scheme replaced a 'voluntary agreement' between the UK Government and the largest six energy suppliers to provide support with energy bills to vulnerable households. The voluntary agreement ran from 2008 to 2011, and largely consisted of social tariffs (extra low tariffs offered to certain types of consumer) and, from 2010, a rebate on electricity bills that was offered to certain pensioners (the latter in effect formed a pilot for what became the WHD scheme).

[8] The WHD scheme also involved 'industry initiatives' (which consisted of a range of measures implemented by energy suppliers to support customers in fuel poverty or at risk of fuel poverty), and 'legacy spend' (which funded a continuation / wind-down of the activities previously delivered via the voluntary agreement). Neither the industry initiatives nor the legacy spend fell within the scope of this evaluation.

[9] The eligibility criteria changed over the scheme years: in year one, the rebate was available for recipients of the Guarantee Credit only, and from year two onwards this was extended to individuals in receipt of both the Guarantee Credit and Savings Credit (who were aged 80+ in year two, 75+ in year three, and 65+ in years four and five)

be named on an electricity account with one of the participating energy suppliers[10]. The Core Group eligibility criteria were defined by the Government, and were the same across all energy suppliers.

- A 'Broader Group', consisting of other vulnerable or low income individuals. In years one to four of the WHD scheme, Broader Group eligibility criteria were defined by participating energy suppliers (subject to approval by Ofgem). In year five of the scheme, the Government introduced a set of mandatory criteria that suppliers had to include in their schemes, though they were still able to apply additional criteria (again, subject to approval by Ofgem). For the most part, the Broader Group eligibility criteria were based on receipt of means-tested benefits (Income Support, Employment and Support Allowance, Job Seekers Allowance, Universal Credit). Individuals also had to have an active electricity account with an energy supplier to be eligible for the rebate.

The WHD scheme was part of the UK Government’s strategy for tackling fuel poverty. There is evidence that low-income households often live in dwellings with poor energy performance, face high energy expenditure, and may not heat their homes to achieve adequate warmth. Fuel poverty is wrapped up in wider issues of poverty, though fuel poverty is seen as distinct because fuel costs (compared to energy use) exist outside the control of the household and because inadequate heating is related to poor health and wellbeing, with a tacit link to excess winter death burden. For England, the Government now uses a metric of fuel poverty that combines low income and having high energy expenditure, known as the ‘Low-Income High-Cost’ (LIHC) indicator of fuel poverty. In Wales and Scotland, the 10% of income spent on fuel remains the measure of fuel poverty.

Quantitative research objectives

The objectives of the quantitative research were as follows:

- To determine the extent to which the WHD scheme was responsible for moving a significant number of households out of fuel poverty, both in relation to the 10% fuel poverty measure under which the scheme was initially set, and also against the Low-Income High-Costs (LIHC) measure of fuel poverty;
- To determine the extent to which the WHD scheme alleviated the distributional impacts of higher energy bills as a result of energy policies and prices – on low-income and vulnerable households;
- To determine the impact that the WHD scheme had on improving thermal comfort in recipient homes.

[10] In years one and two of the WHD scheme this consisted of the ‘big six’ suppliers (British Gas, EDF Energy, E.ON, Npower, SSE and Scottish Power), but expanded to include First Utility and Utility Warehouse in year three, and Co-operative Energy in year four

Scope of quantitative analysis and approach

For the evaluation, an analysis was carried out that examined the drivers of indoor temperatures and explore how internal temperatures vary across different segments of the population, including those estimated to be eligible for the WHD scheme. The changes in energy consumption and internal temperatures of WHD scheme recipients were estimated by combining data on temperature, energy use/spending and energy performance, and used reported household expenditure data to estimate the actual change in fuel expenditure among household eligible and in receipt of the WHD.

Three models were developed to address the above research objectives:

- *Impact of WHD scheme*: Information on energy expenditure was obtained from LCFS, and an effect of receiving the WHD scheme was estimated using Regression Discontinuity Design (RDD).

- *Indoor temperature*: Temperature predictors were derived from the 2011 EFUS and projected onto the wider EHS using data on benefit receipt, dwelling type, size, and other estimators.

- *Energy performance*: Prediction of energy performance was derived from EHS using a number of household characteristics, based on size, type of home and relevant variables. A measure of the thermal characteristic (i.e. E-value, see Oreszczyn et al 2006) was estimated for all EHS dwellings as a measure of energy performance. The model was used to predict the energy performance of LCFS dwelling to generate variables for use in the expenditure modelling.

The three models were linked using a range of household and dwelling characteristics that are observable both in the EFUS, EHS and in the LCFS. Below uncertainties are stated and, where possible, quantified.

Report structure

The Quantitative report comprises three parts:

- Analysis of the impact of the WHD scheme on energy expenditure analysed using Regression Discontinuity design (RDD);

- Analysis of the impact of the WHD scheme on indoor environment and wellbeing; and

- Analysis of the effectiveness of the WHD scheme in terms of identifying vulnerable households.

Impact of WHD on energy expenditure

This chapter presents the results of analysis to estimate the impact of the WHD scheme on rebate recipients' expenditure on heating fuel.

Background

This chapter presents the results of the analysis estimating the effect of the WHD scheme on heating fuel expenditure using Regression Discontinuity Design (RDD) within an Engel Curve framework. An Engel Curve is a commonly used economic relationship which, in its most basic form, describes how a consumer's demand for a good varies as a function of income. Engel curves for a certain good describe the good's income elasticity, and consequently enable one to discern whether the good is an inferior, normal, necessary or luxury good, depending on the value of the income elasticity.

The impact of the WHD scheme in this analysis is measured by the deviation from the pre-existing Engel curve. This impact is additional to ordinary income effects related to the fact that demand for energy is responsive to income levels. In other words, given a certain value of the income elasticity, the WHD scheme is expected to produce a corresponding change in the demand for heating as a percentage of the additional income that is allocated by households to raise heating levels. If energy is a normal good, one would expect an increase in energy consumption when income increases, but the effect this has on the *budget share* of energy consumption as income increases depends on whether energy is a luxury or a necessity good.

This analysis tested for impacts of WHD scheme additional to the movements along the Engel curve as income increased as a consequence of the rebate being granted to eligible households. These additional impacts have been described as a 'labelling effect' by Beatty et al (2014) due to the fact that the payments are explicitly aimed at increasing expenditure in heating.[1] As discussed by Beatty et al (2014), however, traditional economic theory states that any unconditional cash transfer, regardless of whether a label is attached, would be treated by households as general income. This report's quasi-experimental framework followed Beatty et al. (2014), testing the economic rationale that, despite labelling on the cash transfer, households are free to allocate it in a way that maximises their welfare. This was accomplished by testing for a discontinuity in the Engel curve at the level where households become eligible for the cash transfer. Any positive discontinuity in the Engel curve for energy at the eligibility threshold for WHD scheme would be taken as evidence for the labelling effect. That is, it would indicate that the unconditional cash transfer which has been labelled 'Warm Home Discount' causes agents to act differently from economic theory and to concentrate expenditure of the rebate on heating.

Although Beatty et al (2014) found statistically significant evidence of this labelling effect for the UK Winter Fuel Allowance, this effect should be re-examined for the WHD scheme for a number of reasons. As is discussed further below, RDD has low external validity, i.e. its findings should not typically be extrapolated beyond a local effect for the specific policy being tested. Furthermore, the robustness of the Beatty et al (2014) findings were considered for a number of reasons, discussed in more detail below.

It is important to stress that, regardless of the robustness of the results from Beatty et al (2014), their findings would not extend to the WHD scheme for the external validity problem mentioned above. This is discussed in greater detail below.

The following is an explanation of the nuances of the Warm Home Discount Scheme and the data set used, the UK Living Costs and Food Survey. The empirical framework and methods are then described, focusing on the logic behind RDD, the precise causal effect to be estimated, and the various ways in which this effect was estimated. The results are then presented with a discussion of implications for the WHD scheme.

Policy background and data

As outlined in Chapter 1, the WHD scheme is a means-tested, unconditional, labelled cash transfer introduced in 2011. It is a one-time payment of £140 paid through the electricity supplier for households which meet the following criteria:

- the electricity provider is part of the scheme;[11]
- the beneficiaries name or their partner's name was on the bill; and
- the beneficiary was also receiving the Guarantee Credit element of the Pension Credit.

The means testing aspect comes from the Pension Credit Guarantee Credit, which is paid to individuals of pension age (currently 60 for women and 65 for men) who receive less than £151.20 per week for single people or £230.85 per week for couples.

The WHD scheme is separate from the Winter Fuel Payment, another labelled cash transfer to qualifying households. The WHD scheme is not paid directly to households, it is a one-off discount on electricity bills, usually paid between September and March. Eligibility for the following winter is determined by a cut-off date in the preceding summer, currently 12 July.

Data used in analysis

The UK Living Costs and Food Survey (LCFS) was used for the years 2009-2013 to analyse the labelling effects of the WHD scheme. At the time of writing 2013 is the most recent year of data

[11] Electricity suppliers included in the scheme at the time of writing are: Atlantic, British Gas, Co-operative Energy, EDF Energy, E.ON, Equipower (Ebico), Equigas (Ebico), First Utility, Manweb, M&S Energy, npower, OVO, Sainsbury's Energy, Scottish Gas, Scottish Hydro, ScottishPower, Southern Electric, SSE, SWALEC, Utilita, and Utility Warehouse.

available. Because the first payments were made in winter 2011, this means slightly more than two years of data was available during which the policy was in force.

The LCFS, which is considered the primary source of household-level expenditure data in the UK, is an annual national survey of approximately 5,000 households across Great Britain and Northern Ireland[12]. The LCFS is a 'repeated cross-section' data set, which means that a new sample of households is used in each year for which data are published[13]..

The LCFS comprises a range of household-level data, from information about the physical structure of the home through to demographic, socio-economic, and income information as well as expenditure data. Data is collected throughout the year, indicated by sampling month in the data. The survey consists of an expenditure diary and an interview, which is conducted face-to-face by the interviewer (see ONS 2009, 2010, 2011, 2012, 2013).

Fuel budget share was used as the dependent variable, i.e. fuel expenditure divided by the household total expenditure. Fuel expenditure includes gas, electricity, and other heating fuels. It should be noted that the dependent variable includes expenditure on fuel for non-heating purposes, but unfortunately it is not possible in the LCFS to separate expenditure for heating purposes from other fuel expenditure. While, in theory, deviation from the Engel curve is still measurable for this bundle of goods any potential labelling impact of WHD scheme would be smaller as a proportion of the budget share than the effect observed in an Engel curve for heating only. This might also affect the power of the methodology and ultimately the ability to identify the effect of the intervention.

Households with reported age under 18 (N=6) were excluded as these were considered to be anomalous and irrelevant to the study, as the age range of interest was close to retirement age. Households with a weekly income of zero (N=50) or over £1000 (N=2,791) were also excluded, the former due to the existence of potential anomalies and the latter for visualisation purposes. Again, these households are irrelevant for the study, as RDD analysis takes place in a small neighbourhood around the income cut-off. Finally, households with fuel share equal to 0% and greater than 40% (N = 2,015) were excluded. There are a number of reasons why a household would have a 0% fuel share, including misreporting, although it was impossible to discern which one would have caused the observation of a 0% fuel share.

Tables 1 and 2 show summary statistics for the data, for both the entire sample and for WHD scheme eligible households respectively. Figure 1 shows the distribution of all households and WHD scheme eligible households across regions and years. Good regional and temporal coverage was important to avoid any time- or region-specific anomalies leading to spurious findings. Figure 1 shows that the coverage is even across time and space. There are

[12]Data from Northern Ireland was excluded as the policy does not extend there.

[13] The sampling procedure is multi-stage stratified random sampling,with weighting to address non-response. However, the weights were not considered for the purposes of this study as a representative sample is not necessary due to our analysis using a subset of the data near the age and income eligibility thresholds for the WHDS.

significantly fewer WHD scheme eligible households in 2011 as the policy only began to make payments in Winter of 2011.

Because the LCFS does not have a Pension Credit Guarantee Credit or WHD scheme beneficiary indicator, an 'eligibility' indicator was constructed using the eligibility criteria for the Pension Credit Guarantee Credit. This ,is an imperfect indicator, as there was no way to ascertain if all qualifying households received the benefit. However, there are two reasons take-up should be assumed as nearly universal for qualifying households. First of all, many households which meet the criteria receive the credit automatically. In addition, reported take-up is greater than 80% and Brewer et al. (2007) indicate that benefit take-up is typically under-reported in surveys.[2] This applies to the core group only, the subject of this portion of the report. The two tables show that the dataset includes 371 WHD scheme-eligible households out of 21,899 total observed households, about 1.7%. As expected, the mean fuel share for WHD scheme-eligible households (about 12%) is significantly higher than the value observed for all households (around 7%).

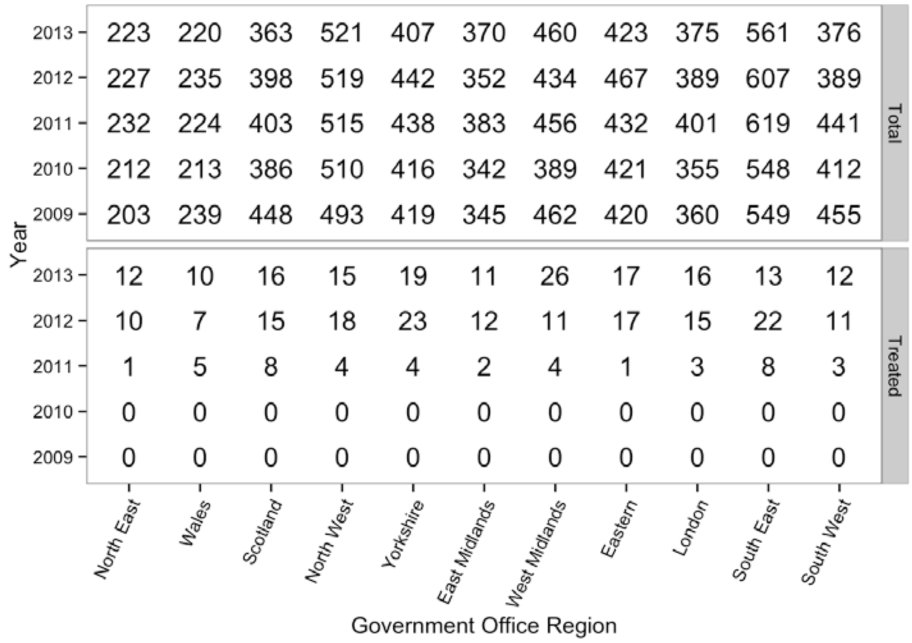
Table 1 – Summary statistics for all LCFS households

Statistic	N	Mean	St. Dev.	Min	Median	Max
Fuel share	21,899	0.071	0.054	0.0003	0.056	0.396
Treated	21,899	0.017	0.129	0	0	1
Raw income	21,899	439.226	231.756	0.040	405.730	999.837
Age	21,899	53.969	16.660	18	54	80

Table 2 - Summary statistics for WHD scheme eligible households

Statistic	N	Mean	St. Dev.	Min	Median	Max
Fuel share	371	0.124	0.080	0.005	0.102	0.396
Raw income	371	144.554	41.335	3.850	143.312	230.304
Age	371	71.399	6.642	60	71	80

Figure 1 - Distribution of WHD scheme eligible and all households across regions and years



Derived Variables

RDD requires a single cut-off in a 'running' or 'forcing' variable, the variable along which there is an arbitrary cut-off which determines treatment assignment. Because the WHD scheme cut-off related to income and age vary depending on whether the household reference person (HRP) is single or in a couple and on their gender, the age and income variables were normalised based on relationship status and gender, respectively. Following standard practice in the literature, the variables were centred on zero so that they can be interpreted as continuous difference from the cut-off point. More formally, for income the derived variable is defined as

$$AdjIncome_i = \begin{cases} income_i - 151.20 & \text{if } single_i = \text{True} \\ income_i - 230.85 & \text{if } single_i = \text{False} \end{cases}$$

and for age,

$$AdjAge_i = \begin{cases} age_i - 65 & \text{if } gender_i = \text{male} \\ age_i - 60 & \text{if } gender_i = \text{female} \end{cases}$$

For example, if the adjusted age is five, age is 70 if the HRP is a male or 65 for a female (negative five would imply 60 and 55 respectively). Accordingly, if the adjusted income is £5, weekly income is £156.20 if the HRP is single or £235.85 if they are in a couple. The centring enables comparison across these different classes with a single cut-off point in each variable, i.e. zero.

The continuous time variable (in monthly intervals) was also centred on zero at the point of the first month of policy implementation, November 2011. This is for ease of interpretation, as all three derived variables now have the cut-off point at zero.

An initial exploration of the data showed little evidence of discontinuity in fuel share when split on age and some evidence when split on income.

Figure 2 shows density plots with dotted lines indicating the mean for fuel share above and below the treatment threshold for each of the forcing variables when the other forcing variables are already over their thresholds.

Figures 3 and 4 show the conditional expectation function (CEF) for frontier RDD within optimal bandwidths before and after WHD implementation. There is some visual evidence of a discontinuity in income after WHD implementation, but little evidence of one for age.

Figure 2 - Fuel share distributions separated by different forcing variable thresholds (within optimal bandwidths) with mean lines

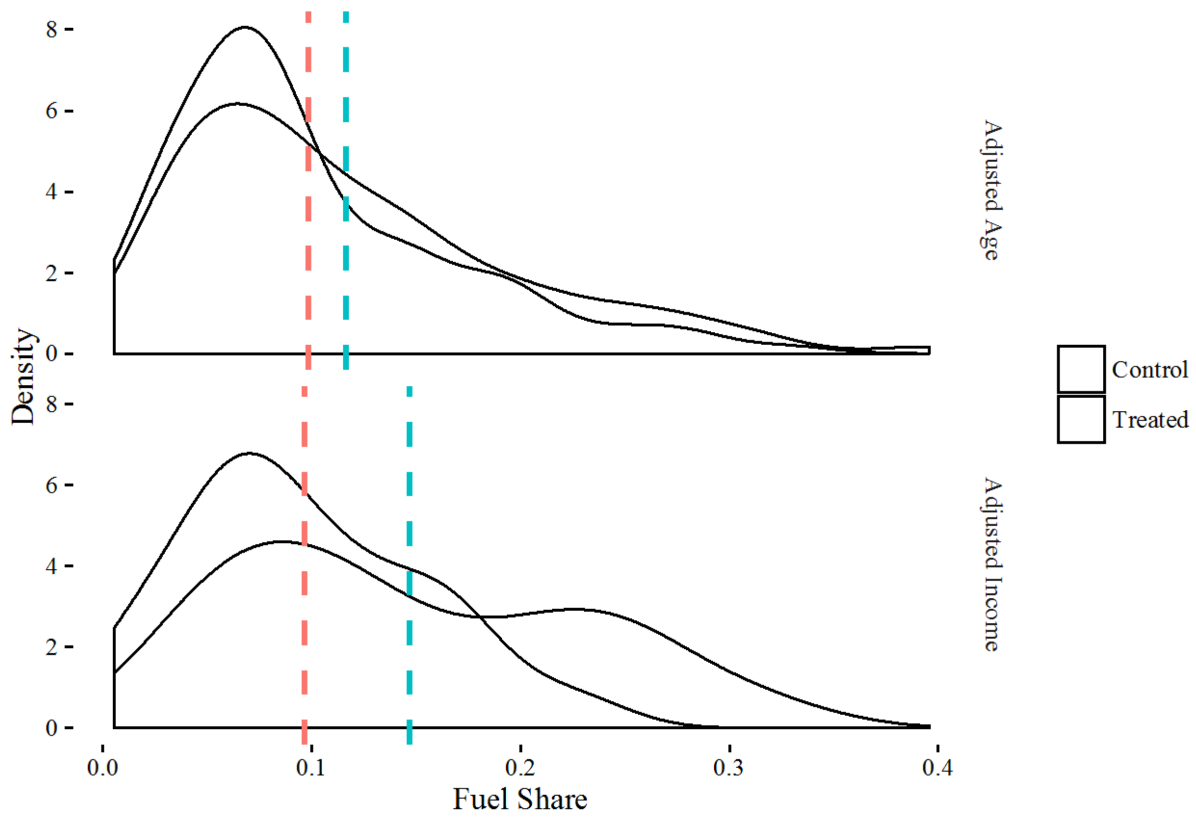


Figure 3 – Conditional expectation functions (CEF) represented by binned means with standard deviation bars

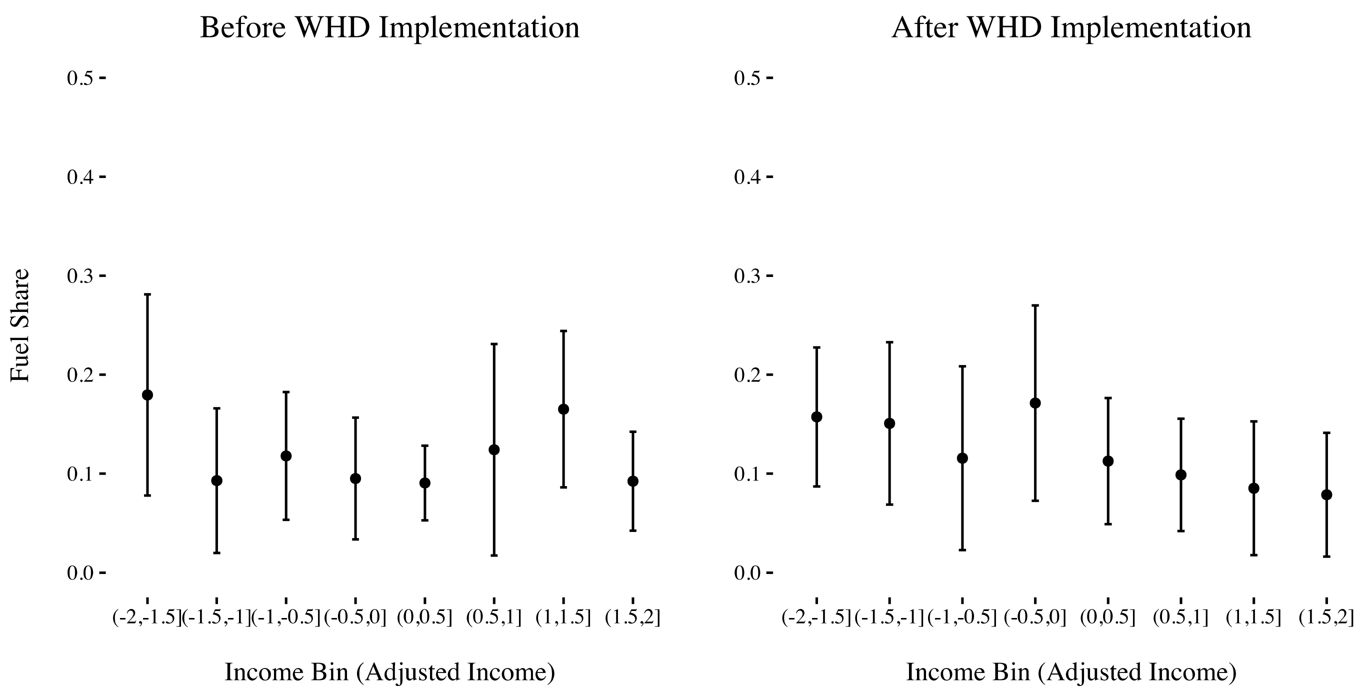
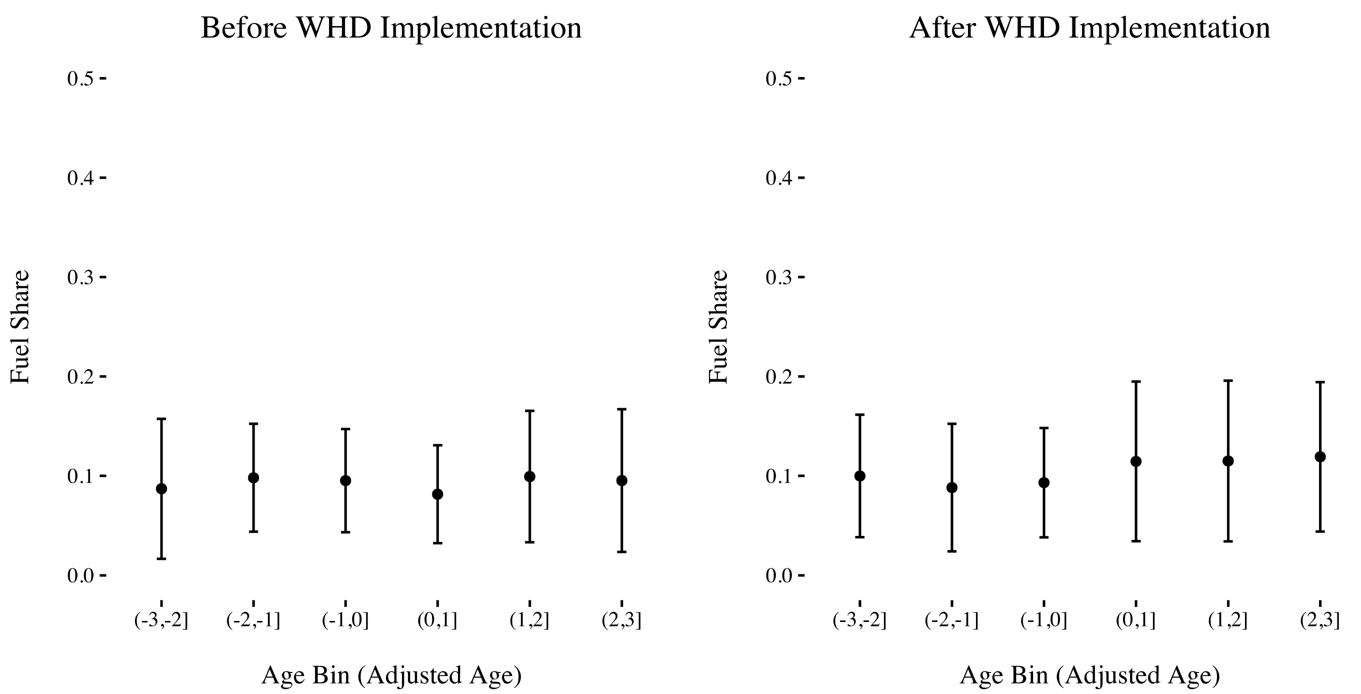


Figure 4 – Conditional expectation functions (CEF) represented by binned means with standard deviation bars



Methods

Regression discontinuity design (RDD)

In order to estimate the labelling effect of the WHD scheme, a Regression Discontinuity Design (RDD) was employed. This is a quasi-experimental method, meaning that observational data was used with statistical methods to make causal inference in a similar way to a randomised controlled trial (RCT). In a RCT, the experimenter randomly assign a treatment to comparable individuals so that they may be considered valid counterfactuals for one another, meaning that the effect of the treatment can be measured directly. Quasi-experimental methods seek to utilise natural sources of randomisation in order to measure the causal relationship between a 'treatment' – in this case the WHDS – and an outcome (fuel budget share). In the case of the WHDS, the measured impact is the difference in fuel budget share between those eligible for the WHDS and their counterfactuals.

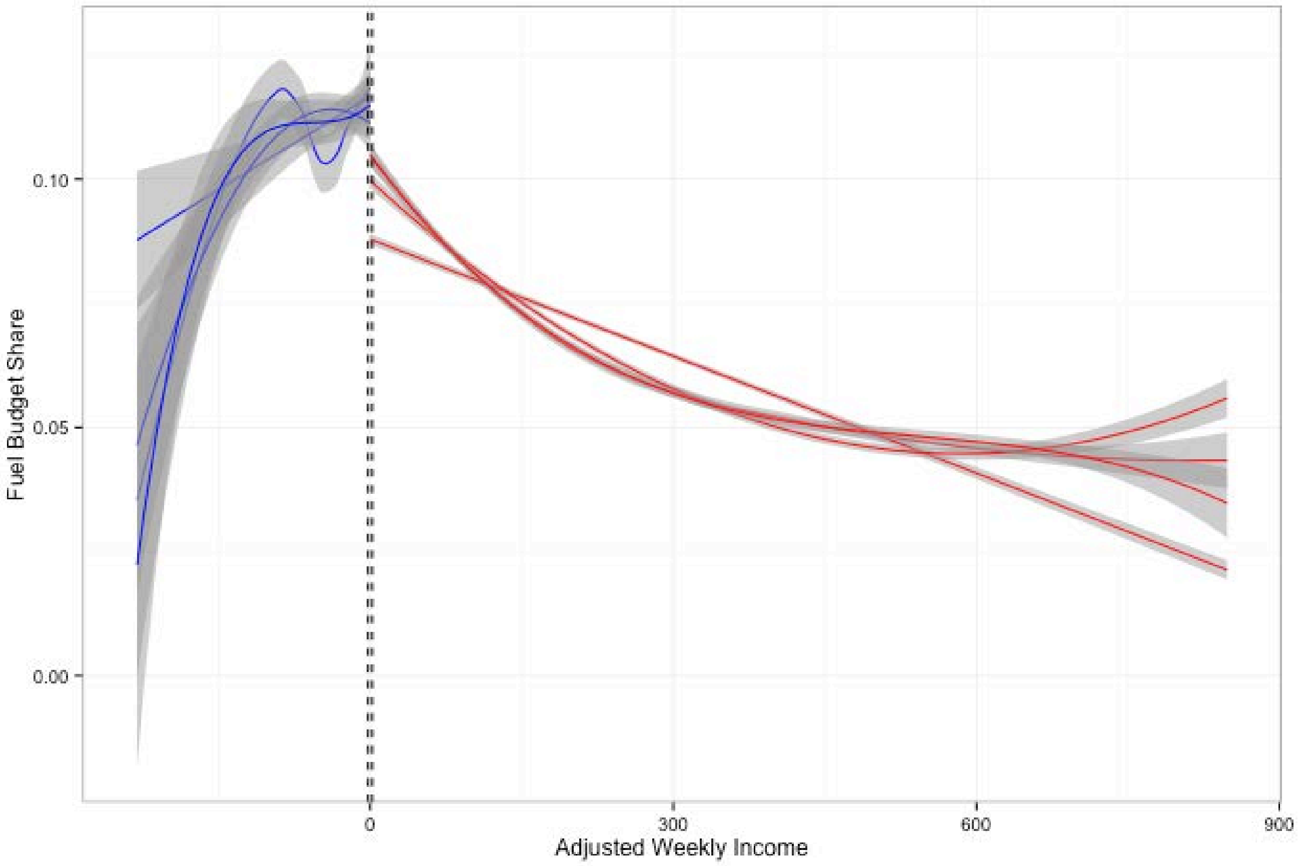
RDD is a particularly reliable method as long as it meets certain assumptions and it is interpreted within its limitations. RDD works by exploiting randomness at an arbitrary cut-off along a continuous variable to identify counterfactuals. A classic example is standardised testing and scholarships, as described in Thistlethwaite and Campbell (1960).[3] If students receive a score above a threshold, they receive a scholarship while if their score falls below a threshold they are not eligible for the scholarship. Reasonably, there is very little difference in students who perform slightly below and slightly above the threshold. This introduces randomisation of assignment through the arbitrary nature of the threshold value and the subject's inability to control the assignment variable around the cut-off (Imbens and Lemieux 2008). The inability to control the assignment variable means that the randomisation of treatment assignment is truly exogenous -- that is, it occurs naturally rather than as an artifice of statistical manipulation. However, this randomisation is only valid within a bandwidth, or a neighbourhood, at the cut-off. The size of this neighbourhood is determined by the data, and an 'optimal' bandwidth can be estimated. These limitations are discussed in greater detail below.

RDD is estimated with ordinary least squares (OLS) using polynomials for the forcing variable(s) or through a semi-parametric approach with local linear regression within a neighbourhood around the cut-off(s). The latter may be more reliable, especially if there is a reason to expect that the relationship between the forcing variable and the dependent variable is non-linear. One of the reasons for this can be seen in Figure 5, which shows estimated fits[14] between adjusted weekly income (a forcing variable) and fuel share. The vertical dashed lines indicate where the optimal bandwidth (i.e. the neighbourhood in which counterfactual comparisons are valid) is. The high degree of uncertainty outside of this bandwidth and the bias introduced from trying to fit the model to data in the tails illustrates why local semi-parametric approaches are considered

[14] Fits include simple OLS, OLS with quadratic and cubic polynomials, and a LOESS smoother.

more reliable. Both approaches were implemented as well as tests for determining the optimal bandwidth around the cut-off points. Given the cross-validation results of the optimal bandwidth calculation, local estimation was relied upon rather than fitting polynomial forms to the entire data set. For this reason, only results from local linear regressions are reported.

Figure 5 - Example of fits between the adjusted weekly income forcing variable and fuel share (Engel curves)



Identification of causal effects

One of the most common approaches to identifying causation is the 'Rubin Causal Model' (RCM), a statistical approach to causal inference derived from the 'potential outcomes' framework (see Rubin and John 2011)[4]. This is a commonly used theoretical approach to causal inference in which the causal effect of an intervention is the outcome for an individual in the state of receiving the treatment and the same individual in the state of not receiving the treatment. The fundamental problem of causal inference is that one can never observe a given individual in these two states simultaneously. Instead, randomised controlled trials or statistical methods are used to identify units similar enough (after controlling for observable factors) so that they can be treated as counterfactuals. RDD accomplishes this by comparing units directly above and below a cut-off value.

Traditionally, in sharp RDD[15], assignment is designated according to a single threshold, where treatment status is a deterministic and discontinuous function of an independent variable X over a known threshold c. The effect of interest is then the Local Average Treatment Effect (LATE) for units with the forcing variable equal to the cut-off value.16

Because very few (or zero) units are exactly equal to the cut-off value, and because comparisons must be made between those receiving the treatment (in this case the WHD) and those who do not, RDD is estimated using regressions within a bandwidth on either side of the cut-off. This bandwidth is determined from the data as described below.

Multiple forcing variables

Means testing introduces an added complexity, which requires generalisation of the RDD framework. This is still an active area of research, but it has been put into practice and there have been some well-regarded methodological papers on the topic (see Reardon and Robinson 2012).[5] In the case of the WHD, two forcing variables are available due to the qualifying mechanism for the Pension Credit Guarantee Credit: income and age. Due to the availability of data before and after the implementation of the policy, this is not collinear with the Pension Credit Guarantee Credit. Three potential measures of the impact of the intervention could therefore be estimated in theory – normally called estimands in the technical literature – in terms of local average treatment effect (LATE):

- the effect at the income threshold for all subjects already over the age threshold;
- the effect at the age threshold for all subjects already under the income threshold; and
- the effect of crossing the multidimensional surface of the union of these thresholds.

Frontier RDD

Frontier RDD simplifies the multiple forcing variable problem described above by only looking at data above all cut-offs but one. For example, individuals can be considered crossing the threshold for income that are past the month and the age cut-offs as described in the first bullet

[15]Sharp regression discontinuity design differs from 'fuzzy' RDD in that the 'sharp' version means that the assignment rule is strict. 'Fuzzy' RDD is where some units may not receive treatment despite being over/under the cut-off. In order to estimate 'fuzzy' RDD, one must have precise data on who receives treatment. Because one must derive the treatment assignment variable according to the assignment rules and because take-up was considered to be very high, 'sharp' RDD was used.

[16] Formally, this is written as $\tau_{RDD} = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = c]$. Following Imbens and Kalyanaraman (2011),[12,13] if one assumes that conditional distributions $F_{Y(0)|X}(y|x)$ and $F_{Y(1)|X}(y|x)$ are continuous in x for all y and that conditional first moments $\mathbb{E}[Y_i(1)|X_i = x]$ and $\mathbb{E}[Y_i(0)|X_i = x]$ exist and are continuous at $x = c$, it follows that the estimand LATE is the difference of two regression functions at the cut-off point: $\tau_{RDD} = \mu_+ - \mu_-$, where e= $\lim_{x \downarrow c} m(x)$ and $\mu_- = \lim_{x \uparrow c} m(x)$.

point above. The estimation procedure can then follow 'normal' RDD estimation using the restricted sample, but with the interpretation that the LATE is local at the cut-off for the specified population. Estimation procedures are described below. An optimal bandwidth was only identifiable for income, but we take reasonable estimates (three and five years) for age and estimate frontier RDD for both income and age.

Response Surface RDD

Response surface RDD allows estimation of the third LATE estimand above, i.e. the effect of crossing the multidimensional response surface at the union of the three thresholds. This is the most 'obvious' way of estimating the LATE. Instead of finding a continuous function for the forcing variable $\mathrm{f}(\mathrm{R_i})$, a continuous function of each forcing variable is identified $\mathrm{f}(\mathrm{R1_i}, \mathrm{R2_i}, \ldots, \mathrm{RJ_i})$.

Box 1 - RDD equation for Warm Home Discount

The following formula was used for Response Surface RDD:

$$y_i = f(age_i, income_i) + \tau D_i + X_i\beta + \epsilon_i$$

where τ is the effect of interest and X_i is a matrix of control variables.

Estimated Models

Following standard practice in the literature, we estimated local linear regression with quadratic terms for the forcing variables (age and income).

For this model, ordinary least squares is a simple linear model of the form $\mathrm{Y_i} = \alpha + \mathrm{D_i}\tau + \mathrm{X_i}\beta + \epsilon$ where $\mathrm{Y_i}$ is the outcome (fuel share), $\mathrm{D_i}$ is a dummy variable indicating treated or not, τ is the causal effect of the WHD scheme, $\mathrm{X_i}$ is a matrix is covariates including the forcing variables, β is a vector of estimated parameters, and ϵ is the error term.

Quadratic (square) terms were introduced for the two forcing variables due to the recognition that the relationships between the forcing variables and the dependent variable are likely to be non-linear. Higher order polynomials than quadratic were not included because these can lead to spurious results (see Gelman and Imbens 2014).[6]

Estimated Specifications

Three different specifications were used to control for sensitivity to the inclusion of covariates. Asymptotically, RDD does not require the inclusion of any other control variables than the forcing variable for unbiasedness. The inclusion of additional control variables may improve the efficiency of the estimator, making hypothesis testing more accurate and less likely to yield type II errors (false negatives). Inclusion of additional control variables is particularly valuable in small samples when statistical power is likely to be an issue, but can also exacerbate that problem by decreasing degrees of freedom. In addition, at smaller sample sizes, the unbiasedness assumption above does not necessarily hold and the inclusion of some covariates may help to avoid biased estimates. The following three specifications were used:

First, in the base specification, the forcing variable(s) and the treatment variable were used, i.e. the dummy variable indicating whether a household is eligible for the WHD scheme or not. Additionally, the total expenditure and its square term were included so that the parameter of interest, the coefficient on the treatment dummy variable, can be interpreted as the deviation from the Engel curve. Under ideal conditions and with enough observations, this specification would give unbiased estimates of the LATE;

Year and region dummies were then included in the second specification to control for cohort effects and regional effects, respectively. The year dummy should account for nationwide factors including the financial crisis and recovery. The regional dummies will control for regional effects invariant across years;

In the third specification, year-region interactions were included to account for regional factors which are not constant across years, e.g. weather.

Limitations and caveats

As discussed by Imbens and Lemieux (2008), it should be noted that RDD can only ever have a very limited degree of external validity. This means that they only provide estimates of the average effect for a subpopulation, i.e. those for whom the forcing variable is equal to the cut-off. This effect can reasonably be extended to those within a small band around the cut-off (the bandwidth) and, indeed, it is on this assumption of homogeneity around the cut-off that RDD relies. In each specification, the forcing variables and their quadratic terms were also interacted with the treatment variable to allow the slopes to vary on either side of the cut-off. In practical terms, this means that the average effect of the policy may be different on a different subpopulation, including those who have received the WHD scheme but who are not near the cut-off.

There were limitations in the data due to uncertainty about when households receive the WHD. Because this can span across the winter and precise data on whether (and when) each household actually received the WHD is unavailable, this means the impact of WHD scheme on energy expenditure was measured with error, e.g. it is possible that energy expenditure which has occurred before the payment took place was included in the estimation.

Additionally, the study’s findings are somewhat constrained by the sample available for the work. It is fundamental to implement the RDD approach with rigorous bandwidth intervals like those suggested by Imbens and Kalyanaraman (2011) in order to claim causality of the WHD scheme for any estimated impact on energy expenditure. However, this implies that the sample used in the estimation is relatively small, as discussed below. This constrains the complexity of models and specifications, in the sense that more complex specifications (like Specification 3) may be less reliable because of degrees of freedom to prevent overfitting.

Results

Determining optimal bandwidths

Determining optimal bandwidths is one of the two most crucial aspects of RDD, along with using the correct functional form for the forcing variable. This is because RDD depends on an assumption of comparability directly above and below the cut-off. The width of the bandwidth determines the maximum distance from the cut-off within which comparability can be assumed. If the bandwidth is too large, RDD becomes questionable, as the economic agents used as counterfactuals are considerably different from those who have received the intervention. However, too small a bandwidth runs the risk of reducing the sample size to the extent that statistical power may be compromised. Imbens and Kalyanaraman (2011) have developed a data-driven method to determine the 'optimal' bandwidth for RDD. Following the Imbens and Kalyanaraman (2011) optimal bandwidth procedure, a bandwidth for income of $\pm$ £2 was identified, but optimal bandwidths for age or month were impossible to calculate due in part to what Dong (2014) calls 'discretisation bias'.[17] An age bandwidth of three and five years was selected to consider the effect of different bandwidths given the uncertainty around the optimal bandwidth for age.

Meeting identification assumptions

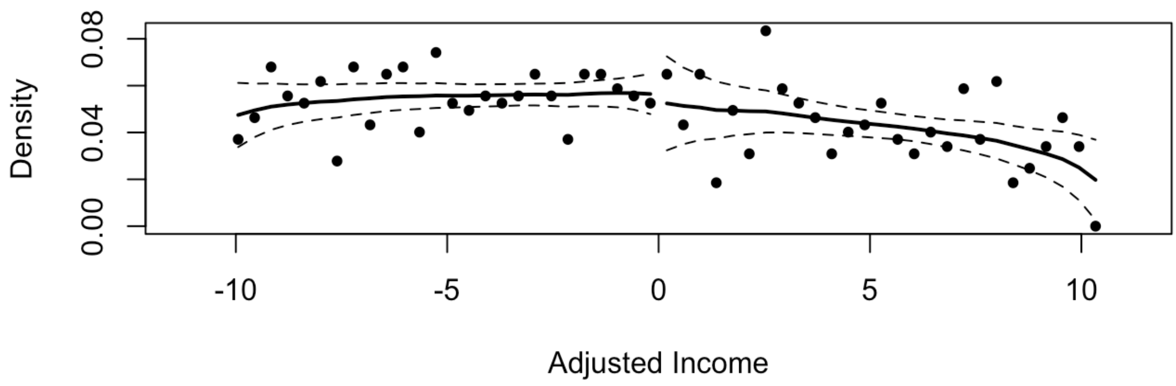
In order to rely on the results in terms of internal validity, the identification assumptions for RDD as discussed above must be met. Fortunately, two of the identification assumptions can be determined from the data, i.e. the assumption of no control over assignment (arbitrary cut-off) and the assumption of no other discontinuities.

McCrary (2008) developed a test to determine whether there is manipulation in the running ('forcing') variable in a RDD.[7] If subjects are able to control whether they fall above or below the cut-off value and do so to be eligible for the intervention, then RDD is invalid because the cut-off does not lead to exogenous randomisation. In other words, there are systematic differences between subjects above and below the cut-off and those just below the cut-off cannot be used as counterfactuals for those receiving the intervention. McCrary's test is a Wald test with the null hypothesis that the discontinuity is zero. The test implies a two-step procedure where in the first step, one obtains a finely gridded histogram and in the second step, one smooths the histogram

[17] Dong (2014) shows 'that standard RD estimation using a rounded discrete running variable leads to inconsistent estimates of treatment effects, even when the true functional form relating the outcome and the running variable is known and is correctly specified' (Dong 2012, p. 422). This is also discussed by Lee and Card (2008), who explain that RDD relies on the continuity of the forcing variable and identification by shrinking the bandwidth to zero at the limit.[15] However, when the observed forcing variable is discrete and rounded, it is not possible to theoretically shrink the bandwidth to zero in order to compare units just above and below the cut-off, meaning that the causal effect cannot be identified. Additionally, it is possible that the optimal bandwidth is less than one year.

using local linear regression, separately on either side of the cut-off (McCrary 2008, 699). Figure 6 shows this smoothing around binned mean points for the adjusted income variable. It is clear that there is no discontinuity, and the Wald test confirms this with a p-value of 0.75. This means that the test failed to reject the null hypothesis of no manipulation.

Figure 6: McCrary sorting test



No other discontinuities

The assumption that there are no discontinuities at the cut-off in pre-treatment characteristics was also tested. This is important because any discontinuities in pre-treatment characteristics would indicate that, for the sample used, observations just above and below the cut-off are not good counterfactuals. This can be interpreted a ‘sorting’ if the variable is readily manipulated or simply as evidence that the covariate is relevant and should be included in the RDD estimation if not.

Tables 3-6 show results from testing this assumption using frontier RDD. Sex of the household reference person, household size, and number of children were tested for discontinuities. Number of children is not shown because all households in the sample had zero children living in the household.

It is found that there is a significant discontinuity in sex of the household reference person within the optimal bandwidth. To account for this, sex of the household reference person was included in the baseline specification.

Figures 7 and 8 show a graphical representation of this relationship within optimal bandwidths.

Table 3 – Results from the Assessment of Existence of Discontinuities in Pre-Treatment Variables. Income Frontier RDD. Bandwidth £2.

	Dependent variable:	
	HRP Sex	Household Size
	(1)	(2)
Treatment	-5.020** (2.277)	-0.321 (0.572)
Observations	103	103
R2		0.101
Adjusted R2		0.014
Log Likelihood	-57.546	
Akaike Inf. Crit.	135.093	
Note:	*=p<0.1;**= p<0.05;***= p<0.01	

Table 4 – Results from the Assessment of Existence of Discontinuities in Pre-Treatment Variables. Income Frontier RDD. Bandwidth £5.

	Dependent variable:	
	HRP Sex	Household Size
	(1)	(2)
Treatment	0.874 (1.238)	-0.477 (0.319)
Observations	265	265
R2		0.053
Adjusted R2		0.019
Log Likelihood	-159.386	
Akaike Inf. Crit.	338.772	
Note:	*=p<0.1;**= p<0.05;***= p<0.01	

Table 5 – Results from the Assessment of Existence of Discontinuities in Pre-Treatment Variables. Age Frontier RDD. Bandwidth 3 Years.

	Dependent variable:	
	HRP Sex (1)	Household Size (2)
Treatment	1.418* (0.779)	-0.097 (0.275)
Observations	397	397
R2		0.037
Adjusted R2		0.014
Log Likelihood	-252.941	
Akaike Inf. Crit.	525.883	
Note:	*=p<0.1;**= p<0.05;***= p<0.01	

Table 6 – Results from the Assessment of Existence of Discontinuities in Pre-Treatment Variables. Age Frontier RDD. Bandwidth 5 Years.

	Dependent variable:	
	HRP Sex (1)	Household Size (3)
Treatment	0.874 (0.669)	-0.477 (0.256)
Observations	265	265
R2		0.053
Adjusted R2		0.019
Log Likelihood	-159.386	
Akaike Inf. Crit.	338.772	
Note:	*=p<0.1;**= p<0.05;***= p<0.01	

Figure 7: Before and after binned CEF plots on pre-treatment characteristics. Income frontier.

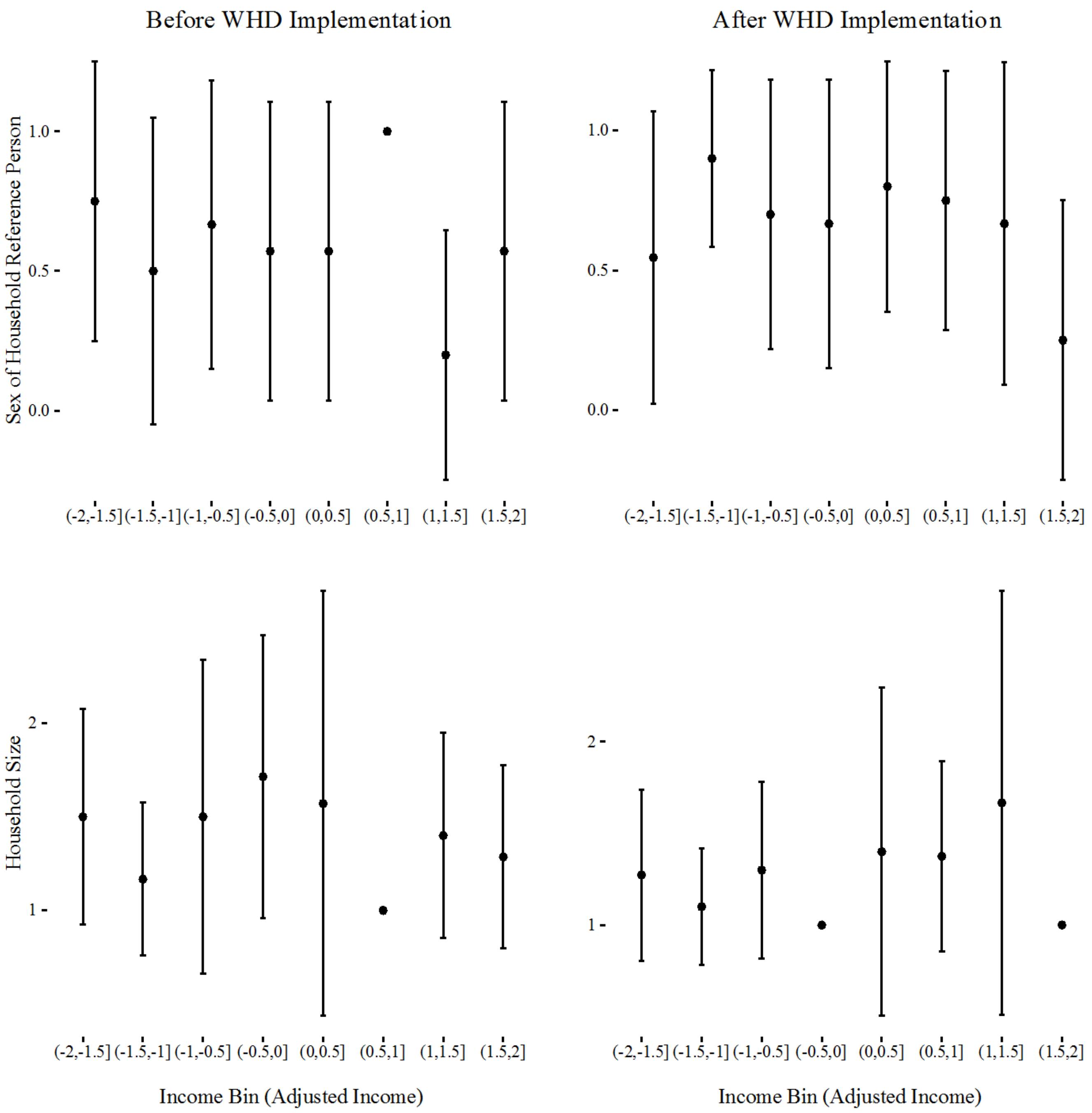
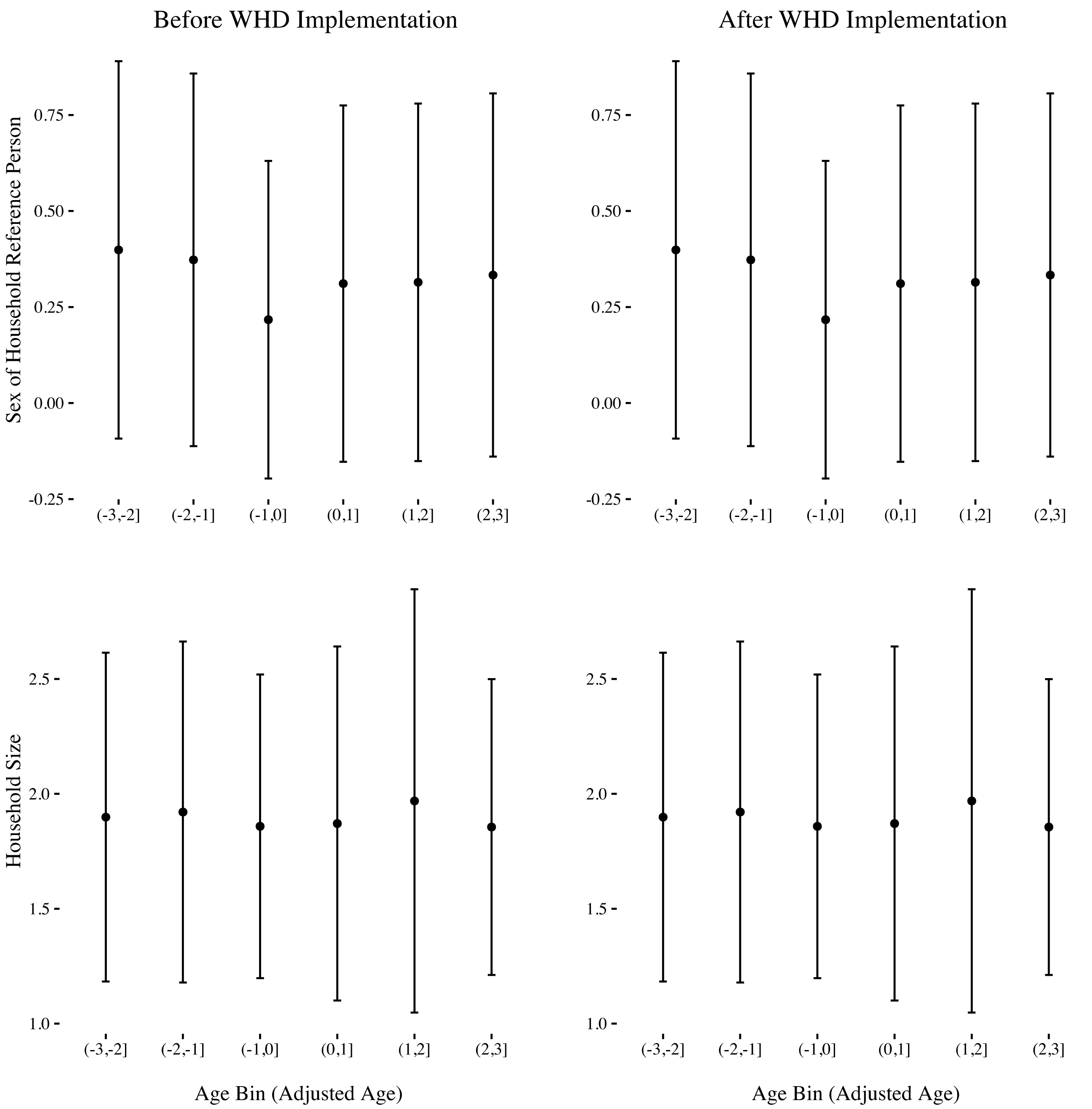


Figure 8: Before and after binned CEF plots on pre-treatment characteristics. Age frontier.



Response Surface RDD

Response surface RDD shows evidence of a labelling effect. By this, it is meant that the energy budget share does deviate from what would be expected with a simple increase in income. Tables 7 through 9 show the results from the models estimated in this study for different bandwidths.

The tables show the estimate of the effect of crossing the multidimensional response surface at the union of the thresholds, as discussed in Section 3.2.

The estimate is interpreted as a percent change in fuel share for being eligible for the WHD versus not being eligible. The magnitude is consistently around a 40% increase but it is also found that the standard error is quite large, around 14% above and below. The effect is consistent in significance and magnitude across specifications and bandwidth selections.

Therefore, it is found that at the multidimensional effect surface, within optimal bandwidths, WHD qualifying households have a greater fuel share than non-qualifying households by approximately 40% (+/- 14%). The translation of this to a pound value is given in the conclusions section.

Table 7 - Response surface RDD. Age bandwidth 3. Income bandwidth 2.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.316* (0.178)	0.259 (0.182)	0.257 (0.182)
Observations	621	621	621
Note:			p<0.1; p<0.05; p<0.01
	Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 8 - Response surface RDD. Age bandwidth 5. Income bandwidth 5.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.419*** (0.138)	0.398*** (0.141)	0.397*** (0.141)
Observations	867	867	867
Note:	p<0.1; p<0.05; p<0.01 Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 9 - Response surface RDD. Age bandwidth 10. Income bandwidth 10.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.301*** (0.113)	0.243** (0.115)	0.242** (0.115)
Observations	1,524	1,524	1,524
Note:	p<0.1; p<0.05; p<0.01 Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Frontier RDD

Tables 10 to 15 show the Frontier RDD, i.e. the effect of WHD scheme when only one of the forcing variables crosses the cut-off. As discussed in Section 3.2, this approach only considers individuals around the threshold for income who are past the age cut-off for income RDD and below the income cut-off for age RDD.

Despite an increased sample size and a local average treatment effect that is simpler to estimate, Frontier RDD only shows a significant effect in a small minority of bandwidth and specification combinations. However, the magnitude of the effect is similar to surface RDD and is consistent across models.

It is found that income frontier RDD is significant if the bandwidth is expanded to £10 on either side of the cut-off. Unlike age, where a span of 10 years on either side of the cut-off corresponds to very different household types, a maximum difference in weekly income of £20 can reasonably be considered not to be associated with major lifestyle differences. Therefore, this is corroborating evidence to the surface RDD, and the magnitude of the effect is very similar.

Table 10 - Income frontier RDD. Income bandwidth 2.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.563 (0.578)	0.446 (0.587)	0.436 (0.592)
Observations	103	103	103
Note:			p<0.1; p<0.05; p<0.01
	Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 11 - Income frontier RDD. Income bandwidth 5.

	Dependent variable:		
	ln(Fuel Share)		
	Specification 1	Specification 2	Specification 3
	(1)	(2)	(3)
Treatment	0.276	0.322	0.333
	(0.337)	(0.346)	(0.347)
Observations	265	265	265
Note:			p<0.1; p<0.05; p<0.01
	Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 12 - Income frontier RDD. Income bandwidth 10.

	Dependent variable:		
	ln(Fuel Share)		
	Specification 1	Specification 2	Specification 3
	(1)	(2)	(3)
Treatment	0.467**	0.474**	0.476**
	(0.225)	(0.232)	(0.232)
Observations	513	513	513
Note:	p<0.1; p<0.05; p<0.01		
	Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 13 - Age frontier RDD. Age bandwidth 3.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.294 (0.198)	0.229 (0.205)	0.225 (0.205)
Observations	397	397	397
Note:	p<0.1; p<0.05; p<0.01 Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 14 - Age frontier RDD. Age bandwidth 5.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.316* (0.178)	0.259 (0.182)	0.257 (0.182)
Observations	621	621	621
Note:	p<0.1; p<0.05; p<0.01 Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Table 15 - Age frontier RDD. Age bandwidth 10.

	Dependent variable:		
	ln(Fuel Share) Specification 1 (1)	Specification 2 (2)	Specification 3 (3)
Treatment	0.318** (0.137)	0.219 (0.141)	0.204 (0.141)
Observations	1,096	1,096	1,096
Note:			p<0.1; p<0.05; p<0.01
	Specification 1 includes adjusted income, adjusted age, and their squared terms. Specification 2 adds year and region dummies. Specification 3 adds year and region dummies and their interactions.		

Winter Fuel Payment Analysis

In order to test the robustness of the findings and compare the mechanisms of cash transfer, the study from Beatty et al (2014) was replicated in this analysis and extended by incorporating best practice from RDD literature. While the results in Beatty et al (2014) show a statistically significant labelling effect, this report found that statistical significance disappeared closer to the ‘optimal bandwidth’ as described in Imbens and Kalyanaraman (2011). This corroborates our analysis of the WHD and indicates that the labelling effect found in Beatty et al (2014) is not robust.

Building on a replication of the work in Beatty et al (2014), the results in Table 16 show what happens to the treatment estimates as the bandwidth shrinks towards the optimal level – loss of significance and a switched sign at the three year bandwidth. It is important to consider that as the bandwidth increases past the optimal size, validity of the study decreases as units on the other side of the cut-off threshold become less comparable to those receiving treatment. The WFP study was replicated exactly, except for choosing not to log transform the dependent variable and running variables. There seems to be no theoretical reason to do so and it raises a slight issue in replication, as it is unclear how Beatty et al (2014) dealt with zero values in the dependent variable as the natural logarithm of zero is undefined and there are zero values in the data.

Table 16 – Results for the WFP analysis when implementing optimal bandwidth

	15 year bandwidth	5 year bandwidth	3 year bandwidth
Estimated Treatment	0.004** (0.002)	0.005 (0.004)	-0.0001 (0.007)
Observations	14024	5211	3338
R2	0.286	0.254	0.271
Adjusted R2	0.28	0.239	0.247
Residual Std. Error	0.034 (df = 13915)	0.034 (df = 5102)	0.035 (df = 3229)
F-statistic	51.589*** (df = 108; 13915)	16.127*** (df = 108; 5102)	11.134*** (df = 108; 3229)

Note: *=p<0.1; **=p<0.05; ***=p<0.01

Conclusions

The analysis finds some evidence of a labelling effect of the WHD scheme in the data. This effect is not robust across bandwidth sizes and specifications, but the consistent magnitude weighs in favour of a real effect. RDD requires strict assumptions for validity and these assumptions reduce sample sizes, which can reduce the power of statistical tests and this may explain a true effect that is difficult to identify.

While the lack of a robust finding may be due to the small sample size, if there is truly no effect, this would corroborate standard economic theory, that unconditional cash transfers are treated as income regardless of label. This would indicate that the policy is economically efficient, as it means that economic agents are allocating income in a way that maximises welfare. That is, households are not disproportionately spending this income on fuel in a way which would crowd out other household expenditures such as food or clothing.

Beatty et al (2014) contradict that theory but building on their framework, this report found that the Winter Fuel Payment labelling effect estimated by Beatty et al (2014) is not robust to model specification or a reconsideration of the design assumptions (i.e. the ‘bandwidth’, or range of data around the eligibility cutoff).

It is possible that measurement error made it difficult to determine an effect, as which households received the WHD and when was impossible to determine precisely from the data, though the proxy eligibility measure can be considered accurate as described above. An additional complication is the fact that age is measured as a discrete variable (which introduces bias for RDD) but this does not apply to the Frontier RDD estimates, as they do not rely on a neighbourhood for age. The results showed a mixed inability to reject the null hypothesis of no labelling effect across several model types and estimation methods. The results are therefore inconclusive, but with some evidence of a labelling effect.

If the significant findings are true and there is a labelling effect, this would be measured as a deviation from the Engel curve. Taking the effect in the base specification of the surface RDD, this would be approximately 40% (+/- 14%). For a marginal increase in income of £140 for the year (or £2.70 per week), the average household in the sample would spend £0.21 per week (or £10.92 per year) if not qualifying for the Winter Fuel Payment. For WHD qualifying households, those amounts are £0.26 per week or £13.52 per year.

Impact of WHD on indoor environment and wellbeing

This chapter presents the results of analysis of the impacts of WHD on the indoor environment and wellbeing of rebate recipients.

Background

The temperature experienced by households is a combination of their choices and needs, which are constrained by their ability to afford certain indoor temperature conditions, the capacity of the heating system performance (including the building), and other institutional drivers (e.g. social housing heating arrangements, or family members). Logically, households seek to achieve a temperature that meets their comfort needs and in any occurrence of a deficit in their comfort, a process that makes energy cheaper would result in actions that seek to reduce this comfort deficit. This is more commonly known as a temperature rebound or 'take-back' (Hamilton et al., 2011).[8]

Bearing this in mind, some customers receiving the WHD scheme may take some benefit in the form of improved warmth within their homes and therefore an increase in fuel consumption, rather than solely reducing overall energy expenditure, by maintaining their current fuel consumption. In both circumstances additional income would be available to the household for either expenditure or savings.

Changes in temperature have been shown to have an appreciable health impact in England (Wilkinson et al., 2001).[9] Their impact have been quantified using an exposure-response model (Hamilton et al., 2015).[10] A rise in indoor wintertime temperatures can result in a reduction in the risk of cardio-respiratory diseases (along with asthma for children) and mental health.

Method

The potential impact of the WHD scheme on indoor temperatures was estimated using an established exposure-response model methodology (Hamilton et al., 2015). The Health Impact of Domestic Energy Efficiency Measures (HIDEEM) model was modified and expanded to examine the potential impact that fuel rebates would have on temperatures and the corresponding impact on occupant health over a specified period.

The modelling comprised:

- Identifying the target groups eligible for the WHD scheme, both for the Core and Broader Group;

- Applying the WHD to those eligible households;
- Estimating the change in indoor temperature during wintertime conditions due to the rebate;
- Estimating the corresponding impact on health for those eligible households

Details of HIDEEM

The HIDEEM model is a building physics model that characterises the indoor environmental conditions of English houses for indoor winter temperatures, exposures to particle pollution, secondhand tobacco smoke (STS), radon, mould growth and energy demand in relation to the energy performance of the dwelling. The model estimates the change in indoor conditions due to energy efficiency or changes in energy expenditure. These changes in conditions are used to estimate health impacts based on a combination of life table methods and directly modelled changes in disease prevalence. More details can be found in Hamilton et al. (2015). In this work, only changes in temperature and mould are considered.

Modelling fuel bill rebates and temperature impact

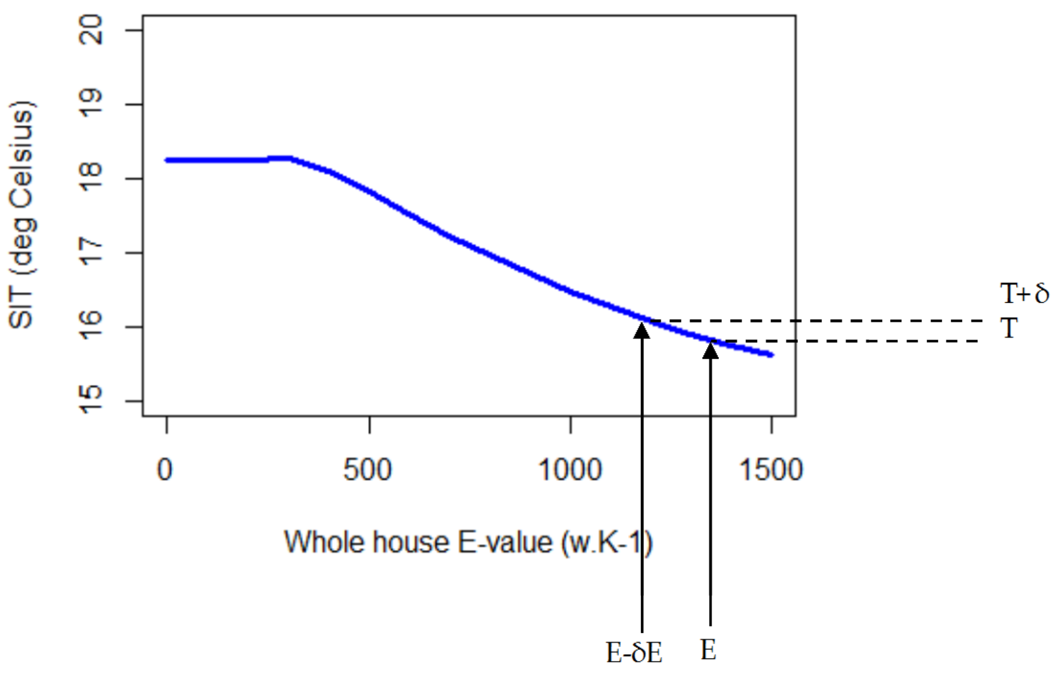
To model the impact of fuel rebates on temperatures, the relationship between energy performance (i.e. E-value18) and indoor temperatures assuming 5 °C outdoors is used - known as the standardised internal temperature (SIT). This relationship shows that dwellings with high E-values, i.e. the least energy efficient homes, have the lowest indoor temperatures (SIT), and that temperatures increase approximately linearly as E-values becomes smaller, i.e. improvement in energy efficiency. The SIT reaches a plateau at E-values near 250 W/K (approximately the stock average E-value). This plateau is used as a basis for making an informed assumption that this is a temperature which the average householder living in a reasonably energy efficient home considers sufficient for comfort.

The curve shown in Figure 11 is used as an indirect indication of householder behaviour. Using the temperature ‘plateau’, the model assumes that householders who are already at the energy efficient end of the E-value spectrum (below around 250 W/K) would not use any share of any fuel rebate to increase internal temperature. The model makes the assumption that the degree to which households heat their home depends on the E-value. In circumstances where the home is relatively energy inefficient (and thus heating costs relatively high), households would on average maintain a lower SIT. The model also assumes that the primary determinant for a

[18] E is the required energy consumption by the principal heating device to maintain a one degree Celsius temperature difference between outside and inside during steady state conditions ignoring incidental gains and ventilation heat losses. $E = \Sigma i(Ui!Ai)/\mu$, where Ui is the heat loss per square meter of surface area per degree temperature difference between inside and outside (W/m2!K) for the ith building element, Ai is its surface area, and μ the efficiency of the main heating device for the dwelling. W is Watts and K is degrees Kelvin.

lower temperature is household choice (based on cost) rather than the physical limitations of the heating system.

Figure 11 – Standardised indoor temperature (SIT) against whole house E-value based on the empirical data function described in Hamilton et al. (2011) and implemented NICE EWD review



The model uses the direct correspondence between E-value and heating cost, with a fixed indoor temperature (SIT) and the same mix/cost of energy sources. The model assumes the households raise the temperature to the maximum they can along the SIT to E-value curve. Households that are below the 250 W/K E-value threshold will use a proportion of the fuel subsidy to increase temperature (with the remainder going to general household income), while households above the 250 W/K threshold would not use any share of the payment to increase temperature therefore the whole amount of the rebate is available to reduce fuel costs.

Household eligibility

Households eligible for the receipt of the Core Group WHD were those defined within the English Housing Survey (EHS) as being in receipt of the Pension Credit. Within the 2012 EHS there were approximately 1.2 million households in receipt of the credit. Households eligible for the receipt of the Broader Group WHD could not be directly identified within the EHS due to the variation in the supplier definitions of the Broader Group. Therefore, as a proxy for those households that were deemed to be comparable to potential Broader Group eligibility were defined as:

- State Pension Credit
- Low income (< £16,190 per year)

- Income benefits (i.e. Income Support, Income Related Employment Support Allowance, or Income Based Job Seekers Allowance) and has a child under the age of 5 or is in receipt of a Disability Benefit (i.e. Long Term Incapacity Benefit or Severe Disablement Allowance or Disability Living Allowance)

Using the above selection criteria, approximately 7.5 million households were identified within the EHS as being eligible for the WHD. Note that it is not possible from the data to determine whether these households would have been eligible or received the WHD as the criteria depends on the household's energy supplier.

Household fuel rebate

The model estimates the effect of a households' receipt of an additional £140 available for expenditure on fuel. Using the relationship between SIT and E-value, the model assumes that households first spend the available amount in meeting their thermal comfort deficit, i.e. increase temperatures, and after their needs are met they allocate the remainder to general household expenditure. If they are not in thermal comfort deficit, the whole amount of the rebate is assumed to be available for general household expenditure. Note that only the addition of £140 is estimated for both Core Group and the Broader Group.

Results

The results are presented separately for the Core Group and the Broader Group analysis. Each set of results show the environmental conditions, the effect on thermal comfort conditions and the corresponding impact on health of being in receipt of an additional £140 is estimated.

Core Group

Table 17 shows that the eligible WHD Core Group tend to live in more energy efficiency dwellings (i.e. lower fabric and ventilation heat losses), resulting in slightly warmer predicted wintertime temperatures. Figure 12 through Figure 14 show that those eligible for the WHD within the Core Group tend to live in new dwellings (i.e. 1945-1980) and in purpose built and mid-terrace dwellings. They also tend to be living in Local Authority or registered social landlord properties, which historically tend to be built to a higher energy performance standard than private dwellings and may have different heating system operations and provision of heat to the dwelling (e.g. centralised heating). The effect of providing an additional £140 to WHD Core Group households is computed to be an average increase of 0.16°C during wintertime conditions. This increase in indoor wintertime temperature suggests that there is a moderate thermal comfort deficit, though some household will see a greater rise due to their living in less efficient dwellings and therefore more likely to have a predicted lower temperature. The potential increase in indoor wintertime temperature is expected to reduce household risk of cardiovascular disease, heart attack, stroke, asthma and common mental disorder. The overall change in health resulting from the increase in predicted temperature is a benefit of 6,093 Quality Adjusted Life Years (QALY) over a 15-year period (i.e. average of 65+ year olds), with a mean benefit of 50.3 QALY/10,000 persons. The improvement in health assumes the temperature increase is maintained over the period. If monetised these benefits could result in approximately £150 per capita of undiscounted health benefit over the 15-year period for the Core Group, assuming a value of £30,000 per QALY.

Table 17 – Table of indoor thermal conditions for English housings stock and WHD scheme eligible group

	Whole Stock	WHD scheme Core Group
	Mean values	
Households (N)	20,789,871	1,210,129
Fabric heat loss (W/K)	300	247
Ventilation heat loss (W/K)	76	60
E-value (W/K)	515	403
Permeability (m3/m2/hr)	16	15
No. Air Change per Hour	0.9	0.9
Average Standardised Internal Temperature (°C)*	17.7	17.9
Heating energy demand (kWh/yr)	23,430	19,300
*Temperature standardised to 5C external		

Table 18 – Mean change in temperature and heating energy demand for WHD scheme eligible households

	Mean change
SIT (°C) – Before	17.94
SIT (°C) – After	18.10
Δ SIT (°C) – Change	0.16

Figure 12 – Dwelling age for WHD scheme Core Group eligible and not eligible households

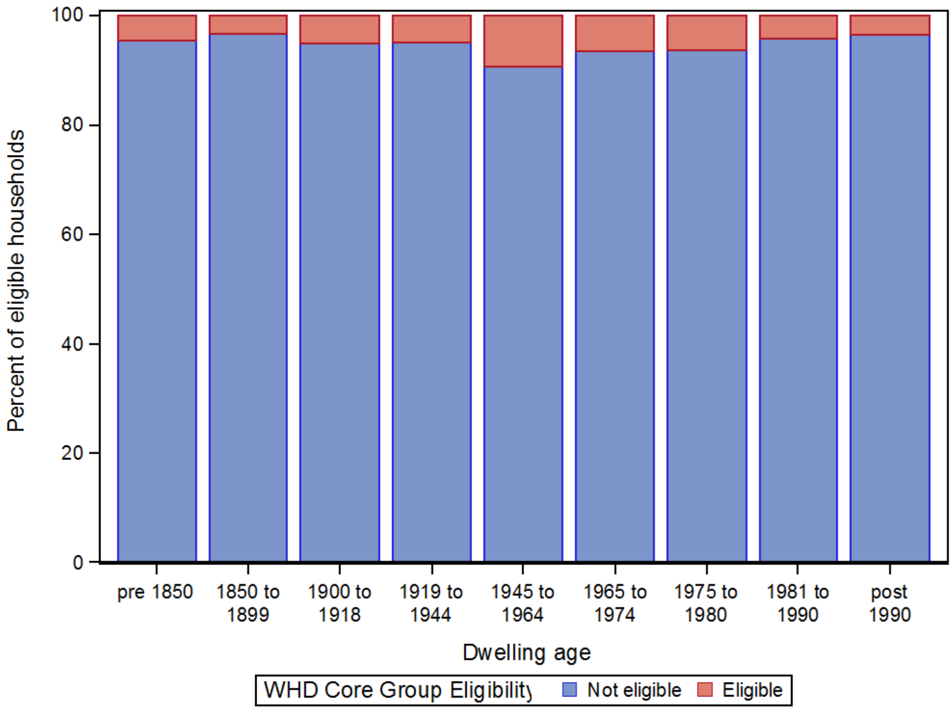


Figure 13 - Household tenure in WHD scheme Core Group eligible and not eligible households

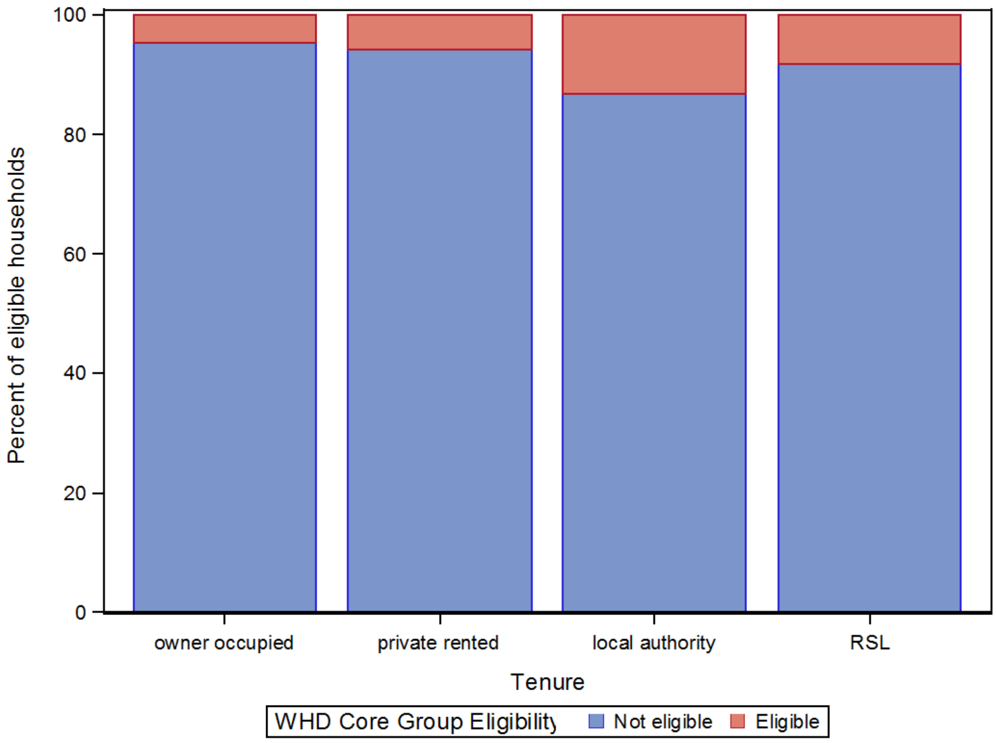


Figure 14 - Dwelling type for WHD scheme Core Group eligible and not eligible households

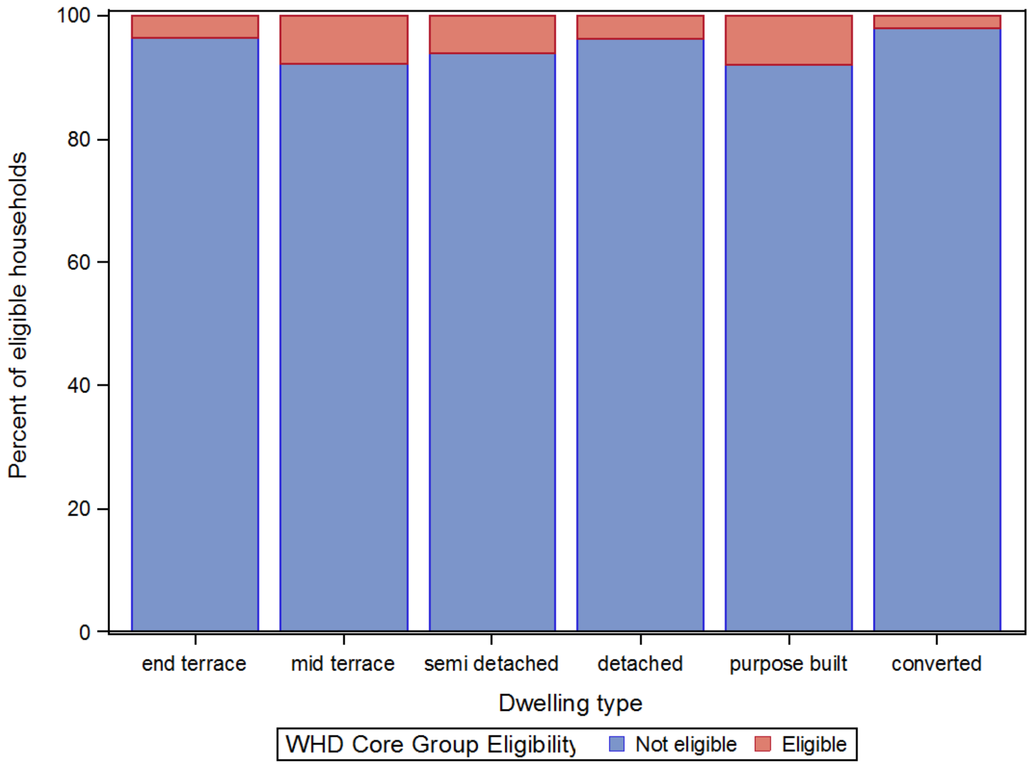


Table 19 – Potential health and wellbeing (mortality and morbidity) change for WHD scheme eligible households

	Potential Health Impact of WHD scheme over 15 year period	
	Total QALYs	**Mean QALY per 10k persons**
Cardiovascular (winter)	3,847	31.7
Heart attack	973	7.9
Stroke	1,127	9.4
Cardiopulmonary	0	0
Lung cancer	0	0
Common mental disorder	146	1.2
Asthma (children)	0	0
Net impact	6,093	50.3

Broader Group

As with the Core Group eligible households, Table 20 shows that the eligible WHD Broader Group tend to live in more energy efficiency dwellings. The average indoor temperature is 0.3°C warmer than those not eligible. Figure 15 through Figure 17 show that those eligible for the WHD scheme within the Broader Group tend also to live in mid-20th Century and newer dwellings (i.e. 1945-1980) and mainly in purpose built and converted flats. Fewer households live in detached dwellings. These dwelling features will play a role in the prediction of higher than non-eligible households wintertime temperatures.

The Broader Group tend to comprise a slightly higher proportion of households living in Local Authority rented dwellings than the Core Group, but also in privately rented dwellings, which corresponds to the converted flats dwelling type. The impact of providing an additional £140 to the WHD scheme Broader Group eligible households would be an average increase of 0.18°C during wintertime conditions, which is approximately the same as the Core Group. However, the implications for health offers a slightly different picture than the Core Group. There is an overall potential net benefit for health of 33,065 QALYs during the 15-year follow up period, reflecting the larger pool of treated households but the per capita benefit is 36.3 QALY/10,000 persons, which is lower than the Core Group.

Although the change in predicted temperatures are approximately the same, the Broader Group will consist of younger households that are less likely to be suffering from disease and whose health benefits may accrue after the 15-year follow up period used in the modelling. As with the Core Group, these benefits assume the temperature increase is maintained for the duration of the modelling period. If monetised these benefits could result in approximately £132 per capita of undiscounted health benefit over the 15-year period for the Broader Group, assuming a value of £30,000 per QALY.

Table 20 – Table of indoor thermal conditions for English housings stock and WHD scheme eligible group

	Whole Stock	**Intervention Group**
	Mean values	
Households (N)	14,414,767	7,585,233
Fabric heat loss (W/K)	325	244
Ventilation heat loss (W/K)	83	59
E-value (W/K)	561	409
Permeability (m3/m2/hr)	17	15
No. Air Change per Hour	0.9	0.9
Average Standardised Internal Temperature (°C)*	17.6	17.9
Heating energy demand (kWh/yr)	25,180	19,540

*Temperature standardised to 5°C external

Table 21 – Mean change in temperature and heating energy demand for WHD scheme eligible households

	Mean change
SIT (°C) – Before	17.94
SIT (°C) – After	18.11
Δ SIT (°C) – Change	0.18

Figure 15 – Dwelling age for WHD scheme Broader Group eligible and not eligible households

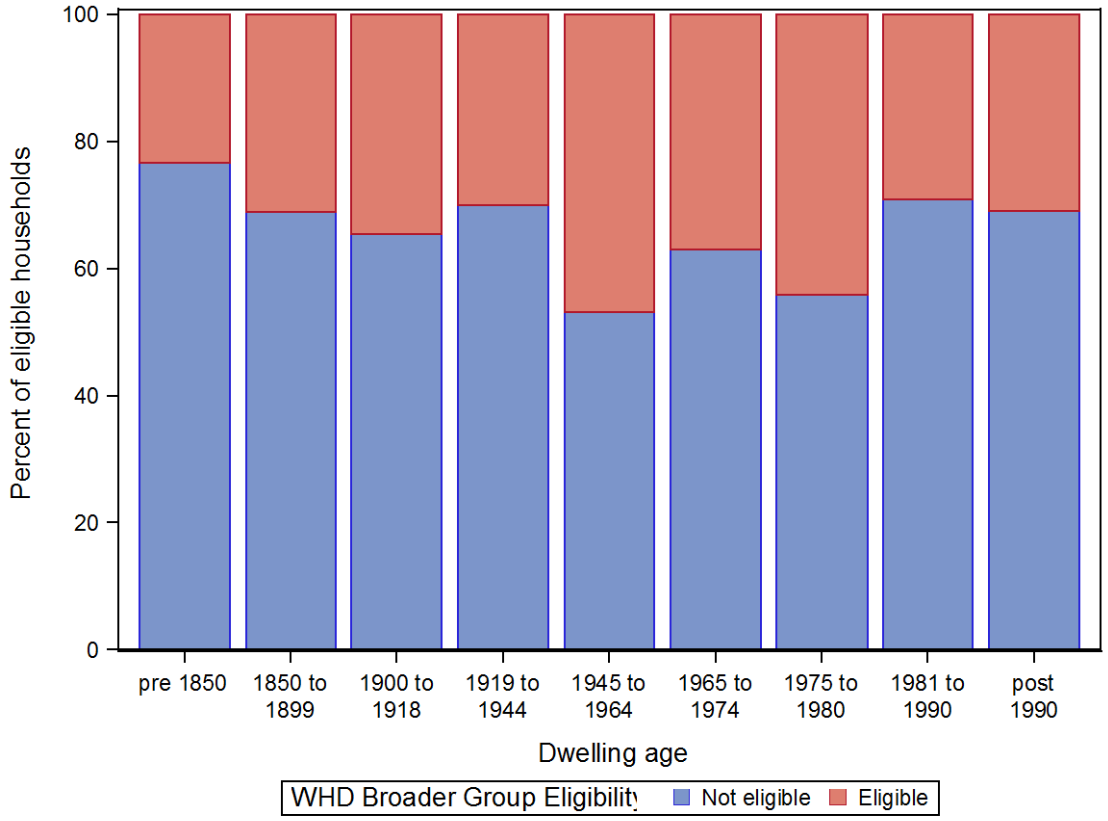


Figure 16 - Dwelling tenure for WHD scheme Broader Group eligible and not eligible households

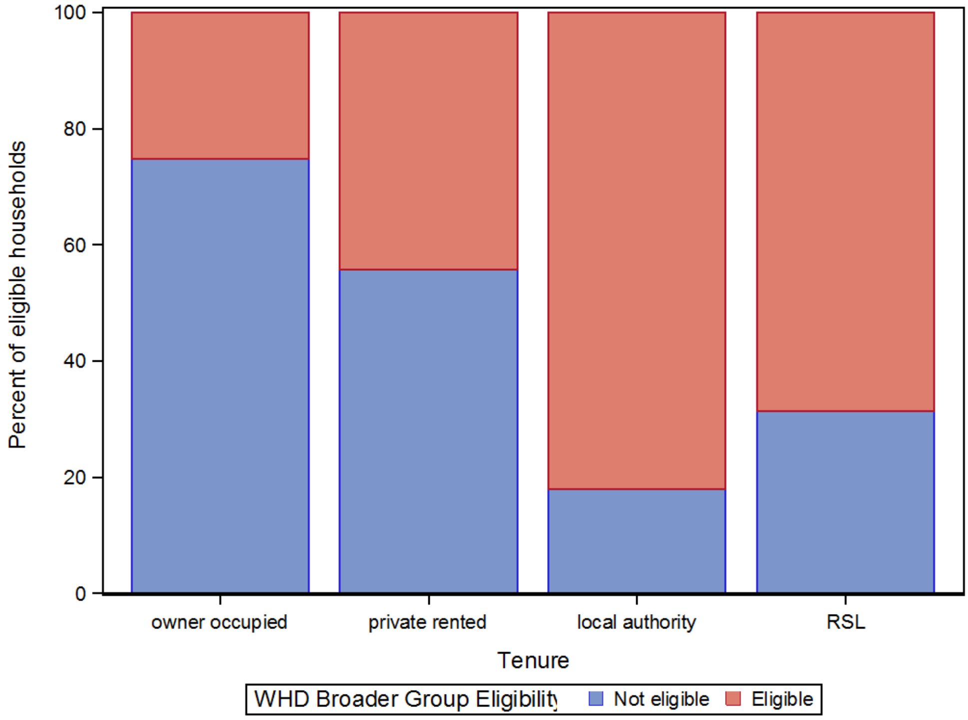


Figure 17 - Dwelling type for WHD scheme Broader Group eligible and not eligible households

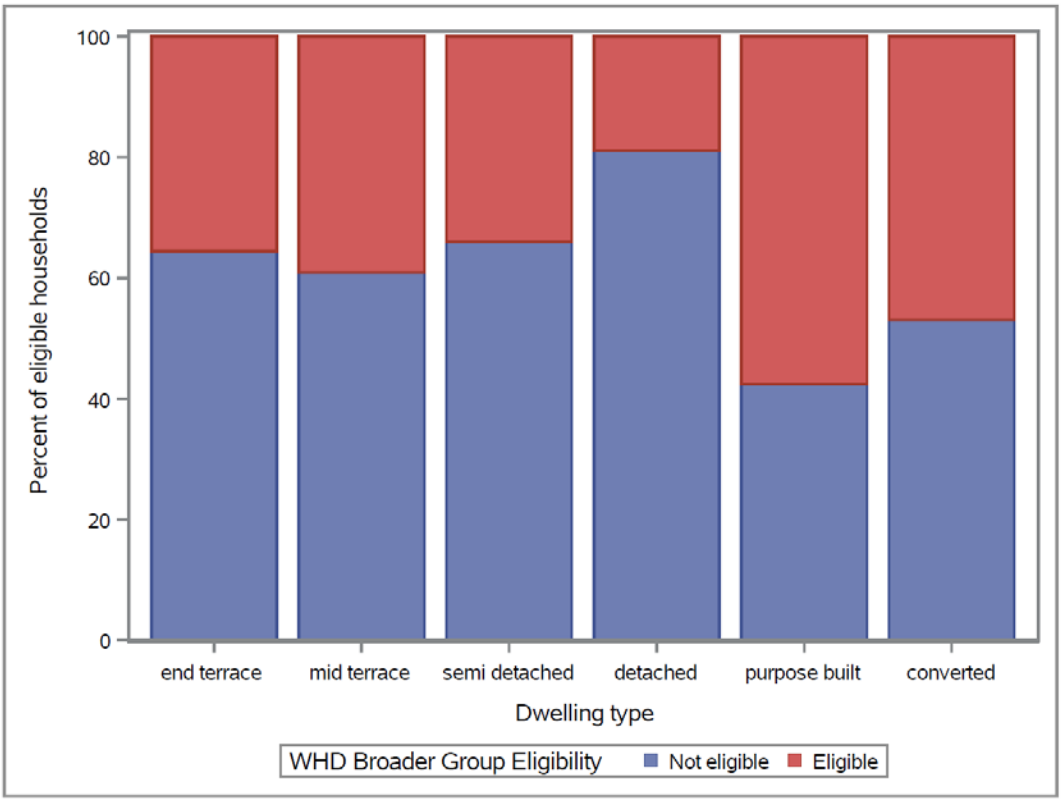


Table 22 – Potential health and wellbeing (mortality and morbidity) change for WHD scheme eligible households

	Potential Health Impact of WHD scheme over 15 year period	
	Total QALYs	**Mean QALY per 10k persons**
Cardiovascular (winter)	20,796	22.8
Heart attack	5,495	6
Stroke	5,779	6.3
Cardiopulmonary	0	0
Lung cancer	0	0
Common mental disorder	994	1.2
Asthma (children)	0	0
Net impact	33,065	36.3

Conclusions

WHD scheme eligible households are predominantly already warmer than their non-eligible households, reflecting the type of dwelling they live in, i.e. newer dwellings that are generally flats and socially rented, i.e. local authority and registered social landlord. On average, the WHD scheme of £140 offered a small change in predicted indoor wintertime temperatures of approximately 0.17°C when outdoor conditions are at 5°C. The increase in predicted wintertime temperatures due to the WHD scheme payments may provide some improvements in cardio-respiratory related health and mental health for those eligible households. The benefits tend to be on average greater among the Core Group than the Broader Group due to their age and their existing underlying risk of diseases affected by temperature. There is a moderate value to these health benefits over the 15-year modelling period of £132 and £150 related directly to the improvement in health. The cost of providing the discount over the 15-year period is approximately £1200 per capita (assuming a 1.75-person household size). However, these figures do not account for potential NHS savings, which could result in further social health cost benefits.

Identifying vulnerable households

This chapter considers how well the Core and Broader Group definitions target the fuel poor, and whether other variables might provide a more robust way of identifying vulnerable households.

Background

The aim of the WHD scheme is to address fuel poverty amongst those households that are at highest risk and severity, defined under the WHD scheme as those being in receipt of Pension Credit, Guarantee Credit. However, the risk of fuel poverty and of living in cold homes may not fully align with the eligibility criteria of WHD. For example, the latest EHS shows that only approximately 20% of households over the age of 60 fall in the bottom two quintiles of household income, and only approximately 8% of these households live in socially-rented housing, which is very likely to be of a higher performance standard. Therefore, the risk and severity of fuel poverty is likely to differ across the recipient group.

This section discusses how well definitions of Core and Broader group target the fuel poor, the neediest part of the fuel poor group and a new proxy which might be developed to better target fuel poor customers under LIHC. A Random Forest classification approach was developed to determine important variables in identifying households that were at risk of living within cold homes (defined as having a wintertime average indoor temperature <18°C). These variables were then used in a logistic regression model and compared the outcome of this model against WHD scheme eligibility and fuel poverty risk (both LIHC and 10%) for predicting being in a cold home.

Method

Random forests are an extension of decision trees for classification and regression tasks introduced by Breiman (2001).[11] Classification trees and versions thereof are one of the most commonly used methods in data mining and machine learning to allocate observations across a number of mutually exclusive classes. In their simplest form, decision trees take a number of predictors and split them along ‘branches’ until they reach the ‘leaves’, which are the points at which all data are in one class or another or there are no further variables on which to split. One problem with decision trees, however, is that they tend to overfit to data (i.e. fit patterns in the data which are the outcome of chance rather than being reflective of the underlying data generation process) and have poor predictive power if the data used in the estimation is not representative of the wider population (see Hastie et al, 2009). Random forests are one way to overcome these problems, at the cost of a small increase in bias and poorer interpretability, but with great gains in predictive power. This is achieved through the estimation of many models which are then recomposed in a more general model so that one can to rank variables in terms of importance. This last quality was utilised to determine important predictors and then a logistic regression model was employed to predict whether or not a home is cold, which was defined as being <18°C.

Random forests are an ensemble learning method, which means that they include several models at once and then combine the results from those models to make a decision. For random forests, that means that several classification trees are estimated, using ‘tree bagging’, a kind of bootstrap aggregating process. Bagging involves estimating several trees (usually hundreds or thousands) on a random sample with replacement of the data set. Random forests take a random subset of the variables (also called features) at each split in the estimation process, a method sometimes referred to as ‘feature bagging’. This means that each decision tree within the random forest is estimated on a subset of the data and a subset of the variables, which avoids the problem of a few very strong features dominating the classification procedure.

One of the benefits of the ‘feature bagging’ is that not all trees in the random forest include the same variables and therefore the variable importance can be determined in two ways. The first is according to prediction strength. For each tree, the out of bag (OOB) sample – the observations not included in the random sample in the bagging process described above – are predicted using the estimated tree. The prediction accuracy is recorded, and this process is repeated with random permutations in a given variable. This takes place across all the trees and the average decrease in accuracy caused by permuting that variable measures its importance in the random forest, expressed as percentage of the maximum. The variables in were used, discretised into dummy variables for each categorical value as included in the EHS.

The Gini impurity index described Box 2 to identify the importance of ‘groups’ of features, or rather the importance of all classes in a particular variable. For example, if household income is split into quintiles, the random forest algorithm treats a dummy variable for each quintile as a separate feature. To determine the overall importance of income the average of each variable’s mean Gini decrease (I) across each of its values is taken. For income, that is the mean of the I for each quintile.

Box 2 – Gini impurity method

Another method of determining variable importance is based on the Gini impurity criterion which is used in the classification of splits during estimation. ‘At each split in each tree, the improvement in the split-criterion is the importance measure attributed to the splitting variable. This improvement is accumulated over all the trees in the forest separately for each variable (Hastie et al 2009). The Gini impurity index is calculated as

$$G = \sum_{i=1}^{n_c} p_i(1 - p_i)$$

where n_c is the number of classes in the dependent variable (in this case whether or not the home is cold) and p_i is the ratio of this class in the split.

Variable importance is then defined as:

$$I = G_{parent} - G_{split1} - G_{split2}$$

which is then averaged over all of the trees in which the feature is present.

In the methodology, after isolating the most important variables to identify whether households are at risk of living within cold homes, these variables were taken and included in a logistic regression.[19] This is in part for interpretability and in part because logistic regression allows for more complicated functional forms. It was means that continuous versions of discretised variables used in the random forests process could be used, as classification trees require discrete breaks in the data. For example, the marginal effect of the physical performance standard of the home may depend on income (and vice versa), with higher income homes less sensitive to poor performance standards and lower income homes more sensitive. This can be expressed in logistic regression through an interaction, while it is difficult to include it in random forests because there is no *ceteris paribus* assumption (holding all other things equal). The final predictive model includes e-values (continuous), length of residency (continuous), household type (categorical), dwelling age (categorical bands), presence of a boiler (categorical), age of the household reference person (continuous), number of people in the home (continuous), household income (continuous), number of bedrooms (continuous), and whether the household reference person is employed (categorical). Income and e-value were also interacted.

Results

The methodology was implemented over the 647 homes for which complete data was available, including 204 cold and 443 normal homes, producing a number of considerably interesting results. Table 23 shows confusion matrices for each of the criteria. The confusion matrices are contingency tables with counts. Columns are predicted classes and rows are the actual classes. The top left and bottom right cells (Normal:Normal, Cold:Cold) are correct predictions. The top right and bottom left cells are incorrect predictions (Normal:Cold, Cold:Normal). For example, there are nine households for which the WHD scheme criteria predicts ‘cold’ which are actually cold. An error rate was calculated for each of the actual classes, which is simply the percentage of incorrect predictions for that class. The false positive rate is the percentage of ‘cold’ predictions which are actually ‘normal’.

In terms of results, it is interesting to notice that the WHD scheme criteria (those who qualify for the Guarantee Credit portion of the Pension Credit) performed worst at predicting cold homes, finding only nine of the 204 cold homes. Another interesting result is that the Low Income – High Cost (LIHC) definition performed considerably worse than the old and much simpler fuel poverty definition based on the 10% income rule. This applies to both the ‘full income’ definition, which includes Housing Benefit, Income Support for Mortgage Interest Relief, and Council Tax Benefit

[19] There is no set rule for choosing the number of predictors at this step. Those felt to be the most important variables and which make theoretical sense and to balance predictive performance with parsimony were included.

in addition to household income, and the ‘basic income’ definition which does not. In case of the LIHC, 179 cold homes were wrongly classified Normal against an average 155 of the two versions of the fuel poverty criteria. Similarly, only 25 of the cold homes are classified as Cold by the LIHC against an average of 50 for the fuel poverty Criteria.

Finally, the estimation procedure was able to identify a model which predicts cold homes more consistently than both fuel poverty criteria. As shown in Table 23, the Logit model outperformed both fuel poverty criteria as it was able to identify 60 of the cold homes against 45 and 55 for the fuel poverty Full and Basic criteria, respectively. A similar improvement in performance is displayed by the Logit model in identifying normal homes, as it is able to identify 409 against 383 and 372 identified by the fuel poverty Full and Basic criteria, respectively. This implies a considerable decrease in the false positive rate, i.e. incorrectly attributing ‘cold’ status to normal households. This occurs in 34 of the Normal homes when using the Logit model, compared to 60 and 71 wrongly classified by the fuel poverty Full and Basic criteria, respectively. This implies a false positive rate of 36% for the Logit against 56% of the versions of the Fuel poverty criteria as the Random Forest model performed slightly worse than the Logit, these results are not discussed here although they are reported in the table.

Table 23 – Confusion matrices for cold home prediction criteria

		Normal	Cold	Error Rate	False Positive Rate
WHD scheme	Normal	413	30	7%	77%
	Cold	195	9	96%	
	Total	608	39		
LIHC	Normal	397	46	10%	65%
	Cold	179	25	88%	
	Total	576	71		
Fuel poverty Basic	Normal	372	71	16%	56%
	Cold	149	55	73%	
	Total	521	126		
Fuel poverty Full	Normal	383	60	14%	57%
	Cold	159	45	78%	
	Total	542	105		
Random Forest	Normal	399	44	10%	44%
	Cold	147	57	72%	
	Total	546	101		
Logit	Normal	409	34	8%	36%
	Cold	144	60	71%	
	Total	553	94		

Conclusions

The WHD scheme is aimed at reducing the cost of fuel expenditure amongst vulnerable households, who may be at risk of living in fuel poverty. Households that are at risk of living in cold homes (i.e. <18C) will be among the most vulnerable of those households living in fuel poverty. For those households, living in temperatures below 18°C may have a higher risk of a range of cardio-respiratory disease and mental health disorders. In order to better determine those households that are at risk of being in cold homes, the research tested the extent to which WHD scheme Core Eligibility Criteria and fuel poverty definitions were able to identify dwellings with low indoor temperatures during wintertime conditions. The use of Random Forest and Logit analysis also sought to identify those dwelling and households features that were important for identifying a risk of living in cold homes.

It was found that the use of the WHD scheme Core Group eligibility criterion is not necessarily a strong indicator that households in receipt of the payment would be living in cold homes (i.e. <18°C). This reflects the predominant type of home that those households occupy, i.e. mid-century flats that are social rentals, and therefore built to a higher energy performance standard. A stronger predictor of coldness was a measure of the dwelling energy performance (i.e. e-value) and other measures of household and dwelling age, which is reflected in other research. While the WHD scheme targets those who are older, it may not necessarily reflect the energy performance of dwellings they reside in.

In addressing the research questions focused on identifying the neediest part of the fuel poor it was found that households in receipt of the WHD may not necessarily have the highest risk of living in cold homes. Using a measure of energy performance (which would also reflect dwelling age) and some form of length of residence within the LIHC definition of fuel poverty could provide a more appropriate proxy to better target households vulnerable to living in cold homes.

Referenced works

1. Beatty TKM, Blow L, Crossley TF, O'Dea C. Cash by any other name? Evidence on labeling from the UK Winter Fuel Payment. *J Public Econ* 2014; **118**: 86–96.

2. Brewer M, Muriel A, Phillips D, Sibieta L. Poverty and Inequality in the UK: 2008. *(IFS Comment C105) Inst Fisc Stud London, UK* 2007; published online July.

3. Thistlethwaite DL, Campbell DT. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *J Educ Psychol* 1960; **51**: 309–17.

4. Rubin DB, John L. Rubin Causal Model. In: International Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 1263–5.

5. Reardon SF, Robinson JP. Regression Discontinuity Designs With Multiple Rating-Score Variables. *J Res Educ Eff* 2012; **5**: 83–104.

6. Gelman A, Imbens G. Why High-order Polynomials Should not be Used in Regression Discontinuity Designs. 2014; published online Aug.

7. McCrary J. Manipulation of the running variable in the regression discontinuity design: A density test. *J Econom* 2008; **142**: 698–714.

8. Hamilton IG, Davies M, Ridley I, *et al.* The impact of housing energy efficiency improvements on reduced exposure to cold - the 'temperature take back factor'. *Build Serv Eng Res Technol* 2011; **32**: 85–98.

9. Wilkinson P, Landon M, Armstrong B, Stevenson S, McKee M, Fletcher T. Cold comfort: the social and environmental determinants of excess winter death in England, 1986-1996. York, UK: Joseph Rowntree Foundation, 2001 http://www.jrf.org.uk/publications/cold-comfort-social-and-environmental-determinants-excess-winter-deaths-england-1986-19 (accessed May 18, 2011).

10. Hamilton I, Milner J, Chalabi Z, *et al.* Health effects of home energy efficiency interventions in England: a modelling study. *BMJ Open* 2015; **5**: e007298–e007298.

11. Breiman L. Random forests. *Mach Learn* 2001; : 5–32.

12. Zajonc GI, Tristan. Regression Discontinuity Design with Vector-Valued Assignment Rules. 2015; : 1–53.

13. Imbens G, Kalyanaraman K. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Rev Econ Stud* 2011; **79**: rdr043–959.

14. Wood SN. Thin plate regression splines. *J R Stat Soc Ser B (Statistical Methodol* 2003; **65**: 95–114.

15. Lee DS, Card D. Regression discontinuity inference with specification error. *J Econom* 2008; **142**: 655–74.