Dynamic time-of-use electricity pricing for residential demand response: Design and analysis of the Low Carbon London smart-metering trial

A thesis submitted to
Imperial College London
for the degree of Doctor of Philosophy

by

James Robert Schofield

April 2015

Department of Electrical and Electronic Engineering
Imperial College London

Abstract

This thesis describes the trial design and analysis of the Low Carbon London (LCL) residential dynamic Time-of-Use (dToU) trial. This trial investigated the potential for dToU tariffs to deliver residential demand response to the Supplier, where it may contribute to system balancing through Supply Following (SF) actions, and to the distribution network operator (DNO), where it may be used for network Constraint Management (CM). 5,533 households from the London area participated in the trial and their consumption was measured at 30 minute resolution. 1,119 of these received the dToU tariff, which subjected them to CM and SF price events that were designed according to the specific requirements of these respective use cases.

A novel, data driven, engagement ranking index was developed that allowed stratification of subsequent results into sets of the most engaged consumers, who may be indicative of a future populace that is more experienced/engaged in home energy management. Demand response (DR) was calculated relative to baseline model that used the dToU group mean demand as an input, with aggregate response levels calculated over a range of time, socio-economic and household occupancy related variables.

Taking a network perspective, the reliability of CM event response was examined and two simple linear models presented as candidate predictors of response level, which was found to be consistent with an 8% reduction in demand. The network capacity contribution of residential DR was theorised to consist of two components: “mean response” and “variance response”, and the real impact of these was investigated using the LCL gathered data. Potential risks to the network from low price induced demand spikes were explored empirically using the SF event data and the times of highest risk were identified.

The extensive metadata set gathered from trial participants was processed into some 200 numerical variables. A correlation analysis was performed which was visualised using weighted correlation network graphs. A number of parameters were found to predict response level, but responsiveness (the level of deliberate engagement) could only be reliably measured by engagement rank.

Acknowledgements

I would like to thank my supervisor, Professor Goran Strbac, for the opportunities and insights provided during the course of this PhD. For support and advice above-and-beyond the call of duty, I thank Simon Tindemans. To the team I worked with on Low Carbon London—and I will single out Mark Bilton, Richard Carmichael, Matt Woolf and Jelena Dragovic—thank you for your ideas, support and humour during the (sometimes challenging) times of this programme. Likewise, I thank the people I worked in EDF Energy and UK Power Networks, without whom the work described within this thesis could not have happened.

To my parents and brother, who have supported me through all my life decisions (and are, for some reason, still proud of me), thank you. To my friends who have raised my spirits even at my lowest moments, thank you. And finally, to those who have quietly helped me through the challenging times (you know who you are), another big thank you.

Declarations

Originality

This is to certify that the content of this thesis is a product of my own work, except where explicitly stated, and that other work is appropriately referenced.

Copyright

Contents

List of figures

List of tables

List of abbreviations

BM Balancing Mechanism.

BST British Summer Time.

CDF cumulative density function.

CLNR Customer Lead Network Revolution.

CM Constraint Management.

CPP critical peak pricing.

DECC the Department of Energy and Climate Change, UK.

DG distributed generation.

DNO distribution network operator.

DR demand response.

DSM demand side management.

DTE Domestic Tariffs Experiment.

dToU dynamic Time-of-Use.

DUoS Distribution Use of System.

EAC estimated annual consumption.

ECDF empirical cumulative density function.

EDRP Energy Demand Research Project.

EMUP Energy Management Unit Project.

EU European Union.

EV electric vehicle.

GB Great Britain.

GMT Greenwich Mean Time.

GSM Global System for Mobile Communications.

HP heat pump.

HV high voltage.

I&C industrial and commercial.

ICT information and communication technology.

IESMT Ireland Electricity Smart Metering Trials.

IHD in-home display.

LCL Low Carbon London.

LCNF Low Carbon Network Fund.

LOLE loss of load expectation.

LPN London Power Networks.

LV low voltage.

MV medium voltage.

NIKP Northern Ireland Keypad Powershift.

nonToU non-time-of-use.

ODS Operational Data Store.

Ofgem the Office of Gas and Electricity Markets.

PDF probability density function.

PMS Participant Management System.

PV photovoltaic.

R2 coefficient of determination.

SBP system buy-price.

SE standard error.

SF Supply Following.

SM smart meter.

SME small and medium sized enterprise.

SMS Short Message Service.

SO system operator.

SQL Structured Query Language.

SSP system sell-price.

SSR sum of the squared residuals.

STOR Short Term Operating Reserve.

TNUoS Transmission Network Use of System.

ToU time-of-use.

TSO transmission system operator.

UK United Kingdom.

VOLL value of lost load.

WLS weighted least-squares regression.

Chapter 1

Introduction

1.1 Motivation

Government legislation responding to the risk from climate change, increasing volatility of international energy markets and energy security, is expected to bring about a dramatic shift in the way the United Kingdom (UK) supplies and consumes its energy. The Climate Change Act of 2008 [1] enshrines in law the target of reducing carbon emissions to 20% of 1990 levels by 2050. Current projections indicate that approximately 30% of Great Britain (GB) electricity demand will be met by renewable generation by 2020 [2], ideally rising to 100% by 2050 [3]. These changes will bring significant new costs under the existing operating paradigm.

On the supply side, in addition to the fixed costs of the new renewable plant and grid connections, the following system integration costs will be borne:

- *Reduction in utilisation of infrastructure.* Though renewable generators can displace energy produced from conventional plant, their ability to replace capacity is limited because they cannot be dispatched reliably. Estimates of the capacity value of wind under the current operating paradigm are around 10% [4]. When combined with the inflexible output of nuclear and Carbon Capture and Storage (CCS) equipped thermal plant, this will lead to increased requirements on generation capacity margin and reduced utilisation of conventional plant.
- *Reduced efficiency of system balancing.* More operating reserve is required to manage the risk of sudden changes in wind output [4]. This increased reserve and lack of flexibility may also reduce the ability of the system to fully utilise renewable generation. Specifically, more wind energy will be curtailed.

On the demand side, the greatest change will be the shift of heat and transport load onto the electricity system. At the national level, full penetration of heat pumps and electric vehicles (EVs) could result in a doubling of peak demand with a disproportionately smaller increase in energy consumption of only 50% [5]. Most of the ensuing network reinforcement costs will fall on the distribution network—up to £4 billion per year additional reinforcement costs by 2040, increasing to £6 billion per year by 2050 [6].

The established solution for balancing is to ensure enough flexible generation capacity exists to meet demand when renewables cannot generate, and provide the necessary reserve requirement when they do. However, more cost effective balancing technologies exist. Alternatives such as transmission interconnection, energy storage and demand response (DR) have together been shown to result in savings of up to £15bn by 2040 [6].

This trial looks at the potential value of residential DR to the distribution network operator (DNO), where it may be used for network Constraint Management (CM), displacing or deferring network reinforcement costs, and to the supplier, where Supply Following (SF) may contribute to system balancing, reducing the need for conventional capacity margin. They are examined here in unison so that the potential conflicts and synergies between the two may be better observed.

The use of residential DR for CM and SF will require a dynamically changing tariff, henceforth referred to as a dynamic Time-of-Use (dToU) tariff. This is in contrast with time-of-use (ToU) tariffs, which are designed to target the predictable high demand periods in the week but are otherwise static, and critical peak pricing (CPP) tariffs, which are not static but used infrequently and targeting only the highest demand periods of the year.

The UK government plans [7] for the roll-out of smart meters to be completed by 2020. As the specifications of the smart meters [8] have been been designed with the express purpose of being real-time-DR-ready, there exists an imminent opportunity, one which could see consumers making very significant savings on their energy bills [6] while supporting a cost effective transition to a low carbon future.

1.2 Scope

This thesis concerns the design and analysis of the Low Carbon London (LCL) residential trial, which took place in the London area during 2013 and was part of the wider LCL programme that ran from January 2011 to December 2014. It is a complete story in that it details first hand experience of all aspects of the trial lifecycle. Work on this trial began in early 2011, with the inception of the LCL programme, and continued through to the beginning of 2015, the official closure of the programme—though the LCL dataset will continue to provide research material for many years to come. The activities performed during this period include:

- Reviewing literature on related trials in order to identify knowledge gaps and learn from past trial designs.
- Analysing preliminary data sets provided by programme partners, such as fault data and annual consumption data, in order to inform the trial design and set expectations for trial results.
- Designing the trial treatment groups and selecting target sample numbers in order to achieve statistically robust results.
- Working with programme partners during the recruitment process in order to ensure the experimental design targets were met.
- Designing the dToU tariff (the design of individual events and their placement in the overall pricing schedule) to achieve the defined research objectives while managing the constraints implied through conducting an experiment on real people and working with industry partners.
- Working with programme partners to ensure access to trial data in parsable formats. This included understanding the data architectures and security standards used by programme partners.
- Cleansing and validating the LCL trial data. Various checks were performed in order to check the validity of the smart meter (SM) consumption data and associated metadata.
- Analysing the LCL data and presenting findings with programme partners and coauthoring three programme learning reports.

Details of the above activities are described where they may contribute to the interpretation or context of the research questions.

The priorities of the LCL programme meant that the research conducted during its course tended towards a network perspective. However, as a real dToU tariff would likely be used to provide services to both supply side and network actors, the Constraint Management use case cannot be considered independently of Supply Following use case. Hence, the analysis described in this thesis addresses the following themes:

Consumer engagement: Various metrics of consumer engagement are examined:

- The shift in the volume of energy consumed at each price level is calculated relative to the non-time-of-use (nonToU) group
- Change in consumers' annual bills as a result of being on the dToU tariff is estimated
- A novel, data driven engagement ranking method is presented.

Response stratification: DR, relative to a linear regression model calculated baseline, is then investigated using multiple stratifications of the response signal in addition to engagement rank stratification:

- By season, month, day of week and time of day
- By socio-economic group
- By household occupancy level

Reliability and risk: Reliability of CM event DR and risk to the system from a network perspective.

- Predictability of CM event response level is analysed
- Network capacity contribution of residential price induced DR is calculated
- Risk to the network from low price signal induced demand spikes is determined

Metadata analysis: Correlation analysis of survey collected household metadata.

- Spearman's rank correlation of 25 chosen root variables against a full set of 200 variables
- Abstraction of results into weighted correlation network graphs for interpretation

1.3 Structure and original contributions

This section describes chapter structure of this thesis. The research described in chapters 5–9 was performed under the supervision of Dr Simon Tindemans. All other contributions are explicitly stated in the below text where relevant.

Chapter 2: Primer on demand response. This chapter describes the general principles and context of DR:

- The theory of social welfare maximisation; the high level motivation for implementing DR.
- The different methods that may be used to measure DR.
- Roles that DR may play within the electricity system.
- The benefits that DR may bring to each sector of the industry.
- System implementation approaches.
- System integration considerations.

Chapter 3: Background. This chapter provides context for LCL dToU trial by reviewing literature on the results of past trials:

- Global overview of the landscape of residential DR trials.
- Detailed review of six trials that were closely related to the UK context.
- Comparison of chosen, closely related trials, through use of normalised response and price metrics.
- Review of ongoing trials and anticipated results.
- Gaps in the understanding of residential DR in the UK.

Chapter 4: Trial design. This chapter describes the design and implementation of the LCL dToU trial:

- Trial design objectives are defined with respect to the gaps in current understanding that were identified in the previous chapter.
- High level trial design. This was performed in consortium with the LCL trial partners and, on the Imperial side, Dr Mark Bilton and Dr Richard Carmichael. As well as joint contributions, my specific contribution was the calculation of the target group populations necessary to achieve statistically robust results. These calculations also underpinned the choice of the number of experimental groups.
- Trial recruitment. The strategy and design of the recruitment process were arrived at in consortium with the other LCL trial partners. Specifically EDF Energy and UK Power Networks. On the side of Imperial College, Dr Mark Bilton, Dr Richard Carmichael and myself helped steer the process to achieve research objectives.
- The design of the dToU events is described including their placement in the overall dToU pricing schedule.
- Information and communication technology (ICT) architecture: A description of the end to end data collection process. All elements described as occurring at the Imperial end of this chain were my own contribution.

Chapter 5: Analysis basis. This chapter describes the analysis basis that was used in the analysis of subsequent chapters.

- Summary of amount and type of data collected from the LCL trial.
- Description of the data validation and cleansing process.
- Measurement of DR: A linear model for the prediction of baseline demand is developed and presented.
- Baseline model design and validation: The rationale behind the model design and its validation are described.

Chapter 6: Consumer engagement. This chapter presents two simple aggregate measures of trial engagement as well a novel method for ranking individual consumers' engagement with the dToU tariff.

- Measurement of consumption shift: The shift in proportions of energy consumed at each of the tariff's price levels is calculated by using the nonToU group as a reference point.
- Change in consumers' bills: The distribution of change in consumers' annual bills, as a result of being on the dToU tariff, is estimated.

- Engagement ranking index: A novel method for ranking dToU consumers according to their engagement in the tariff is introduced. This model later serves as a means for stratifying the measured DR signal according to household responsiveness.

Chapter 7: Response stratification. This chapter provides an overview of the primary results of the LCL dToU trial. The DR signal, calculated as described in Section 5.4, is examined over a number of different stratifications of the response signal.

- High level full trial results.
- CM event response: Exemplar CM event traces are examined and peak demand reductions as a result of the high price signal are presented.
- SF event response: Exemplar SF event traces are examined and aggregate response was examined over a range of signal stratifications. To a high level, these stratifications were based on time segmentations, socioeconomic groupings and household occupancy levels.
- LCL in context: The results of LCL were compared with those of closely related trials through use of normalised response and price metrics. This section builds on the trial comparison work described in chapter 2.

Chapter 8: Reliability and risk. This chapter presents an analysis of the reliability of residential DR and describes the risk to the network from low price induced demand spikes that may be introduced in the SF use case.

- Predicability of CM event response: The predictability of CM events is analysed and two simple predictor models are presented.
- Network capacity contribution of residential DR: The contribution to network capacity of residential DR is analysed and the effective contribution decomposed into two components which we call "mean response" and "variance response". Using LCL data, the real effect of the newly defined network capacity contribution components is computed and it is show that this contribution can outperform the mean response.
- Risk to the network from the SF use case: The risk to the network from low price induce demand spikes is analysed empirically using trial data, and the times of highest network risk identified.

Chapter 9: Metadata analysis. This chapter looks at the relationships between primary demand related metrics (such as DR) and the large set of metadata collected through survey responses from the LCL trial participants. I have been liaising with Dr Richard Carmichael for this research and a joint paper is in preparation that combines my quantitative analysis with his social science interpretation.

- Data sources: The data sources are described and the validation, cleansing and encoding process is detailed.
- Analysis approach: The basic correlation approach is described followed by the introduction of weighted correlation network graphs as a means of visualising and clustering high dimensional data sets.
- Results: Correlation and graph results are presented and discussed.

Chapter 10: Summary and conclusions. This chapter presents the overall conclusions of this thesis. It begins with a summary of the findings of the work performed and ends with a discussion on the avenues for future development.

1.4 Publications

The following publications are based on the work described within this thesis.

Published or accepted:

- J. Schofield, R. Carmichael, S. Tindemans, M. Woolf, M. Bilton, and G. Strbac. *Residential consumer responsiveness to time-varying pricing.* Report A3 for the "Low Carbon London" LCNF project: Imperial College London, 2014.
- R. Carmichael, J. Schofield, M. Woolf, M. Bilton, R. Ozaki, and G. Strbac. *Residential consumer attitudes to time-varying pricing.* Report A2 for the "Low Carbon London" LCNF project: Imperial College London, 2014.
- UK Power Networks. *DNO Guide to Future Smart Management of Distribution Networks.* Summary Report for the "Low Carbon London" LCNF project, 2014.
- J. Schofield, R. Carmichael, S. Tindemans, M. Woolf, M. Bilton, and G. Strbac. *Experimental validation of residential consumer responsiveness to dynamic time-of-use pricing.* 23rd International Conference on Electricity Distribution (CIRED), 2015.

In preparation:

- J. Schofield, S. Tindemans and G. Strbac. *A non-parametric consumer engagement ranking method for use with dynamic time-of-use tariffs.* To be submitted to IEEE Transactions on Smart Grid.
- S. Tindemans, J. Schofield and G. Strbac. *The effective network capacity contribution of residential demand response.* To be submitted to IEEE Transactions on Smart Grid.
- J. Schofield, R. Carmichael, S. Tindemans and G. Strbac. *The drivers of demand response: A metadata analysis of the Low Carbon London dynamic time-of-use tariff trial.* In preparation.

Chapter 2

Primer on demand response

This chapter gives a brief introduction to the theory surrounding demand response (DR) and thus provides context for the following chapters. It begins with an overview of methods of measurement, then moves on to the roles DR can play within the electricity system, the resultant benefits to the high level segments of the system, a summary of implementation approaches, and finishes by touching on some of the key considerations associated with its system integration. The descriptions aim to be general, but the UK system is used for illustration where specific details are required.

2.1 Motivation: social welfare maximisation

At the highest level, the motivation for the implementation of DR is rooted in the theory of social welfare maximisation. In this section we show that the maximum social welfare is achieved when the retail price equals the marginal cost of generating and delivering energy [9, 10].

Let us imagine that Fig. 2.1 represents the supply and demand curves for an arbitrary settlement block (i.e. at a point in time).

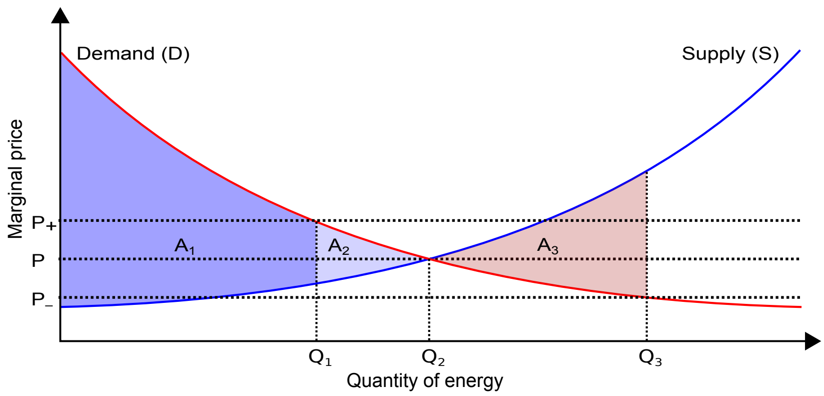

Figure 2.1: Illustration of the marginal price of supply and demand vs quantity of energy traded.

The marginal cost of generating each subsequent unit of energy is plotted against the total amount generated and illustrated by line S. Likewise, the marginal value of each unit of energy—assumed to be the amount a rational consumer is willing to pay—is plotted against the quantity procured and illustrated by line D.

Let us say suppliers are constrained to only offer flat rate tariffs and the rate is chosen to represent the average cost of procuring energy over a year (which would see them break even operationally). At a particular point in time, the actual marginal cost of supply may be equal to, greater than, or less than the flat tariff rate; illustrated by prices P, P_+ and P_- respectively.

Taking the full system view, we may say the net social benefit is equal to the total value the consumers take from the quantity of energy delivered, minus the total cost of procuring this quantity. Therefore, graphically, the benefit to the consumer is equivalent to the area under line D and the cost of production is equivalent to the area under line S, both bounded on the x-axis between zero and the quantity produced.

It is assumed that a rational consumer will not purchase energy from a supplier when the price of this energy is higher than the value she would gain from it. So, if the flat tariff rate were set to price P_+, the resulting quantity consumed would be Q_1. In this case, the net social benefit is represented by the area between lines S and D, bounded on the x-axes between 0 and Q_1. We shall call this area A_1. Following a similar procedure for price P and P_-, we can summarise the cases as:

- The tariff is priced equal to the unit cost of energy delivery, price P. The net social benefit is equivalent to area $A_1 + A_2$.
- The tariff is priced above the unit cost of energy delivery, price P_+. The net social benefit is equivalent to area A_1.
- The tariff is priced below the unit cost of energy delivery, price P_-. The net social benefit is equivalent to area $A_1 + A_2 - A_3$.

Therefore it is clear that the net social benefit is maximised when the tariff is priced equal to the cost to the supplier of procuring the energy.

It is worth making a note here that, while the net social benefit is achieved when the tariff price equals the cost of procuring energy, suppliers are not incentivised to deliver this without competition. Suppliers will generally act in a way that maximises their net profit and, with sufficient competition in the marketplace, it can be shown that such self-serving actions align to maximise social benefit [10]. However, if the number of market players is too few, some of the strategic players may be able to manipulate the market prices through their decisions. In such a situation, social welfare may not be maximised by the market alone. Ensuring the existence of sufficient competition and regulating where there is not is the role of the regulator—the Office of Gas and Electricity Markets (Ofgem) in Great Britain (GB).

2.2 Measurement

The US Department of Energy provides a comprehensive definition [11] of DR:

> "Changes in electric usage by demand-side resources from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized."

Actual consumption is easy to define and measure, but the hypothetical normal consumption pattern is more subjective. We shall call this, the counterfactual normal consumption pattern, the *baseline demand*. If baseline demand is known, DR is calculated as the actual demand minus the baseline demand. The task of measuring DR is therefore one of "baselining". The remainder of this section will give a brief overview of the main baselining approaches.

2.2.1 Baselining approaches

Much of the literature on demand baselining is related to industrial and commercial (I&C) load and is based on the experiences of the US market where DR services have been operating at scale for over two decades [12]. Here, a recent interest in the creation of DR standards by both industry and regulatory authorities resulted in a number of white papers [13, 14, 15] being produced to inform the debate.

For this section, it is assumed that the load in question is *interval-metered*—that is, a recording of the average consumption over a given time interval (typically hourly or half-hourly) is taken for each subsequent time interval. Some of the main baselining approaches that use data of this kind are detailed below. In their basic form, they are all simple to understand by the consumer—a significant advantage in a commercial context. However, the delineation between the methods given below is neither complete nor absolute; for real applications it is possible for combination approaches to be used.

Previous days. This approach [15] calls on recently recorded no-event days; typically within 30 days of the event. When an event is called, the consumption data for a number of recent historical days of the same type (e.g. business days) are selected. In the simplest case, these may then be averaged to construct the baseline profile. A commonly used adaptation of this approach, known as "high X of Y" [14], involves selecting the X days with the highest demand from the previous Y no-event days of the same type as the event day. The consumption profiles of the X days are then averaged to create the baseline.

Using previous days (i.e. recent history) increases the chances that the historical data contains similar usage patterns to the event day, particularly where weather and seasonal influences are concerned. As such methods typically use a small number of aggregated historical days to form the baseline, they usually require some form of adjustment (see next section) to account for changes in key external variables like temperature.

Day matching. In contrast to the *previous days* method, this method calls on a larger set of historical days; potentially as much as year's worth. Segments of the consumption profile on the event day are used to find the closest matching no-event days in a defined set of historical consumption data. A number of the closest matching day-profiles are then used to construct the baseline profile via an aggregation approach (usually averaging). Variations on the matching criteria may include the use of additional variables in the matching, for example, weather, day type or abstractions of the consumption data such as average daily consumption values.

The use of a typically larger set of aggregated historical days means that some of the variability caused by (approximately random) environmental variables is reduced. However, depending on the details of this process, there still may be a need for further baseline adjustment.

Regression. The data from historical no-event days is used to train a linear model to predict the baseline consumption on event days. Additional variables, for example weather or day type, may be used to improve the accuracy of the prediction. The complexity in this approach is in choosing the correct variables and avoiding overfitting—a situation in which, given enough variables and a limited set of historical data for training, the model will tend towards predicting the training data and not the counterfactual demand that is desired.

Load modelling. Also known as "engineering algorithms" [15], load modelling differs from the previous methods, which might be described as data driven, in that the use of historical consumption data is not absolutely necessary (though may be informative). Rather, an understanding of the individual load's dynamics are used to construct an additive bottom-up model. This has the advantage that, should a large individual load be powered down, it would be reflected in the overall baseline prediction. Such bottom-up modelling of loads is currently more suited to larger I&C sites where the scale of the individual loads may justify the modelling investment.

2.2.2 Related considerations

Baseline adjustments. Both *previous days* and *day matching* approaches allude to the need for a baseline adjustment—a form of post-processing of the baseline. Such adjustments are necessary to account for the effects of changing environmental variables or exceptional usage patterns [14]. An example might be the use of a regression model to correct for the temperature differences between the *previous days* used for constructing the baseline and the day of the event.

Pooled vs unpooled. In certain situations it may be the case that not all the loads in a DR programme are metered, but data from a meter that aggregates several sites, or that is representative of a group of sites, is available. When several loads share one meter, a baseline may be created for that meter (using one of the above methods), and then an appropriate method applied to apportion the calculated DR to each of the individual sites.

The use of a pooled approach for calculating DR is uncommon for I&C programmes as the ratio of the cost of metering to the value of DR is small. Though historically this was a more common situation in the residential sector, today, with smart meter (SM) rollouts being a common policy of many developed nations, such situations are less frequently encountered.

Where individual site metering is available, calculating a baseline for each site—as opposed to aggregating the sites then calculating a collective baseline—can allow for more fine grained analysis of the DR resource.

2.3 Roles

This section provides a brief introduction to the key roles in which DR may function, with a focus on a description of the service. It begins with a brief introduction to the high level theoretical basis of maximising social welfare, then moves on to provide a discussion of the potential value of DR from the perspective of the system segments.

2.3.1 Balancing services

In GB, the process by which the system operator (SO)[1] compensates for mismatches between the generation contracted by suppliers and the actual demand, or unforeseen outages of contracted generators, is know as the Balancing Mechanism (BM) [16]. The approaches to dealing with system imbalances are determined by the time period in which a response is necessary in order to maintain system stability. This section begins at the shortest response timescale—frequency regulation—then proceeds to a more general discussion of reserve capacity, finishing with a description of the financial settlement process as relevant to the supplier.

Frequency response

Frequency response services are the responsibility of the SO and are necessary in order to maintain the system frequency within the specified limits of 50 Hz$\pm$1%. There are two general categories of frequency response service; dynamic and non-dynamic. Dynamic frequency response deals with the normal second by second fluctuations in system frequency, typically delivered by generators fitted with a governor that allows them to react automatically to frequency changes within these short time scales. Conversely, non-dynamic frequency response is triggered at a defined frequency deviation and tends to be called upon after an unforeseen system outage. This typically requires a larger capacity response (MW range) to be delivered within a few seconds in order to save the system from collapse.

[1]In GB the SO and transmission system operator (TSO) are the same entity; National Grid.

DR is currently able to participate in frequency response within the UK. Participants are required to be able to respond within 2 seconds with a minimum of 3 MW for a minimum duration of 30 minutes. This service is normally called upon ten to thirty times per year [17]. However, in future, it may be possible for home appliances to automatically deliver frequency response services [18].

Reserve

A number of different *reserve* capacity service types are defined by National Grid [19] in the UK. These are listed below with a brief description of their defining requirements:

- Frequency response by DR (repeat of previous section for completeness): able to respond within 2 seconds with a minimum of 3 MW for a minimum duration of 30 minutes.
- Fast reserve: able to respond within 2 minutes with a minimum of 50 MW (ramp rate of 25 MW/minute) for a minimum duration of 15 minutes.
- Short Term Operating Reserve (STOR): able to respond within 240 minutes with a minimum of 3 MW for a minimum duration of 2 hours.
- BM Startup: provides access to additional generation that would not otherwise have run. Generators should be able to prepare for synchronisation (hot standby) within 89 minutes of a call and maintain this state of readiness for an agreed period of time.

DR is currently used for balancing within the UK system, provided mainly by large energy consumers (>25 MW), though smaller I&C loads may be aggregated into virtual units [20] in order to reach the minimum capacity (3 MW) for entry into the programme.

Due to the short response times required, some reserve generators must be running partly loaded so they can be ramped up at short notice upon demand. To reflect the cost to the generator of running at part load, both generator and DR units that provide reserve are compensated both for their availability and utilisation [17].

Using DR as reserve can be more competitive with respect to operating cost as it does not incur a fuel cost to maintain readiness. In the long run, using more DR for reserve may also displace the need to invest in the plant used to fulfil this role[6].

2.3.2 Supply following

We define *supply following* as the process by which retail prices are enabled to better reflect the marginal cost of energy procurement and therefore allow more efficient matching of supply and demand in a system with less flexible demand. The motivation for doing this is based on the principle that maximum social welfare (Section 2.1) is achieved when the cost of procuring energy equals the retail price offered by the supply companies.

If the supply and demand curves were invariant with time, a flat rate tariff could be set to reflect the cost of procuring the energy and market optimality would be achieved. This is not the case; for example, the value of lighting in winter is greater than in summer. Until recently, the supply curve had been relatively stable; changes in generating capital occurred over long run timescales and the cost of fuel could be locked in via specialised financial products (e.g. futures or contracts for difference). Now however, with increasing penetrations of renewable generation, the supply curve is becoming more variable due to the impossibility of scheduling the output of most renewable sources, such as wind power in the United Kingdom (UK). This is leading to greater mismatches between flat tariff rates and the cost of supply, and therefore reduced social welfare.

DR implemented via dynamic electricity pricing may be able to increase market efficiency by increasing the exposure of consumers to the true cost of production. This may lead to increased consumption of low operating cost renewable energy when it is in high supply, or reduced consumption of high marginal cost (and often inefficient) generation during peak load periods. By increasing the saleability of variable output renewable energy, the value of such generating capital would increase, potentially spurring further capital investment and reducing the need for government subsidies. At the same time, the need for investment in backup generation would decrease as increased prices during low renewable supply periods incentivises reduced consumption.

In the UK retail supply companies enter into contracted positions with generators in order to meet the demand of their customers. In the longer term such transactions tend to be conducted via bilateral forward contracts, while in the shorter term, via standardised, anonymised energy exchanges. The markets are closed one hour before delivery and suppliers can no longer alter their positions. Deviations from their contracted positions and actual demand turnout are dealt with through the BM.

The costs to generators and suppliers of deviations from their respective contractual positions are apportioned according to the Balancing and Settlement Code [21], administrated by Elexon. The day is divided into 48 half-hour settlement blocks and the energy spot market closes one hour before each settlement block is due for delivery. All contractual positions are then fixed in the system. Imbalances between the contractual positions of the generators and suppliers, and the actual demand turnout must then be bought from or sold to the SO at the system buy-price (SBP) or system sell-price (SSP) respectively. Exposure to the BM prices is usually undesirable in that, relative to buying and selling in the power markets, the SBP is usually higher and the SSP lower than those which can be obtained in the marketplace.

If consumers were able to react to price signals at close to real time, supply companies may also be able to use DR to reduce undesired exposure to the BM. However, this is likely to require automation and it is also unknown whether the consumers would find such price volatility acceptable.

2.3.3 Constraint management

The power flow through network components must sometimes be constrained lest engineering limits be exceeded. Most relevant to this work are *thermal* and *voltage* limits, though other limits exist within real systems. Below follows a brief description of thermal and voltage constraints, and how they apply to the network, ending in a discussion of the role DR can play in their mitigation.

Thermal constraints

As current flows through components within the network they are heated due to their natural electrical resistances to this current. When materials are heated, their physical properties may change. Effects may include thermal expansion, increase in electrical resistance or, ultimately, the thermal breakdown of components.

Thermal expansion is a particular problem for overhead lines, where overheating may cause them to sag dangerously low. Increase in resistance leads to increasing energy losses and can also lead to the voltage dropping below statutory limits on long lines. The breakdown of components due to excessive temperature is the main current limiting factor in transformer substations. Substation capacity is expensive and therefore active cooling is often economically justifiable in order to increase current limits.

Network components are given a "static rating" for the maximum current they can carry. As cooling rates are strongly linked to ambient temperatures, current limit values are often defined for both summer and winter. As this is usually based on worst case scenario operating conditions, for most normal situations, it is an underestimate of the actual safe capacity limit. With more detailed understanding of the thermodynamics of the component, combined with real time monitoring data, it is possible to safely extend the static rating to what is called a “dynamic rating” [22]. This allows temporary overloading of components above their static rating, relying on thermal inertia to provide some leeway before critical temperatures are reached. To this end, real time monitoring of transformer health is now widespread in primary substations (33 kV to 11 kV in the UK), and becoming increasingly more so at the secondary level (11 kV to 415 V in the UK).

Voltage constraints

Voltage limits are determined differently according network location. At the transmission level, voltage limits are determined principally by system stability requirements. At the distribution level the principal determinant is more often the quality of the consumer’s supply. In the UK, on low voltage (LV) networks—the connection point of residential properties—the statutory quality of supply guidelines [23] determine that voltage should be maintained within the limits of +10% and -6% of the 230 V nominal.

Voltage profiles along lines will change according to power injection/consumption from generators and loads. For example, along a distribution feeder line servicing only loads, there will be a gradually decreasing voltage profile, decreasing roughly proportional to the amount of load on each line segment. The voltage on such a line would typically be kept within its statutory limits by changing the tap positions of its parent transformer. This has the effect of raising or lowering the whole line voltage profile by a constant amount. Historically voltage profiles have not been so steep that both the transformer-end and the open-end cannot be kept simultaneously within limits via this method.

As we are seeing an increase in distributed generation (DG) connection, such as domestic solar panels, the profiles of the lines are therefore becoming more complex than the monotonically decreasing profile of a load only line. Effectively the converse of loads, generators increase the line voltage at point of connection. Indeed, if enough generation is connected to a line, as has been observed to be the case where a large capacity of domestic photovoltaic (PV) generation is installed on a single feeder, the voltage profile may be reversed, meaning that the automated tap changes, which are historically programmed only to deal with decreasing voltage profiles, may soon exacerbate voltage problems [24, 25].

The role of DR

DR and DG are currently used to alleviate constrained sections of network by temporarily reducing the power flows through components that are close to being overloaded. Within the transmission grid, the use of DR from large industrial consumers is standard practice though, through programmes such as Low Carbon London (LCL), it is now being investigated as a means of managing distribution level network constraints. By contributing to the management of network constraints, DR and DG usually also contribute towards *security of supply.*

2.3.4 Security of supply

Security of supply refers to the robustness of the network against unforeseen outages of network components. Security is ensured in part through redundancy of network assets, and as such comes at significant cost. To make best use of available resources, security of supply standards are set according to the magnitude of the load served. The transmission grid is therefore subject to security supply requirements which are universally more stringent than those of the distribution systems.

In the UK transmission grid, an “n-2” security standard is adhered to [26]. This means that any two lines may be lost before the supply to any area is affected. Thus, the transmission grid requires large capacity margins in order to be able to reroute the power flows which might result from up to two simultaneous major line or large generator faults. A correspondingly low average utilisation factor of less than 50% [27] is therefore typical on the UK transmission system.

In the distribution system, when network group demand is greater than 1500 MW, security of supply standards are in line with those of the transmission system security of supply requirements [28]. As group loading of network segments reduces (generally in line with reduction in voltage levels), so do the security standards: Between 1 MW and 12 MW there is a statutory requirement to reconnect load within 3 hours of a fault, reducing to within the best effort repair time on sections of network serving group loads of less than 1 MW.

The guidance document [29] that accompanies Engineering Recommendation P2/6 [28] has recently been amended [30] so as to recognise contributions from DR in delivering security of supply.

2.4 Benefits by industrial sector

The previous section described some of the roles that DR may be able to fill within the system. This section develops this by discussing the potential value to the high level segments of the electricity system: generation, transmission and distribution. It is assumed that the benefit to the consumer is manifest in reduced bills and therefore does not warrant further description.

2.4.1 Generation

Generation and transmission capacity are sized to be able to provide a *reliable* supply to the end user. To a high level, reliability can be split into two interrelated elements; *adequacy* and *security* of supply. Adequacy refers to the long run ability of generation to meet peak demand, which, in a market context, involves ensuring a sufficient pipeline of capital investment. Security refers to the ability of the system to withstand sudden unforeseen changes in supply, demand or transmission (as described in Section 2.3.4), over the day or hour timescale, such as the loss of a large generating unit. Security is the principal driver of *generation margin*; generating capacity in addition to the expected maximum demand to be used as contingency in the case of a major disturbance in the system.

Before the UK electricity market liberalisation in 1990, the late Central Electricity Generating Board sized this margin so that loss of load would have a probability of occurring nine times in every 100 winters. This typically led to a generator margin of greater than 20%. New technology and updated operating procedures have allowed this margin to be reduced. The measure of reliability has now switched to loss of load expectation (LOLE) [31], a probabilistic estimate of risk, and the target is now set to a maximum of 3 hours of LOLE per year. In the winter of 2013, the UK had a generator margin of approximately 6% [32], which corresponded to a LOLE of approximately 1 hour per year (based on the reference scenario) [33]. Generator margin and the variability of the load profile combine to give an average utilisation factor over all plant of less than 50% [34]. With more low *capacity credit* wind generation being developed, this could reduce still further.

Use of DR to shift or curtail load on the occasions when a generator unexpectedly fails may be more cost effective than building reserve generating capacity. While this is routinely practiced with large industrial loads, the integration of many distributed smaller loads into a DR scheme may prove more challenging - control mechanisms, commercial arrangements and risk profiles are likely to be different from those of larger DR loads. Proving the reliability of distributed DR assets and their respective control mechanisms will be the key issue.

2.4.2 Transmission

According to the security of supply regulations [26] in GB, the transmission system is required to be secure against a fault in:

> 5.1.1 a single transmission circuit, a reactive compensator or other reactive power provider; or
> 5.1.2 the most onerous loss of power infeed

where transmission circuits are inclusive of double lines. This requires substantial excess capacity to be built into the grid.

As an alternative to building excess transmission capacity, reserve generation and DR can also be used to ensure security of supply. By locally offsetting demand, the power import required to balance a constrained area can be reduced. A simplified illustration is given in Fig. 2.2. Here, pre-fault, Area B imports all 11 GW of demand from Area A. With the loss of a double line (10 GW of capacity), Area B is balanced by 10 GW of imported power plus 1 GW of either local reserve generation or DR.

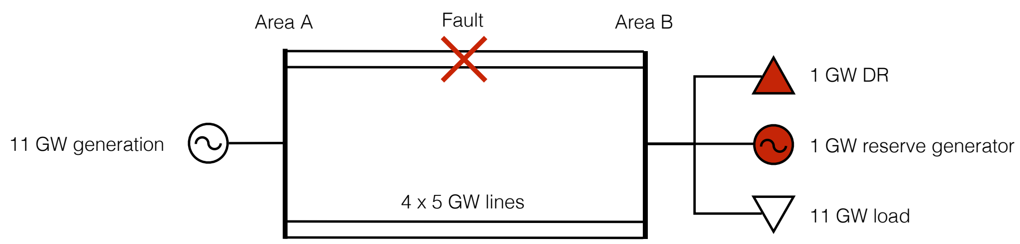

Figure 2.2: Simplified illustration of a double transmission line outage between two transmission areas. Post fault, Area B demand cannot be met by imports alone, but reserve generation or DR may be used for balancing.

DR currently makes a significant contribution to system balancing with approximately a third of frequency response and Short Term Operating Reserve (STOR) being provided by the demand side [35].

In response to concerns over the future availability of reserve capacity, efforts have been made to allow more effective exploitation of DR provided system services. In 2014, National Grid launched a tender for a new Demand Response Balancing Service [36]. This scheme is viewed by some as a temporary measure until a new Capacity Mechanism, as part of the more general Electricity Market Reform [37], reaches a state of functional maturity.

The Capacity Mechanism was envisaged to offer payments to both generators and DR providers on a technology neutral basis, to be tendered via auction in the newly established Capacity Market. The market held its first auction in December 2014 with 5% of the awarded units being won by DR providers [38]. There is an ongoing (at time of writing) debate as to whether the Capacity Market, in its current guise, is enabling fair competition for DR.

2.4.3 Distribution

Security of supply standards within the distribution network are tiered according to the number of consumers served [28]. In general, this means that they are highest at high voltage (HV), reducing through medium voltage (MV) and LV. For example, for group demands of over 1500 MW, the security standards mirror the equivalent for the transmission system. At intermediate levels of 60–300 MW, a first circuit outage should result in an immediate reconnection of the full group demand minus 20 MW, then the full group demand within the subsequent 3 hours. From 1–12 MW, reconnection of group demand minus 1 MW should be made within 3 hours. At the lowest level, group demands of less than 1 MW should be reconnected within reasonable repair times.

For the highest group load network, DR may be used in an analogous way as in the transmission example; by providing effective network capacity through reliably offsetting peak demand. At lower voltages, where security of supply standards are less rigorous, benefit from DR may still be taken. Here, involuntary disconnections due to faults or maintenance may be reduced in duration or avoided, thus enhancing the performance metrics of the distribution network operator (DNO). As well as providing a virtual network capacity contribution, use of DR may further improve the security of supply by enabling the prioritisation of load curtailment.

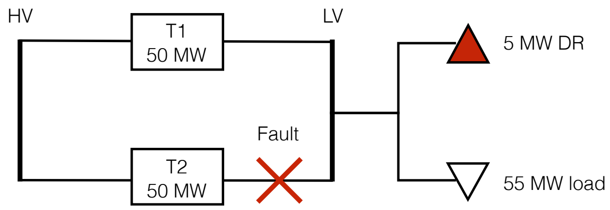

Figure 2.3: Simplified illustration of transformer fault in a distribution network substation. The fault reduces the substation capacity from 100 MW to 50 MW. To prevent overloading of T1 and involuntary disconnection, 5 MW of DR is used to bring load down to the 50 MW safe limit.

Figure 2.3 depicts a distribution level DR scenario. The substation loses one of its two transformers due to a fault. In a "business as usual" scenario this would result in all load on this transformer being shed until the fault could be fixed or power flow rerouted. With DR or DG, the transformer load may be balanced, avoiding involuntary disconnection and thus deferring investment in capacity upgrades.

DR can also be used to smooth the predictable daily peaks in consumption and thus effect a more long term alteration in the demand profile. Smoothing peaks in the demand profile can reduce resistive losses [39]. As losses are greatest at LV, typically 4% of the near 7% distribution losses [40] occurring at this voltage, savings here would have the greatest impact.

With increased penetration of DG, supply balancing at the distribution level, enabled by DR, may be beneficial. This would reduce use of system further up the network; sidestepping the use of system costs and line losses associated with a longer supply chain, and may allow consumers to pay close to wholesale prices for power generated locally. Such local level management of both demand and supply would also help to mitigate voltage and thermal constraints, allowing greater penetration of distributed renewable resources before network reinforcement is necessary.

2.5 Implementation approaches

DR can be implemented in a number of different ways. Programmes can be differentiated into two major categories; *incentive based* and *price based* [41, 27]. The difference lies in the method by which the signal to reduce demand is given. In the former, participation is negotiated, with the signal and conditions for demand reduction all agreed upon in advance. In this way DR can be considered to be dispatched deterministically. In the latter, though participation may still be negotiable, the price is both the incentive and the signal.

The DR types outlined below are meant to serve as examples of the key differences that may exist between tariffs, and not as an exhaustive list of available tariff designs. As such, variations and combinations of the tariff structures outlined below are possible.

2.5.1 Incentive based

Direct load control programmes allow the utility to directly shut down and switch on loads. Domestic loads such as water heaters and air conditioning units, where lack of service will not be immediately missed, are usually chosen. In the UK, a radio teleswitching service [42] is used to implement a tariff known as "Economy 7". The tariff involves two price bands with the option for direct control over water and space heating.

Interruptible or curtailable load control programmes operate similar to direct load control, with the difference that participating consumers are asked to reduce their loads to a predefined level. Participants who do not respond can face penalties. Large consumers may have bilateral contracts with the transmission or distribution network operator, though small and medium size consumers are more often aggregated via a third party business—simplifying the communication and management for the network operator.

Demand bidding programmes involve consumers submitting bids detailing how much load they are willing to curtail at a given price. The utilities will then make up the capacity they need from bids that are lower than the marginal price. Such programmes typically operate in the short run at day or hour ahead intervals. They are seen as a low risk way for consumers to engage in the energy markets.

Capacity markets involve consumers offering load curtailments in lieu of reserve generator capacity. Capacity markets differ from demand bidding in that they operate over medium and long run time scales. Closely related to capacity markets are *emergency demand response* programmes that substitute for reserve plant when shortfalls arise. Emergency DR programmes operate in a similar way to capacity markets but with more stringent dispatch conditions. Both programme types usually involve contracts that lock the participants in for a given period of time—typically years. The UK has recently introduced a Capacity Market that allows both demand and generation to bid for the provision balancing services [37].

Ancillary service markets are similar to capacity markets with the difference that the focus is on operating reserve services rather than energy balancing. Participants usually consist of large and regular energy consumers, where the type of operating reserve they can supply is determined by how quickly and for what duration they can respond.

2.5.2 Price based

Use of system charges. Although not usually considered as a DR implementation, they are mentioned here as they fit the definition of being a financial incentive applied to the consumption profile. Use of system charges are levied by both the transmission and distribution networks, and usually apply only to wholesale consumers.

In GB, Transmission Network Use of System (TNUoS) costs are distributed between wholesale consumers proportional to their respective consumption on the three highest half-hour settlement periods of the winter (November to February), know as the "triads". These days are chosen retrospectively at the end of each year. As such, wholesale consumers often try to predict when they will fall so that they may reduce their exposure to TNUoS by reducing their consumption on these days. In this way, the triads incentivise a reduction in annual peak system demand.

In GB, each regional DNO may design its Distribution Use of System (DUoS) charges within the parameters of the regulatory framework defined by Ofgem. A typical bill [43] may be broken into a number of components of which two effectively incentivise a reduction in peak demand: a capacity charge, levied according to the maximum import capacity of the network connection; and a time-of-use unit rate, set over a weekly schedule, with the highest rates covering the times of peak network load.

Time-of-use (ToU) tariffs. These differ from the pre-existing static tariff in that they have more than one price point for energy, determined by the time at which the energy is consumed. These tariffs are static in that the schedule changes infrequently so that consumers know, from days to months ahead of delivery, what their electricity prices will be. The prices are designed to represent the average cost of generating and delivering electricity at those times of the day. As such, the highest price points tend to be used around the predictable peak demands of each day. UK examples of such tariffs include the allocation of distribution use-of-system charges to suppliers and, for residential consumers, the Economy 7 tariff; a two tier tariff designed to shift load from peak to off-peak times.

Dynamic Time-of-Use (dToU) tariffs. Differentiated from time-of-use (ToU) tariffs by price change notifications that occur at a higher frequency (dynamically), typically one day ahead of delivery, though notification periods can be as short as one hour ahead. Reducing the notification period of price changes allows the tariff to better reflect the more rapid fluctuations in the price of production and delivery of electricity. This additional flexibility comes at the cost of consumer foresight, and therefore may hinder their ability to respond. Understanding the reliability and level of the response to dynamic electricity pricing is still the subject of much work [44]. It is hoped that technology innovation will make this trade-off less onerous.

Critical peak pricing (CPP) tariffs Similar to dynamic Time-of-Use (dToU) tariffs but with more infrequent events and a much greater price differential between the default and high price bands. The higher event prices are used to incentivise greater and more reliable reductions in demand for provision of system critical services. Such tariffs have not yet been trialled in the UK, though France, due to its heavy dependency on inflexible nuclear generation, has been utilising a critical peak pricing (CPP) tariff known as “Tempo” since 1993.

2.6 System integration considerations

The previous sections have discussed the roles, sector benefits and implementation approaches of DR. This chapter now concludes with some general considerations surrounding system integration.

2.6.1 Demand profile

While demand profiles do change gradually over the long term, in the short term the profile for a given day in the year is quite consistent. In the UK minimum demand is observed on summer nights and the highest peak of the year observed on a winter evening (as illustrated in Fig. 2.4). Maximum demand of the year is almost three times greater than minimum demand. The greater the peak to average ratio of demand, the lower the load factor of network and generation assets. From a network perspective, the most economically efficient load profile is a flat one, though this may not be the case for generation. Here marginal cost may vary with the supply of variable renewable resources such as wind.

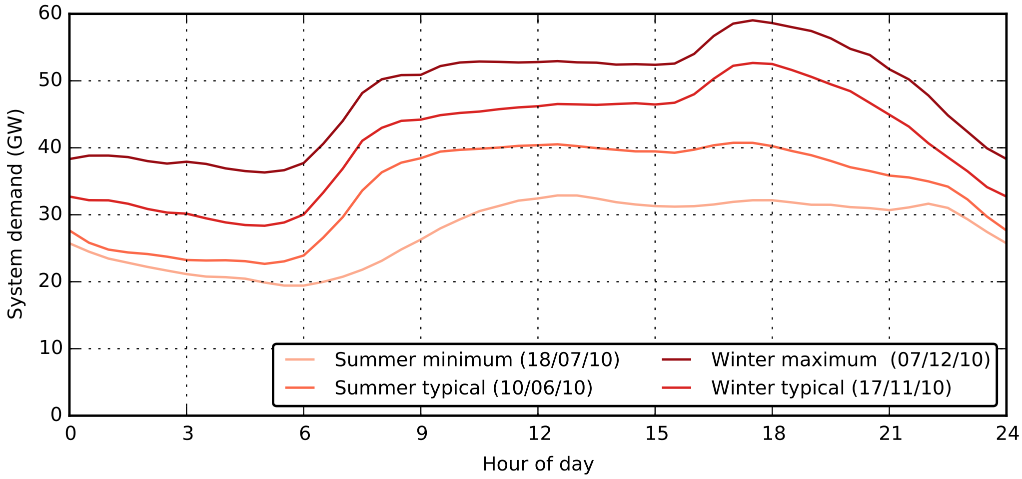

Figure 2.4: System demand profiles for the UK [45].

2.6.2 Value of load

The concept of the value of loads on the electricity system has historical significance in the determination of the adequacy and reliability investment. In this context the objective of industry is to match the marginal cost of preventing loss of supply to the marginal benefit for the customers of avoiding that loss of supply [46]. It is implicit from this definition that some measure of the value of load is required. Historically, in the case of reliability, this metric has been called the value of lost load (VOLL); defined as the average cost incurred by all consumers as a result of unexpected interruptions to their electricity supply.

As can be seen from Table 2.1, VOLL varies considerably between consumer types and with duration of outage. From the perspective of DR, the most immediately striking observation is that these values are several orders of magnitude greater than the cost[2] of the energy supplied. Initial inspection would therefore suggest that DR will not work because consumers place too high a value on the availability of their supply. But VOLL does not always accurately represent the the cost of customer interruptions.

Table 2.1: Value of lost load for different consumer classifications. Data reproduced from Kirschen et al, 2003 [47]. Values measured in £/kWh.

Duration	Residential	Commercial	Industrial	Large users	Typical bus
1 min	-	153.00	776.40	606.60	258.21
20 min	1.35	29.17	85.62	30.87	28.75
1 h	1.62	26.63	50.52	10.77	18.37
4 h	2.79	24.40	36.11	3.32	14.14
8 h	2.99	24.58	30.03	1.82	12.55
24 h	3.12	10.41	12.53	0.83	5.60

VOLL is an aggregated value of the cost of unexpected interruptions across all consumer types and all their respective load types. Electricity consumption, however, has a different value for different use types and, indeed, even the same use type may be ascribed a different value by a different consumer. Furthermore, should notice of an interruption be provided, the consumer may well be able to reduce or negate the cost of that interruption. Herein lies the opportunity for DR: If the value of a demand reduction to the system is greater than the value to the consumer, there exists the potential to construct a business case for said transaction.

[2]Retail price at time of writing was approximately 0.12 £/kWh.

For example, a refrigerator pump may have its cooling cycle postponed without loss of *service* so long as the interior temperature does not exceed a certain value [48]. This constitutes a potential demand flexibility that may be exploited at little or no cost to the consumer. The addition of technology to enable rapid, optimal rescheduling of loads could significantly enhance the availability of low cost demand flexibility.

2.6.3 Elasticity

Price elasticity is defined as the amount by which a marginal change in price will elicit a change in demand. In the long run an increase in average prices will stimulate a reduction in demand and/or investment in efficiency measures to increase the unit value of consumption. However, in the short run, most consumers only have the option to curtail or reschedule their demand in response to an increase in price. If DR is to be delivered via price signals, some knowledge of the elasticity of demand will be necessary for integration [49].

While many of the practical considerations for the implementation of DR will depend on understanding short run elasticity [50], long run elasticity will have a significant impact on the business case for DR. For example, the greater the volatility of prices, and the greater the differential between high and low prices, the stronger the business case for storage will be—the most direct competitor to DR [51].

2.6.4 Intrinsic storage capacity

The flexibility of an electric load is dependent on the intrinsic storage capacity present in the service it provides. The use of intrinsic storage capacity allows loads to be shifted in time without the negative impact of a temporary service reduction being felt by the consumer. Storage can be divided into two broad categories; *service storage* and *energy storage*; described below:

Service storage is when the output of the electric load has some intrinsic storage capacity. Take the dishwasher as a simple example. Here, the duration that a wash may be deferred depends on the amount of clean kitchen utensils available. Clean kitchen utensils are therefore the service that is stored in this case.

Energy storage is when the energy to power the load is stored so that the load may provide its service without the need to draw power from the network. Obviously any device with a battery will fall into this category, though batteries are not the only form of energy storage. If the purpose of the electric load is heating, energy may be stored thermally, for example in a hot water tank or high heat capacity material. In the UK, electric storage heaters used in conjunction with the Economy 7 tariff are common examples of this.

A key point to note is that, in all situations where energy is stored, charge-discharge cycle efficiencies that are always less than unity[3] and overall energy consumption is increased.

2.6.5 Load recovery

As explained in the previous section, for a load to be shifted forwards or backwards in time without any reduction in the service that it provides, it must have some form of storage capacity. When such a load is subject to a DR event it will maintain service levels without consuming network power by consuming its stored energy or service outputs. In order to ensure sufficient storage capacity to span a DR event, a load must recover the energy that it would have consumed during

[3] An implication of the second law of thermodynamics.

the event at some other time [9]. This time can either be before or after the event. We call the increase in demand surrounding a DR event *load recovery*.

Without proper control, load recovery effects might result in a spike in energy consumption before and/or after a DR event. It is plausible that such consumption spikes may be comparable in magnitude to the load peak that the DR event sought to mitigate. Care must therefore be taken to ensure the dispatch of a DR event does not merely shift the peak demand issue to a different point in time.

Methods to mitigate the load recovery effects that could emerge from the usage of DR include staggering event start and end times, direct load control with optimal scheduling [52] and distributed market based mechanisms [53].

2.6.6 Diversity of demand

Over a defined time period (usually a year is chosen) and for a group of consumers, the *diversity factor* is defined as the sum of the maximum demands of individual consumers, divided by the maximum concurrent demand of the consumers [54]. Mathematically this can be written:

$$F_D = \frac{1}{D_g} \sum_{i=1}^{n} D_i \tag{2.1}$$

where D_i is the maximum demand of individual household i and D_g is the coincident maximum demand of the group. The inverse of this $(1/F_D)$ is known as the *coincidence factor*.

The action of demand diversity can be explained intuitively. Let us say that an individual household has its own generator, such that it is self sufficient for its own power needs. It is clear that in this case the generator must have a capacity equal to or greater than the maximum possible demand of the household. Next, suppose that the generator is shared between several households. Now it must be sized, not for the sum of the maximum possible demands of each household, but for the maximum concurrent demand of the group of households.

Intuitively, the probability that all households will be consuming their maximum possible demand at the same time becomes lower with increasing group size. Of course that probability exists, so an acceptable value for the probability that the maximum demand will not exceed the generator capacity is chosen. The choice of this value could be based on the consumers' collective appetite for risk, but in real systems, this decision is taken by the SO: the generating capacity is sized to at least match the maximum possible after-diversity load on the system, plus some margin to account for unforeseen plant outages.

For a given confidence level, the maximum possible load on the system can be calculated either empirically from historical data [55], or analytically, via computer modelling [27]. In the UK it has been shown that a group of 10,000 houses will exhibit a load coincidence factor of just 0.1 [27]. That is to say, the required generating capacity per household for this group is just one tenth of the capacity which would be required if each household generated and consumed power in a self-sufficient manner.

There is a risk that loads may be synchronised by DR events. If a load were to be shifted to the extent that its inherent *storage capacity* were exhausted, it would need to consume power immediately after the load shifting operation lest service quality be reduced. Such an event applied to a group of normally unsynchronised loads would effectively synchronise their post event consumption as they would be linked by a common need to replenish their intrinsic storage. Usage cycles may remain synchronised for some time after the event, potentially enhancing *load recovery* effects.

2.6.7 Persistence of response

Persistence of DR refers to the extent by which the magnitude of the response remains unchanged over time. It may be divided into long run and short run effects.

Short run effects are based in the day-to-day running of the household and may occur over time scales of days to months. For example, if a DR action is called upon for a day, it may be possible for the consumer to shift a load to the day after or before the event. However, if the DR action is called upon for several consecutive days, intrinsic storage capacity limits may be reached leading to reduced service for the consumer. In turn, this may reduce the consumer's appetite to respond to the DR event the next day.

Long run effects typically occur over time scales of months and years and involve more lasting changes to the ability of the consumer to respond. This could include reduced response as novelty wears off, or increased response as learning and automation make for more effective load management. Extending the short run example to the long run, if such events looked likely to be a lasting feature of the electricity tariff, consumers might invest in automation equipment or storage devices to enhance their ability to respond without loss of service.

Chapter 3

Background

This chapter begins by introducing the context of the Low Carbon London (LCL) programme within which the trial analysed in this thesis took place. An overview of the global landscape in residential demand response (DR) is provided, moving on to a more focused review of trials that are closely related to the UK context. Several upcoming trials of interest that were in progress at time of writing are listed and briefly described. This is all brought together in a discussion of the research opportunities that currently exist in the area of residential DR. The chapter concludes with a summary of the knowledge gaps that exist in the field of residential DR in the UK, and those that may be filled with this work.

3.1 Low Carbon London

In response to the challenges facing the UK electricity system, the UK electricity markets regulator, the Office of Gas and Electricity Markets (Ofgem), created the Low Carbon Network Fund (LCNF) with the objective of incentivising research and innovation within the distribution network sector. The LCL programme was funded in 2010 under the LCNF tier 2 scheme by the amount of £21.7 million with an additional £6.6 million of funding contributed by programme partners [56]. It was commissioned to demonstrate and gather performance data on a number of innovative 'smart grid' technologies, of which the residential dynamic Time-of-Use (dToU) pricing trial analysed in this thesis was one [57, 58].

It was conducted by a partnership of industry stakeholders organisations and academia. Key partners in the design and implementation of this trial included:

- UK Power Networks: The London distribution network operator (DNO) and the lead programme partner.
- Imperial College London: Trial planning and data analysis.
- EDF Energy: Retail energy supplier.
- Siemens: Information and communication technology (ICT) framework.
- Logica (now CGI): Smart meter head-end.

3.2 Overview of global demand response trials

This section provides a review of related trials and a summary of the state of knowledge as applicable to residential dToU in the UK. High level results are summarised by meta analysis with references to primary literature where greater detail is deemed useful.

The LCL trial constitutes the UK's first dToU pricing initiative and, as such, it was not possible to find literature relating exactly to such schemes within the UK context. Nonetheless, literature of closely related tariff types, time-of-use (ToU) and critical peak pricing (CPP), from countries with comparable climates to the UK, are informative.

In order to gain an understanding of the general landscape of trials, the following meta-analysis were used as a starting point, with specific trial reports referenced where greater detail was sought:

- The Department of Energy and Climate Change, UK (DECC) (2013) [44]: The most recent and also the most relevant by virtue of it being commissioned specifically to inform the design and implementation of future DR trials. Some 30 domestic DR trials were compared. All trials considered were completed after the year 2000 and had a focus on measures to shift (rather than to reduce) demand. Included trials were from North America, Australia, and Europe.
- Faruqui and Palmer (2012) [59]: Focused on 9 of the "best designed", most recent experiments to identify the relationship between peak to off-peak price ratio and reduction in peak demand. Included trials were from North America and Ireland.
- Ehrhardt-Martinez (2010) [60]: Examined 56 residential sector feedback studies to provide insight into the factors that influence feedback-induced energy savings. Included trials were from North America, Europe, Australia and Japan.
- Faruqui and Sergici (2010) [61]: Surveyed the, then, 15 most recent trials with dynamic pricing of electricity. Included trials were all from North America.

Time-of-use tariffs: In general, the literature indicates that economic incentives are effective in changing consumer behaviour, though results have been highly varied.

The effect on total energy consumption is typically small compared to the effect on peak demand. The Energy Demand Research Project (EDRP) results show an approximate 4% reduction [44] in weekday peak energy consumption for consumers with in-home display—though significance was low due to the low number (170) of participants. In contrast, the Customer Lead Network Revolution (CLNR) showed a reduction in peak consumption of 6% [62] from 600 households, and the Ireland Electricity Smart Metering Trials (IESMT) a reduction of 7-12% [44] from 3,000 households. Interestingly, there appears not to be a strong relationship between the size of the difference between peak and off-peak prices, and the size of the consumer response across studies. Results from North American trials show similarly large variations in peak consumption reductions.

Table 3.1 shows the variation, in peak reductions across ToU trials covered by the DECC literature review [44]. Also given, when available, are details of the price differentials used (relative to the median price).

Dynamic Time-of-Use tariffs: There is little evidence relating to the efficacy of dToU pricing in the UK context. Geographically closest to the UK, the EFFLOCOM trial in Norway[1] found larger peak reductions when the tariff depended on the spot price of power than when it did not, though this was based on only 81 households.

Moving to the USA, in Illinois, the Energy-Smart Pricing Plan trial [63] tested a dToU tariff in conjunction with automatic cycling of air conditioning load, which was found to result in a peak demand reduction of 9.8%. Elasticity was also found to increase as price increased, and with decreasing household income.

Pacific Northwest examined the use of DR in conjunction with distributed generation for constraint management [64]. Like the Illinois trial, they augmented price signals with home automation

[1] Project report no longer available online, but reviewed by DECC [44].

Table 3.1: Summary of past ToU trials reviewed by DECC [44].

Index	Trial	Country	N	Min price ratio	Max price ratio	Min average reduction in peak demand (%)	Max average reduction in peak demand (%)
1	Ontario Smart Price Pilot (2006-2007)	Canada	124	1.40	1.40	0.0	0.0
2	Idaho DSR trial (2005-2006)	USA	85	1.84	1.84	0.0	0.0
3	Missouri CPP trial (2004-2005)	USA	91	3.49	3.49	0.0	0.0
4	CL&P Pilot (2009)	USA	188	2.08	4.08	2.0	3.0
5	PSE's ToU trial (2001-2002)	USA	300,000	-	-	5.0	5.0
6	California State-wide Pricing Pilot (2003-2004)	USA	226	2.00	2.00	1.0	6.0
7	myPower Trial (2006-2007)	USA	379	1.87	1.87	3.0	6.0
8	Norway EFFLOCOM Trial (2001-2004)	Norway	237	-	-	10.0	10.0
9	Xcel Energy Trial	USA	2,900	-	-	5.2	10.6
10	PG&E's Trial (2008-2010)	USA	86,222	-	-	11.0	11.0
11	Ireland Electricity Smart Metering Behaviour Trials (2009-2010)	Ireland	2,920	1.43	2.71	7.0	12.0
12	Energy Demand Research Project Trials (2007-2010)	UK	1,546	1.65	1.65	-	-
13	Northern Ireland Powershift trial (2003-2004)	Northern Ireland	100	2.67	2.67	-	-

which was concluded to be “helpful” in increasing responsiveness. Both ToU and dToU tariffs were tested. Price responses were observed to be stronger for the ToU tariff, though with only 112 households, evenly split between ToU, dToU and control groups, statistical significance was low.

Critical peak pricing tariffs: There is strong evidence that CPP tariffs have a greater impact on peak reduction than ToU tariffs. A reduction in peak demand was achieved in all CPP trials reviewed by DECC, with reduction ranges typically lying between 5% and 38%. Evidence on the effect of peak to off-peak price differentials is mixed, though generally point towards larger price differentials resulting in greater peak demand reductions. Notifications for such trials were typically sent the day before, with peak periods spanning around 5 hours. For trials with longer duration peak periods, a slight reduction in responsiveness was noted.

Perhaps the the most relevant to the UK is the use of the Tempo tariff, in France. Driven by the high penetration of inflexible nuclear generators, Electricité de France (EDF), implemented the tariff to allow smoothing of both the annual and daily electricity load curves [65]. The tariff includes six rates for electricity with day ahead price notifications. The highest price band may only be used on 22 days of the year, designed to reduce the system critical peaks. The tariff also includes regular ToU load profile smoothing elements. Tempo has led to a reduction in consumption of between 15% and 45% depending on price band and 90% of customers are reported to be satisfied with the tariff.

Direct load control: A recent meta-analysis by VaasaETT [66] found that automation augmented the response to price signals by over 100% in many cases. This chimes with the findings of Faruqui and Palmer [59] which showed the significant impact of enabling technologies for enhancing the peak reducing effect of residential consumers on ToU pricing tariffs. Such results allude to the long run potential of DR, when market based pricing is augmented by home automation.

Radio teleswitching of loads has been practiced within the UK since 1984 [42], allowing up to 200 MW of load reduction for use by the transmission network operator.

Information provision: Information provision was seen to be the most effective lever to enhance the DR effect after direct load control and automation [44]. There are many vectors for such information, typically by paper bills, but more recently via in-home displays (IHDs). In the Ireland Electricity Smart Metering Trials (IESMT), a combination of ToU with bi-monthly bills, a bespoke energy statement and an electricity monitor, were found to reduce peak demand by 11.3%. This is in contrast to ToU tariffs without such additional stimuli, where the average was an 8.8% peak reduction [67]. It is difficult to compare this figure with other key European ToU trials as the communication method was never an isolated experimental variable. The EDRP trial, which included two ToU trials, did not measure the effect of IHDs alone. North American trials have found mixed results while trying to isolate the effect of IHDs, some even showed a slight negative impact on peak reduction [44].

Household occupancy: Results are mixed for impact of household occupancy.. The EDRP trial and selected North American trials found that smaller households were more responsive to price changes than larger households. In the case of EDRP, consumption during the peak price period was seen to increase by approximately 4% per additional household member. This increase in peak consumption was slightly lessened to an approximate 3.5% when the additional household members were under the age of 16. Though with only 170 households in this study, such findings are not conclusive. In contrast, the IESMT found that households with children under the age of 15 were more responsive to their ToU tariff by 10.7% compared to 6.5%.

Low income consumers: Though there is little evidence from the UK, a meta study [68] looking at five trials in the USA suggests that low income consumers will benefit from ToU pricing. Indeed, because of flatter than average load profiles, they may benefit even without changing their usage profile. Despite this, evidence from these trials also suggests that low income consumers may have a lower than average response to ToU pricing. Three of the five trials investigated showed this. Various reasons have been proposed [44], though evidence is thin: Lower overall electricity use may mean there are fewer discretionary loads to manage, flatter initial loads may reduce scope for load shifting, or other characteristics such as housing standard or appliance ownership may be at work.

Change in bills and consumer attitudes: In general attitudes towards variable pricing tariffs were positive, with over 80% of trial participants responding positively[2]. However, this is likely to be strongly related to the fact that nearly all trials were designed to be revenue neutral (the bill for an unchanged average demand profile remains the same as on the incumbent tariff). In this situation, people who respond appropriately to price signals, even minutely, are guaranteed to make a net saving from their previous tariff. As a result, most consumers across such trials saved money.

Electrification of heat and transport: Penetration levels for electric vehicles (EVs) and heat pumps (HPs) within Europe are low and, as such, there is little evidence from trials for the flexibility of these new loads with respect to DR. Early data from EV trials suggests that consumers may be content to charge their vehicles overnight rather than at peak times. However, as numbers of participants have been low, results lack quantitative robustness and therefore estimates of the DR

[2] Survey questions were not consistent across trials, though for this analysis, responses were interpreted as positive if words to the effect that the consumer would willingly choose to remain on the tariff should the option be provided, were used.

flexibility of EVs are still largely qualitative. As heat pumps are effectively air conditioning units in reverse, it may be possible to assume similar flexibility characteristics, though there are many local variables which could affect this.

Persistence: As trials are by nature temporary, there is little evidence available on the long run persistence of DR beyond two years (most trials tend to last for one year). Of the four US trials examined by DECC with durations greater than a year, only one [69] showed lower peak demand reductions in the second year relative to the first. Closer to the UK, the IESMT saw peak reductions increase in the second half of the year relative to the first. With respect to UK applicability, evidence for long run persistence is thin.

3.3 Focus on related trials

This section gives a more detailed description of five trials that were considered to be of particular relevance to the context of UK residential dynamic electricity pricing. All trials took place in the UK or Ireland and are presented in chronological order.

3.3.1 The Domestic Tariffs Experiment (1966–73)

Dynamic electricity pricing is not a new idea. The Domestic Tariffs Experiment (DTE) [70] ran from 1966 to 1973 within Great Britain (GB) and was managed by the then extant Electricity Council. The objective of the programme was to determine whether any of three experimental domestic tariff structures might be more efficient (economically) than the pre-existing options. Much like today, the main tariff options available then were the "block tariff", where the unit cost falls in blocks of increasing energy consumption; and a "restricted consumption tariff" with an off-peak rate designed for use in conjunction with timed appliances such as water heaters—the precursor to the existing Economy 7 tariff.

The experimental tariffs consisted of two that were time related, "Seasonal" and "Seasonal Time-of-Day" (STD), and one that was demand related, "Load Rate" (LR). The Seasonal tariff was priced at 1.5 times the standard rate during the months from December to February and 0.7 times standard for the rest of the year. Load Rate imposed a high rate on energy consumed when demand was over a certain power limit, but was cheaper otherwise. It was priced at 0.6 times standard when demand was below this limit and 2–3 times standard when above. The Seasonal Time-of-Day tariff was of the ToU type with structure summarised in Table 3.2. All three experimental tariffs were designed so that a consumer with average demand profile would have the same bill as on the standard flat rate tariff.

Table 3.2: The Domestic Tariffs Experiment's Seasonal Time-of-Day tariff structure. Prices are in units of the local standard rate tariff at the time. Data from [70].

Price band	Times	Price
Peak	08:00 to 13:00, Mon to Fri, Dec to Feb 16:30 to 19:30, Mon to Fri, Dec to Feb	3.0
Night	23:00 to 07:00	0.4
Off-peak	All other times	0.8

As smart meters (SMs) were not yet available, magnetic tape demand recorders capable of measuring consumption on a half-hourly basis were developed especially for the trial.

Trial recruitment was on an opt-in basis sourced from all GB. Based on the assumption that any practical application of the tariff would only apply to larger consumers, only households with an annual consumption of over 3,000 kWh were selected. Incentive payments were used to encourage enrolment and overall attrition from the experimental groups was around 25%. After the recruitment phase, the trial ran for five years from 1966–72.

Results were reported in the form of the averaged difference between tariff groups and the control group over the five year trial period. It was observed that consumption increased in all three tariff groups: Seasonal by 6.9%, STD by 1.7% and LR by 3.2%. This was partially achieved by increasing consumption outside of the peak time: relative to the control group's load factor (peak to average power demand ratio) of 50%, Seasonal was 60%, STD was 57% and LR was 51%.

The STD tariff was the most effective at shifting consumption away from peak times with an average reduction of 116 kWh (25%). However, the increase in consumption at other times of the day was slight, at 1.4%, given the price at these times was only 20% less than the standard rate. The Seasonal tariff reduced the fraction of overall consumption that fell within the winter months by 2.1%. LR was the least effective at improving system load factor and there was little impact to the average consumer load curves for this group.

Consumer attitudes to the tariffs were positive. Survey results found that consumers on all tariffs thought them worth continuing. Appliance usage was not significantly affected with the exception of the STD tariff. Here, a 38% increase in total installed storage heater capacity was noted relative to the control group.

An approximate calculation of consumer surplus was made in order to determine whether such tariffs might be of net benefit to the operation of the electricity system and therefore consumers. It was assumed that the tariffs' price rates were representative of the average marginal cost of supply. With the additional assumption of a linear relationship between consumption and prices (elasticity), the consumer surplus relative to the control group was calculated as:

$$\Delta\text{Consumer surplus} = \frac{1}{2}\sum^{\text{Price bands}}(\Delta\text{Price}\cdot\Delta\text{Consumption})$$

The consumer surplus was compared to the cost of implementing the new tariffs, which were dominated by the cost of the specialised meters developed for the trial. A summary of the findings is shown in Table 3.3.

Table 3.3: Results of the consumer surplus and implementation cost estimates performed by the Domestic Tariffs Experiment [70].

(£p.a.)	Seasonal	STD	LR
Gross consumer surplus	1.2	2.3	0.8
Cost of implemetation	2.0	4.0	1.6
Net benefit	-0.8	-1.7	-0.8

The STD tariff was found to have the greatest gross benefit for the consumer and the LR tariff the least. However, for all tariffs, the benefits were more than offset by the cost of implementation. It was concluded that the benefits of ToU tariff structures might be realised in future should the cost of metering fall sufficiently. The programme therefore recommended that the cost of metering be continually monitored for such an eventuality.

Today, with the developments made in electronics and manufacturing, the cost of metering has fallen to sufficiently low levels that time varying tariff structures are a feasible proposition for the domestic consumer. Such meters have become known as "smart meters" (SMs).

3.3.2 The Energy Management Unit Project (1989–90)

The Energy Management Unit Project (EMUP) [71] ran from 1989 to 1990 within GB and was managed by the then extant Electricity Association. The main objectives of the programme were to field test a multi-rate meter and associated IHD, and to quantify changes in consumption patterns and attitudes as a result of the two tested experimental ToU tariffs.

Table 3.4: Structure of the two experimental ToU tariffs used in the Energy Management Unit Project. Prices are in units of the standard rate tariff at the time (£0.0616/kWh). Data from [71].

	Tariff A			Tariff B	
Price bands	Nov to Feb	Dec to Jan	Mar to Oct	Nov to Feb	Mar to Oct
Weekdays					
00:30 to 07:30	0.36	0.36	0.36	0.36	0.36
07:30 to 16:00	1.23	1.58	0.85	2.56	0.71
16:00 to 19:00	3.10	6.59	0.85	2.56	0.71
19:00 to 20:00	1.23	1.58	0.85	2.56	0.71
20:00 to 00.30	0.60	0.60	0.60	0.70	0.70
Weekends					
00:30 to 07:30	0.36	0.36	0.36	0.36	0.36
07.30 to 00:30	0.60	0.60	0.60	0.70	0.70

The ToU tariffs were designed so that an average consumption profile would result in the same bill on the experimental tariffs as on the standard flat tariff at the time. The structure of the two ToU tariffs is summarised in Table 3.4.

503 households were recruited from five regional electricity companies (formerly area boards) within GB and divided between the two experimental tariff groups. Consumption was recorded at half-hourly intervals in 146 households that were fitted with specialised demand recorders. A control group of 75 households was formed from existing load research projects.

Results were divided between customers who were previously on the standard (flat) tariff and those who were on Economy 7. The ratio between these groups was approximately 9:2 meaning that, for each ToU tariff, the sample numbers would have been approximately 205 and 45 from standard and Economy 7 respectively.

For households selected from the standard tariff, tariff A saw peak demand on winter weekdays between the times of 16:00–20:00 reduce from 0.95 kW to 0.80 kW relative to the control group. However, the maximum demand of the day remained the same and was shifted forward in time to the period of 20:00–21:30. Tariff B, on the other hand, was able to reduce the average peak demand on winter weekdays over all hours from 0.95 kW to 0.80 kW. That the peak was not shifted forward in time was perhaps because the high price period persisted from the hours 08:00–20:00 rather than the much shorter 16:00–20:00 high price period used in tariff A.

Control households on the Economy 7 tariff were characterised by much higher demand during the night. The impact of tariff A on this demand curve was much the same as it was on the standard tariff group in that the peak demand normally seen between 16:00–20:00 was shifted forward to the period 20:00–21:30. In contrast, tariff B had additional impact on the amount of energy normally consumed in the night, increasing the peak demand from approximately 3.2 to 4.2 kW.

In general, consumer attitudes towards the experimental tariffs were positive, with 77% saying they would like to remain on the tariffs if available and 85% claiming that the tariffs were easy to understand.

3.3.3 Northern Ireland Keypad Powershift (2003–04)

The Northern Ireland Keypad Powershift (NIKP) [72] ToU trial ran from October 2003 to September 2004 and was managed by Northern Ireland Electricity. The trial took place in Northern Ireland and involved 200 households. 100 households were placed in a “Price Message Group” that received a ToU tariff, while the remaining 100 were used as a control and received the standard flat rate tariff that was given to Keypad (prepayment meter) customers at that time. The structure of the ToU tariff can be seen in Table 3.5.

Table 3.5: The ToU tariff structure used in the Northern Ireland Keypad Powershift (NIKP) trial. Prices are in units of the standard tariff rate at the time (£0.1094/kWh). Data from [72].

Price band	Times	Price
Peak	16:00 to 19:00, Mon to Fri	1.60
Off-peak	00:00 to 08:00	0.60
Day	All other times including weekend	0.90

Overall consumption in the ToU groups was found to be 2% greater than in the control group. Significance of results were not reported, though with 100 households in each group, this value is unlikely to be statistically significant. Peak consumption was reported to have reduced by 12% relative to the control group, with increases in consumption during Day and Off-peak times of 2% and 4% respectively.

Annual bills for the ToU group were found to be 5.5% less than those of the control group on the standard flat tariff. However, if the control group bills were calculated using the ToU tariff, the reduction would only have been 1.5%. This suggests that some of the savings may have been passive.

3.3.4 The Energy Demand Research Project (2007–10)

The Energy Demand Research Project (EDRP) [73] ran from 2007 to 2010 within GB and was managed by Ofgem on behalf of DECC. Though principally a demand side management (DSM) programme with a directive to investigate the effect of information provision on long run consumption, it also contained two load shifting trials implemented via ToU tariffs. Overall, the trial involved some 60,000 households, 18,000 of which were fitted with SMs. The project was half funded by government to the sum of £9.75m, with a matched contribution split between the four participating energy suppliers: EDF Energy Customers Plc, E.ON UK Plc, Scottish Power Energy Retail Ltd and SSE Energy Supply Ltd (EDF, E.ON, Scottish Power and SSE).

Energy reduction interventions took a number of approaches, chief of which was information provision. "Real time displays"—which we henceforth call IHDs in line with the nomenclature of this thesis—provided live data on household consumption, access to historical consumption data including various aggregates thereof (e.g. consumption by time of day), as well as information on the carbon emissions and cost as a result of consumption. SMs together with IHDs were found to result in a 3% reduction in overall consumption. Of the information provided by the IHD, surveys indicated that consumers may be more responsive to cost information than relating to units of consumption or carbon emissions.

Energy efficiency advice was delivered via generic written letters or, more immediately, via a website. A reduction in consumption of around 5% was seen in some of the trial groups, but others showed no significant change. Persistence of consumption reduction was reported by EDF where their particular combination of SM and energy efficiency advice measured a 2.3% reduction in consumption in the first trial year and 4% in the second. No effect was measured for web based feedback of historical, one day delayed consumption data, though survey results suggested that this may have been more effective if real time data was provided.

Financial incentives were offered in the form of prizes if consumption was under a given target value over a defined period. Results were mixed and concluded that no significant change was measured.

Two suppliers, EDF and SSE, used ToU tariffs in order to incentivise demand shifting from peak times. In both trials, home automation was not included in the trial intervention and use of such devices was assumed to be low. In addition, no data was gathered on the types of appliances that were involved in the load shifting.

Table 3.6: EDF's ToU tariff structure for EDRP. Prices are in units of Off-peak rate (0.0841 to 0.0903 GBP/kWh). Data from [73].

Price band	Times	Price
Night	23:00 to 06:00	0.56 to 0.65
Off-peak	All other times	1.00
Peak	16:30 to 19:30	1.61 to 1.69

The EDF ToU trial ran from January 2009 to September 2010, within the area of London and south-east England, using the tariff structure shown in Table 3.6. It included 170 households (after exclusions) with the measurement of load shifting being determined by comparison to both a "control group" of 135 households and a "wall panel" (IHD) only group of 141 households. Participants were limited to those who were on an electricity only tariff but not the Economy 7 tariff[3]

A maximum of 4.4% reduction (at 90% confidence) in energy consumption, relative to the control group, was observed during the tariff's peak hours on weekdays relative to the wall panel group. A greater response was observed at the weekend with 8% reduction (at 95% confidence) in peak consumption relative to the wall panel group. It was reported that there was no significant overall difference in consumption between the control and wall panel group at the 95% confidence level.

Of the other variables measured, the most significant interaction with response was observed to be that of the number of people aged 16–64 in the household. Here, for each additional person aged 16–64 in the ToU and wall panel groups, the proportion of consumption in the peak period was observed to increase by 4.3% and 4.8% for weekdays and weekends respectively.

Table 3.7: SSE's ToU tariff structure for EDRP. Prices are in units of the Off-peak rate (ranging from 0.0841 to 0.0903 GBP/kWh from March to October and 10.87p/kWh to 11.46p/kWh from November to February). Data from [73].

Price band	Times	Price from Mar to Oct	Price from Nov to Feb
Night	00:30 to 07:30	0.50 to 0.60	0.50 to 0.60
Off-peak	All other times	1.00	1.00
Peak	16:00 to 19:00	1.80 to 1.90	1.80 to 2.10

The SSE ToU trial ran from November 2008 to September 2010, with trial participants selected from across GB, using the tariff structure described in Table 3.7. In contrast to the EDF trial, the ToU group of 1,418 households was divided into 10 subgroups so as to create a *two-factor* experiment with a *complete factorial* design [74, Chapter 14]—each subgroup had a different intervention applied to it in addition to the main intervention of the ToU tariff. Change in demand as a result of the ToU tariff was measured relative to households with SMs that were not on the ToU tariff. This reference group totalled 1,936 households and was also divided into 10 subgroups, each subgroup having a different intervention applied, but crucially not including the ToU tariff. The 10 subgroups on the ToU tariff were paired with the non-ToU group with the same secondary intervention. Comparisons were then made pairwise between ToU and non-ToU groups.

[3]Economy 7 is a simple ToU tariff with two rates and includes the facility for radio tele-switching of loads. See Section 2.5 for more details on tariff types.

Results showed that the percentage of consumption in the ToU tariff was reduced, but only by a small amount: a 1.5% reduction on weekdays and a 2.6% reduction on weekend days. Interestingly it was noted that this reduction was significantly larger when the comparison was made between households without a “real time display” (IHD). It was suggested that this might be an interference effect—too many changes for the consumer to assimilate at one time. Furthermore, though peak consumption levels did vary with Mosaic [75] demographic group, no could be identified which responded significantly better.

3.3.5 The Ireland Electricity Smart Metering Trials (2008–11)

The Ireland Electricity Smart Metering Trials (IESMT) [67] programme ran from 2008 to 2011 and was conducted by the Commission for Energy Regulation within the Republic of Ireland. The broad objective of the programme was to establish the potential of SM technology when combined with ToU tariffs and DSM stimuli. The trial included small and medium sized enterprise (SME) and residential consumer types, though this summary focuses on residential only.

Table 3.8: Groups and ToU tariff structure for IESMT. Prices are in units of the standard rate tariff at the time (€0.141/kWh). Data from [67].

	Prices by tariff group					
Price band	A	B	C	D	W	Control
Night (23:00 to 08:00)	0.85	0.78	0.71	0.64	0.71	
Day (All other times)	0.99	0.96	0.92	0.89	0.99	1.00
Peak (17:00 to 19:00, Mon to Fri)	1.42	1.84	2.27	2.70	2.70	
Weekend (All weekend)					0.71	
N (households)	1,368	511	1,370	509	100	1,170

The duration of the programme was split into an initial period for benchmarking, where no interventions were made and normal consumption data was collected, and the trial period, where a number of different interventions were trialled. These included four ToU tariffs, labeled A–D, of different structure and four different DSM stimuli. The DSM stimuli were: bi-monthly billing, monthly billing, bi-monthly billing with an electronic energy monitor, and bi-monthly billing with a demand reduction incentive. The energy monitor provided a view of current, historical and by-tariff-band consumption. Demand reduction incentive targets were set to 10% less than the historical (from the benchmark period) daily average consumption for that household. If a customer was successful in reducing demand by this amount over an 8 month period, a €20 reward was given.

Each ToU tariff and DSM combination were combined so as to create 16 trial groups in a *two-factor* experiment with *complete factorial* design [74, Chapter 14]. In addition, a weekend tariff (labeled W) and a control group were also included. No interventions were applied to the control group and they were allowed to remain on the tariffs they were on before trial enrolment. Table 3.8 summarises the tariff groups, tariff structures, and the number of household in each group (N). For the ToU tariffs A–D, the number of households on each were split approximately equally between the four different DSM intervention groups.

At the time of the trial, all households in Ireland were customers of Electricity Ireland. Recruitment onto the trial was opt-in and demographic profiling was used to ensure the samples were representative of Ireland as a whole. Only standard electricity tariff consumers were recruited as those on the existing Night Saver tariff (similar to Economy 7 in the UK) had effectively already experienced a form of ToU intervention. The sample numbers targeted were chosen so as to enable a 2% change in overall consumption to be detectable between groups A and C at a 90% confidence level. All households were given the same flat tariff during the benchmarking period.

Table 3.9: Consumption change by IESMT trial group. Results are quoted at 90% confidence. Data from [67].

Tariffs groups:	All	A	B	C	D	W
Overall consumption reduction (%)	2.5	2.7	3.4	1.9	2.4	3.7
Peak consumption reduction (%)	8.8	7.2	9.8	9.0	10.9	11.6

The analysis was conducted at group level, with average changes to the experimental groups' consumption profiles being measured relative to the average consumption of the control group. Two key metrics were considered: change in overall consumption and change in peak consumption. The results for the experimental groups are summarised in Table 3.9. All results were reported to at least 90% confidence level.

A 2.5% reduction in overall consumption and 8.8% reduction in peak consumption were measured when aggregating all ToU tariff groups. While it was possible to say that tariff B was more effective than C, and that tariff A was the poorest performing of all at reducing peak, other intra-tariff comparisons were not statistically significant.

The DSM intervention combining bi-monthly billing, a detailed energy statement and an electricity monitor were found to be the most effective at reducing overall consumption, resulting in an average reduction of 3.2%. However, it was not possible to isolate the effect of the electricity monitor as it was not trialled alone.

An attempt to measure the persistence of response was made by comparing consumption between the first and second halves of the trial year. While peak consumption reduction was found to be consistently strong in all groups during the second half of the trial, overall consumption showed mixed results. Arguably, this analysis approach is ineffective when used on one year of data as the observed differences may conceivably be accounted for by weather effects alone.

There was some evidence of a relationship between the peak price and the level of peak reduction observed for each tariff. As expected, increasing the peak rate enhanced the peak reduction. If this trend is real, it would imply that electricity consumption at peak has weak but non-zero price elasticity.

Some evidence of demand shifting was observed. As there was an overall reduction in consumption as a result of the interventions, an approach whereby the percentage of total consumption that fell in each half-hour block of the day was compared between the control and ToU groups. A slight increase in the fraction of consumption that occurred after the peak period was observed in the ToU group average between the times of 19:30 and 01:30.

Households on the ToU tariffs on average saved 2.5% on their bills as compared to the control group. The amount saved and the associated reduction in consumption were observed to be enhanced with the increasing wealth (UK Market Research Data was used for classifications) of the participants. No significant difference was observed between vulnerable and non-vulnerable customers.

Interestingly, the reduction in peak consumption was observed to increase for households with children. This is in line with the EDRP trial (Section 3.3.4) findings that response reduced with the number of 16–64 year olds living at the premises, but in conflict with the findings of the CLNR trial (Section 3.3.6) where it was concluded that households with people younger than 5 or older than 60 had reduced response at peak relative to those that did not.

3.3.6 The Customer Lead Network Revolution (2010–15)

The Customer Lead Network Revolution (CLNR) [76] programme ran from 2010 to 2015 and was conducted by Northern Powergrid in the north of England, UK. Funded by the Ofgem's LCNF, it sought to trial low carbon technologies including photovoltaics (PVs), HPs and EVs, as well as monitoring demand levels for various consumer types—residential, SME and industrial and commercial (I&C)—in order to assess the impact of a number of interventions. Included in these interventions was the use of a ToU tariff in the residential sector.

The ToU tariff was conducted with a sample of some 600 households over the course of two years, though data used in the analysis of the trial spanned October 2012 to September 2013. SMs and IHDs were installed in all participating households. The tariff consisted of three price points corresponding to the times of day they were in effect. These can be seen in Table 3.10.

Table 3.10: CLNR ToU tariff structure. Prices are in units of the British Gas standard rate tariff at the time. Data from [76].

Price band	Times	Price
Day	07:00 to 16:00, Mon to Fri	0.96
Peak	16:00 to 20:00, Mon to Fri	1.99
Off-peak	All other times including weekend	0.69

The effect of this tariff on demand was determined by comparison against the average demand of a control group consisting of some 9,000 households [77] with SMs. To justify this direct comparison, the control group was compared with the ToU group across a number metrics, including socio economic group (provided by Mosaic, Experian), household occupancy and number of rooms. It was determined that the groups were sufficiently representative of each other to be considered as having been sampled from the same population.

On average it was found that the ToU tariff reduced peak demand, between the hours of 16:00 and 20:00, by 6.39% [62]. This reduction was accompanied by an increase in consumption during the other times of the day, indicating that some demand shifting was occurring.

Consumers with higher income showed increased demand reduction. Furthermore, those with greater income had more home appliances and commensurately greater overall consumption. This supported the idea that response is linked to overall consumption. Higher income consumers were also found to have been users of timers for appliances since before the start of the trial, suggesting the alternate hypothesis that a greater socio-technical preparedness may also have played a part in the greater response to the tariff. However, in general, Experian's Mosaic socioeconomic classifications were not found to be a good predictor of responsiveness to the tariff.

39% of those on the ToU tariff would have paid more if their consumption was billed at the standard flat rate tariff provided by British Gas at the time, implying that nearly two thirds of those on the trial made a saving.

Households including people younger than five and older than sixty were found to be correlated with reduced response to the tariff, perhaps related to less flexible daily routines. From customer survey responses, the household tasks of laundry and dishwashing were reported to be the easiest to shift.

3.3.7 Comparison

Six key ToU trials were reviewed. These trials took place within the last 4 decades and were situated within the UK, including Northern Ireland, and the Republic of Ireland. Ireland was considered to be close to the UK in both culture and climate. A summary of the results from these trials can be seen in Table 3.11.

In general, the results are in line with the peak reductions reported by DECC in Table 3.1. Results and tariff designs varied considerably over the trials examined, with peak to standard price ratios ranging from 1.42 to 6.59, and peak consumption reductions ranging from 6.4% to

Table 3.11: Summary of ToU trial results from those considered closely related to the UK. For each trial, the table gives the peak to standard price ratio, the reduction in peak consumption achieved and the number of households in both the ToU and control groups (when available).

Index	Trial	Year	Location	Price ratio	Reduction (%)	N_tou	N_control
1	DTE [70]	1973	GB	3.00	25.0	-	-
2	EMUP [71]	1990	GB	6.59	16.0	250	75
3	NIKP [72]	2004	N. Ireland	1.60	12.0	100	100
4	EDRP EDF [73]	2010	GB	1.69	8.0	170	135
5	IESMT A [67]	2011	Ireland	1.42	7.2	1,368	1,170
6	IESMT B [67]	2011	Ireland	1.84	9.8	511	1,170
7	IESMT C [67]	2011	Ireland	2.27	9.0	1,370	1,170
8	IESMT D [67]	2011	Ireland	2.70	10.9	509	1,170
9	IESMT W [67]	2011	Ireland	2.70	11.6	100	1,170
10	CLNR [76]	2015	GB	1.99	6.4	600	9,000

25%. Significance of the reported values also varied considerably due to the number of households in the respective ToU and control trial groups ranging from 75 to 9,000. It can be seen that the more statistically significant trials tend to group together at more conservative price ratios of 1.5–2.7 and there is a lack of robust experimental evidence at higher peak to standard price ratios.

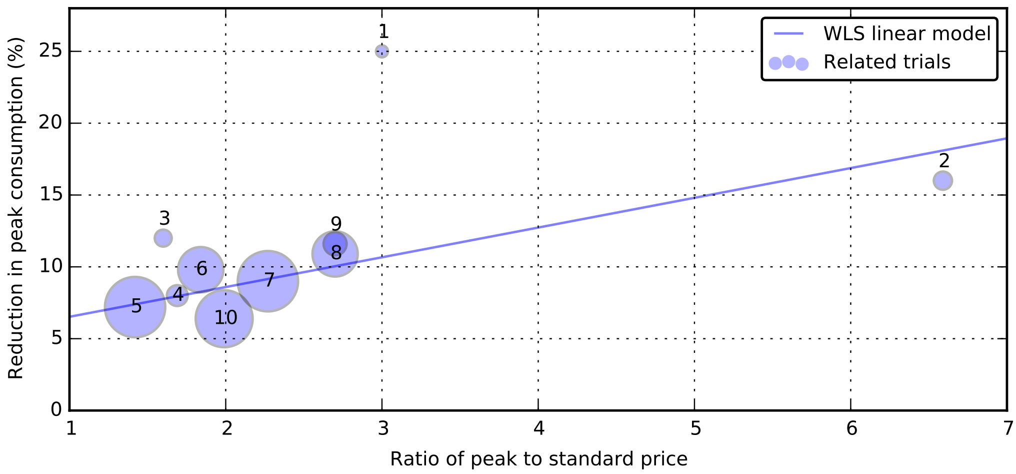

Figure 3.1: Summary of ToU trial results related to the UK context. For each trial, the size of the point is proportional to its weighting, which is a function of the sample numbers. Each trial marker is referenced to the corresponding trial in Table 3.11 via its index number.

To gain a sense of the distribution of peak reductions relative to price ratios, they are plotted together in Fig. 3.1. The number of households in each trial group were used to determine a statistical weighting of each result. Assuming that each trial group had the same standard deviation of demand—this number was not provided in most trial reports—the variance of the response measurement, S_R^2, can be shown (proof omitted) to be proportional to the below function of the sample numbers:

$$S_R^2 \propto \frac{1}{N_{\text{tou}}} + \frac{1}{N_{\text{control}}} \tag{3.1}$$

The data point weightings were set so as to be proportional to $1/S_R^2$, as is the standard in weighted least-squares regression (WLS). These were depicted in the scatter plot via the size of each point's marker—larger markers having greater statistical weight. In the case of trial index 1, each group was assumed to contain 50 households. This was considered a conservative estimate (below the minimum reported group population of 75) and allowed the point to feature in the graphic with a small weighting. In addition, a best WLS regression model consisting of a constant and a proportional parameter was fitted to the points to give a sense of whether a trend might exist.

Evidence for or against a relationship between price ratio and change in consumption was weak. The WLS regression resulted in parameter p-values each greater than 0.10, a level normally considered too weak for the model to be adopted. This highlights the need for more data to be gathered on the elasticity of electricity demand to changes in price level. Future trials might also target higher peak to standard price ratios where there is currently little experimental data.

3.4 Upcoming work of interest

The below lists related trials that were ongoing at time of writing and have not yet published results.

EcoGrid EU (2011–15): This European Union (EU) funded project aims to demonstrate a working prototype of a future smart grid concept where real-time energy markets allow the demand side to react to changes in supply. The trial is taking place within the Danish island power system of Bornholm, where wind generation accounts for more than 50% of the local capacity. Some 2,000 households, a significant proportion of the 28,000 on the island, will participate in flexible demand, of which many will be fitted with home automation and 'smart appliances'. Real time prices are calculated by the island transmission system operator (TSO) at five minute intervals in order to balance the system. [78, 79]

Smart Energy GB and UCL (2014–15): This research project, between Smart Energy GB and the University College London (UCL) Energy Institute, is investigating the acceptability of ToU tariffs to consumers. Two randomised control trials involving some 4,000 people have been conducted to test how the design and marketing of tariffs impacts their acceptability. Results are due to be presented in 2015. [80]

Project SoLa BRISTOL (2013–16): Also funded by the LCNF, this trial is being conducted in the Bristol area by local DNO, Western Power Distribution. The project's primary objective is to enable high density PV generation to connect to the low voltage (LV) network more efficiently through use of in-situ battery storage and variable tariffs. The trial aims to include thirty houses, ten schools and an office. The variable rate tariff will encourage electricity use at times of high PV generation and to use electricity stored by the battery when the network is heavily loaded. The DNO will be able to control the times when the battery charges and discharges in order to assist in network management. [81]

Vulnerable Customers and Energy Efficiency (2014–17): This LCNF funded project, run by (DNO) UK Power Networks, focuses on "fuel poor" residential consumers and seeks to quantify the benefits of ToU pricing, energy management devices and energy advice to both the network and consumer. The trial will include some 550 households from within the London borough of Tower Hamlets, which has a high percentage of fuel poor households, social housing and tower blocks. [82]

3.5 Summary and conclusions

This chapter provided an general overview of global residential demand response (DR) trials then focused on the details of six time-of-use (ToU) trials that were closely related to the United Kingdom (UK) context. The below summarises the key areas where more knowledge was found to be needed regarding the implementation of dynamic pricing in the UK. For each area, a brief description of potential research options is given.

Knowledge gaps

Present gaps in the understanding of residential dynamic electricity pricing in the UK are:

- dynamic Time-of-Use (dToU) tariffs have not yet been trialled in the UK and evidence from comparable countries (that may be extrapolated to the UK) is thin
- Data on the types of loads shifted is thin
- Evidence on the effect of household occupancy is mixed
- Evidence and reasons for differing responses from vulnerable or low-income consumers is thin
- Evidence on persistence of DR is thin
- Evidence on the effect of information provision and its vector were mixed
- Data on the response of consumers to different price levels is thin with little significant experimental data beyond a peak to standard price ratio of 3
- Data on the effect and extent of the electrification of heat and transport is thin
- Models of DR for use by network planners and operators

Research options

The following paragraphs provides a brief descriptions of potential research options for improving the understanding of each of the knowledge areas listed in the previous sections.

Dynamic Time-of-Use (dToU) trials. The complexity and subjectivity of residential demand makes it unlikely that theory alone can inform on the effect of dToU signals. Understanding this must include an experimental approach. Attention to statistically significant sample sizes, appropriate social stratification and good experimental design will be crucial in ensuring the quality of results. This has been the primary contribution of the Low Carbon London (LCL) residential dToU trial.

Data on the types of loads shifted. The most direct approach would be to sub-meter appliances at high resolution within the household. This would give unambiguous confirmation of the loads that were contributing to the measured response for a particular event. A less intrusive alternative might be identification and disaggregation of load signatures [83] from the power measurements of the primary meter, though this approach requires high resolution measurements and more complex analysis techniques. However, both these approaches do not differentiate between actions that are the result of deliberate engagement with the DR programme, and coincidence. To gain more information here, surveys may be used to ascertain the appliances that the consumer finds easiest and most difficult to respond with.

Data on the effect of household occupancy. Statistically significant data on the effect of household occupancy can be obtained by ensuring the number of samples in future trials are such that, after grouping by the number of occupants in each household, each group still contains a significant number of households. The data necessary to determine the number of occupants in each household will necessarily have to be gathered via a survey.

Evidence and reasons for differing responses from vulnerable or low-income consumers. More data on low-income and vulnerable consumers may be obtained by targeting DR trial recruitment such that statistically significant numbers are selected from low income and vulnerable consumer classes. This could be augmented by consumer surveys to obtain data on the types of appliances that consumers are shifting. Data on the effects of household occupancy should be obtained in a similar manner, by choosing trial areas and targeting recruitment until statistically significant numbers of the chosen occupancy levels are obtained. This is now being studied in a related Low Carbon Network Fund (LCNF) project, Vulnerable Customers and Energy Efficiency, discussed in Section 3.4.

Evidence on persistence of DR. Data on the persistence of peak load reductions can only be obtained from longer trials or commercial offerings. Increasing trial duration or designing trials that may transition into commercially viable operations would provide valuable long term data.

Evidence on the effect of information provision and its vector. If the experiment is designed to test information interventions, all other variables should be held constant to as greater degree as reasonably possible. If the experiment is not designed to test such interventions, the information provided to consumers, and the vector by which it is provided, should remain a constant of the experiment.

Data on the response of consumers to different price levels. As can be see in Fig. 3.1, trials that were considered closely related to the UK context have tended to be conservative in their choice of peak to standard price ratio—there is currently little experimental response data beyond a price ratio of 3. Future trials should therefore aim to obtain data beyond this point.

More data on the effect of price differentials may be obtained by designing experiments with significant numbers of participants, using multiple tariff bands by using a sufficient range of price ratios to make it possible to deduce trends (should they exist) from results. As there are many time related variables which cannot be controlled in trials, ensuring the pricing schedule is designed to minimise noise while appropriately spanning seasons will be important.

Data on the effect and extent of the electrification of heat and transport. Until penetration levels for electric vehicles (EVs) increase, it will be difficult to obtain statistically significant experimental results on their DR potential, or indeed the network issues they may pose, from individual trials. In order to make best use of the information available, current efforts are geared towards increasing penetration levels and the international sharing of research data. Green eMotion [84] is the largest such programme in Europe. Combining international data with local data such as driver patterns and consumer preferences may be the best approach at this stage in EV roll-out. In the UK, plans to test small fleets of instrumented EVs in order to obtain information on driver patterns and charging preferences, as well as channeling investment into EV charging infrastructure, are underway in major cities and, indeed, were part of the LCL programme [85].

Data on the effect of heat pumps on the network and their potential to engage in DR is in a similar condition to that of EVs, though their situation is somewhat different. As heat pumps will be replacing conventional heating systems with little to no difference in service dynamics (unlike electric vehicles which have considerable reduced range and increased charging times relative the the incumbent), usage patterns can be derived from existing data. Furthermore, as these systems are effectively reversed air conditioning units, their aggregate load characteristics are relatively well understood. As load cycles may look similar to that of cooling loads in the USA, it may even be possible to infer their DR potential from existing data. Nevertheless, assumptions like this should not be made without experimental verification. It should be noted that the scale of such trials should not need to be as large as those necessary to test consumer responsiveness to prices.

Chapter 4

Trial design

This chapter describes the experimental design of the Low Carbon London (LCL) dynamic Time-of-Use (dToU) pricing trial that took place in London over the year of 2013. Informed by the the review of literature described in Chapter 3, the objectives of the trial, experimental groupings, recruitment procedures and design of the experimental dToU tariff are described.

4.1 Objectives

This trial was concerned with the implementation of a dToU tariff in order to inform its potential future use for Constraint Management (CM) and Supply Following (SF). Through implementing both use cases in the same tariff it was believed that the trial would be more realistic: In a smart future, demand response (DR) is likely to be used to achieve multiple objectives and by multiple actors. Simultaneous implementation would therefore better inform on potential conflicts and synergies between the two chosen use cases.

The below lists the main trial design objectives including a brief description of how each was met:

- Ensure statistically robust results: The experiment was designed to detect a minimum difference between the mean group demands of 5% to at least a 90% confidence level—the level considered by convention to be borderline statistically significant [86, chapter 8].
- Demonstrate the use of the dToU tariff for constraint management: Events were designed to achieve substantial demand reductions during the identified annual peak loading periods.
- Gather data on the distribution of the DR resource over time to inform SF: Events were designed to cover a range of season and day types, and to span all times of day with different start times.
- Gather data on the effect of event duration on response magnitude: Both high and low price events were designed with varying durations and distributed in the schedule using a factorial experimental design.
- Gather data on the effect of household occupancy on response magnitude: Overall experimental group size was targeted so that it would be large enough to contain sufficient numbers of households of different occupancy levels.
- Ensure the applicability of results to London: Acorn consumer classifications were used to ensure trial groups were reasonably representative of London as a whole.

- Gather information on the types of loads shifted: Surveys and interviews were used to obtain extensive metadata on household and appliance composition, including consumer attitudes to the dToU tariff.
- Gather data on the effect social classification on response magnitude: Acorn classifications were used to stratify consumers during analysis.
- Gather data on the effect of the tariff on vulnerable and low-income consumers: Acorn consumer classifications (a non-invasive measure) were used to guide sampling, and ensure London representative numbers of each group type.

4.2 High level design

This section lists the key design decisions made in consortium with programme partners. A brief rationale is given where appropriate.

4.2.1 Experimental groups

Households in the trial were split into two groups:

- Dynamic Time-of-Use (dToU): This group received the experimental dToU tariff.
- Non-time-of-use (nonToU): This group received a standard flat tariff, one of EDF Energy's existing commercial offerings.

4.2.2 Resource and technical design decisions

Availability of resources inevitably imposed some constraints on the trial design. In addition, technical constraints were imposed by the capabilities of the systems available for use in the trial. Key decisions relating to resource and technical constraints were:

- Up to 6000 smart meters available for the trial.
- 1 year trial duration, from the beginning to the end of 2013.
- Consumption measurement made to a resolution of 30 minutes: This was the highest measurement resolution possible given the available information and communication technology (ICT) and smart meter specifications [8]. Landis and Gyr E470 smart meters were installed in all households recruited into the trial. A picture of this meter and associated in-home display (IHD) can be seen in Fig. 4.1.
- 3 price bands—named "High", "Default" and "Low"—to be used within the experimental dToU tariff: The smart meters had 3 price registers, thus allowing a maximum of 3 price bands.
- Notifications were made via the mobile network Short Message Service (SMS). The cost and complexity required to develop other solutions, for example phone applications or a web service, was considered uneconomical for the programme. Using SMS, it was possible to send notification messages to customer's mobile phones and to the smart meter in-home display, as show in Fig. 4.1. An example message reads, "From 5am Thurs 21st to 5am Friday 22nd your rate is LOW except HIGH 7am-10am".

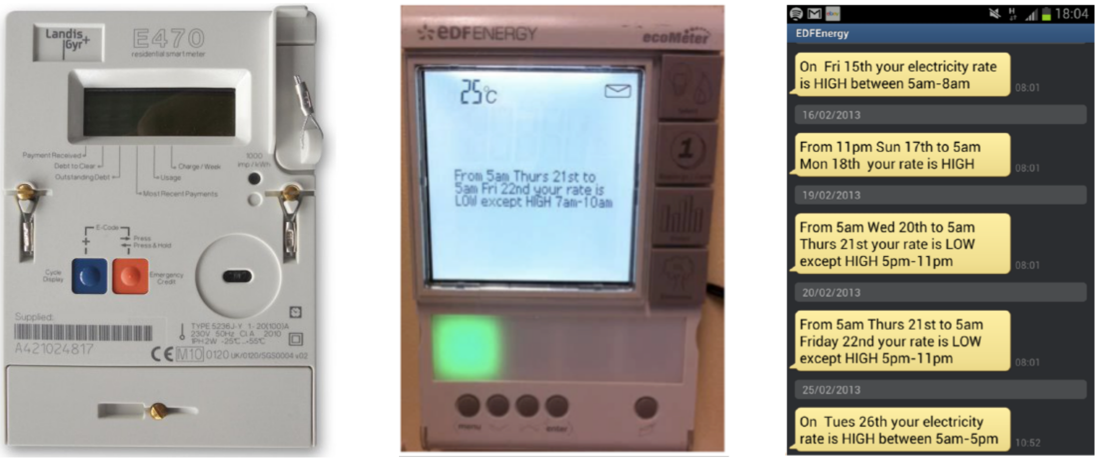

Figure 4.1: Example picture of the Landis and Gyr E470 smart meter (left) and its associated in-home display (centre) showing a rate change notification and indicating a low current load via the green light. A mobile phone screenshot (right) shows an example of the SMS price change messages.

4.2.3 Social and business design decisions

Social and business constraints also influenced design decisions. Social constraints arose from the ethical obligations relating to intervening in people's lives and collecting their personal data, and business constraints from both the industrial partners' business processes and the imperative not to damage their respective businesses—trial participants were also the customers of the partnership organisations. Particular care was taken to ensure that customer relationships were not damaged by the trial and that any personal data was kept securely and in accordance with a data privacy strategy approved by the Office of Gas and Electricity Markets (Ofgem). Key decisions in this area were:

- Opt-in trial: Consumers had a choice in whether their personal details could be used in the trial. This was true both for the treatment dToU group and the non-time-of-use (nonToU) group.
- London area only: Recruitment was restricted to the London distribution network administration area (LPN) only. This was principally because the trial was designed to give information about urban areas, and in particular the London area, but also because of the availability of network data within this area.
- Up to 3 individual low or high price event days in a week: This was chosen in order to limit inconvenience to customers as a result of being on the trial. We define an *event day* as any day containing a price band other than the default price.
- Up to 1 low-high-low event per week, but this could extend to consecutive days: There was some concern that consumers would be confused or irritated by an overly complex tariff structure. This was considered to be the maximum acceptable complexity for the tariff. The detail which could be communicated via SMS was also a key consideration in this decision.
- Notifications were made a minimum of 24 hours ahead of delivery: Though any consumer notification period was technically feasible, to limit customer dissatisfaction and allow people a reasonable time to react to price changes, it was decided that notifications would be sent at 8:30am on the days preceding events.
- Revenue neutral price bands: The tariff was designed so as to result in no annual change in

bill for Elexon[1] Profile Class 1 consumers who did not react to the dToU tariff. This was in accordance with the rules for trials as laid out by Ofgem.

Household metadata was collected from a number of additional sources. The first three of the below items were used directly in the work of this thesis. These data sources were:

- Acorn [87] socioeconomic group types for each household.
- Appliance survey: Issued at the beginning of the trial, this survey collected data on the numbers of appliances in the home as well as details relating to the characteristics of the home and its occupants (e.g. number of rooms, occupants etc.).
- End of trial survey: Relating to the attitudes of consumers towards the tariff and engagement with the tariff.
- Consumer interviews: To facilitate a longitudinal study of consumer attitudes [58].

4.3 Experimental units

There are two major sources of epistemic uncertainty in consumers' response to the dToU tariff: One stems from differences between average behaviour of households and the other from differences in the conditions surrounding each day. In addition to these, there is unavoidable aleatoric uncertainty arising from the unpredictable nature of human behaviour.

We wish to gather data on residential DR that will help us form a view of the true variation in response over households and over time. Depending on the analysis viewpoint, the *experimental unit*—the thing that is being sampled—changes. If we wish to learn how response varies with time, it is periods of time that are the experimental units, with response variations due to differences in households considered a source of noise. If we wish to learn how response varies across households, the opposite is true; variation in conditions over time are considered a source of noise. In both these cases it is necessary to ensure the source of noise is truly noise in the statistical sense: it should have a mean of zero. This is achieved by aiming for representative samples of both.

The basic experimental units of this trial were defined as:

- Households: The individual metered premises occupied by the people recruited into the LCL trial.
- Trial days: The period of 24 hours beginning at 05:00 (clock time) during the trial year of 2013.

While the reasons behind the choice of *households* is self evident, the choice of *trial day* requires some justification: The objective was to define a unit of time within which price events could exist somewhat independently of each other. With sleep being the natural divider of human days, it was desirable from an analysis perspective to define the start and end of a day according to the typical sleep cycle. Perceptually, a day begins before people wake in the morning and ends after they go to sleep at night. Assuming that the lowest point in demand in the night is when most people are asleep, we may use the Profile Class 1 data to choose an appropriate start time for the trial day. From this profile, shown in Fig. 4.8, it can be seen that the lowest demand in a 24 hour period is between the hours of 3am and 5am, with demand increasing shortly after 5am, most likely as people begin to wake. 5am was chosen as the beginning of a trial day.

[1]Elexon is responsible for administering the Balancing and Settlement Code within the Great Britain (GB) electricity system.

4.4 Household sampling

The salient details of the trial recruitment process are described below. For more detail an interested reader should view the companion report on consumer attitudes [58].

4.4.1 Recruitment process

Recruitment onto the trial was managed by EDF Energy, working closely with Imperial on points concerning experimental design. The trial recruitment steps can be summarised as:

1. *Primary recruitment* into the LCL residential trial programme. Participants were recruited on an opt-in basis from EDF Energy customers within the London Power Networks (LPN) area. There was a need to ensure the demographic spread of individuals was similar to that of London. As a non-intrusive measure, this was monitored using the Acorn consumer classification groups described in Table 4.2. If a class of consumer was found to be underrepresented, recruitment was intensified within this group until the correct ratio was achieved.
2. *Install smart meters* and IHDs in participant's homes. At this point the technician would explain the operation of the meter and IHD. For a variety of technical reasons it was not possible to install smart meters into all the homes recruited. In this case, primary recruitment was repeated until recruitment targets were met.
3. *Secondary recruitment* into the experimental groups. Participants were recruited, again on an opt-in basis, into the dToU group from the set of existing primary programme participants. Those not recruited into the dToU group formed the nonToU group.

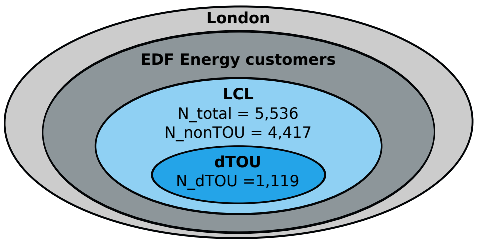

Figure 4.2: Venn diagram illustrating sample selection including final recruited numbers.

Figure 4.2 illustrates the sampling population subsets. It should be noted that, because secondary recruitment was also opt-in, and taken from the group of households created by the primary recruitment drive, it was assumed that self selection had created a natural difference between the populations of the dToU and nonToU groups.

Partly to mitigate the issue of disparities between the groups, and to enhance the collection of meta-data, customers were offered several incentives. These are discussed in detail in LCL report on consumer attitudes [58]. The salient points are given below:

- A guarantee that they will be reimbursed at the end of trial if they are worse off on the dToU tariff than they would have been on their previous tariff.
- £20 for returning the appliance survey.
- Assurances regarding how many hours would be charged at the high price band.
- £100 for signing up to the dToU tariff .
- Another £50 for staying on the dToU tariff until the end of trial.
- £20 for returning the consumer dToU tariff survey at the end of the trial.
- Entry into a prize draw after completion of the post trial survey.

4.4.2 Selection criteria

In order to be offered the dToU tariff, customers had to meet a number of conditions. Readers interested in the motivation for these conditions should refer to the companion report on consumer attitudes [58]. The conditions were:

- Be an existing LCL trial customer.
- Have had their smart meter (SM) installed more than 1 month before the trial start date.
- Have an IHD installed and working.
- Not be on the Economy 7 tariff.
- Not be on a Dual Fuel tariff. Approximately 5.4% of the EDF Energy customers within the LPN area were dual fuel.
- Have the trail standard smart meter type: Landis and Gyr E470.
- Not have a prepayment meter: No smart meter with a pre-payment facility compatible with existing infra-structure was available at the time of recruitment.
- Not have micro-generation. This is because, clearly, micro-generation affects net demand profiles, but also micro-generation is known to affect energy use behaviour.
- Not have debt or any special conditions associated with their account.
- Have completed the meta-data survey.
- Not have plans to move home within the first 6 months of the trial.

4.4.3 Group populations

Statistically robust results were a trial design objective. Specifically, the target here was to be able to detect a minimum difference between mean group demands of 5% to at least a 90% statistical confidence level.

A significant part of the experimental design was to ensure that differences between averaged grouped measurements would be detectable to within a reasonable level of statistical confidence. Before the decision that the trial would be composed of two groups (dToU and nonToU), consideration was given to experimental designs with multiple intervention groups. It was envisaged that these intervention groups would be compared to a single control group. In this scenario, the optimal number of samples to allocate to each experimental group can be calculated. Readers interested in these equations can find them in Appendix A.1.

Using informed assumptions as to the dispersion of group demand and level of change that might be observed (details in Appendix A.2), it was concluded that it would not be feasible to achieve the desired measurement resolution and significance if the available samples (estimated to be 6,000 households) were split between more than one treatment group. It was therefore decided that there would be one treatment group and one control group (which later became the nonToU group).

Given two trial groups, in order to maximise the statistical confidence of group comparisons, the optimum population numbers for each of the groups would be an equal split of the total number of experimental units (households) available. However, it was not possible to achieve this due to the opt-in nature of recruitment. Guide targets for minimum group populations were set in order to achieve the trial's statistical objectives. Taking into consideration an estimated group attrition rate of 20% and assuming equal group populations, it was calculated (details in Appendix A.2) that 1,521 households should be recruited into each trial group in order for statistical robustness objectives to be achieved. It was also recognised that population shortfall in one group may be compensated for by increase in another group.

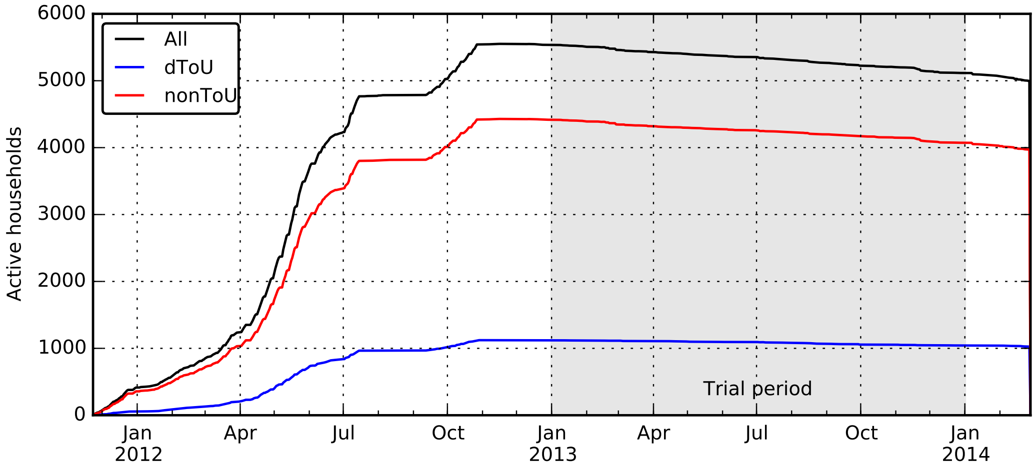

Figure 4.3: Population of trial groups with time.

Figure 4.3 shows the group populations against time with final group numbers given in Table 4.1. While the dToU group fell below the above mentioned guide population number, the statistical objectives were still achieved due to the larger size of the nonToU group.

	nonToU	dToU	Total
Beginning of trial	4,417	1,119	5,536
End of trial	4,073	1,043	5,116
Attrition rate	7.8%	6.8%	7.6%

Table 4.1: Trial group population numbers and attrition rates.

4.4.4 Group demographics

Recruitment took place within the LPN area. The approximate[2] locations of the households in the trial are show in Fig. 4.5. As a non-intrusive measure, Acorn group classifications were used to monitor demographic spread. Figure 4.4 shows the breakdown by Acorn group name of the nonToU and dToU groups as compared to EDF Energy customers within the LPN area. In general, both trial groups were shown to be reasonably representative of EDF Energy customers in the London area, who were in turn shown to be representative of London, though this data is proprietary and cannot be published.

[2]Noise was added to the locations in order to protect the anonymity of the trial participants.

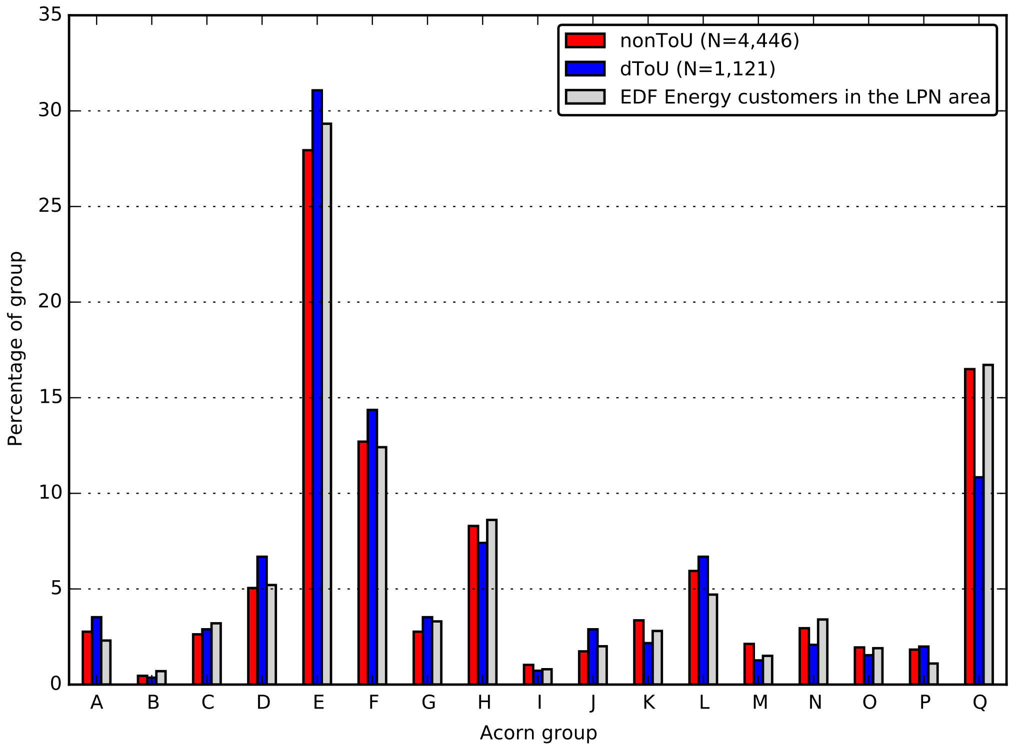

Figure 4.4: Proportions of Acorn groups for the dToU group the nonToU group and EDF Energy customers in the LPN area. See Table 4.2 for label definitions.

Label	Acorn group	Acorn category
A	Wealthy executives	Wealthy achievers
B	Affluent greys	Wealthy achievers
C	Flourishing families	Wealthy achievers
D	Prosperous professionals	Urban prosperity
E	Educated urbanites	Urban prosperity
F	Aspiring singles	Urban prosperity
G	Starting out	Comfortably off
H	Secure families	Comfortably off
I	Settled suburbia	Comfortably off
J	Prudent pensioners	Comfortably off
K	Asian communities	Moderate means
L	Post industrial families	Moderate means
M	Blue collar roots	Moderate means
N	Struggling families	Hard pressed
O	Burdened singles	Hard pressed
P	High rise hardship	Hard pressed
Q	Inner city adversity	Hard pressed

Table 4.2: Acorn group names and categories [87].

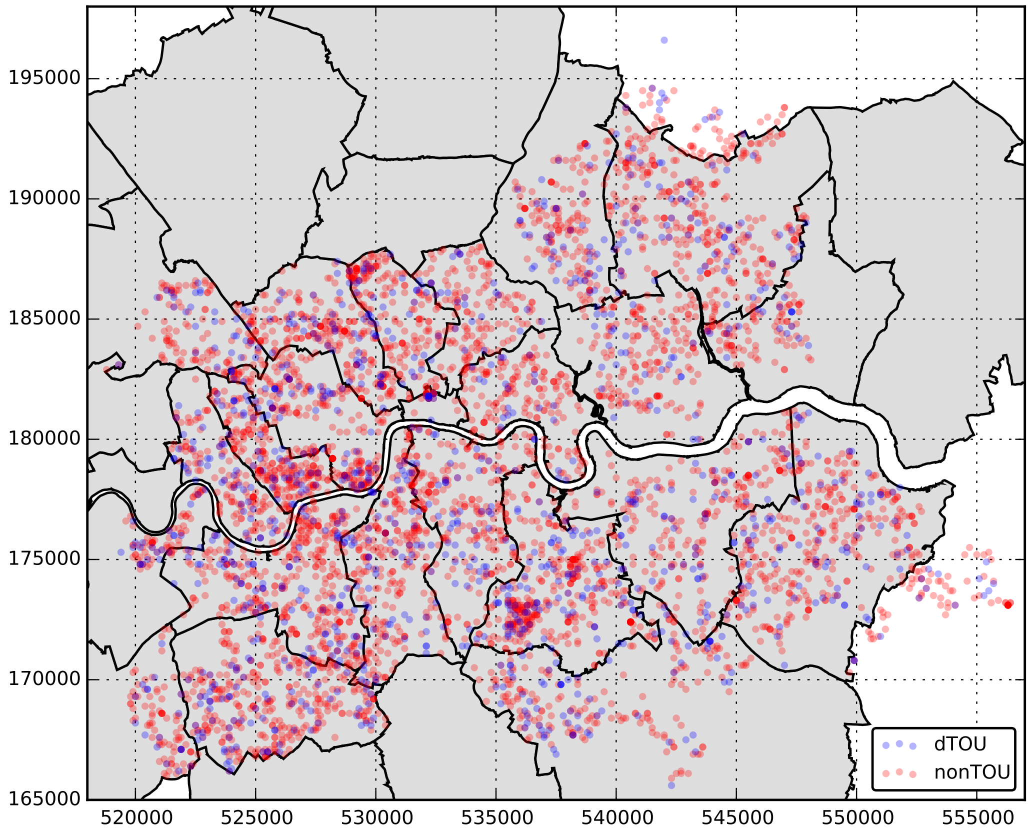

Figure 4.5: Trial household sample locations overlaid on the borough boundary map of Greater London. Map data from the Greater London Authority [88]. Ordnance Survey [89] coordinate system.

4.5 Parametrisation of demand response

This section describes the parameters through which a DR action in the dToU tariff can be quantitatively described. As the objective of the trial is to test consumer response to the tariff, we are only concerned with those attributes of the tariff that are observable by the consumer.

The salient parameters are listed below. They are not considered to be exhaustive, indeed many can be subdivided into finer details. Rather, they are those that are considered necessary to inform the design of the tariff.

- *Notice* period given in advance of event: This may range from almost no notice, in the case of fast reserve for an unpredicted plant failure, to months ahead, in the case of regular peak shaving for predictable system load.
- *Price* at which electricity is sold at time of use: For this trial, three price bands were available, referred to as the *high*, *default*, and *low* price bands.
- *Duration* for which a DR action is required: This could range from just a few seconds per event, as in the case of frequency response by demand management (FCDM) [17], to hours, as might be the case for use in peak shaving actions.
- *Timing* of events: Time of day, day of week and season of year were considered.
- *Frequency* of events: How often the events should occur within the schedule.
- *Persistence* of events: the number of consecutive days over which an event should occur. It was a design decision to limit this to a maximum of 3 consecutive days.

In this trial, price events were considered to be the application of the high or low price bands instead of the default price. Designing events required picking appropriate DR parameter values to achieve the objectives outlined in Section 4.1. Notice period was fixed at 24 hours in advance of the event and the design constraint to have revenue neutral pricing meant that the levels of the three price bands were dependent on the structure of the tariff. This left the duration, timing, frequency, persistence and price as the parameters with which to design events.

4.6 Supply following events

One of the key drivers for the investigation of dynamic pricing for Supply Following (SF) is the increased penetration of renewable generation. Within the context of the UK, renewable generation is predominantly wind, and is likely to remain so for at least the next decade. Though market price will depend on many factors, including generation outturn and demand at the time, as a rough approximation, it is assumed that wind output will be a strong determinant. The basic assumption was that higher supply will result in lower prices and vice versa. The design of supply following events was therefore informed to some degree by data and literature on the variability in wind power output.

4.6.1 Parameters

Prior work [90] shows that, seasonally, the proportion of annual wind power is biased towards the winter months, and by time of day, towards the afternoon. In both these partitions, the difference between highest and lowest output was less than 50%, which justifies obtaining data on all times of day and seasons of year.

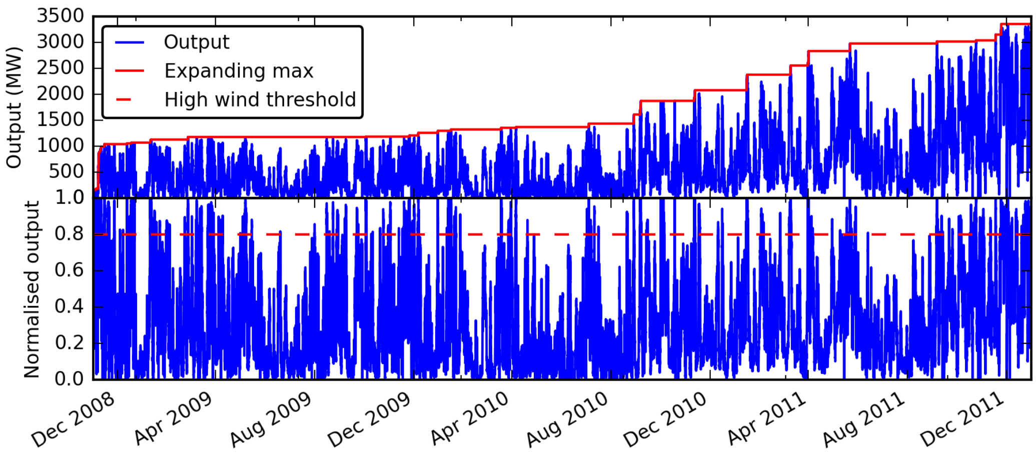

Figure 4.6: Wind output, and normalised against expanding maximum. The green line illustrates a conceptual high-wind threshold.

To understand the durations of high-wind events, analysis of Elexon's wind power output data [91] was undertaken. Figure 4.6 shows historic wind output data (above) and normalised against the expanding maximum (below) in order to account for capacity increases. The dashed red line indicates a conceptual *high-wind* threshold (arbitrarily chosen). The duration of a high-wind event is defined as the continuous time period over which the normalised output is above a chosen high-wind threshold value. The distribution of high-wind durations was measured for a range of threshold values, from 0.1 to 0.9 in increments of 0.1. The following observations were made:

- The choice of threshold has little impact on the distribution of the durations of the events.

- 70% of high-wind events are below 3 hours in duration.
- Durations of between 3 and 20 hours occupy the next 20% of cases.

From this it was decided that SF event durations of 3, 6 and 12 hours would be used for high and low events, with a 24 hour duration event for low price alone—it was considered unreasonable to impose a 24 hour high price on consumers. These numbers were chosen because they are factors of 24 and thus cleanly subtend the day while also covering the majority of wind event durations.

4.6.2 Experimental design

It was identified in the previous section that SF events should collect data across a range of event durations and times of the day. As it is not possible to test more than one event in a day, it is implicit that inter-day effects are also present. A factorial experimental design [74] is used to allow straightforward isolation of the effect of variables *duration* and *time-of-day*.

Factorial design theory

For illustration, let us assume the below first order (ignoring potential variable interactions), linear (meaning variable effects are additive) response model:

$$y_{i,j,k,n} = \mu + A_i + B_j + C_k + \epsilon_{i,j,k,n} \tag{4.1}$$

where $y_{i,j,k,n}$ is an observation of the response, μ is the overall population mean response, A_i is the effect of duration indexed by i, B_j is the effect of time-of-day indexed by j and C_k is the inter-day effect indexed by k. $\epsilon_{i,j,k,n}$ is the noise term that measures the deviation of observation $y_{i,j,k,n}$ from the true population mean response $\mu_{i,j,k}$, where n is the observation count for the set of conditions $\{i, j, k\}$.

While the indices of the variables *duration* and *time-of-day* are easily definable, it is difficult to identify, yet alone quantify the variables that are changing in the inter-day case. Furthermore, even if they could be quantified and measured, they would likely not be controllable and therefore they would not be replicable—a necessity for experimental design.

Though inter-day variations are not random—we know there are underlying causes and structures in their effects—it is possible to approximate them as such so long as they are sampled at random. This means we must select event days, our experimental units, randomly from the trial year. The noise from this new random variable is assumed to be indistinguishable from the existing random noise term, $\epsilon_{i,j,k,n}$, and as such they are modelled as one. Therefore, so long as we randomise the event days throughout the trial year, we may use the below, simplified response model:

$$y_{i,j,n} = \mu + A_i + B_j + \epsilon_{i,j,n} \tag{4.2}$$

$$\mu_{i,j} = \mu + A_i + B_j \tag{4.3}$$

where $\mu_{i,j}$ is the population mean response at level indices i, j,

$$\sum_{i=1}^{a} A_i = \sum_{j=1}^{b} B_i = 0 \tag{4.4}$$

$$\mu = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \mu_{i,j} \tag{4.5}$$

and a and b are the number of different levels of variables A and B respectively.

This response model is used as the basis of a *two-factor* experimental design in order to reduce the uncertainty in variable effect estimates. This is achieved by pairing each level of variable A with each level of variable B in such a way that the effects of variable B sum to zero when averaging over the levels of variable A. The effect of variable A may therefore be estimated without the need to consider the influence of variable B, and *vice versa*.

$$\begin{array}{c|cccc} & B_1 & B_2 & \cdots & B_b \\ \hline A_1 & \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,L_d} \\ A_2 & \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,L_d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_a & \mu_{1,b} & \mu_{2,b} & \cdots & \mu_{a,b} \end{array}$$

Table 4.3: Theoretical two-factor treatment matrix.

This design is depicted in Table 4.3. Here, the mean response to a particular factor level index i is given by:

$$\mu_i = \frac{1}{b} \sum_{j=1}^{b} \mu_{i,j} = \mu + A_i \tag{4.6}$$

Using this design the analysis of the experiment is simplified and the confidence in estimates of parameter effects are enhanced.

Implementation

For the above discussed two-factor design to be implemented, the time-of-day and duration variables must be defined independently of each other, however this is not straightforward. For example, the period of the day during which an event is in effect is dependent on both its start time and duration, so the start time of the event is not a perfect measure of time-of-day.

The design chosen involved arranging events of each duration so as to sequentially cover all times of day. In this way a large proportion of the effects caused by time-of-day should cancel when averaging response over all blocks of a particular duration. Likewise, when averaging over a particular time-of-day, the effects of event duration should largely cancel.

Price band	Duration (hours)	Number of unique start times	Repeats	Total
High	3	8	3	24
Low	3	8	3	24
High	6	4	3	12
Low	6	4	3	12
High	12	3	3	9
Low	12	3	3	9
Low	24	1	3	3
			Total	93

Table 4.4: List of SF event numbers by duration, price band.

The described experimental design approach was repeated for both high low price events. The list of unique SF events created by this process is summarised in Table 4.4 and depicted in Fig. 4.9. Each of the event types listed were repeated 3 times and all events were randomly distributed throughout the days of the trial year.

4.7 Constraint management events

The Constraint Management (CM) use case is focused on the mitigation of network constraints at the distribution level. Network constraints are most likely to occur during the regular peak demand periods. While the cause of constraints, even in a residential area, may be other than residential load, for the purposes of designing this trial, the focus was on mitigating residential load-dominated constraints.

4.7.1 Parameters

Peaks in residential load were used as a proxy for the likely timing and duration of constraints. Elexon's Profile Class 1 data was used to inform this. Figure 4.8 shows the Profile Class 1 data for 2010, split by the Elexon defined seasons. High demand periods are identified as light red shaded areas. These were the time periods during which it was considered feasible for a constraint management event to occur. Peaks are labeled in the form "Px", where x is a number. The targeting of peaks was prioritised in order of the peak magnitudes with attention given to ensure a mix of day types. This inevitably meant that they were clustered around the winter months. Peaks $\{1, 3, 4, 6, 8, 9\}$ were chosen for testing.

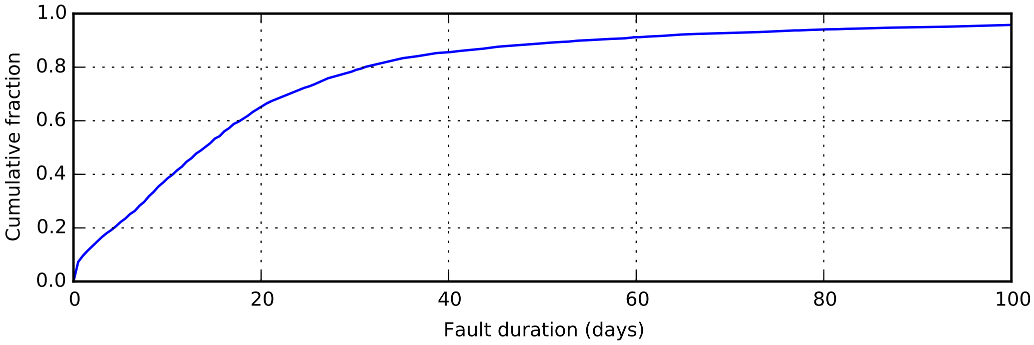

Figure 4.7: Empirical cumulative distribution plot of fault durations for 4,233 recorded faults in the LPN area from 2003 to 2011. Data courtesy of UK Power Networks.

To determine the persistence of network constraints, programme partner UK Power Networks provided historic data on network faults. The durations of some 4,000 network faults (time of reporting to resolution), are plotted in Fig. 4.7. It can be seen that, in the vast majority of cases, fault durations greatly exceed the "3 event days per week" constraint of this trial. Nevertheless, as a proof of principle, it was deemed useful to simulate the use of a dToU tariff for constraint management. In a commercial offering, greater event persistence may be possible by interleaving subgroups of consumers so as to provide DR on alternate days, or it may be that DR is only one component to the resolution of a sustained outage. It was decided that events would be varied in persistence, from one and up to the maximum of 3 consecutive days.

Due to the potentially high cost of failing to manage a network constraint, it is likely that a future dToU tariff would be structured to incentivise the maximum possible DR during the critical period. In this sense it is likely to follow the design logic of a critical peak pricing (CPP) tariff. Within the trial's design constraints, the maximum possible DR is likely to be achieved when a high price period is enclosed on both sides by low price periods, so as to both disincentivise use during the critical period as well as incentivise load shifting into the periods before and after. CM events were structured accordingly.

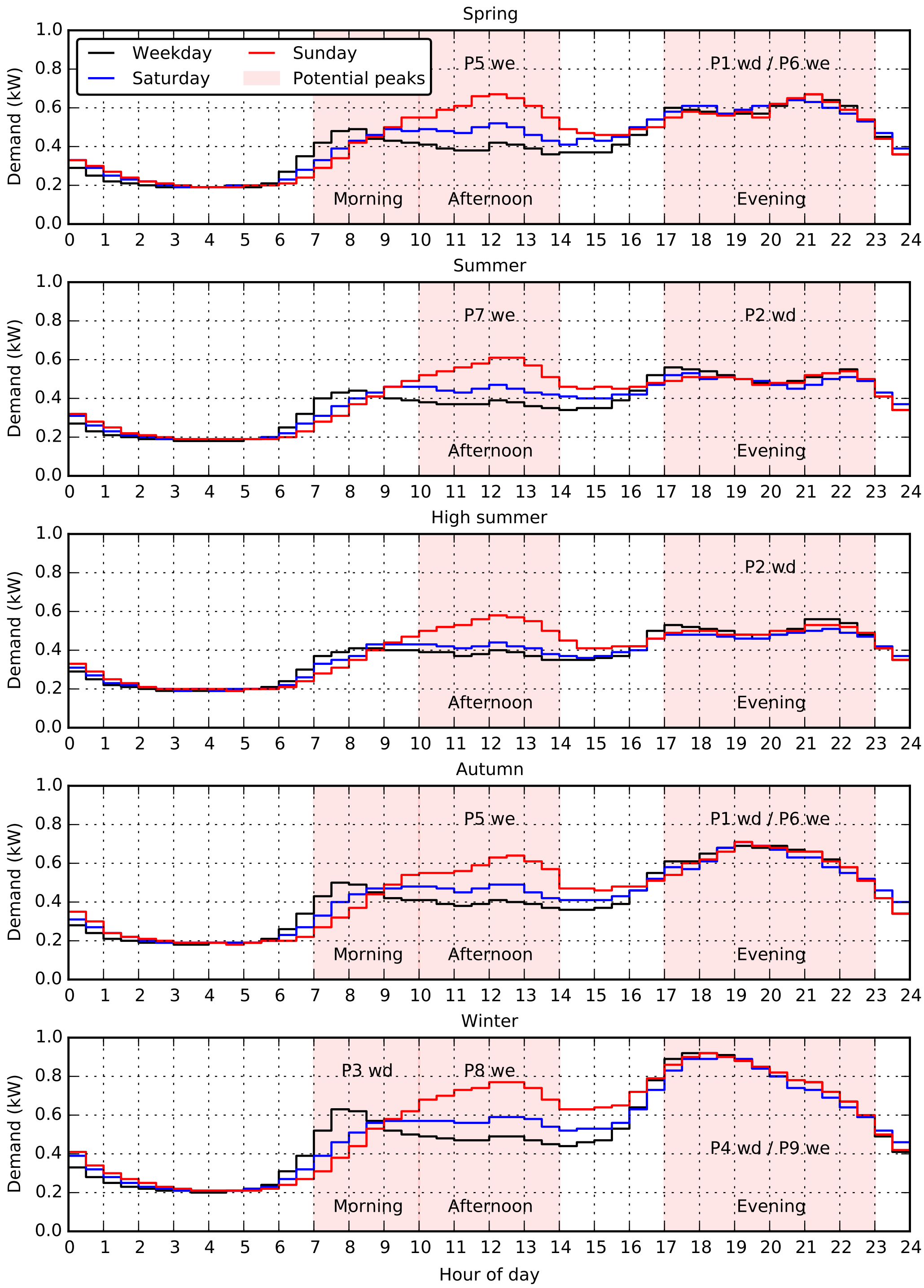

Figure 4.8: Elexon's profile class 1 data, 2010 [92]. Light red shaded areas indicate the time spans over which it was considered feasible for peaks, and therefore CM events, to occur. Labels of form "Px" are assigned to the peaks for both weekdays ("wd") and weekends ("we").

4.7.2 Experimental design

The division of experimental units (trial days) between the SF and CM events was driven by their respective requirements. SF events were designed to gather information by sweeping across all times of day at a variety of different durations. This placed a significantly greater demand on the number of samples needed relative to the design of CM events. Here the use case was more tightly defined—events should target the highest peaks in the year across a variety of day types (weekday and weekend). After deduction of the of the event days required for SF, 26 event days remained to test CM events.

Table 4.5 lists all CM events designed for the trial. Events are named according to "Px_yD", where x is the peak number as defined in Fig. 4.8 and y is its persistence—number of consecutive days this event acted over. The 'peak from' and 'peak to' columns indicate when (measured in wall clock time) the high price band is applied.

Event index	Event name	From	To	Consecutive
1	P9_2D	Sat 19 Jan 17:30	Sat 19 Jan 23:00	⎤
2	P9_2D	Sun 20 Jan 17:30	Sun 20 Jan 23:00	⎦
3	P3_1D	Tue 29 Jan 07:30	Tue 29 Jan 10:00	
4	P8_1D	Sat 09 Feb 10:30	Sat 09 Feb 14:00	
5	P4_2D	Wed 20 Feb 17:30	Wed 20 Feb 23:00	⎤
6	P4_2D	Thu 21 Feb 17:30	Thu 21 Feb 23:00	⎦
7	P9_2D	Sat 16 Mar 17:30	Sat 16 Mar 23:00	⎤
8	P9_2D	Sun 17 Mar 17:30	Sun 17 Mar 23:00	⎦
9	P3_1D	Thu 21 Mar 07:30	Thu 21 Mar 10:00	
10	P4_3D	Wed 27 Mar 17:30	Wed 27 Mar 23:00	⎤
11	P4_3D	Thu 28 Mar 17:30	Thu 28 Mar 23:00	⎥
12	P4_3D	Fri 29 Mar 17:30	Fri 29 Mar 23:00	⎦
13	P1_1D	Tue 23 Apr 17:30	Tue 23 Apr 23:00	
14	P1_3D	Wed 01 May 17:30	Wed 01 May 23:00	⎤
15	P1_3D	Thu 02 May 17:30	Thu 02 May 23:00	⎥
16	P1_3D	Fri 03 May 17:30	Fri 03 May 23:00	⎦
17	P6_1D	Sun 13 Oct 17:30	Sun 13 Oct 23:00	
18	P6_1D	Sun 20 Oct 17:30	Sun 20 Oct 23:00	
19	P1_2D	Tue 26 Nov 17:30	Tue 26 Nov 23:00	⎤
20	P1_2D	Wed 27 Nov 17:30	Wed 27 Nov 23:00	⎦
21	P9_1D	Sun 15 Dec 17:30	Sun 15 Dec 23:00	

Table 4.5: List of CM event days in indexed order of appearance in the trial. "Wall clock" time is used.

4.8 Overall pricing schedule

The SF events were randomly distributed throughout the year, with CM targeted on the remaining viable event days (not breaching design rules listed in Section 4.2). Due to the nature of their use case, they inevitably targeted the winter months, making this time of year significantly busier than the summer months with regards to price events.

Figure 4.9 gives a graphical representation of the unique events in the trial. The first letter of each event name refers to the type of event: "H" and "L" for high and low price SF events respectively, and "P" the for peak targeting CM events that are named as described in Section 4.7.2. SF events are named according to "Hx_y", where x is the duration of the event in hours and y is the the starting hour (in "wall clock" time).

The full pricing tariff schedule can be seen in Fig. 4.10. Care was taken to ensure that events did not occur on the day of the change from Greenwich Mean Time (GMT) to British Summer Time (BST) due to the complications this might present in analysis.

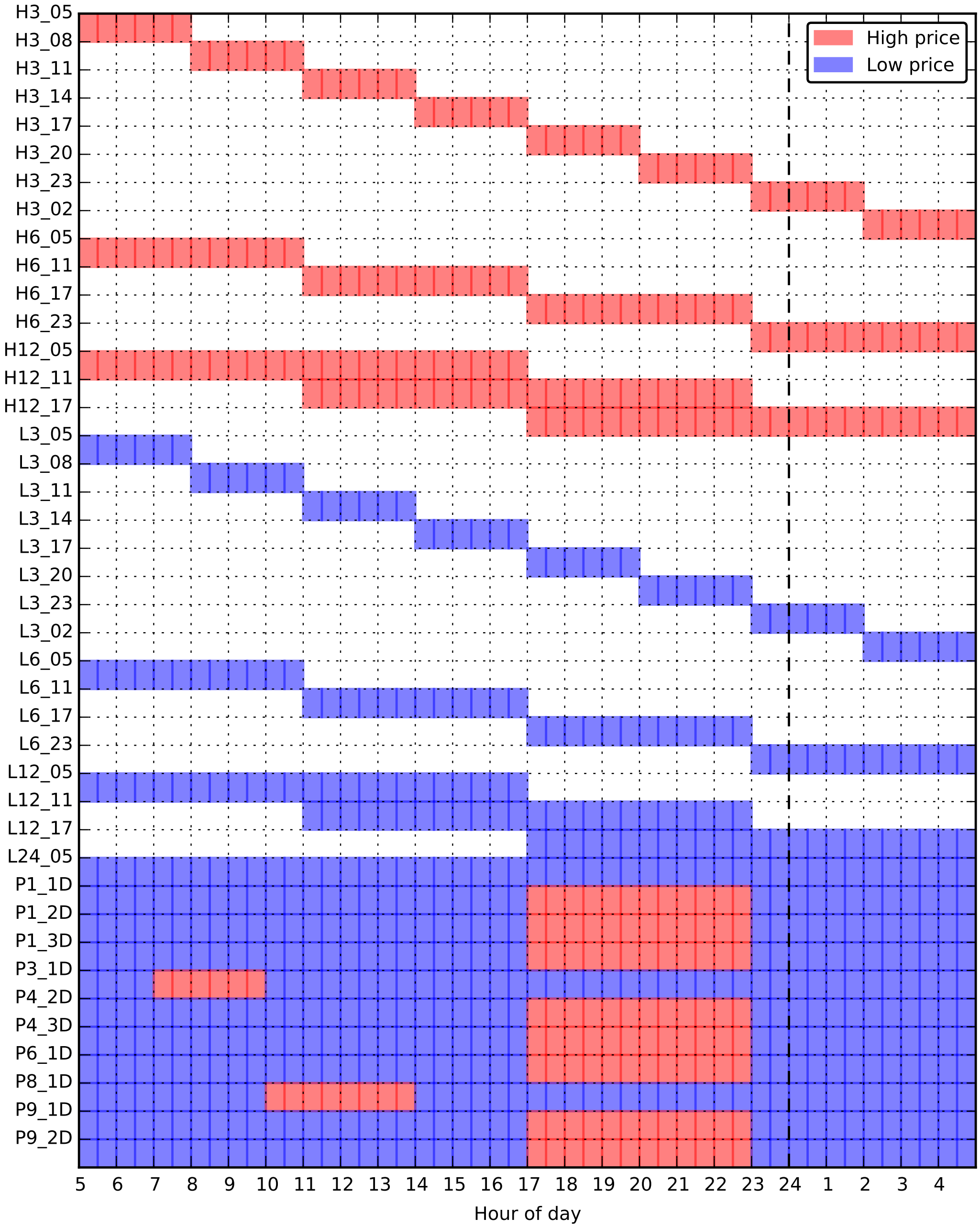

Figure 4.9: Graphical representation of unique event types. 'H' and 'L' prefixes are used for SF events, 'P' for CM events. Hours of the day are arranged in the order in which they occur in the *trial day*. Further details for the CM events ('P' prefix) can be found in Table 4.5.

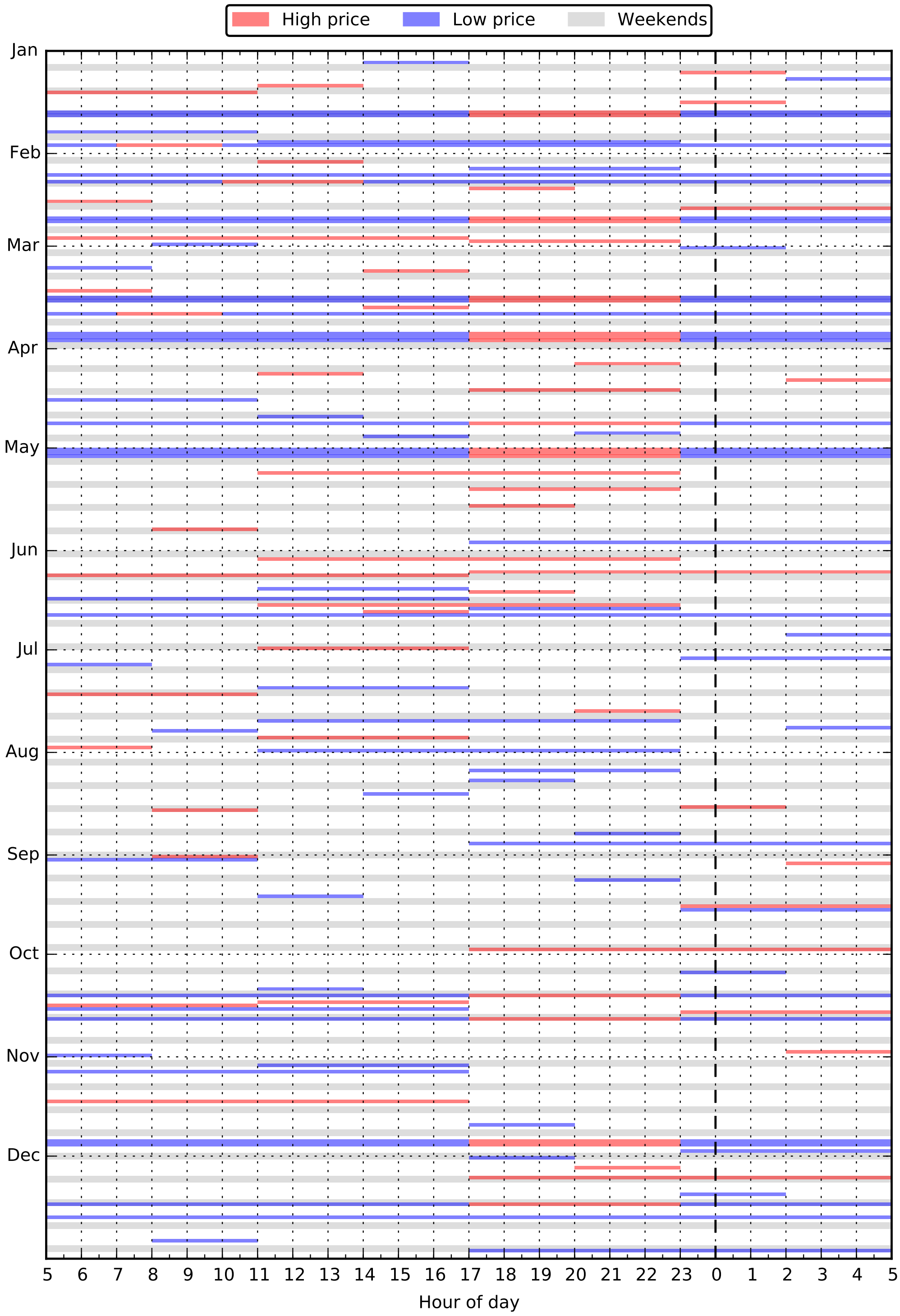

Figure 4.10: Event schedule in “wall clock” time.

4.8.1 Price bands

As the first trial of its kind in the UK, there was some uncertainty as to the level of response, if any, that dynamic pricing would have on consumption. It was therefore considered prudent to set price bands so as to maximise the expected response, and therefore the chance of detecting a change in consumption. Furthermore, as related trials of a static time-of-use (ToU) type have, for a variety of reasons, tended towards more conservative price band ratios, it was of interest to enter new territory with a greater price ratio. This also suited the CM use case, where network critical constraint situations suggest closer analogy with CPP tariffs.

As consumer bills should be revenue neutral in the case that they did not respond to DR events, the values of the high, low and default price bands were somewhat dependant on each other. The calculations of the price bands was carried out by EDF Energy. The rates for the price bands were:

- High: 67.2 pence/kWh
- Default: 11.76 pence/kWh
- Low: 3.99 pence/kWh

The nonToU tariff group was charged at fixed standard rate of 14.228 pence/kWh.

4.9 ICT architecture

Responsibility for the design and operation of the trial's information and communication technology (ICT) architecture was distributed between the programme partners associated with this trial. To a high level and in approximate order of information flow these roles are summarised below:

- EDF Energy were responsible for installing the SMs and administering the tariff pricing schedule (sending prices to meters) once the trial was in operation. As the party in direct contact with the consumer, they also managed the communication with customers, billing, and administered the collection of the survey results that provided the majority of the metadata.
- Logica (now CGI) were responsible for the infrastructure to communicate with the smart meters, known as the "smart meter head end".
- Siemens were responsible for designing the Operational Data Store (ODS), a database for the smart meter and network data collected from all trials within the LCL programme. This task included the system integration aspects necessary to ensure the data was delivered to the ODS.
- UK Power Networks, as the lead programme partner, were responsible for overall coordination. They were also the custodians of the data collected in the trial. This involved hosting and administering the ODS database once it was operational and designing then administering the Participant Management System (PMS) database, a database for household metadata including survey results. As the distribution network operator (DNO) they were also responsible for supplying network operational data for the ODS, though the network data was mainly used in other LCL trials.
- Imperial College London were responsible for the experimental design of the trial and the subsequent analysis of the data collected. This entailed providing a secure data link between the ODS site and the Low Carbon London Learning Laboratory (hosted by Imperial) where the data analysis was undertaken. Secure servers were used for analysis of data containing personally identifiable details.

Figure 4.11 shows a high level diagram of the ICT architecture with arrows indicating information flows. The supplier (EDF Energy) is able to send price signals via the CGI head end (Logica) to the smart meter (SM) via the GSM network. The customer may then view the prices for the upcoming 24 hours on the in-home display (IHD), which uses the Zigbee protocol for low power wireless communication to communicate with the SM. In addition, the supplier can also send details of the upcoming day's schedule to customer's mobile phones via SMS service.

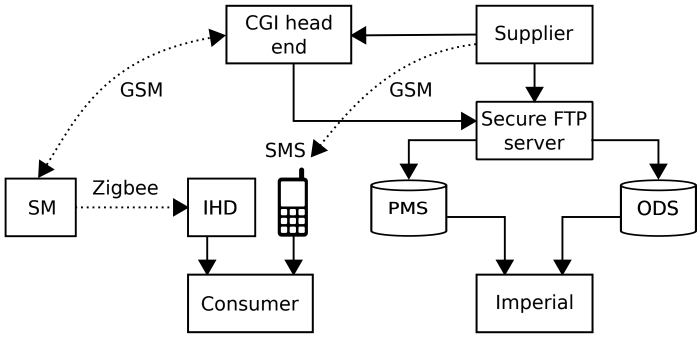

Figure 4.11: ICT architecture with arrows indicating information flows.

Consumption data is fed back from the SM to the head end where it is transmitted to a secure file transfer protocol (FTP) server. This server then updates the ODS with the consumption information. The supplier may also access the ODS consumption data in order to inform billing.

The PMS was updated as and when new information became available. This was usually when transcribed survey results became available.

The ODS was composed of two parts. For the time-series consumption data, an OSIsoft PI server [93] was used, while network operational data was stored in an associated Microsoft SQL Server [94]. The PMS also used the Microsoft SQL Server platform. Both databases allowed read-only Structured Query Language (SQL) queries to be executed from a securely linked server at the Imperial College site.

At the Imperial end, ODS data was accessed via SQL scripts. As the desired data sets were often large (some 2GB including consumption data), Python scripts were developed to manage data extracts. These split large data extracts into batches allowing lengthy requests (sometimes up to 12 hours) to be monitored for progress as well as performing some basic validation tests to avoid corruption. Data was parsed, using the Python language Pandas library, [95] into DataFrames, python native data structures, then serialised to disk using the Python language Pickle protocol. This offered fast access to an analysis ready format. In general, the PyData [96] ecosystem (of which Pandas is a part) provided the mainstay of the analysis tools used for this thesis.

Chapter 5

Analysis basis

This chapter describes the basis of analysis that is used throughout subsequent chapters. It begins with a summary of all data collected for use in the analysis described in this thesis. Next, smart meter (SM) data integrity is examined and discussed, followed a description of the data cleansing process and the definition of the data set used in subsequent analysis. The final section concerns the measurement of demand response (DR) and centres on the design and validation of the baseline demand model.

5.1 Summary of data collected

At a high level, raw data sources can be summarised as:

- Consumption data: Measured at 30 minute resolution totalling over 168 million measurements between 1,126 PI Server tags in the experimental dynamic Time-of-Use (dToU) group and 4,486 tags in the the non-time-of-use (nonToU) group. Tags should ideally have a one-to-one mapping with households, though exceptions exist. This is discussed further in Section 5.3.1.
- Acorn consumer classification data [87] for all 5,567 households within the trial.
- Appliance survey: Issued at the beginning of the trial, this consisted of appliance ownership numbers, energy relevant physical details of the premises (e.g. insulation) and basic details of its occupants. 1,870 submissions were received from the nonToU group and 990 from the dToU group.
- Post meter installation survey: Questions were compiled by EDF Energy and mainly concerned customer satisfaction with the service provided by the supplier. It is listed here for completeness, though results were not used in the analysis described in this thesis. 209 submissions were received from the nonToU group and 49 from the dToU group.
- Attitudes survey: Targeted at those on the dToU group, it was designed to assess attitudes and behaviour change related to the tariff. Additional focus was given to the factors that enabled and hindered responsiveness to the dToU tariff. 714 submissions were received in total.

5.2 Smart meter data integrity

This section focuses on the significance of dropouts (missing values) in the SM data feed. The first validation involved inspection of the SM load data. It was confirmed that the daily and seasonal load profiles were of the expected shapes (similar to those shown in Fig. 4.8) and with annual consumption values within expected ranges (similar to those shown in Table A.1).

Figure 5.2 shows a heat map of the fraction of missing consumption data by measurement block for 5,103 households during the trial year of 2013. Only households where the first and last records spanned the trial year were included so that results were not skewed by people deliberately leaving the trial. In general, no more than 3.4% of data was missing in any one 30 minute measurement period, the mean missing data per period being 0.17%, with a standard deviation of 0.11%.

The missing data is characterised by apparently random moments high intensity dropouts (the darkest dots) with an underlying low level of structure dropouts. The structured dropouts appear to span days as they manifest as horizontal lightly shaded lines. Analysing causes is not in the scope of this work.

Data dropouts did not raise concerns for analysis as, in general, they affected a small fraction of the total households on the trial.

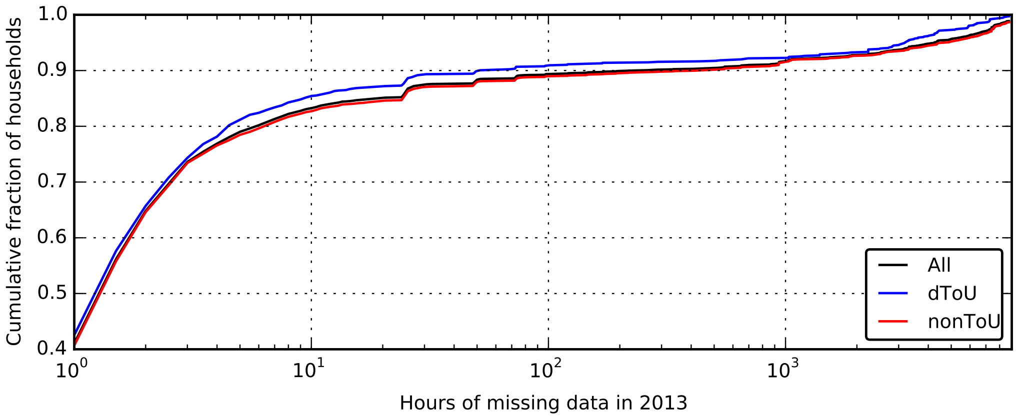

Figure 5.1: Cumulative fraction of households with hours of missing data in the trial year of 2013.

The cumulative distribution of the total number of hours of missing data for each household is shown in Fig. 5.1. As can be seen from the individual group traces, there was not a significant difference between the groups with regards to missing data. Overall, the median missing data period was 1 hour, though the distribution has a long tail, with small numbers of households having large numbers of missing data hours. 80% of households had 5.5 hours or less of missing data during the trial year.

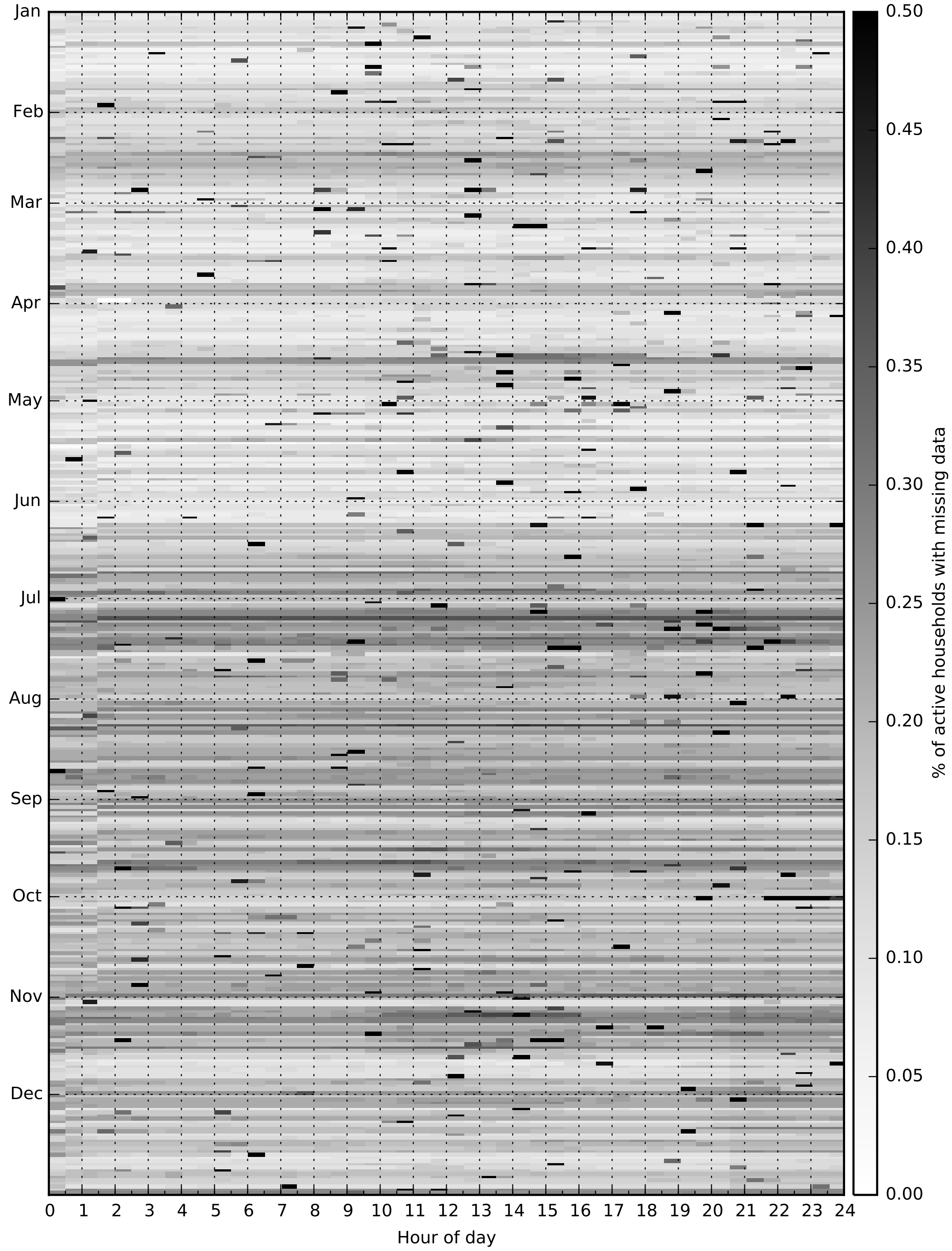

Figure 5.2: Heat map of the fraction of missing consumption data by measurement block for 5,103 households that completed the full trial year.

5.3 Data cleansing

This section describes the means by which data was selected for analysis. As described in Section 4.3, this trial effectively sampled data in two independent dimensions—households and time—and the experimental unit being sampled depends on the context of the research question being answered. Here, both dimensions are dealt with independently. Exclusions of corrupt or inaccurate data are detailed and the resulting set of valid data defined for use in subsequent analysis.

5.3.1 Households

Not all households within the trial were suitable for analysis. Households were excluded from the analysis set for the following reasons:

- Withdrawals: Householders who made a specific request to be withdrawn from the trial were not included in the analysis set. 292 households were removed for this reason.
- NonToU households on a non-standard tariff: To provide a baseline for the dToU group, nonToU group household had to be on a standard flat rate tariff. Some households were later found to be on non-standard tariffs such as Economy 7. 135 households were removed for this reason.
- Multiple meter feeds: Time-series consumption data was stored in an OSIsoft PI database. Each series was identified by a PI tag. Ideally this means that there should be a one to one mapping between tags and households, however there were exceptions. For example, if there was a fault with the original meter a new tag was sometimes created for the replacement meter. Duplicate tags were found for 44 households on the trial. Because of the ambiguity surrounding relinking their data sets and the likely large hole in measurements as a result of a meter change, it was decided better to remove them from the analysis set.
- Missing consumption data: In order to ensure analysis was not skewed by missing values, households with more than 50 hours of missing consumption data over the trial year were excluded. 650 households were removed for this reason.
- Miscellaneous: In addition to the main exclusion reasons listed above, some special cases existed. For example, the removal of specific people who had inside knowledge of the trial and households with generation capability (as this would offset demand). 36 households were in this category.

It is worth noting that the exclusion numbers given above are not mutually exclusive—some households may have been excluded for multiple reasons. This means the total of these numbers does not equal the total number of households omitted from the analysis set. After exclusions, the analysis set populations stood at:

- dToU: 988 households.
- nonToU: 3,768 households.

It was possible that the exclusions had skewed the relative representation of social demographics between trial groups. A comparison of the Acorn group breakdown at recruitment, shown in Fig. 4.4, and again after exclusions, shown in Fig. 5.3, found that no significant changes had occurred.

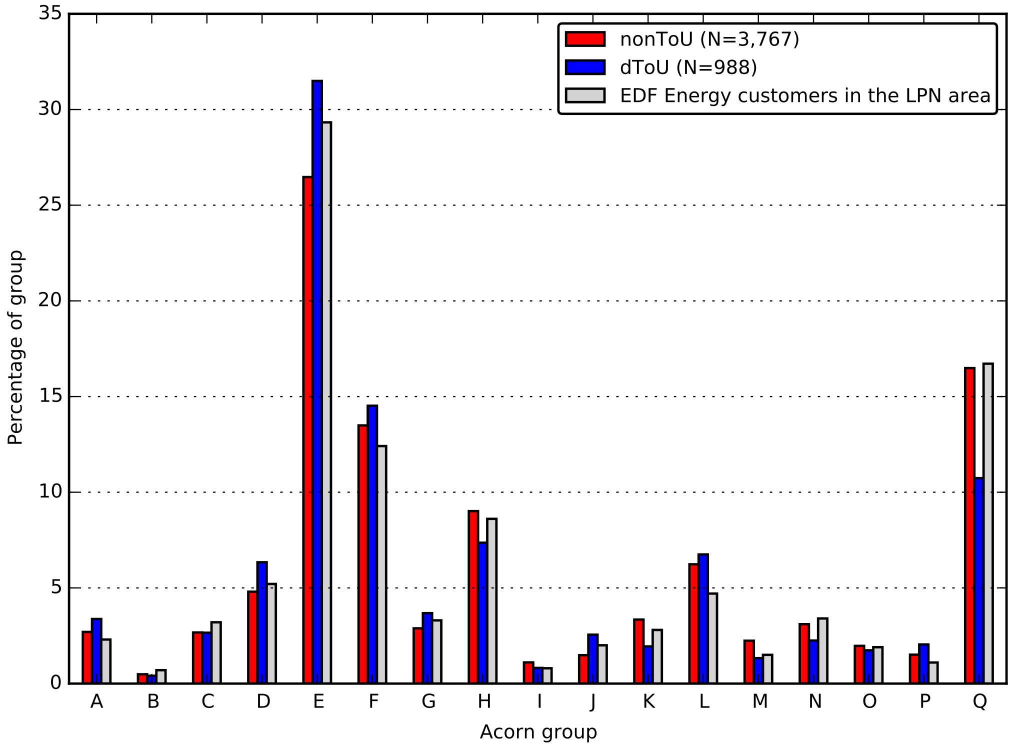

Figure 5.3: Proportions of Acorn groups for the dToU group the nonToU group and EDF Energy customers in the London Power Networks (LPN) area for households that were not excluded from analysis. See Table 4.2 for label definitions.

5.3.2 Time horizon

The measurement of DR requires a baseline—a reference from which to measure any changes. This was the purpose of the nonToU group. A simple approach to constructing a baseline is to use the average demand of the nonToU group directly. In this case, the DR signal would equal to the difference between the average demand of the dToU and nonToU groups. Ideally there should be no DR signal during the time prior to the trial as no price events occurred. In reality there will noise in this signal, particularly so using the simple baseline described here. Nevertheless, the relative noise level in the DR signal prior to the commencement of the trial may be used as an indicator of the validity of the underlying data.

Both the mean demand difference and standard error (SE) on this value are calculated as follows. First, mean demand at the settlement block at measurement index m is calculated as:

$$\overline{A}_m = \frac{1}{N_m} \sum_{h}^{N_m} A_{m,h} \tag{5.1}$$

where $A_{m,h}$ is the actual demand (average power measured during the 30 minute settlement block) of the dToU group for household h in measurement index m and N_m is the number of measurements return from the households in measurement block m. Group demand difference D_m is then calculated as:

$$D_m = \overline{A}_m - \overline{A'}_m \tag{5.2}$$

where prime ($'$) signifies the equivalent value for the nonToU group. By ignoring potential autocorrelations between measurement indices, we may approximate the variance at measurement index m as:

$$S_m^2 = \frac{1}{N_m}\sum_h^{N_m}(A_{m,h} - \overline{A}_m)^2 \tag{5.3}$$

and therefore calculate the SE of D_m at measurement index m as:

$$S_m^{D_m} = \sqrt{S_m^2 + S_m'^2} \tag{5.4}$$

where prime again signifies the equivalent value for the nonToU group.

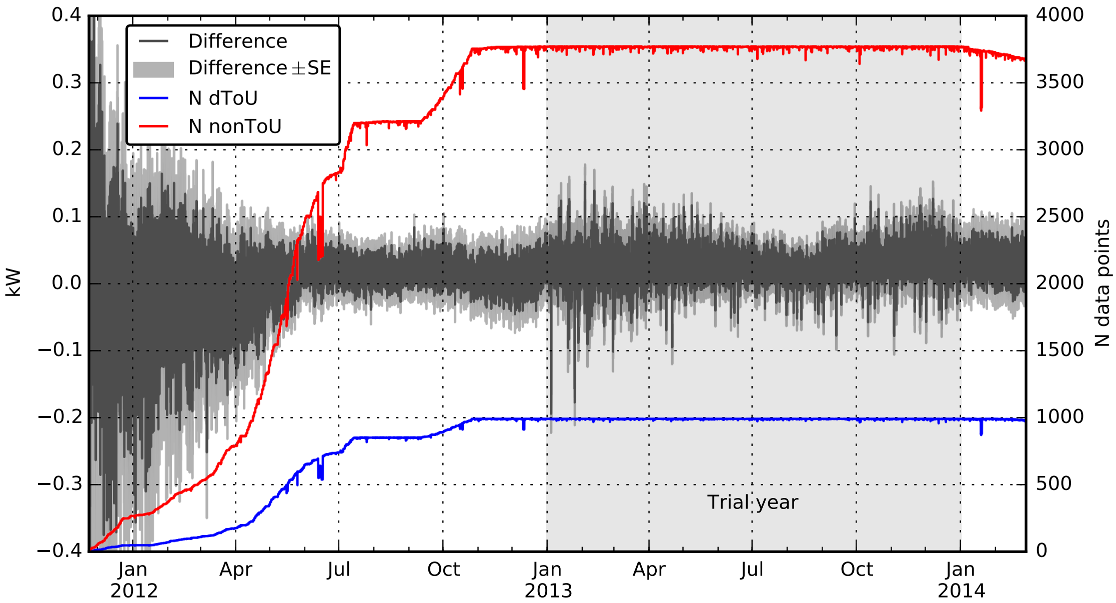

Figure 5.4: On the left axis, the difference between nonToU and dToU group mean demand with associated SE (SE), and the right axis, the number (N) of data points available for each trial group, both plotted against time for each 30 minute period recorded in the trial.

Figure 5.4 shows the difference (D_m) between the mean nonToU and dToU group demands plotted against time for each 30 minute period recorded in the trial. The lighter grey shaded area indicates $D_m \pm S_m^{D_m}$, the extent of the SE of the DR signal. The red and blue lines indicate the number of measurements take at each measurement index for the nonToU and dToU groups respectively.

Large fluctuations are visible during the first six months of data and are consistent with statistical noise from low sample numbers. This underlines the necessity to use large samples in trials. By July 2012, the difference in average group demand had stabilised to a relatively constant and significantly reduced range. This was considered to be representative of the natural difference in average group demand. As such, data before the beginning of July 2012 was discarded from the analysis data set.

Data from the two months after the trial year was also considered to be representative of the natural difference in average group demand. The valid time span of time-series data was therefore from the beginning of 2012-07-01 to the end of 2014-02-28.

As an aside, it is promising to note that DR events can be seen as spikes in D_m, both above and below the normal range, within the trial year.

5.4 Measurement of demand response

DR is defined as change in demand as a result of a price signal, relative to the hypothetical demand that would have been observed had the price signal not been sent. We call this the baseline demand.

The aim was to create a bespoke baseline model for each household in the dToU group. At the individual household level these results were noisy, which was to be expected as it is not possible to predict an individual person's behaviour with great accuracy. However, as the majority of the subsequent analysis involved taking the mean of grouped measurements, much of this random noise cancelled.

This approach had the benefit of allowing for intuitive error checking. By allowing DR to be calculated as a time-series, in the same format as the actual demand measurements taken by the smart meters, an appreciation of the accuracy of the model can be gained by looking at days without price events, where one would expect DR to be close to zero.

This approach also made it possible to calculate DR metrics at the individual household level, a crucial feature of the analysis because it allowed for linking between the *power system* and *social* sides of the trial. The former being more concerned with grouped response metrics, the latter, in part, looking in depth at the relationship between trial observations and the individual consumers [58].

DR from a single household h at time index m was estimated as:

$$R_{m,h} = A_{m,h} - B_{m,h} \tag{5.5}$$

where $A_{m,h}$ is the actual demand and $B_{m,h}$ is the baseline demand for household h at time index m. Furthermore, mean DR at measurement index m is calculated as:

$$\overline{R}_m = \frac{1}{N_{\mathcal{H}}} \sum_{h \in \mathcal{H}} (A_{m,h} - B_{m,h}) \tag{5.6}$$

and mean DR for household h is calculated as:

$$\overline{R}_h = \frac{1}{N_{\mathcal{M}}} \sum_{m \in \mathcal{M}} (A_{m,h} - B_{m,h}) \tag{5.7}$$

where $\mathcal{M}$ and $\mathcal{H}$ are sets of measurement indexes and households respectively, and $N_{\mathcal{M}}$ and $N_{\mathcal{H}}$ represent the cardinalities these sets.

5.4.1 Relationship between groups

It was expected that there might be some natural differences in group profiles owing to self selection during trial recruitment. To examine this, average group profiles were compared in the absence of price events. Figure 5.5 shows a comparison of the average load profiles for all consumption data excluding event days, stratified by the Elexon defined seasons and day types. It can be seen that, while the group profiles are similar, the dToU group tended to exhibit lower demand than the nonToU group. Also, the difference in profile between groups appears not to be linear, in that it was neither a constant difference nor a proportional difference. For example, the difference between group demand is greater on winter Saturday evenings than on winter weekday evenings, despite both days having similar magnitudes of demand.

It was concluded that a natural difference exists between group consumption profiles and therefore the nonToU group should not be used directly as the baseline for the dToU group. However, as both the dToU and nonToU groups were subject to the same external variables, such as weather and public holidays, the nonToU group presented a way in which to "black-box" many highly complex external variables into one measured variable. The task of baselining therefore becomes

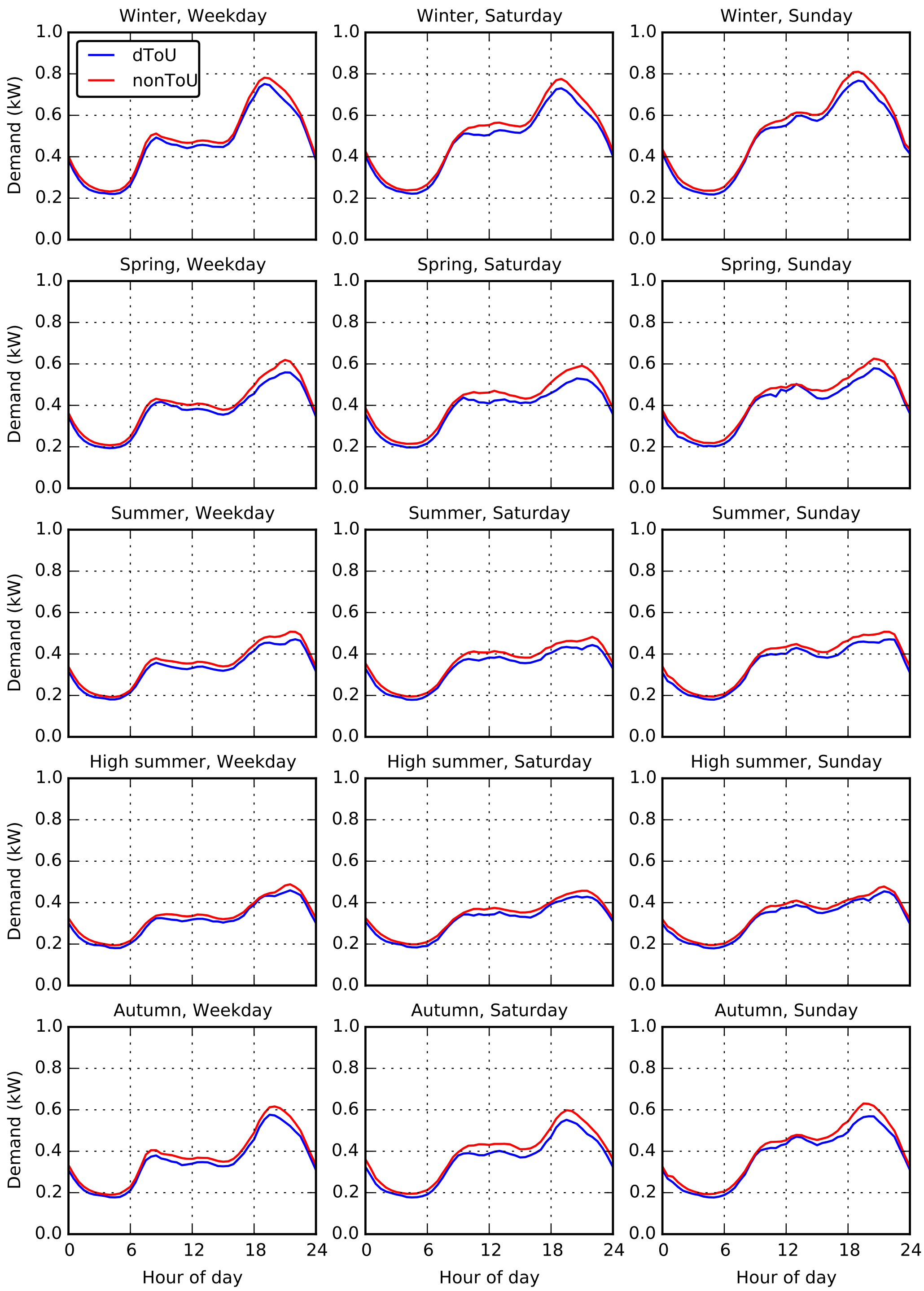

Figure 5.5: Comparison of mean nonToU and dToU group load profiles for all recorded demand data except event days.

that of defining the relationship between the nonToU group and the dToU group. The remainder of this section describes the approach for constructing a baseline demand model and justification thereof.

5.4.2 Baseline demand model

The model designed to calculate the baseline demand for each household in the dToU group is defined as:

$$B_{m,h} = \sum_{w=1}^{W} (\alpha_{w,h} d_w + \beta_{w,h} \overline{A'}_m d_w) + \gamma_h m \tag{5.8}$$

where $B_{m,h}$ is the baseline demand of household h at half-hour measurement index m and household h; $\overline{A'}_m$ is the mean actual demand of the nonToU group at measurement index m; d_w are binary dummy variables, one for each hour w in the $W(= 168)$ hours of the week; and the Greek letters $\{\alpha_{1,h}, \cdots, \alpha_{W,h}\}$, $\{\beta_{1,h}, \cdots, \beta_{W,h}\}$ and γ_h are parameters to be determined by the regression solver. This model effectively relates each half-hour measurement index linearly to the nonToU group mean demand, for each hour of the week, with an overall trend line.

The model was fitted using the Scikit-learn, Python [97] library, which implements a least-squares algorithm. A baseline was calculated for each household in the dToU group.

This model was used as the basis for measuring DR in the subsequent chapters of this thesis. As an initial visual validation, Fig. 5.6 shows the mean DR (actual mean minus baseline mean, across households) as calculated by this model for the full trial year. Demand reductions and increases can be seen to correspond to respective high and low prices shown in the tariff schedule in Fig. 4.10.

5.4.3 Training data

In order to characterise the relationship between the typical consumption profiles of the nonToU and dToU groups, consumption data where this relationship was unperturbed by DR events was used.

Six months worth of viable data was collected before the trial commenced with a further two months after the trial. This was supplemented with days from the trial year that did not contain price events. It is possible that such days could have been influenced by price events on other days, though to assume this would have left too little training data. Training data was therefore defined as all data between July 2012 and February 2014 (inclusive) that did not contain a price event within the trial day.

It was decided that nonToU group demand would be aggregated into a single number by taking the mean across households for each time index. This aggregate value was viewed as a “black box” variable that represented the average effect on demand of all external variables that were experienced by all households in the Low Carbon London (LCL) trial. Which is to say, it was believed that the average effects of global variables such as weather, special days (e.g. public holidays) and large scale events, should be captured in this way.

The effect of missing values was considered. Due to the number of households in this group being large (3,768) and the exclusion criteria ensuring that all households had a maximum of 100 half-hour measurement missing during the trial year, it was highly unlikely that a significant number of households would be missing data concurrently. Indeed, that this was the case can be seen in Fig. 5.4 from the traces of the number of consumption measurements in each group against time. For the whole time period over which training and prediction took place, the number of households from which the nonToU mean was calculated did not fall below 500, and for the most part was over 900. The process of taking the mean effectively filled in the missing data.

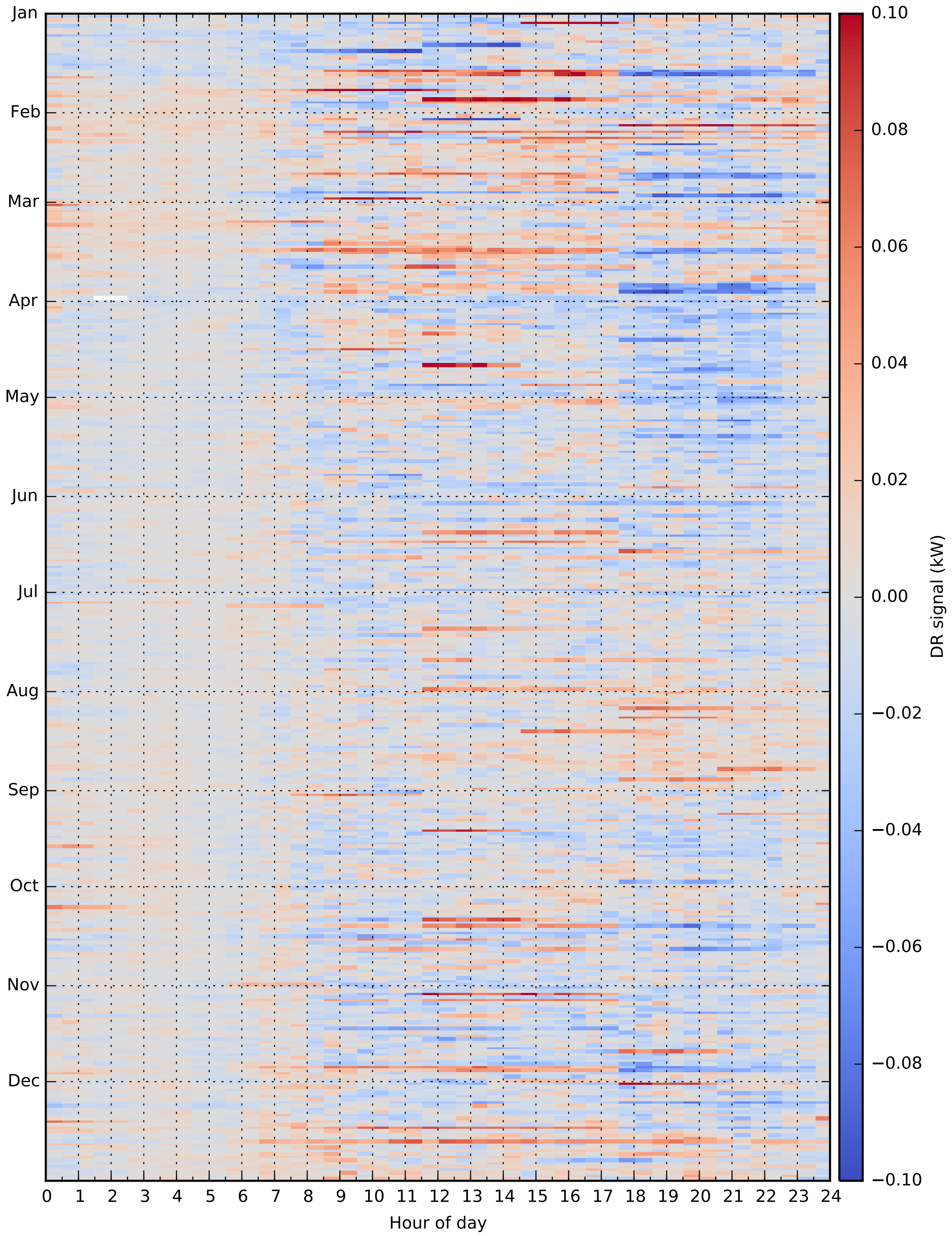

Figure 5.6: Heat map of the overall DR signal as calculated by the chosen baseline model, for the full trial year. Demand reductions and increases can be seen to correspond to respective high and low prices shown in the tariff schedule in Fig. 4.10.

5.5 Model design and validation

This section provides a description of the design process that lead to the baseline model described in Section 5.4.2. Model structure is discussed, followed by the cross-validation of three nested models. Finally, the accuracy of the model in the context of predicting the baseline for an individual household is examined, and an estimate individual and group baseline resolution provided.

5.5.1 Structure

To illustrate the high level relationship between nonToU and dToU demand, they are plotted against each other in Fig. 5.7. It is immediately clear that the general trend is well described by a linear model. This is confirmed by the linear and second order polynomial best fit lines being superimposed over each other. Despite the high coefficient of determination (R2) value for the linear model, it was believed worthwhile to attempt some improvement via a higher order model based on other (temporal) aspects of the load as estimation of DR depended solely on the quality of this model.

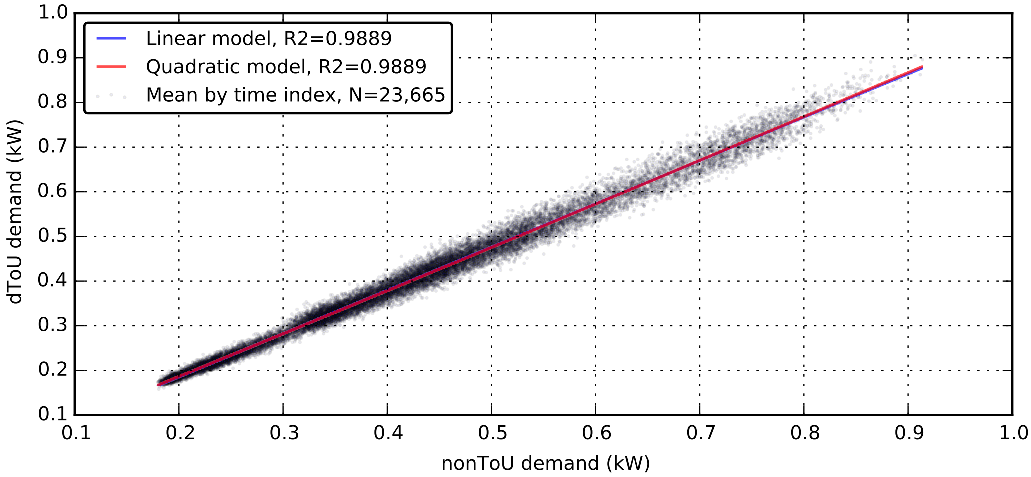

Figure 5.7: Mean demand of nonToU vs dToU for each measurement block in the training data set with linear and second order polynomial fit lines overlaid.

The linear model displayed in Fig. 5.7 explained overall group variance well (with an R2 value of 0.9889), but it could not capture non-linear variation in the relationship such as differences caused by day of week. To examine non-linear temporal structure, the residual of the linear model shown in Fig. 5.7 were aggregated over different time periods and their averages were plotted.

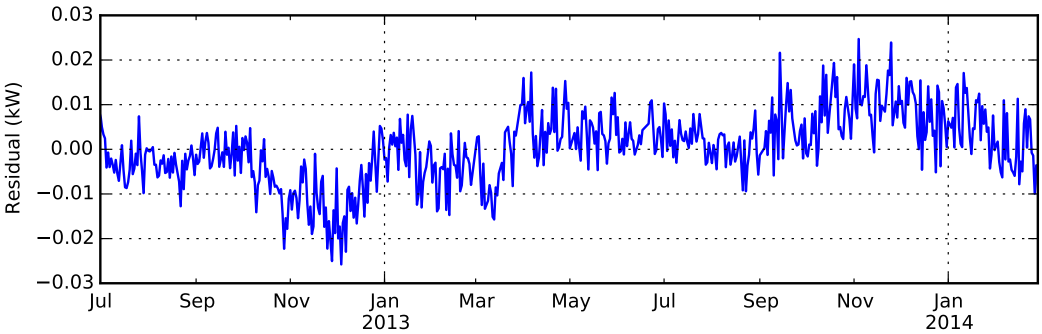

Figure 5.8: Mean residuals of the linear model shown in Fig. 5.7, aggregated by date.

Figure 5.8 shows the mean residuals aggregated by date (i.e. by individual days). This aggregation was necessary to filter out the shorter timescale fluctuation in residual that may have prevent a trend at the longer timescales from being noticed. While fluctuation is observed, it does not appear to be strongly related to the time of year. It was concluded that temporal features in the residuals that correspond to the time of year would not be targeted for capture in the baseline model structure.

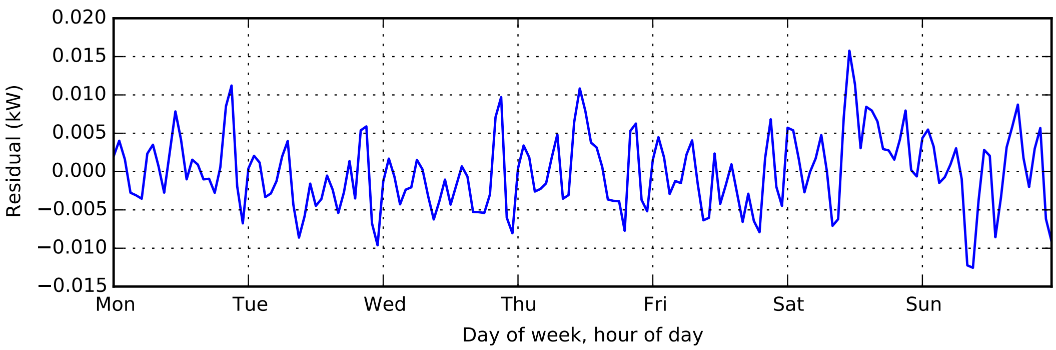

Figure 5.9: Mean residuals of linear model show in Fig. 5.7, aggregated by day of week and hour of day.

Figure 5.9 shows the mean residuals aggregated by hour-of-week. Some cyclic variation can be seen. There is evidence of a relationship with time-of-day and perhaps also with day-of-week, at least where the distinction between weekdays and weekend days are concerned. It was concluded that it would be worthwhile conducting further investigation into whether day-of-week and time-of-day structure in the model might significantly improve its performance.

A linear approximations of such cyclic variations can be made if there are enough of each kind of measurement in the training data set. As the training data set contained 23,665 measurements spanning 2012 to 2014, this was a viable option.

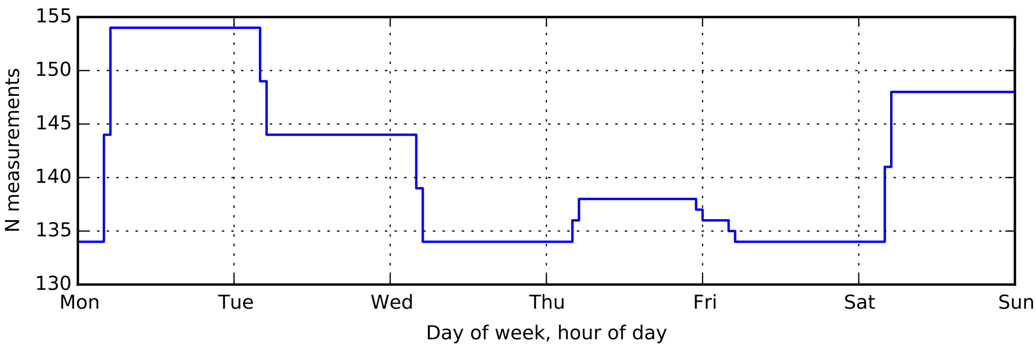

Figure 5.10: Number of measurements in the training set at each hour of the week.

Figure 5.10 shows the number of measurements in the training set for each hour of the week (assuming all data for the households was valid). The median number of measurements available for training in each group was 138 and did not fall below 134. This is to say, the hour-of-week time group contained measurements from at least 134 different weeks in the training index. These numbers suggest that it should be possible to construct an independent linear model for each hour of the week by introducing a commensurate number of binary dummy variables.

5.5.2 Validation

Multiple model constructions were attempted though this section only reports on the validation of the set of nested models that lead to the chosen baseline model. The model may be considered valid if it can be shown the number of parameters it contains does not result in overfitting—the situation whereby the model describes random error or noise instead of the underlying relationship.

In this section, the model is validated against the dToU group mean demand. Ultimately, the objective is to have a baseline for each individual household, however, the below described validation process was computationally intensive and it was believed that fitting a baseline to each individual households would be unnecessarily time consuming. As the dToU mean demand time-series should be representative of the features of the individual household, validating a mean demand model against overfitting was considered to implicitly validate the equivalent per-household model.

Three nested linear models were validated so that the effect of adding extra parameters may be seen relative to the previous model. The first model (M1) was used as a reference point and consisted of two parameters, a constant and proportional term (as used in Fig. 5.7). The second model (M2) included 336 binary variables, so as to effectively create an independent M1 model for each hour of the week. The third model (M3), added a load growth (or reduction) term to capture the average change in consumption with time. M1, M2, and M3 are defined below as Eqs. (5.9) to (5.11) respectively:

$$\overline{B}_m = \alpha + \beta \overline{A'}_m \tag{5.9}$$

$$\overline{B}_m = \sum_{w=1}^{W} (\alpha_w d_w + \beta_w \overline{A'}_m d_w) \tag{5.10}$$

$$\overline{B}_m = \sum_{w=1}^{W} (\alpha_w d_w + \beta_w \overline{A'}_m d_w) + \gamma m \tag{5.11}$$

where $\overline{B}_m$ is the mean baseline demand of the dToU group at measurement index m; $\overline{A'}_m$ is the mean actual demand of the nonToU group at measurement index m; d_w are binary dummy variables, one for each of the $W(=168)$ hours w in the week; and the Greek letters α_w, β_w and γ are parameters to be determined by the regression solver.

The models were evaluated using a k-fold cross-validation procedure implemented by the Scikit-learn library [97]. This non-parametric test involved "folding" the training index into k equal size sets without sorting. Each fold was sequentially omitted from the training index and the model was fitted using a least-squares algorithm.

This validation approach was considered particularly appropriate as it closely resembled the intended usage of the baseline model—to backcast over event days. The training data set contained 579 unique days and so a choice of $k = 500$ folds meant that each fold was a little over a day in duration. As the training index was maintained in chronological order and folds were sequential along the index, this approach ensured that folds consisted of contiguous sets of measurements that represented continuous, approximately day length periods. As price events typically occurred in clusters of one to three consecutive event days, this validation approach was believed to closely reflect the true performance of the models in question.

For each of these k iterations, the goodness of fit is measured by its R2 score, calculated using only the omitted data (i.e. the test set), and recorded. Histograms of the resulting scores are shown for each model in Fig. 5.11. The models can be seen to perform almost equally well, with only small incremental improvements in the mean score. This may seem a small difference at first, though it should be put into the context of the reduction in residuals as a consequence of the additional model parameters. The sum of the squared residuals was reduced by 7.6% between M1 and M2, and by a further 5.1% between M2 and M3. Performing a Z-test using the parameters displayed in the figure shows that these differences are significant to p-values of less than 1% and 5% respectively.

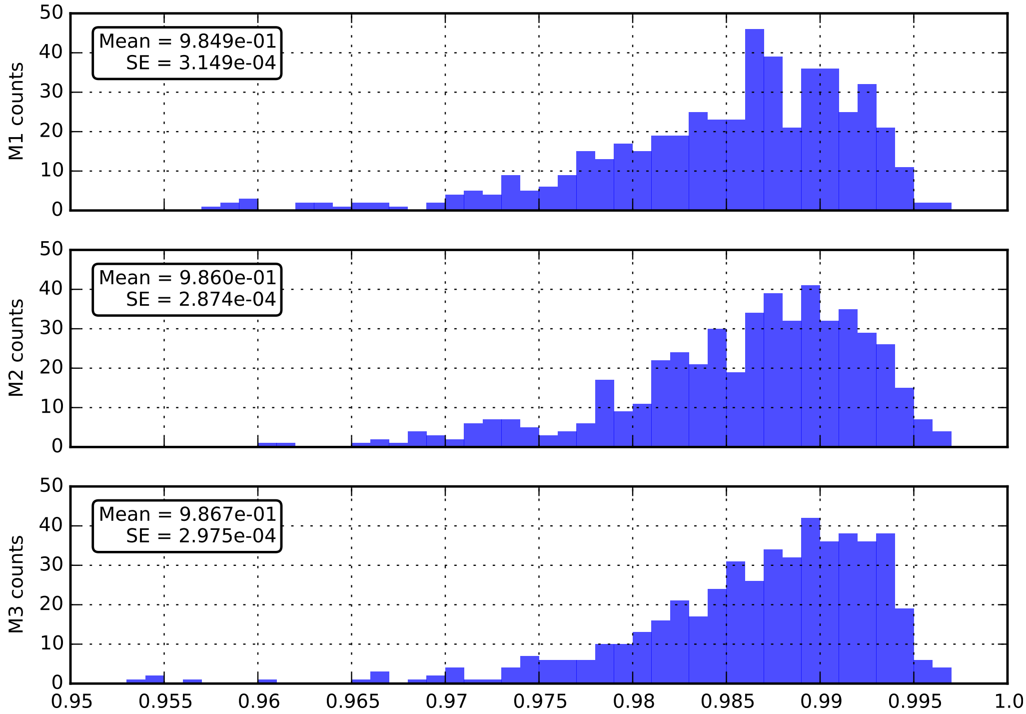

Figure 5.11: Histograms of k-fold ($k = 500$) cross-validated R2 scores for models M1, M2 and M3 with mean and standard error (SE) given for each distribution.

Furthermore, the observation that the dispersion of the R2 distributions did not increase (SE can be used as a proportional indicator for this) as more parameters were added was considered additional evidence that overfitting had not occurred. If a model had fitted to noise features rather than the true features of the demand profile, the k-folds cross-validation would have showed increased dispersion due to the score of each fit becoming more dependent on chance.

The following conclusions were drawn: M2 offered a significant improvement on M1, and M3 a slight improvement on M2. Combined with the observation that the k-fold score dispersion did not increase, this was considered strong evidence that overfitting had not occurred.

5.5.3 Per-household performance

The use of mean-demand time-series for model validation (described in the previous section) had the effect of reducing the difference between the demand profiles and thus understating the benefit of the hour-of-week and load growth parameters. This is due to eccentric individual household consumption profiles averaged towards a central mean. This can be seen clearly in Fig. 5.12 where individual household routines are visible yet highly eccentric relative to the overall group means. Given this, M2 and M3 may be expected to perform better when used to calculate per-household baselines.

Due to the high computational load of k-fold validation when calculating for 988 dToU group baselines, a different approach to performance evaluation was taken. With concerns of overfitting assuaged, model performance was gauged using the sum of the squared residuals (SSR) calculated over the full training index.

Models M1, M2 and M3 were used to calculate per-household baselines and the SSR over all households was calculated for each. As expected, M2 performed markedly better than when used to calculate an average baseline, showing an 18.7% reduction in SSR relative to M1. The additional parameter in M3 took the overall reduction to 20.0%.

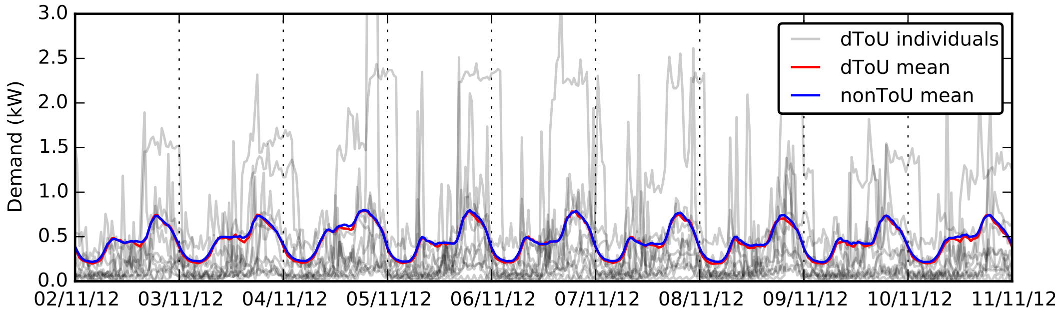

Figure 5.12: Demand traces for 10 randomly selected households in the dToU group shown against the mean demand traces of the dToU and nonToU groups. Arbitrary days were selected during the pre-trial year to avoid the interference of price events.

Model M3 was chosen (Section 5.4.2) for use as the baseline model for use in the calculation of DR.

5.5.4 Accuracy of the baseline model

For each household, the mean residual over the training index was effectively zero ($< 10^{-10}$ kW), as expected. In addition, no clear structure could be found to indicate that further refinement of the model might be possible. Figure 5.13 shows the baseline residuals over the whole training index and, below, a zoomed segment of this period in order to demonstrate the lack of structure at the day scale.

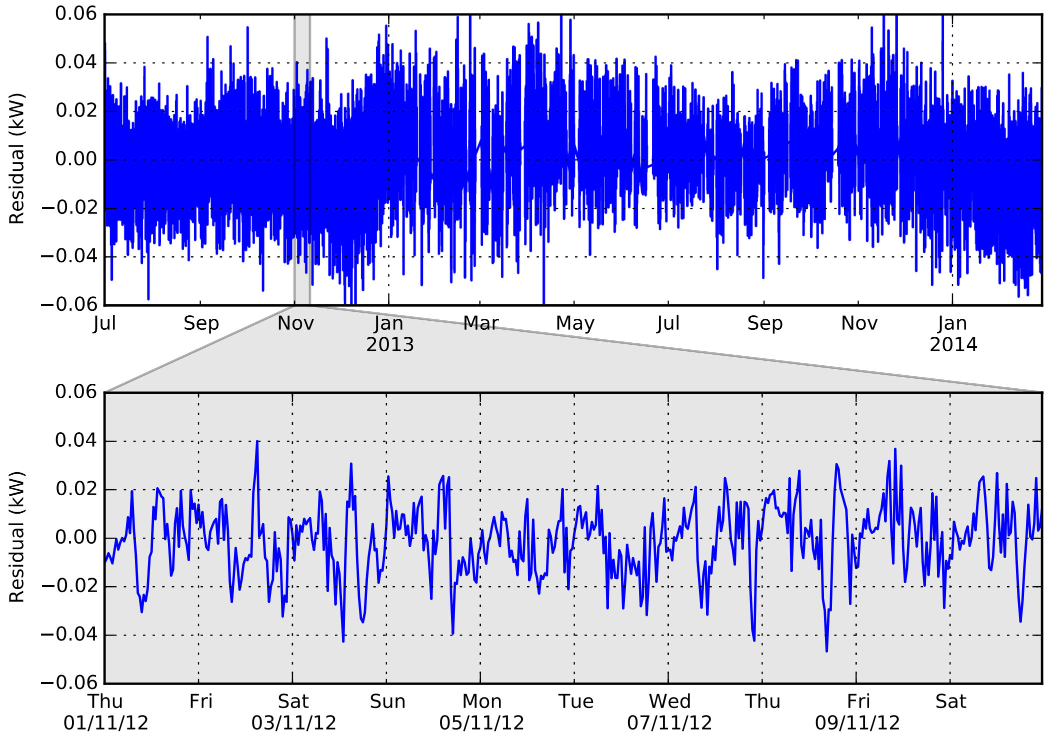

Figure 5.13: Baseline model residuals for the whole training index and for a zoomed segment of this period.

Across households, the baseline model will vary in its ability to predict demand depending on the routine of that household's occupants. For example, if the occupants for a household are away, the baseline model would over predict demand during this period.

We make an empirical estimate of the cumulative distribution of baseline residuals for both an individual household and for the group mean across all households. In both cases, residuals are calculated by taking the difference between actual demand and baseline demand across all measurement indices in the training set. This is analogous to calculating DR as described in Eq. (5.5), though in this case, with no price interventions, we assume we are measuring noise. The distribution of the residuals is assumed to be representative of the probability of observing an error on the baseline demand calculation at a random measurement index.

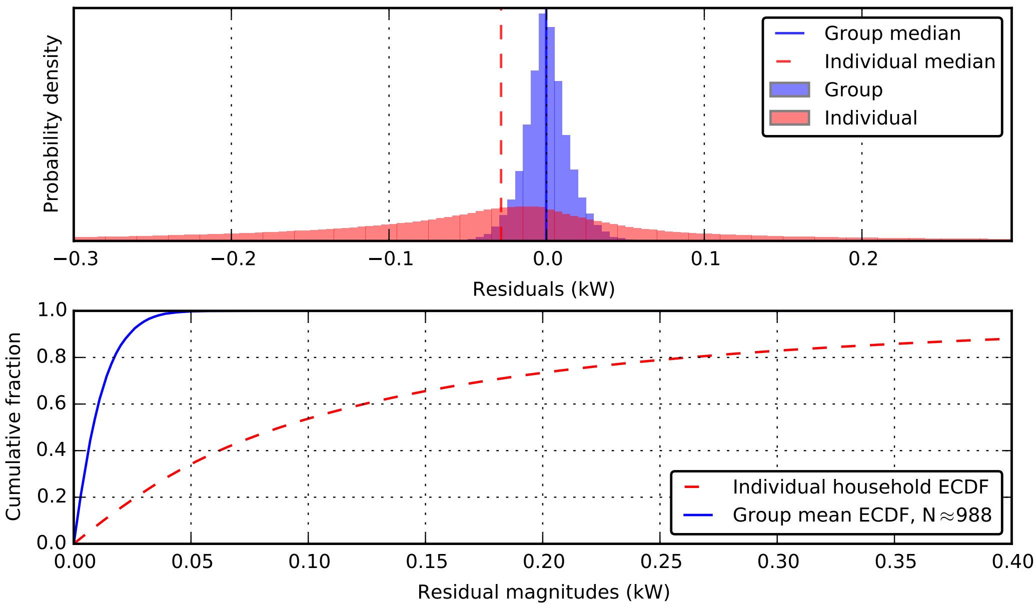

Figure 5.14: Histogram of residuals (upper) and their absolute empirical cumulative density functions (ECDFs) for all dToU households in the training index and the group mean across the 988 households.

Figure 5.14 shows the empirical distribution of residuals (upper) and the empirical cumulative distribution of the absolute residuals (lower), for both an individual households and for the group mean across all households (calculated as Eq. (5.6)).

The distribution of the individual households is, as expected, much wider than that of the group mean. The errors for the individual households have a mean close to zero (by design), but their distribution is asymmetric. This is because actual consumption has a lower limit of zero (at least in this trial as homes with generation were excluded), but individual measurement blocks can display very high consumption. The distribution therefore has a long tail to the right and a median less than zero.

The role of sample numbers in increasing result resolution can be seen in the difference between the two ECDF lines. At the 90% confidence level, an individual household will have up to 0.456 kW of error in the baseline calculation, while the aggregate group response has just 0.024 kW.

It should be noted that this analysis is based on the assumption of choosing a random household and a random measurement block. In reality, it is unlikely that errors will be distributed so randomly and therefore error values mentioned here should not be used as a measure of the accuracy of the results reported herein. Rather, they are a measure of the overall accuracy of the baseline model at predicting demand on no-event days. Error values, when used, are calculated in the context the research questions being asked. Further discussion of errors is given in the results chapter in Section 7.1.3.

Chapter 6

Consumer engagement

This chapter presents high level results of the Low Carbon London (LCL) dynamic Time-of-Use (dToU) tariff trial with a focus on indicators of consumer engagement. In the first section the effect of the dToU tariff is examined from the perspective of the network operator or supplier, where the basic objective is to reduce consumption during high price periods and increase consumption during low price periods. The extent to which this objective is achieved is examined through comparison of the proportions of energy consumed at each price level. In the next section, the perspective shifts to the consumers, where the main motivation for switching to a dToU tariff is the opportunity to make savings on their annual energy bills. Examination of the changes in annual bills therefore provides a first indicator of the level of consumer engagement with the tariff. In the third section this principle is developed into a per-household consumer engagement ranking index through use of a data driven, non-parametric technique. This technique did not require the use of the demand baseline described in Section 5.4 and is therefore not subject to any assumptions regarding the baseline procedure. This engagement ranking index was used as the basis for the per-household stratification of results in subsequent chapters.

6.1 Energy consumption shift

From the perspective of the distribution network operator (DNO) and retail energy supplier, the function a dToU tariff is to incentivise a reduction in consumption during the high price periods and an increase in consumption during the low price periods, relative to the normal consumption levels that would be present at those times. In this context, the most basic indicator of consumer engagement with the tariff would be the observation of a change in the proportion of energy consumed at each price level. This section presents a simple approach, using the non-time-of-use (nonToU) group as a reference point, to quantifying such a shift.

6.1.1 Method

For each group, the average energy consumed at each price level is calculated for each individual household. To illustrate exactly how this was calculated, the equation for the mean demand at high price is presented for household h in the dToU group:

$$H_h = \frac{1}{N_h} \sum_{m \in \mathcal{M}} A_{m,h} \tag{6.1}$$

where $A_{m,h}$ is the actual measured demand of the dToU group at measurement index m and household index h, $\mathcal{M}$ is the set of measurement indices in the trial year of 2013, and N_h is the total number of non-null demand measurements for household h over the same time period.

Equivalent mean demand values were calculated for each of the three price levels in each of the dToU and nonToU groups. We name these variables as below:

	dToU	nonToU
High price	H_h	H'_h
Default price	D_h	D'_h
Low price	L_h	L'_h

The above variables were then converted into normalised units of the nonToU group mean consumption at each respective price level. These per-unit consumption levels were calculated as:

$$\hat{H}_h = \frac{H_h / D_h}{a} \tag{6.2}$$

$$\hat{L}_h = \frac{L_h / D_h}{b} \tag{6.3}$$

$$\hat{H}'_h = \frac{H'_h / D'_h}{a} \tag{6.4}$$

$$\hat{L}'_h = \frac{L'_h / D'_h}{b} \tag{6.5}$$

where

$$a = \frac{1}{N_{\mathcal{H}}} \sum_{h \in \mathcal{H}} (H'_h / D'_h) \tag{6.6}$$

$$b = \frac{1}{N_{\mathcal{H}}} \sum_{h \in \mathcal{H}} (L'_h / D'_h) \tag{6.7}$$

$\mathcal{H}$ is the set of household indices in the nonToU group and $N_{\mathcal{H}}$ is the number of households this set.

As the above variables have been normalised relative to the nonToU group, it is clear that $\hat{H}'_h$ and $\hat{L}'_h$ will have a mean of 1. If there was no shift in consumption in the dToU group, then $\hat{H}_h$ and $\hat{L}_h$ would also have a mean of 1. A reduction in high price consumption relative to the nonToU group equivalent consumption fraction would be indicated by $\hat{H}_h$ having a mean of < 1. Likewise, a relative increase in consumption at low price would be indicated by $\hat{L}_h$ having a mean of > 1.

6.1.2 Results and evaluation

Figure 6.1 shows a scatter plot of the above defined per-unit consumption variables at both high and low price for the households in the dToU (blue) and nonToU (red) groups. Here, the x-axis represents per-unit consumption at low price, and the y-axis represents per-unit consumption at high price. Surrounding the scatter chart are four histograms of the scatter points, pertaining to each price and group combination respectively.

By design, the nonToU group points are centred at 1 on both axes. Relative to these, the dToU group points can be seen to be drifting towards the lower right corner of the graph. The direction of drift indicates both a reduction in the fraction of energy consumed during the high price periods, and an increase in the fraction consumed during low price periods. The mean per-unit consumption of the dToU group at low price has a 95% confidence range of 1.11–1.14, which implies an 11–14% increase in consumption relative to the nonToU group over the whole trial year. For high price, this range is 0.91–0.93, which implies a 7–9% reduction in consumption over all high price periods in the trial year.

This analysis shows that the dToU tariff has likely had a significant impact on the average fraction of annual energy consumed at each price level. However, this approach is less suitable for quantifying the engagement of individual households.

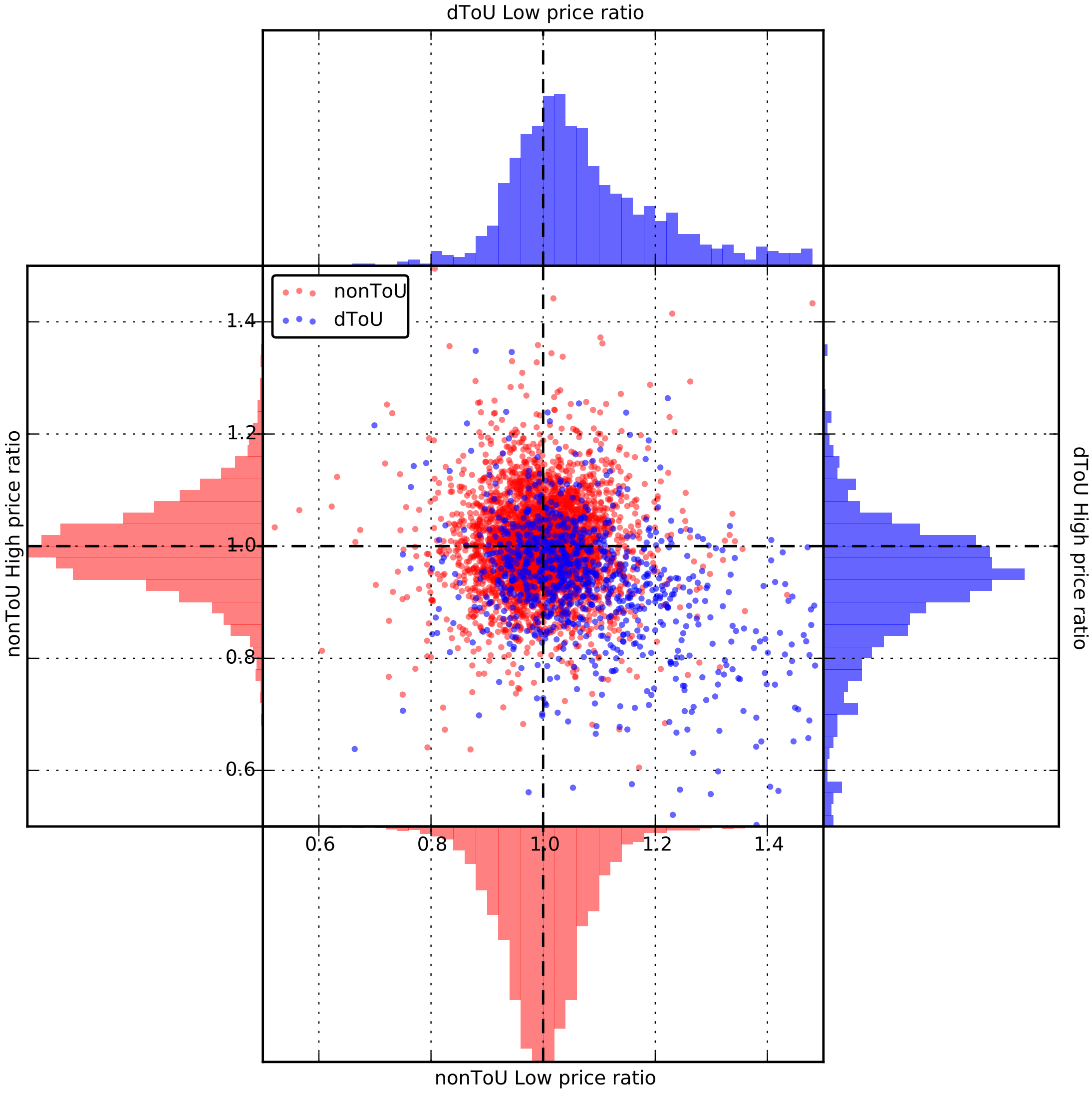

Figure 6.1: Ratios of high and low price consumption to default price for both dToU and nonToU groups, normalised against the nonToU group.

6.1.3 Controlling for group population difference

Constraint Management (CM) events were designed to reduce peak consumption levels by targeting the times of the day when peak demand is likely to occur with a high price signal. To incentivise maximum load shifting, these high price periods are flanked on either side by low price periods for the remainder of the trial day. As a consequence of this design, CM events are characterised by greater durations of time at the low price level than at the high level. For a valid direct comparison between dToU and nonToU groups, of the fractions of energy consumed at each price level, it is a necessary assumption that these fractions are the same when both groups are on the same tariff—which is to say, that both groups are samples from the same population. Due to the opt-in nature trial group recruitment (describe in Section 4.4), this may not be a valid assumption.

In contrast, Supply Following (SF) events have close balance between the total number of hours spent at the high and low price levels, as well as the times of the day over which they were scheduled. Furthermore, all event days were randomly placed throughout the trial year. For each price level, this balanced design combined with the randomisation of their placement means that any natural difference (when both groups are hypothetically assumed to be on the same tariff) between normal group load profiles should average towards zero when integrating consumption over a large number of events. Therefore, in order to minimise the introduction of systematic error caused by population differences between trial group load profiles, only SF events were considered in this analysis.

6.2 Change in consumers' bills

From the perspective of the residential energy consumer, a substantial motivation for switching to a dToU tariff is the opportunity to make savings on their annual energy bills. Examination of the changes in annual bills can therefore provides a first indicator of the level of consumer engagement with the tariff.

The actual annual bill for each household in the dToU group is calculated by integrating the product of the time-series price and demand vectors over the trial year of 2013. The annual energy bill for household h in the dToU group is therefore given by:

$$b_h = \sum_{m \in \mathcal{M}} p_m A_{m,h} \tag{6.8}$$

where $A_{m,h}$ is the actual measured energy consumption of the dToU group at time index m and household index h, $\mathcal{M}$ is the set of time indices in the trial year of 2013, and p_m is the tariff price at time index m.

However, measuring a change in energy bill implies the need for a reference point that represents what the bill would have been in the absence of the dToU tariff. The dToU tariff introduces two kinds of change that impact the energy bill. The first is the difference in energy pricing that defines the dToU tariff. The second is the change in consumption pattern that the new pricing structure incentivises. In order to calculate a reference bill, a reference for both the price and consumption time-series vectors are necessary. The reference price is taken as the standard flat rate tariff of 0.14228 £/kWh that was given to the nonToU group during the trial year. For the consumption profile reference, we use the per-household baseline model that was introduced in Section 5.4.

As this model was designed to represent the hypothetical demand that each household would have consumed had they not been subject to the dToU high or low price signal, it was considered particularly appropriate as a demand reference point for this analysis. While this model cannot predict the intrinsic variability of an individual household's demand, on average, over a large enough time horizon, it should provide a good approximation of the average consumption levels at each price level.

The reference bill b'_h is therefore calculated as:

$$b'_h = 0.14288 \sum_{m \in \mathcal{M}} B_{m,h} \tag{6.9}$$

where $B_{m,h}$ is the calculated baseline demand of the dToU group at time index m and household index h (defined in Section 5.4). The percentage change in annual bill can now be expressed for each household h in the dToU group as:

$$c_h = \left(\frac{b_h}{b'_h} - 1 \right) \times 100 \tag{6.10}$$

Figure 6.2 shows the change in annual electricity bill relative to what each household would have paid on the standard flat rate tariff of 0.14228 £/kWh and with the hypothetical consumption profile of that household on a standard flat rate tariff, as calculated by the baseline model. The average change in annual bills amounted to -4.9%, with 85% of customers achieving lower bills on the dToU tariff than on a hypothetical flat rate tariff.

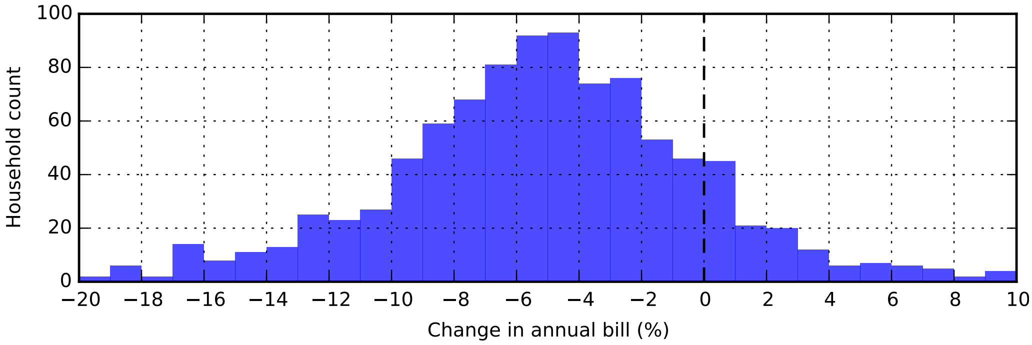

Figure 6.2: Empirical distribution of the change in annual bill for all households in the dToU group.

The dToU tariff had been calibrated previously such that it would be cost-neutral to an average non-responsive consumer assuming Elexon's Profile Class 1 [92] (this calculation was performed by programme partner EDF Energy), so the overall decrease in bills should be consistent with response to the dToU tariff. However, this comparison serves only as a group statistic. At the individual level, some households may have inadvertently had consumption profiles that lead to lower bills on the dToU tariff even without active engagement, or higher bills with active engagement.

It is worth noting that the above calculated decreases in annual bill were the result of a limited number of targeted experiments. Commercial implementation of a dToU tariff would likely have different prices and schedules, and may not be subject to the constraint of cost-neutrality for the average consumer. They may therefore provide households with additional possibilities to lower their bills.

6.3 Engagement ranking

This section presents a novel method of ranking individual households according to their engagement with a dToU tariff signal. By designing this engagement rank index to be largely independent of baseline demand, magnitude of demand and of regular daily patterns in demand, the statistical error of selection bias that would be introduced if results were stratified by the same signal (demand response (DR)), is avoided. In doing so it differs from direct DR metrics and thus provides a rank index against which stratification of DR measurements may be justifiably performed.

6.3.1 Problem statement

The previous two sections introduced the metrics of consumption shift and change in annual bills as a means of detecting the level of engagement with the dToU tariff. While these are useful metrics for quantifying aggregate results, changes in bills for individual households are affected by a number of noise sources, which also affects DR metrics.

To illustrate, Fig. 6.3 shows traces of the measured actual demand and calculated baseline demand for an arbitrary individual household, against the mean of these taken over all households in the dToU group. DR is calculated from actual and baseline measurements according to Eq. (5.5), which implies that it is represented by the area between the red and the blue lines, positive if red is higher than blue, negative if blue is higher than red. It can be seen that, even in the absence of the high price signal (indicated by the red shaded area), fluctuation in the difference between actual and baseline demand are visible, though to a much smaller degree in the group mean than in the individual case. We consider such fluctuation to be the source of intrinsic noise in the DR signal. This noise stems from the irregularity of human activities. Over many samples, this noise will tend towards a mean of zero. When averaged over all households in dToU group, the level

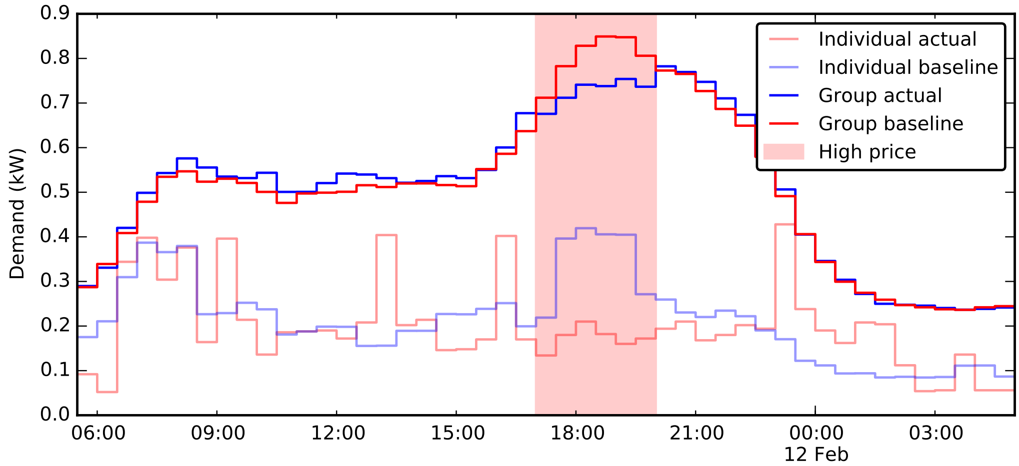

Figure 6.3: Traces of the measured actual demand and calculated baseline demand for an arbitrary individual household, shown against mean actual and baseline demand as calculated over all households in the dToU group.

of noise is reduced proportionally to $1/\sqrt{N}$, where N is the number of households in the group (assuming independence between household deviations from regular patterns).

The implications of noise in the DR signal are that, for an individual household, it is impossible to determine with certainty whether the measured response was the result of a deliberate choice by the consumer, or the result of chance alone. For example, if the consumer was on holiday when a high price signal was applied, a strong DR signal would be measured because the baseline demand model, which predicts the average load profile of the household, would calculate a non-zero value, while the actual measured demand would be at or close to zero. In addition, households differ in their daily schedules. Accidental interactions between this schedule and the dToU price signal also lead to bill changes and implied DR, even in the absence of engagement.

We wish to stratify response magnitudes according to how good households are at responding to the price signal. Performing this stratification with the DR metric alone would mean that a household which had a particularly high DR signal by chance alone would still be classified as a good responder, even when the occupants of that household made no conscious effort to respond. In this way, a *selection bias* towards households where occupants happened to be away during price signals, would be introduced into the stratification of consumer engagement.

In the remainder of this section, we develop and validate an engagement rank index with the aim of quantifying the responsiveness of individual households to dToU signals. As well as allowing the stratification of results in subsequent chapters, the per-household engagement rank complemented the more qualitative approach described in an associated LCL report [58].

6.3.2 Design of the engagement rank index

We aim to rank the participating households according to their engagement with the dToU trial. As in the previous section, we assume that the this is related to the degree in which their annual bills are modified by the price signal. To determine this ranking we compute, for each household, the likelihood that the actual annual electricity bill came about by chance (the null hypothesis). If a random response to prices provides a good match for the actual annual bill, it is unlikely that the household has engaged with the trial. Whereas, a bill that is significantly lower signals a conscious attempt to follow the price signal.

The likelihood of a chance realisation of the actual bill is estimated using a bootstrap approach. The dToU price signal for 2013 was divided into the 365 trial days and shuffled randomly, and the

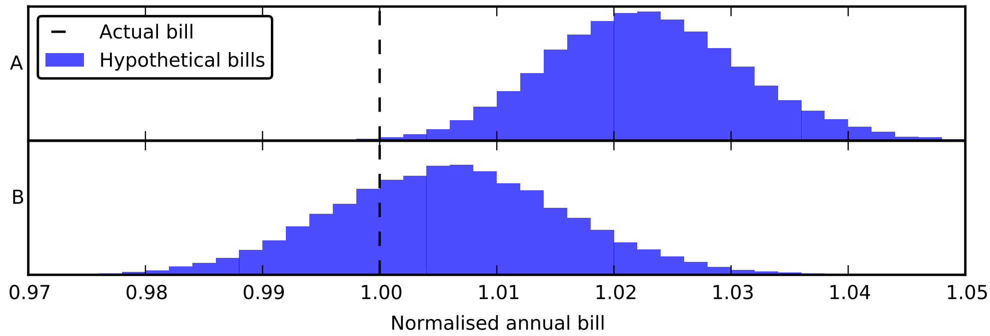

Figure 6.4: Illustration histograms of the hypothetical bills resulting from randomised pricing schedules. The likelihood that household A achieved its actual bill by chance alone is smaller than that of household B.

resulting annual bill was computed for each household. In shuffling by trial day, it was ensured that the price signal had the same salient features as the actual signal. This procedure was repeated 30,000 times to provide a large range of hypothetical annual bills. These annual bills collectively represent the hypothetical situation where the customer is unaware of the (random) price signal.

Each household's *actual* 2013 bill was then compared to the simulated distribution of outcomes and the corresponding percentile was determined. If the percentile is very low (e.g. $< 1\%$), it is highly unlikely that the bill came about by chance, and conversely, if the percentile is moderate (e.g. $\approx 30\%$) the household is likely not to have acted on price signals. This process is illustrated in Fig. 6.4. The upper panel shows the process for a responsive household, where the distribution of hypothetical annual bills for random price signals clearly deviates from the realised value. Conversely, the lower panel is generated for a household that is thought not to have responded significantly to the dToU signal, because the random hypothetical bill distribution adequately explains the realised bill. The computed percentiles were used to establish a responsiveness ranking across all households in the dToU group.

Even with 30,000 bootstrap iterations, 117 households outperformed *all* generated hypothetical bills, indicating a very significant response. In total there were 193 households with tied percentiles. A larger number of bootstrap iterations could in principle be used to increase the resolution and thus differentiate between tied households, but small differences in ranking are unlikely to be relevant considering the unavoidable sources of noise affecting the results (sample selection, limited number of events). When ranking, tied percentiles were assigned the same rank ordinal where that ordinal was equal to the average of their positions in the ascending order of the values. In the next chapter, engagement ranking is used to classify households into four groups according to their level of responsiveness, a measure which is not sensitive to small changes in ranking that may arise from the tied values.

6.3.3 Validation

Figure 6.5 demonstrates the relation between engagement rank and the observed DR. DR was calculated according to Eq. (5.7), where the set of time indices $\mathcal{H}$ over which it was measured were the sets of time indices during which each respective price level (high, default and low) applied. The panels depict the time-averaged response to high (left), default (centre) and low (right) price signals respectively, and each dot represents a single household with its engagement rank index on a range of 1 to 988 shown on the x-axis.

As expected, highly engaged households (low rank index) tend to decrease their consumption in response to high price signals and increase their consumption in response to low price signals,

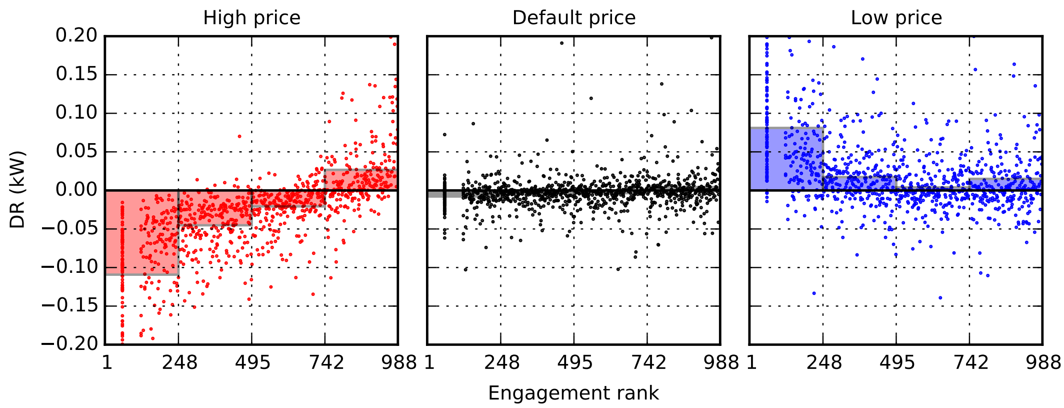

Figure 6.5: Household performance rank against measured DR, by price band.

and the magnitude of the response generally decreases with increasing rank index. Note that the trend is more pronounced with respect to high price signals than to low price signals, which may be explained by the higher price differential used to incentivise demand reduction. Also, as expected, the observed deviation at default price periods is very low and unrelated to the ranking. These general trends are quantified using Spearman-rank correlations between the engagement rank and the DR metrics in Table 6.1.

	High price	Default price	Low price
Responsiveness	0.78	0.18	-0.43

Table 6.1: Correlation between household's responsiveness rankings and their year-round averaged DR, computed using Spearman's rank test.

In Fig. 6.5, we have included bars that represent the average responses of the households in each quartile. The first bar reflects the behaviour of the households with the highest engagement, the second bar of the households with slightly lower engagement, etc. It may be expected that as people become more familiar with the DR concept, and particularly with the proliferation of DR automation, future households will become more responsive to dToU signals. In this line of reasoning the behaviour of the first quartile(s) may be considered indicative of the *potential* for DR.

There is an apparent antagonistic response to high price signals for the worst-responding quartile of households. This results from using the same data set both for determination of the ranking attribute and computation of the resulting response. Inevitably, some of the incidental noise will be picked up by the ranking algorithm, i.e. an accidental good response cannot reliably be distinguished from a conscious response, and an accidental bad response cannot be distinguished from antagonistic behaviour. Nevertheless its effect is much smaller than the observed sympathetic DR and it may be considered an upper bound for the error on the estimated DR for the 25% best responders.

6.3.4 Relationship to DR and annual consumption

Figure 6.5 also illustrates an interesting feature of the engagement ranking method: the highest ranked households are not necessarily the ones with the highest absolute change in demand in response to price signals. This is because the method does not quantify directly the magnitude of the absolute (kW) response to price signals, but its consistency and the degree to which it can be ascribed to chance. This means households with limited means of DR may still rank highly if fluctuations in the consumption are clearly linked to the dToU signal.

Figure 6.7 provides an alternative visualisation of relationship between DR and engagement rank with the addition of a third variable; annual consumption. Low price DR is depicted on the x-axis, with high price DR on the y-axis. Engagement rank is indicated by the colour of each data point; darkest blue for the most engaged ranging to lightest red for the least engaged. Annual consumption is depicted by the relative size of the data points; larger for greater annual consumption.

In the upper chart, showing nearly all households in the dToU group, a clear correlation can be seen between higher engagement and the magnitude of DR, for both high and low price. Likewise, there is an apparent correlation between annual consumption and DR, though this trend is weaker than that of engagement rank.

An interface between good and bad responders appears to exist close to 0 kW DR on both axes. Zooming into this interface, shown in the lower chart, it can be seen that higher annual consumption households that achieve moderate DR magnitudes are more often ranked as having low engagement (greater rank index ordinal) than households with an equivalent response magnitude but lower annual consumption. This observation is consistent with the intrinsic noise in the DR signal being greater for households with higher annual consumption.

6.3.5 Overall performance of engaged households

To gain a sense of the power of the engagement ranking index, Fig. 6.6 shows the mean DR values for the full trial year, separated into responses to high and low price points. The darkest bars represent the average response of the full dToU group, while the lighter bars represent the mean observed DR for the subpopulations of the 75%, 50% and 25% most engaged households, respectively.

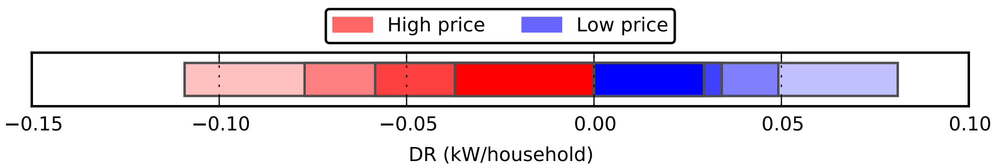

Figure 6.6: Mean DR by price level calculated over all events in the trial year. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

The most engaged households (lightest shade) produced an average response that is much larger than all households combined (darkest shade). Averaged across all trials and households, high price signals resulted in an average decrease in demand of 0.04 kW/household relative to the default price signal, and the low price signal resulted in an increase of 0.03 kW/household. The most responsive 25% of households outperformed these values by a factor of three, with a decrease of 0.11 kW/household and an increase of 0.08 kW/household, respectively.

Finally, and in order to illustrate the filtering ability of the engagement rank index, we compare the average DR signal for the whole dToU group against the average signal for the 25% most engaged households (247 households). Figure 6.8 shows this plotted as two adjacent heat-maps, the left showing the average DR signal for the full dToU group and the right, the signal for the most engaged quarter of this group. At this point it may be useful to contrast this figure against the price schedule given in Fig. 4.10, or Fig. 7.2, which is the same as this figure, except the left panel has been replaced with a price heat-map for comparison.

Looking at Fig. 6.8, it is clear that in restricting the analysis from the full dToU group to the most engaged quarter of households, the signal to noise ratio has been significantly enhanced.

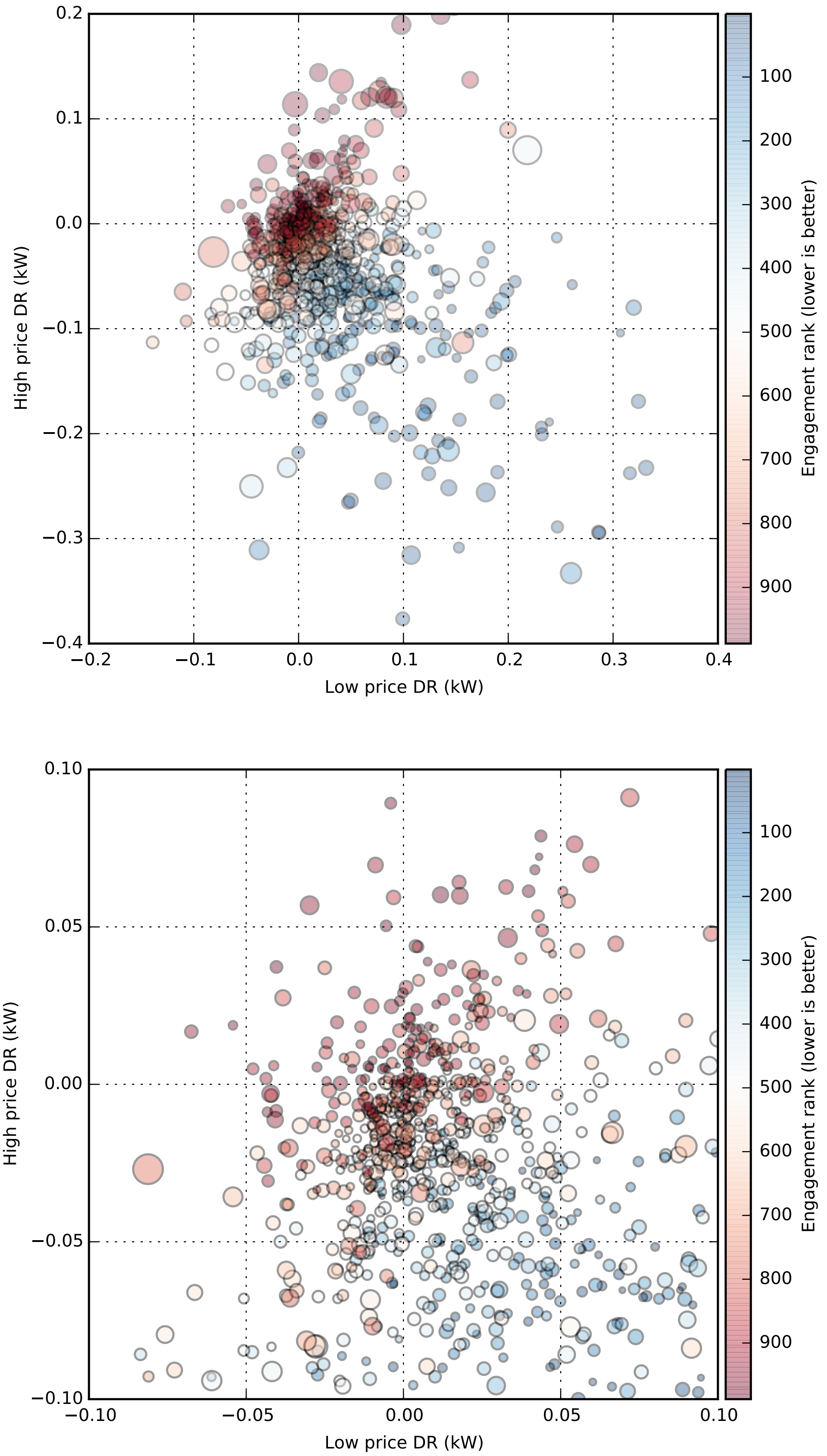

Figure 6.7: Two scatter plots of the high and low price DR for all households in the dToU group. Annual consumption during 2013 is indicated by the size of each point and engagement rank indicated by colour. The lower scatter plot provides a zoom in on the centre region of the upper scatter plot.

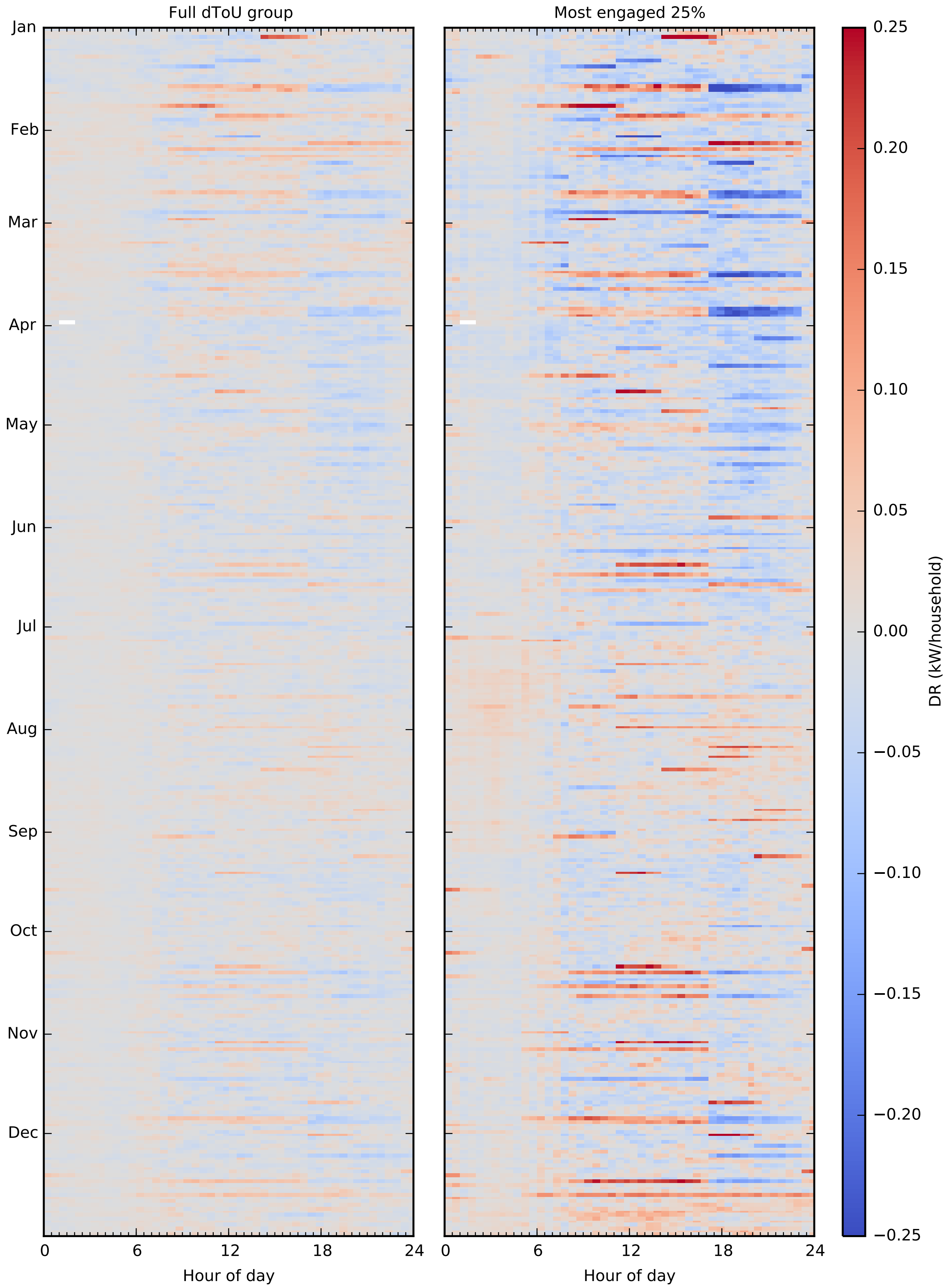

Figure 6.8: Heat-map image plots of the DR signal for the full trial year, (left) averaged over the whole dToU group and (right) averaged over the most engaged 25% of that group.

High and low price DR events are now clearly defined against the background of the default price. While, in this thesis, we use the ranking index as an analytical tool, it may also have commercial application as a means of targeting rewards for participating in DR schemes, or to select households that would be well suited for particular tariff types.

6.4 Summary and conclusions

This chapter presented three approaches to measuring the level of engagement of residential consumers with a dynamic Time-of-Use (dToU) tariff. The first two were aggregate measures and examined engagement from prospectives of the energy supplier or network operator, and the residential consumer respectively. The third method provided a means to quantify the level of engagement of the individual household. This was developed into an engagement ranking index, which may be used for the stratification of the demand response (DR) response signal.

Consumption shift. From the perspective of the distribution network operator (DNO) and retail energy supplier, the function of a dToU tariff is to incentivise a reduction in consumption during the high price periods and an increase in consumption during the low price periods, relative to the normal consumption levels that would be present at those times. In this context, the most basic indicator of consumer engagement with the tariff would be the observation of a change in the proportion of energy consumed at each price level in the dToU group. To this end, a simple approach was developed that used the non-time-of-use (nonToU) group consumption as a reference point from which to measure change in consumption at the high and low price levels. Natural differences between average group load profile were accounted for by focusing only on Supply Following (SF) events. These events had the desirable features of being approximately uniformly distributed across high and low prices, and the times of day in which they took place; then were also randomly distributed throughout the days of the trial. These features mean that natural differences between group load profiles should average towards zero over increasing time scales. Overall it was found that the dToU group had reduced average consumption during high price periods by 7–9%, and increased average consumption during low price periods by 11–14% (95% confidence ranges).

Change in annual bills. From the perspective of the residential energy consumer, the main motivation for switching to a dToU tariff is the opportunity to make savings on their annual energy bills. Examination of the changes in annual bills can therefore provides a first indicator of the level of consumer engagement with the tariff. Over the trial year of 2013, we calculated that 85% of households on the dToU tariff received lower annual bills than they would have had on the standard flat tariff of the nonToU group, with the mean reduction in bill being 4.9%. Although the overall decrease in annual bills is a first indicator of overall engagement, it does not necessarily extend to individual households. For example, consumers that are often away during the evening are likely to have missed the Constraint Management (CM)-type evening peak trials, resulting in a lower average bills without deliberate engagement with the tariff.

Engagement rank index. To classify the engagement of individual households with the trials a measure of responsiveness to dToU signals was developed. It determines the likelihood that the realised annual bill came about by chance, if the household had paid no attention to the dToU signal. If this likelihood is very low, it is assumed that the household has actively responded to the signal, whereas a high likelihood is consistent with a lack of engagement. The likelihood measures were used to rank all households according to their perceived responsiveness to dToU signals.

As expected, highly engaged households (low rank index) tend to decrease their consumption in response to high price signals and increase their consumption in response to low price signals, and the magnitude of the response generally decreases with increasing rank index. An important characteristic of the responsiveness ranking is that it does not strictly select for those households with the largest absolute DR, which tend to be the largest consumers of electricity. By measuring the statistical properties of a household's energy consumption the method also identifies consumers that deliver small but consistent DR contributions.

Averaged across all trials and households, the high price signal resulted in a decrease in demand of 0.04 kW/household relative to the default price signal, and the low price signal resulted in an increase of 0.03 kW/household. The most responsive 25% of households outperformed these values by a factor of three, with a decrease of 0.11 kW/household and an increase of 0.08 kW/household, respectively.

The responsiveness ranking also plays a key role in the extrapolation of results in subsequent chapters where highly ranked households are assumed to be indicative of future, "business as usual" consumers who are increasingly responsive to dToU signals, either manually or mediated by home automation devices and services. As well as use as an analytical tool, the engagement ranking index may also have operational application as a means of targeting rewards for participating in DR schemes, or to select households that would be well suited for particular tariff types.

Chapter 7

Response stratification

This chapter presents the overall demand response (DR) results of the Low Carbon London (LCL) trial as well as various stratifications of them. It beings with a brief primer on the interpretation of the results presented within, with the remainder laid out in four main sections. The first two sections focus on the analysis of Constraint Management (CM) and Supply Following (SF) event types respectively. Both begin with the presentation of overall response levels and graphs showing demand traces on chosen event days, followed by analysis of the specific features and objectives of each event type. In the case of CM events, the ability to reduce peak demand was the focus, while for SF events it was the ability to respond across different time stratifications. The third section examines the the effect of two high level social variables; socio-economic group and household occupancy level. In the final section, the results are compared against those of the closely related trials discussed in Chapter 3.

7.1 Interpretation of results

Demand changes in response to a dynamic Time-of-Use (dToU) signal can be measured and reported using absolute or relative measures. For the purpose of this chapter, DR is measured in terms of absolute deviations in power consumption (kW/household). This approach directly quantifies the effect of DR on the network and is therefore well-suited to the SF use case, where both distribution network operators (DNOs) and suppliers are concerned with the absolute volumes of power transported and consumed. It is also a natural candidate for CM applications, where the DNO must meet or exceed a certain load reduction. When metrics other than kW/household are used to represent DR, it is highlighted in the relevant section.

7.1.1 After diversity demand

Diversified demand values, calculated using Eq. (2.1), may be used to place DR into the context of its relative impact on load profiles. For the non-time-of-use (nonToU) group, the maximum diversified peak demands during 2013 were 0.99 kW/household in winter and 0.69 kW/household in summer. In this context, DR level of 0.05 kW would be equivalent to a 5% demand reduction in winter or a 7% demand reduction in summer.

Further discussion of the effect of diversity in the context of the LCL residential dToU trial can be found in LCL report “Quantifying Demand Diversity of Households” [98].

7.1.2 Engagement stratification

As explained in Chapter 6, each household participating in the trial was assigned a ranking according to its apparent engagement with the dToU price signal. Where possible throughout this chapter, results are presented by four stratifications of engagement rank:

- average across all households
- average across the 75% most engaged only
- average across the 50% most engaged only
- average across the 25% most engaged only

These stratifications are depicted by progressively lighter shading in both bar charts and demand profile traces.

The results for more engaged subpopulations could be interpreted as indicative of the DR potential in future scenarios where dToU signals are increasingly familiar to users and appliance automation is able to take advantage of dynamic tariffs.

When the number of households in a particular analysis is insufficient to make statistically valid statements regarding the behaviour of its subpopulations, only the population mean response is shown. In such cases, error bars representing the standard error (SE) of the mean are also included.

7.1.3 Significance

The analysis of DR data is generally concerned with detecting significant changes in behaviour in response to dToU signals, and furthermore significant changes in the response between household categories. The term significance is used here in the statistical sense, which implies that the observed difference is unlikely to have come about by chance. This likelihood is quantified using a confidence level, where a common 90% confidence level implies that the likelihood of observing a reported value by chance is 10% or less.

The per-household DR calculation introduced in the Chapter 5 also forms the basis for quantification of confidence in the results. Any quantity that is to be analysed, say the average response to all dToU events in the winter months, is computed independently for each household. This results in a set of values $\overline{R}_h$ representing all households ($h = 1, 2, \ldots$). This set forms a distribution with a mean $\overline{R}$ and a standard deviation σ_R. For a large number of households, the Central Limit Theorem implies that the observed mean $\overline{R}$ is a normally distributed random variable centred on the true mean with a standard deviation of $\sigma_R/\sqrt{N}$, where N is the number of households involved in the calculation. The quantity $\sigma_x/\sqrt{N}$ is known as the *standard error (SE) of the mean* and its magnitude is shown explicitly for all results in Section 7.5, where as a results of stratifying by across households, sample numbers are relatively limited.

Using the mean and standard error, the ubiquitous Z-test can be employed to determine the confidence in the results. For example, a response to a dToU signal is deemed significant at a 90% confidence level if it differs from zero by at least 1.65 SE units. When statements regarding significance are made, this thesis adopts a 90% confidence level unless otherwise stated.

A slightly more elaborate approach is needed to determine the significance of a difference between two measured quantities, such as the responses of household categories a and b. Let the households in category a have a mean response $\overline{R}_a$ with an associated SE e_a, and those in category b a mean response $\overline{R}_b$ with a SE e_b. When considering whether there is a significant difference between $\bar{x}_a$ and $\bar{x}_b$, this is equivalent to asking whether the difference $\delta_{a-b} = \overline{R}_a - \overline{R}_b$ deviates significantly from zero. Assuming independence of the results between the categories, the SE associated with the random variable δ_{a-b} is $e_{a-b} = \sqrt{e_a^2 + e_b^2}$. In this frame of reference, the Z-test can again be used to determine the level of confidence in the results.

Note that comparisons between household categories require larger samples than those between a single category and a fixed reference value. This is because both the observed value and the reference value are subject to noise, reducing the significance of results for any given sample size. In the example above, assuming identical SEs for the populations a and b would result in $e_{a-b} = 1.41e_a = 1.41e_b$. Therefore, the values $\overline{R}_a$ and $\overline{R}_b$ would need to be separated by at least $2.3 \times e_a$ to be considered significantly different at a confidence level of 90%.

7.2 High level results

This section presents high level results that pertain to all events in the trial year. The mean DR values for the trial, separated into responses to high and low price points, can be seen in Fig. 7.1. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders. The most engaged households (lightest shade) produced an average response that is approximately three times larger than all households combined (darkest shade). However, as will be seen in the remainder of this chapter, DR signal vary considerably depending on situation, and so these full year aggregate results should not be used as the expected response level for an individual event, but to serve as a reference point against which other results may be compared.

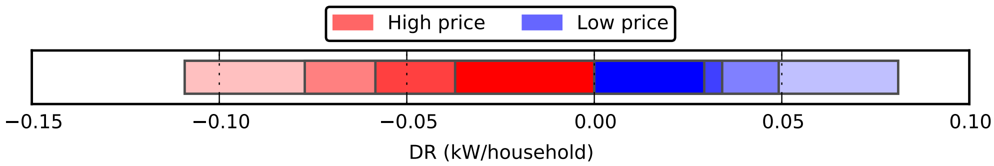

Figure 7.1: Mean DR by price level calculated over all events in the trial year. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

7.2.1 Trial year overview

Figure 7.2 provides a visual illustration of the response to dToU price signal for the most engaged 25% of households over the full trial year. The left panel shows the electricity price signal by day (rows) and hour of day (columns) and the right panel displays the observed DR with identical axes. The colour scale is clipped to an upper and lower bound of ±0.25 kW/household to ensure visibility of low-amplitude events, but some events far exceeded these visual bounds.

The minimum and maximum measured DR values over the year were -0.32 kW/household and 0.49 kW/household for high and low price events respectively, though this is quoted for a single time index averaged over 247 households, so intrinsic variability in response signal will be significant. A comparison between both panels shows that the high prices correspond well with the blue demand reduction periods, and the low prices with the red demand increase periods.

As expected, it is apparent from Fig. 7.2 that the largest responses occur during the hours when people are generally awake, approximately between seven in the morning and midnight. Furthermore, from a seasonal perspective it appears that DR is lower during the summer months, likely because people are out more or have less loads to shift. Furthermore, the responses at the beginning of the year, during winter, appear to be of greater magnitude than those in the following winter. This is most likely related to the cold spell in the early months of 2013, but we cannot exclude the possibility that this is related to the novelty value of the tariff wearing off.

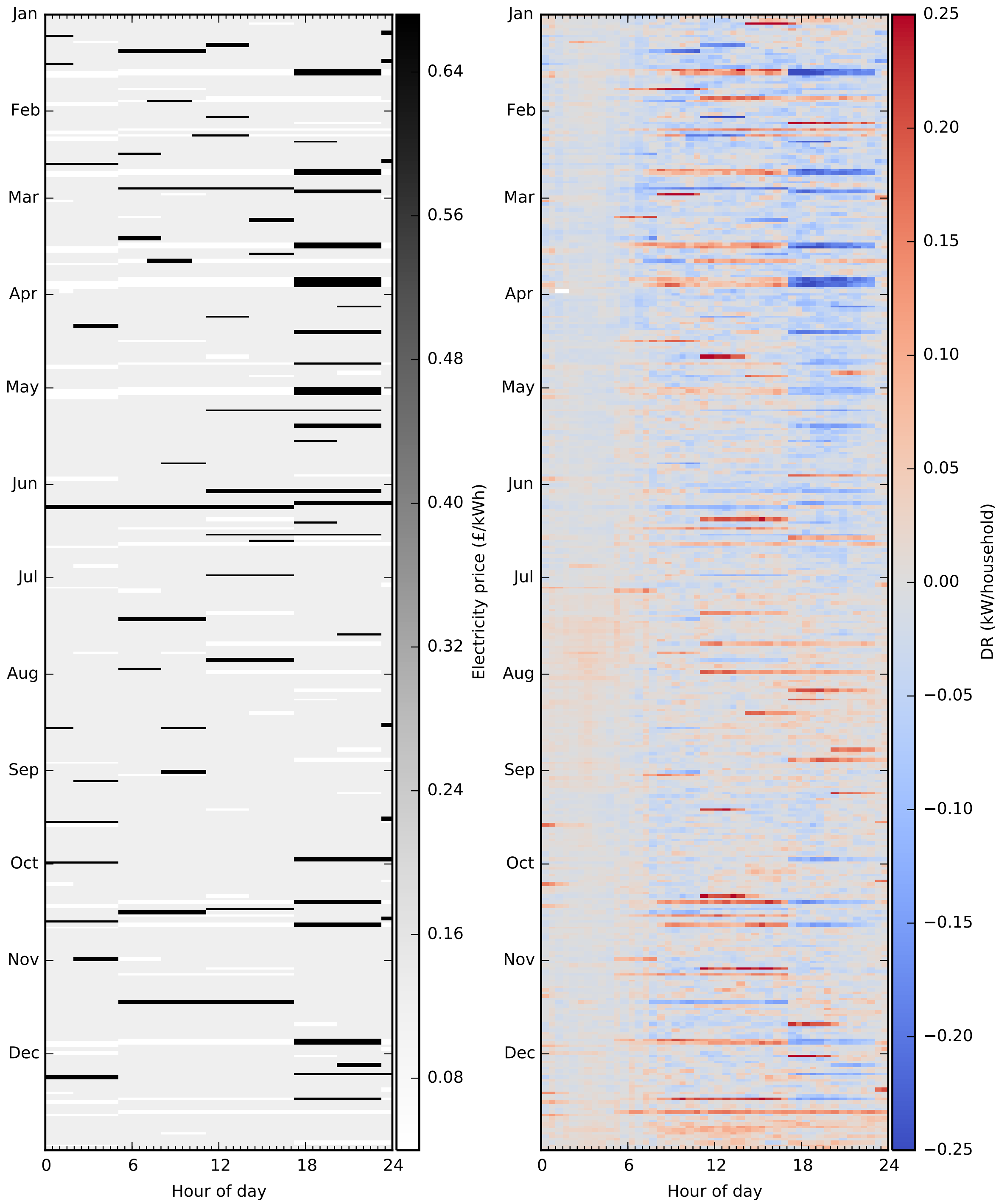

Figure 7.2: Heat map of (left) price and (right) DR signal for the full trial year of 2013. DR is calculated for the most engaged 25% of households only.

7.3 Constraint Management events

The Constraint Management events are tailored to the management of constrained distribution networks, by incentivising households to reduce their electricity consumption at peak times. This is achieved using a high electricity price during peak hours, coupled with a low price in the surrounding hours, thus enhancing both reduction and shifting of load.

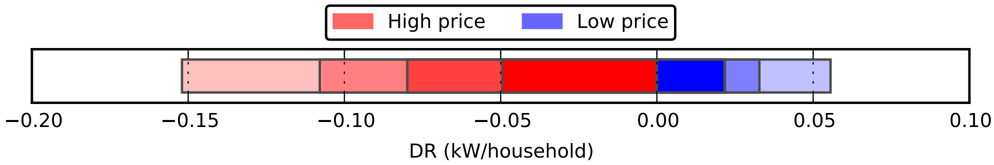

Figure 7.3: Mean change in consumption in response to CM events. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

Figure 7.3 shows the observed deviations in demand averaged over all CM event hours. Recall that CM events are inherently imbalanced between low and high price durations in order to achieve the highest possible demand reduction. The mean reduction during peak hours amounted to 50 W/household, which tripled to 150 W/household for the 25% best responders. From a network perspective, both values indicate the potential for significant impact on load profiles when considered against the measured after diversity peak demand of 1 kW/household.

Any incentivised load increase in the adjacent low price periods was spread over a larger number of hours, resulting in an average increase in consumption over all time indices of only 20 W/household.

7.3.1 Reductions in peak and mean demand

CM events were analysed individually to quantify their success in reducing the peak load of participating households. Figure 7.4 shows the mean load reduction across all high price time indices. The full group means, indicated by the darkest red bars, have a mean over all events of 50 W per household, while the most engaged 25% of the group, indicated by the lightest bars, regularly exceed 100 W/household reductions. No significant trends are visible with respect to event persistence (number of consecutive days) for either the peak or mean demand reduction graphs.

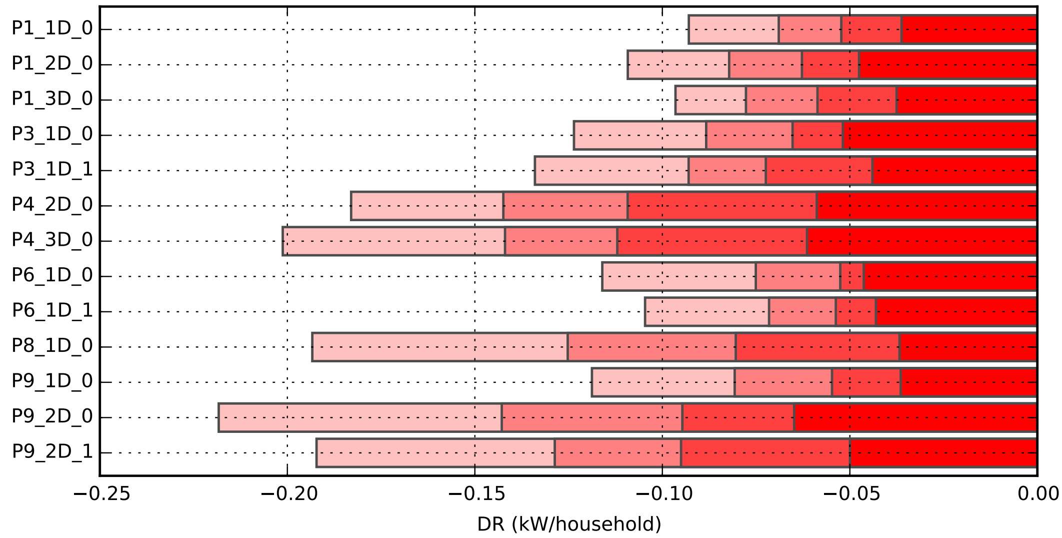

Figure 7.4: Mean change in demand over the high price period of the CM events. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders. Details of the individual events are given in Table 4.5 and Fig. 4.9.

7.3.2 Highlighted Constraint Management events

This section illustrates the observed response to CM events by analysing load profiles for three events that took place respectively on a weekday morning, a weekend afternoon and a weekday evening.

Morning, weekday

Figure 7.5 shows a CM event that took place on a weekday morning between 7am and 10am. The black line depicts the actual power consumption and the inferred increase or decrease in power consumption relative to the computed baseline is shown in red and blue respectively. The semi-transparent curve plotted alongside shows the result of the same analysis applied to the 25% most engaged households.

A clear reduction in demand during peak hours can be seen, accompanied by subsequent increase in the period spanning approximately five hours after the event. In this case there is no anticipatory demand increase before the high-price event, probably because the event took place early in the morning.

Although the observed behaviour is consistent with load shifting, it is not possible without detailed per-household analysis to distinguish shifting of loads from independent load reduction during the high price period and load increase during the low price period. For the subpopulation of the 25% best responders both the demand reduction and the subsequent demand increase are seen to be significantly enhanced.

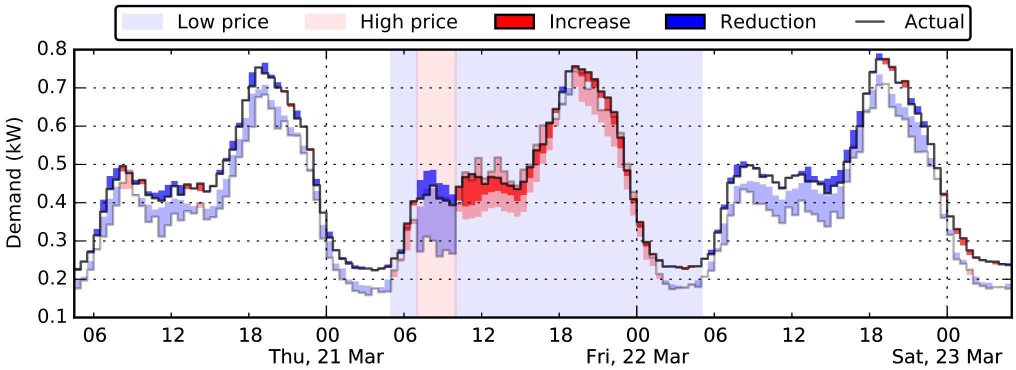

Figure 7.5: CM event P3_1D_1. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

Afternoon, weekend

Figure 7.6 shows a weekend afternoon event. Demand decrease during the high price period is again accompanied by demand increase during the surrounding low price periods. This time, with people having been awake for several hours before the high price period, load increase is visible both before and after the event.

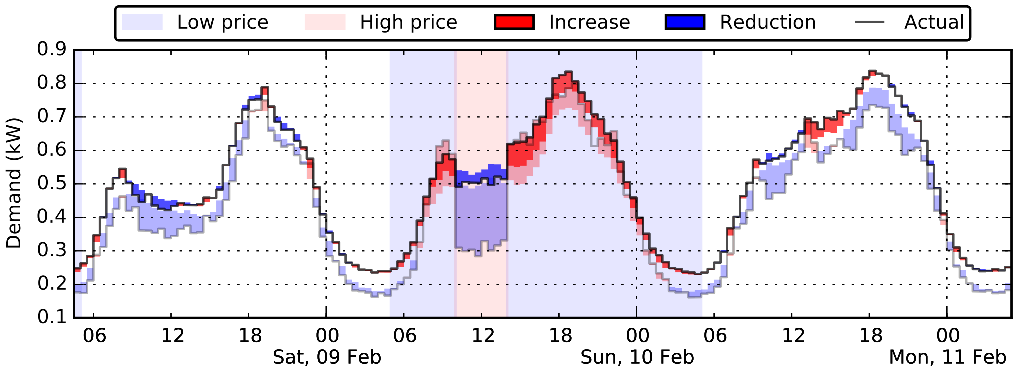

Figure 7.6: CM event P8_1D_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

Evening, weekday

The evening peak event spanning 2 trial days, shown in Fig. 7.7, demonstrates the ability to reduce peak loads on consecutive days. On both event days significant load shifting into the period before the high price event is visible. As was the case for the morning event in Fig. 7.5, the asymmetric nature of the observed shift in load is probably the result of people being asleep after 11pm, reducing the ability to modify demand after the high price period.

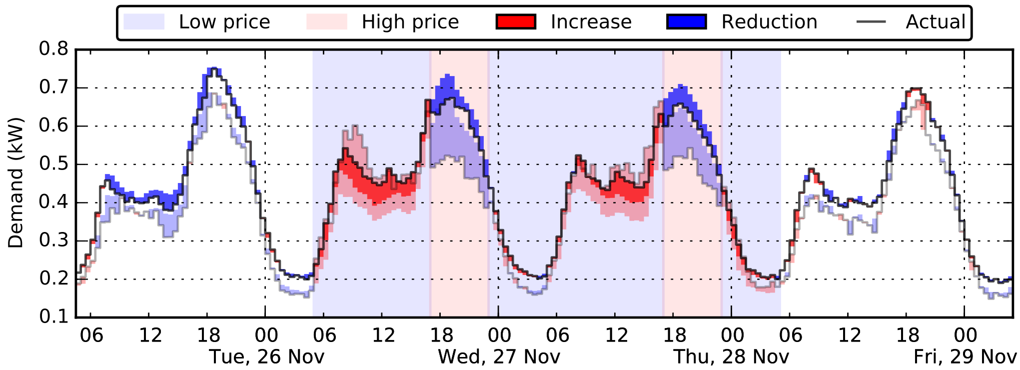

Figure 7.7: CM event P1_2D_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

7.4 Supply Following events

While CM events were aimed at the specific use case of peak load reduction, the SF events are designed to provide a more general insight into consumer response to pricing signals, with supply demand balancing as a representative use case. The SF trials make use of a single high or low price signal that is maintained for durations in the range 3–24 hours.

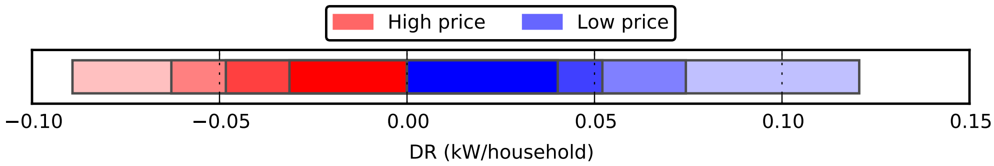

Figure 7.8: Mean DR over all SF events. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

Figure 7.8 shows the average change in demand for all SF events except the 24 hour low price event (L24_05). This event was excluded to ensure an equal number of low and high price hours were aggregated, therefore enabling a fair comparison of the mean response magnitudes.

The response to low price signals of 40 W/household (increase) was slightly larger in magnitude than the response to high price signals of 30 W/household (decrease). As was the case for CM events, the mean change in demand of the 25% best responding households is approximately three times larger.

The greater magnitude of response from low price events is perhaps surprising considering the asymmetry in the price incentives: consumption during high price periods was penalised by £0.55/kWh relative to the default price, while shifting consumption to a low price periods was rewarded (compared to default price) by only £0.08/kWh. This fits with the narrative that consumers respond to the concept of a high and low pricing period, rather than the numerical value of the price itself.

7.4.1 High price events

Figures 7.9 to 7.11 depict the observed load profiles and inferred changes in demand level for three high price events, with a duration of 3, 6 and 12 hours, respectively. A reduction in demand is visible in all cases.

The early morning event shown in Fig. 7.10, with high price beginning at 5am and ending at 11am, shows response increasing as people awoke. As with the CM events, the best responding

25% of households shows a much more pronounced demand reduction in all cases; typically three times greater in magnitude than the full group mean.

Curiously, although there is a significant reduction in demand during the high price periods, there does not appear to be a corresponding increase in demand in the surrounding default price periods. This is qualitatively different from the CM events where signs of load shifting were visible. This may be the result of the difference in the structure of CM events for which the high price period was flanked by low price periods.

7.4.2 Low price events

Figures 7.12 to 7.14 depict the observed load profiles and inferred changes in demand level for three low price events with a duration of 3, 6 and 24 hours respectively. All events show appreciable levels of increased demand, especially at those times where the demand was already anticipated to be high.

As expected the performance of the 25% highest ranked households outstrips the performance of the general population in all three examples. In fact, in Fig. 7.13 the inferred increase in demand is so high that the resulting morning demand peak for this subpopulation would exceed the evening peak as a result of the low price signal. This is an important finding for a future in which households are set to become increasingly responsive to dynamic pricing signals. A signal meant to stimulate demand at a system level (supply following) may thus trigger unexpected local network constraints. This highlights the need for an integrated approach to generation and distribution system operation and planning.

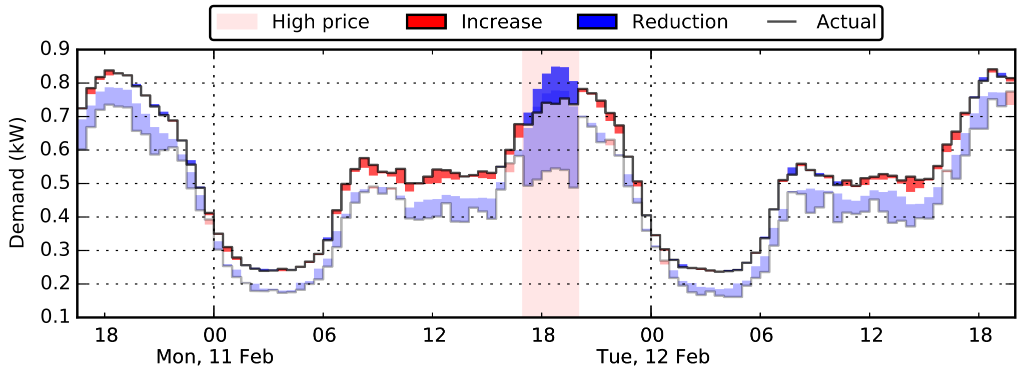

Figure 7.9: SF event: H3_17_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

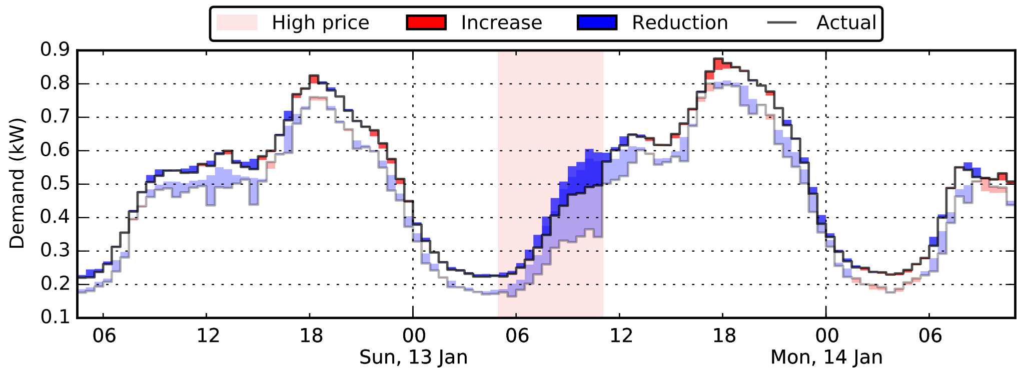

Figure 7.10: SF event: H6_05_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

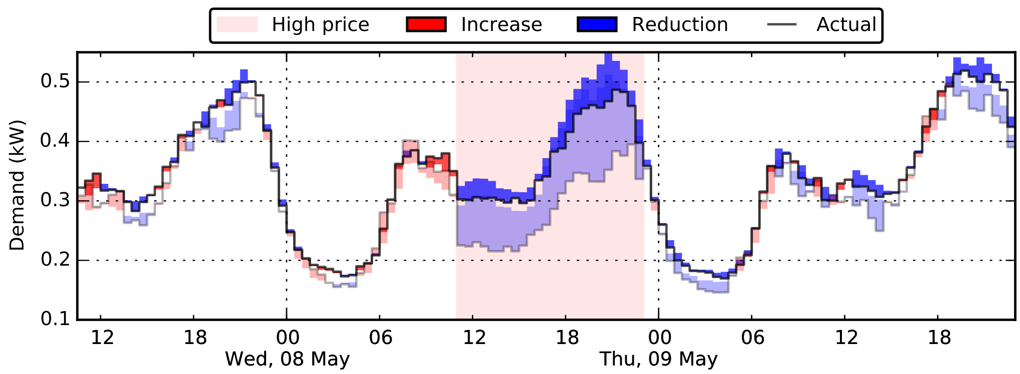

Figure 7.11: SF event: H12_11_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

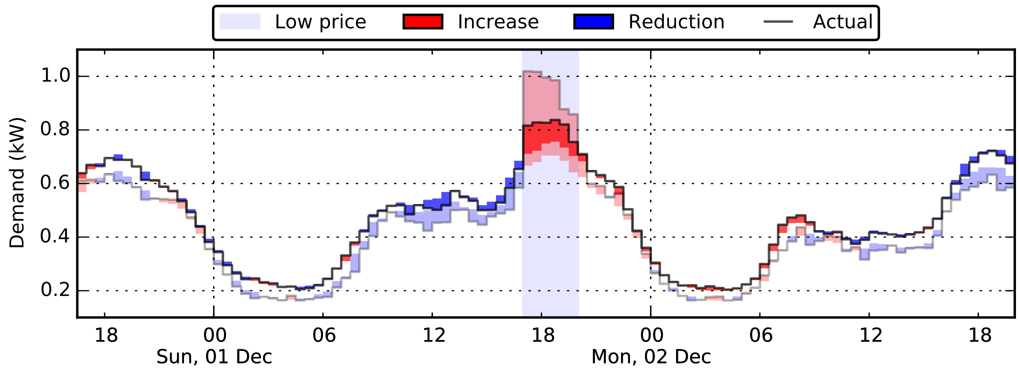

Figure 7.12: SF event: L3_17_2. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

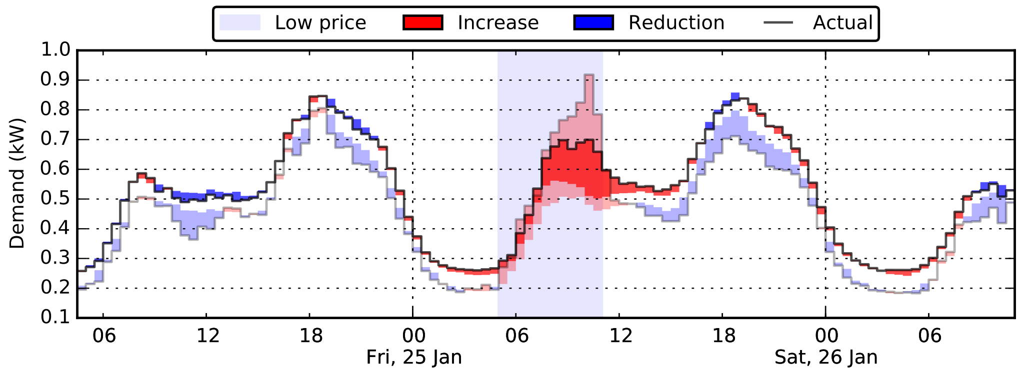

Figure 7.13: SF event: L6_05_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

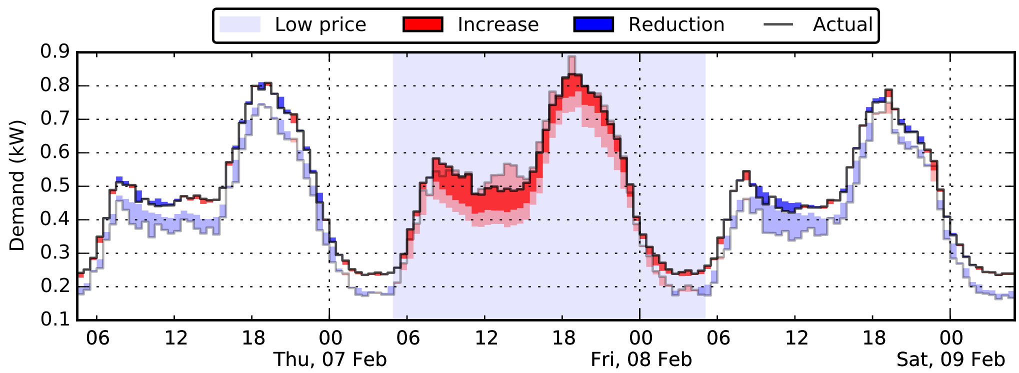

Figure 7.14: SF event: L24_05_0. The lighter shaded Increase, Reduction and Actual indicate the response from the most engaged 25% of households.

7.4.3 Temporal determinants of response magnitude

The results in Figs. 7.9 to 7.14 suggest that consumer responsiveness to dToU pricing is highly dependent on the timing and duration of the event in question. In this section the observed changes in demand are analysed as a function of event duration and their timing by season, month, day-of-week and hour-of-day.

Duration of events

Figure 7.15 shows the breakdown of DR by event duration. Mean demand reduction in response to a high price signal is approximately constant at a full group mean of 30 W/household for all households, increasing to approximately 80 W/household for the best responders.

For low price signals a trend is visible in that shorter duration events result in greater average demand increases. At 12 and 24 hour durations, response level is around 30 W/household, increasing to to 50 W/household for 3 hour duration events. This suggests that consumers have a limited ability to make opportunistic use of energy (power × time). In other words, there may be a limit to the value gained from the additional use of appliances (e.g. washing machine, dishwasher, etc.).

For the most engaged 25% of households, and for both high and low price signals, response levels are again seen to be approximately 3 times that of the full group mean.

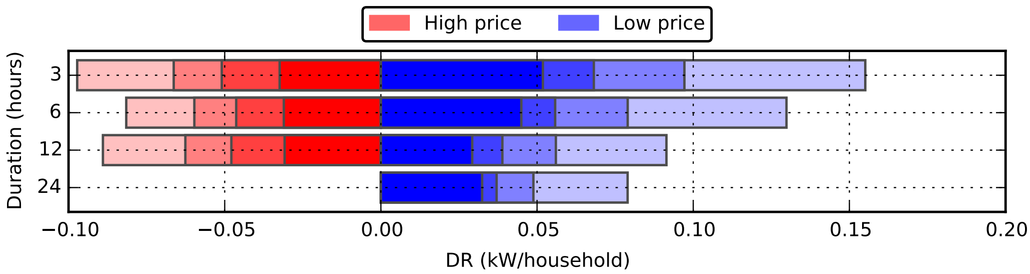

Figure 7.15: Effect of event duration on DR. N is the number of events over which the mean was taken. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

Season and months

Figure 7.16 shows mean DR by the Elexon defined seasons and Fig. 7.17 provides a breakdown by month. As expected, the ability to reduce demand is highest in the colder and darker winter months, decreasing dramatically during the high summer and early autumn months. However, the ability to increase demand in response to a low price signal is impacted to a lesser extent.

As high price DR constitutes a reduction in demand, ability to respond is limited by the level of discretionary demand available for reduction. Low price response, being an increase, is not limited in this way. This may explain why the low price response is limited by seasonality to a lesser extent than high price response.

Day of week

Figure 7.18 shows the DR breakdown by days of the week. No clear trend can be seen, except perhaps for a slightly enhanced response to low price signals on Fridays and Sundays.

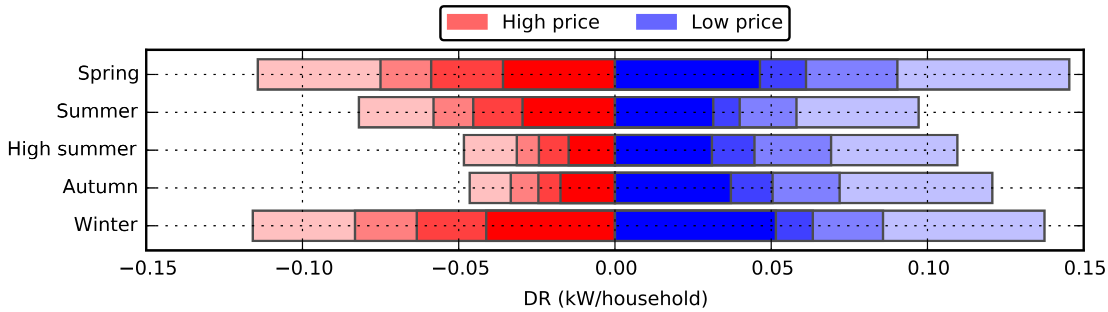

Figure 7.16: Mean DR by Elexon seasons. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

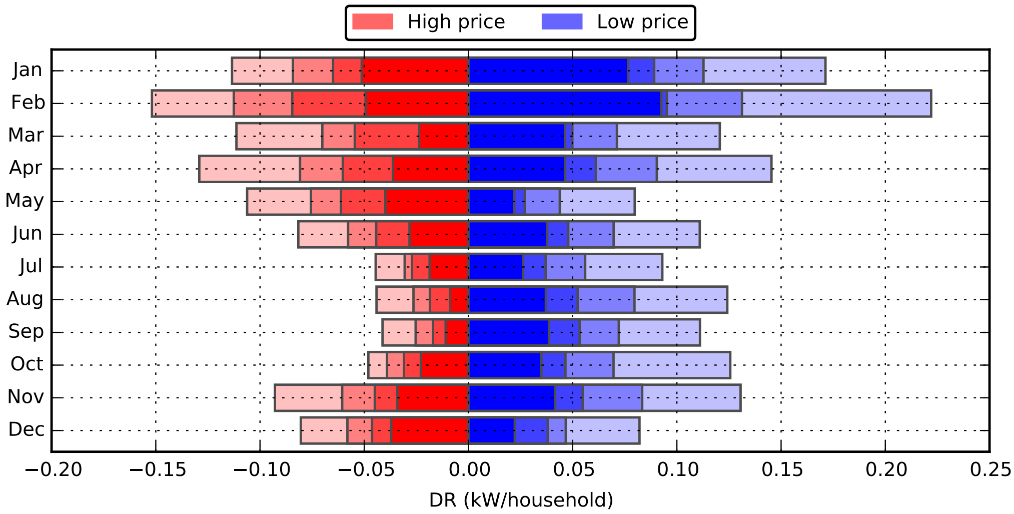

Figure 7.17: Mean DR by month. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

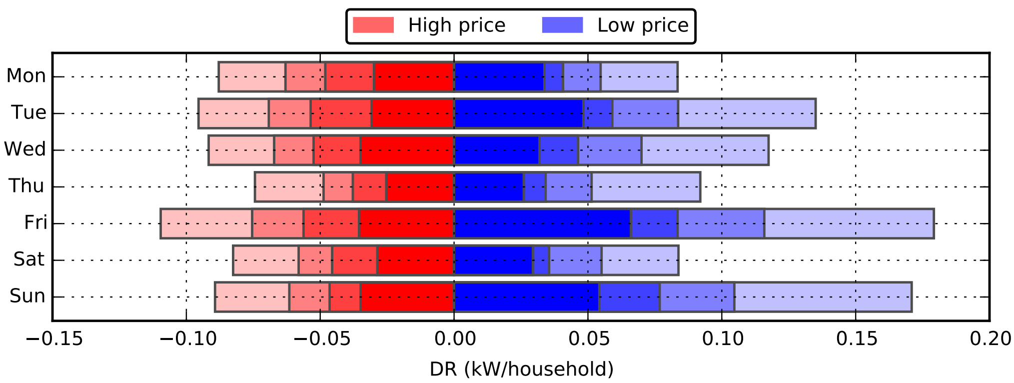

Figure 7.18: Mean DR by day of week. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

Time of day

Figure 7.19 shows a breakdown of DR by time-of-day. This is further refined in Fig. 7.20, which adds a subdivision by Elexon seasons. The gaps in the graph indicate an absence of events for those particular season/hour combinations. Note that even when relevant events are available, numbers of events are generally low, leading to increased statistical noise in the reported figures. Nevertheless, the results clearly demonstrate that price-driven changes in demand are most abundant during the waking hours of the day and in the winter season.

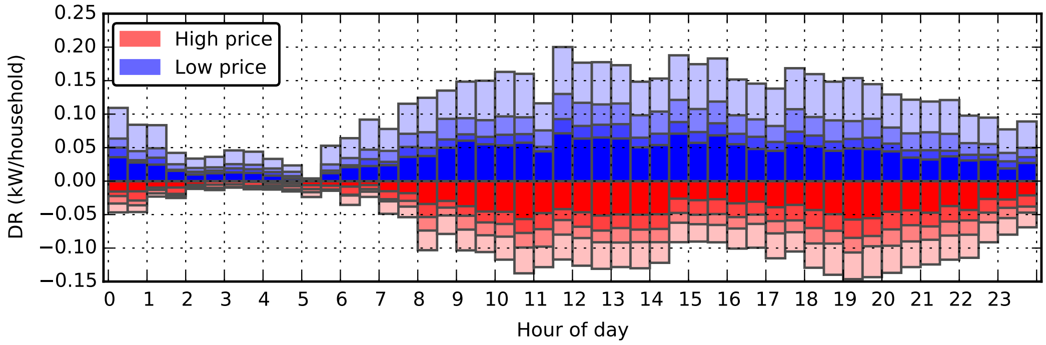

Figure 7.19: Full year mean DR by hour of day. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

Crucially, the availability of the DR resource is most prevalent during those times when it is most likely to be needed for constraint management: during the winter high demand periods. As can be seen in Fig. 7.20, in winter the 50% best responding households can achieve demand reductions of 100 W/household during the periods of the morning and evening peaks. A potential concern for DNOs is the limited ability to reduce residential power consumption during high summer afternoons. With the increased dependence on air conditioning in office buildings in particular, high summer conditions are joining the winter peaks in defining binding network constraints.

From a supply following perspective, the potential to increase or decrease power consumption during the nighttime hours looks limited, particularly during the summer. This is likely due to the inconvenience of manually controlling appliances at this time. This may change with increased penetration of home automation technology.

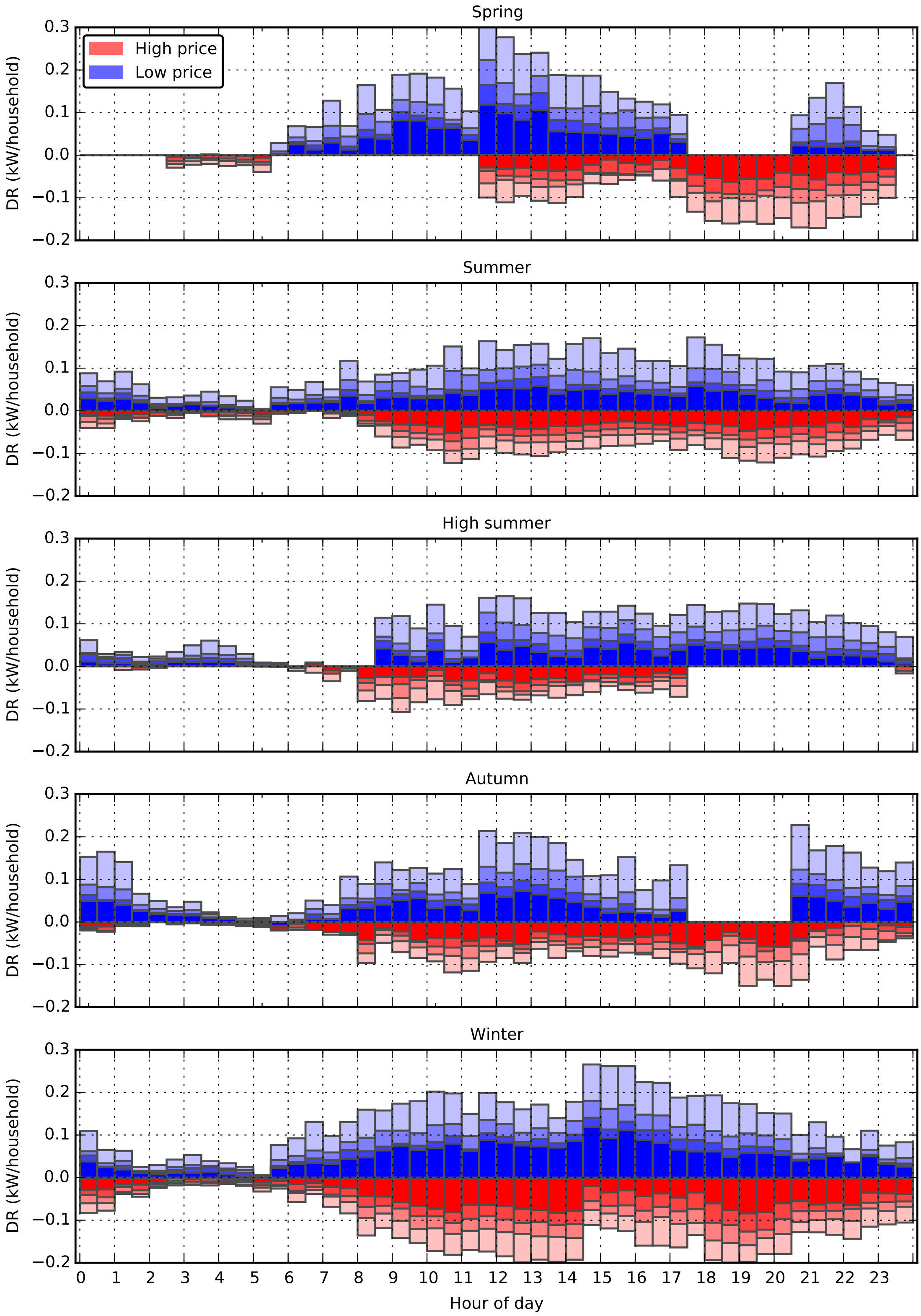

Figure 7.20: Seasonal mean DR by hour of day. Bars, from lighter to darker shading, represent the average for subgroups of the most engaged 25%, 50%, 75% and 100% of responders.

7.5 Social stratification of response

The previous sections have mainly been concerned with the time dependancy of DR. This section presents a complementary analysis that considers the properties of the receiver instead of the signal. High level social indicators are used to quantify and understand the observed measures of DR. Acorn group and household occupancy are first examined separately and then combined in an aggregated format using the three aggregate Acorn group classes and three occupancy classes that were defined for use in LCL.

In the case of SF events, all events except the 24 hour low price events (L24_05) were included in the analysis. This ensured that an equal number of low and high price hours, that were taken from the same hours of the day, were examined. This allowed a direct comparison of the high and low price response magnitudes. For CM events, this was not possible due to the asymmetric nature of the event design; low price hours greatly outnumbered high price event hours.

7.5.1 Acorn group

Figure 7.21 and Fig. 7.22 show a breakdown of DR by Acorn group for SF and CM events respectively. While certain groups performed better than others on average, these differences are hardly statistically significant and there is only a minima correlation with the albeit loose Acorn trend of decreasing wealth from A to Q. It should be noted that results from groups with low sample numbers may not be representative of their populations, as indicated by the error bars.

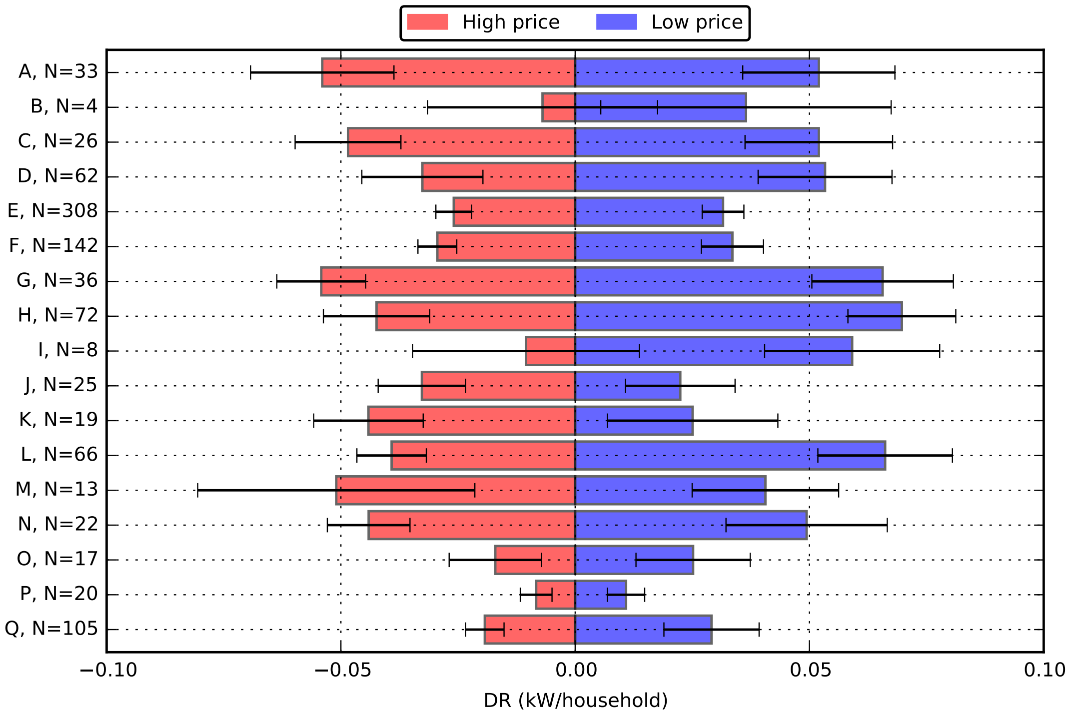

Figure 7.21: Mean DR by Acorn group. SF events only. N is the number of households in each group, and the SE of the mean is indicated by black bars.

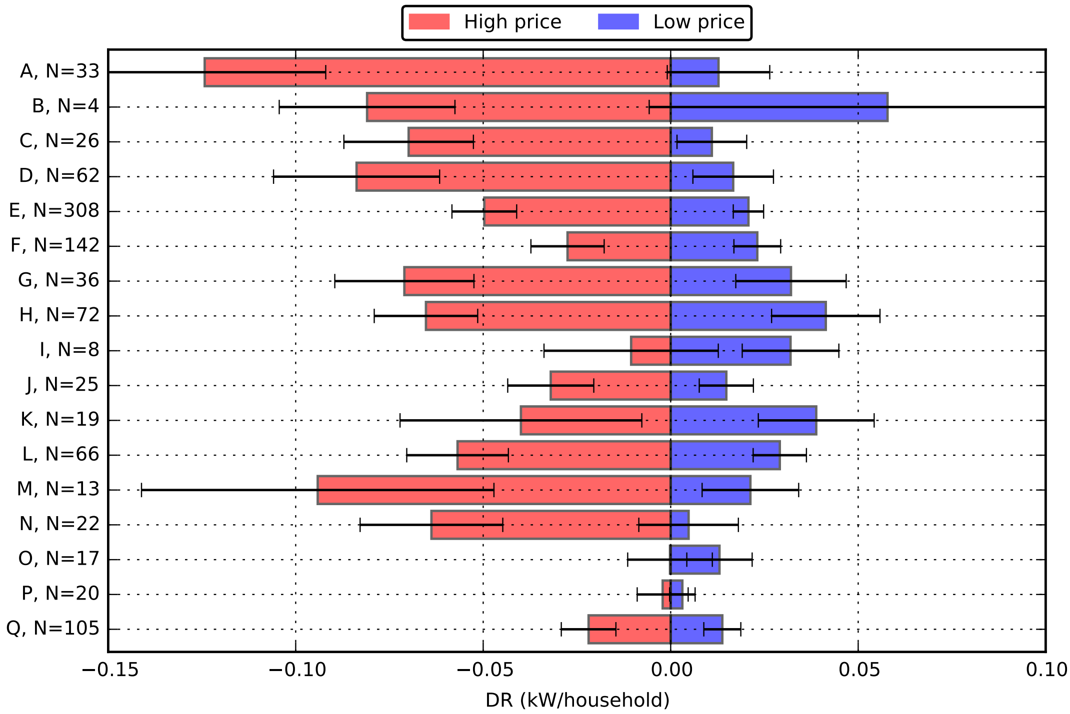

Figure 7.22: Mean DR by Acorn group. CM events only. N is the number of households in each group, and the SE of the mean is indicated by black bars.

7.5.2 Household occupancy

Figure 7.23 and Fig. 7.24 show a breakdown of DR by number of household occupants for SF and CM events respectively. The results show a qualitative difference in the response to high and low price events.

For the high price signal and for both CM and SF event types, the response clearly increases in magnitude with occupancy counts up to three. Results for higher occupancy counts are insufficiently precise to extrapolate this trend beyond three occupants. The demand increase resulting from low price events appears to show a similar trend, but less pronounced and with less significance (see the error bars).

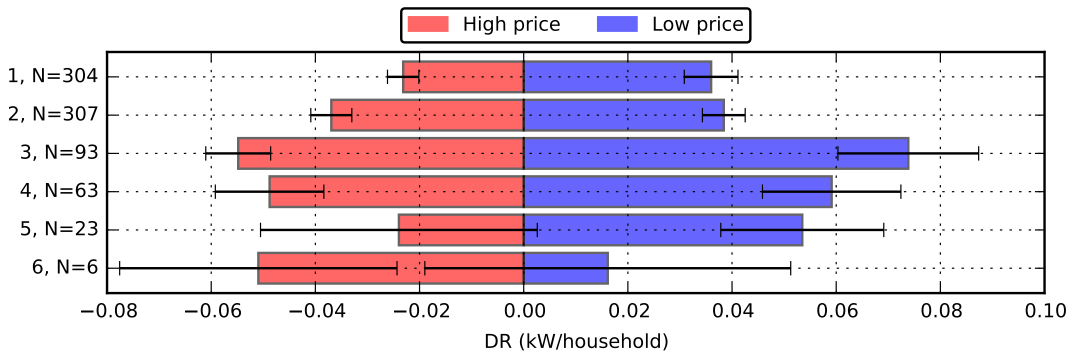

Figure 7.23: Mean DR by number of household occupants. SF events only. N is the number of households in each group and includes only those for which survey data was available. The SE of the mean is indicated by black bars.

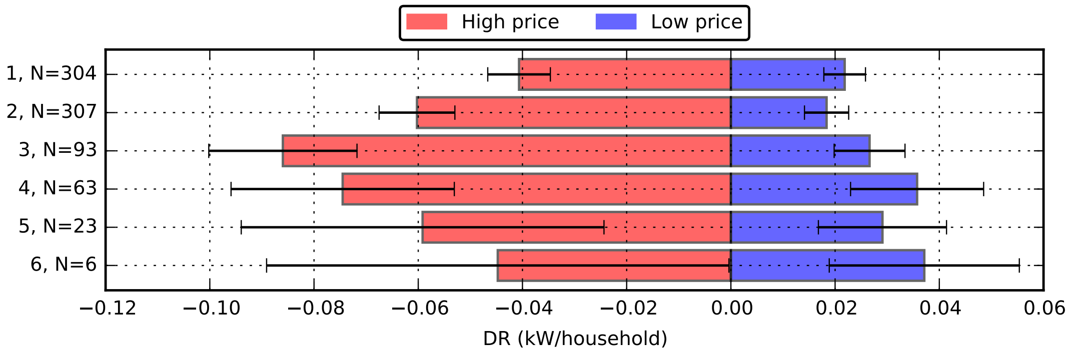

Figure 7.24: Mean DR by number of household occupants. CM events only. N is the number of households in each group and includes only those for which survey data was available. The SE of the mean is indicated by black bars.

7.5.3 LCL Acorn and occupancy classifications

LCL specific classifications were created from the Acorn group and occupancy variables by aggregating each into three discrete groups. These classifications were created in order to standardise analysis between LCL programme partners and across output reports. The LCL Acorn classes are:

- Affluent: Acorn groups {A, B, C, D, E}.
- Comfortable: Acorn groups {F, G, H, I, J}.
- Adversity: Acorn groups {K, L, M, N, O, P, Q}.

And the orthogonal occupancy classes are:

- 1 person.
- 2 people.
- 3 or more people (3+).

Demand response

A breakdown of DR by the above LCL Acorn and occupancy classes are presented for SF events in Fig. 7.25 and for CM events in Fig. 7.26.

Both figures clearly demonstrate a trend of increasing response with occupancy level at high price, though this trend is less pronounced for the Adversity class. The same trend is visible for the low price signal applied to SF events, though only for the Affluent and Comfortable Acorn classes. No clear trend is visible for the low price signal applied to CM events, for either Acorn group or occupancy level. Large (3+) Affluent households clearly outperform all other groups with respect to high price signal response, with a demand reduction 110 W/household for CM events at an occupancy level of 3+. The load profiles presented in the next section in Fig. 7.27 suggest that the greater observed response signal may be related to greater overall consumption levels.

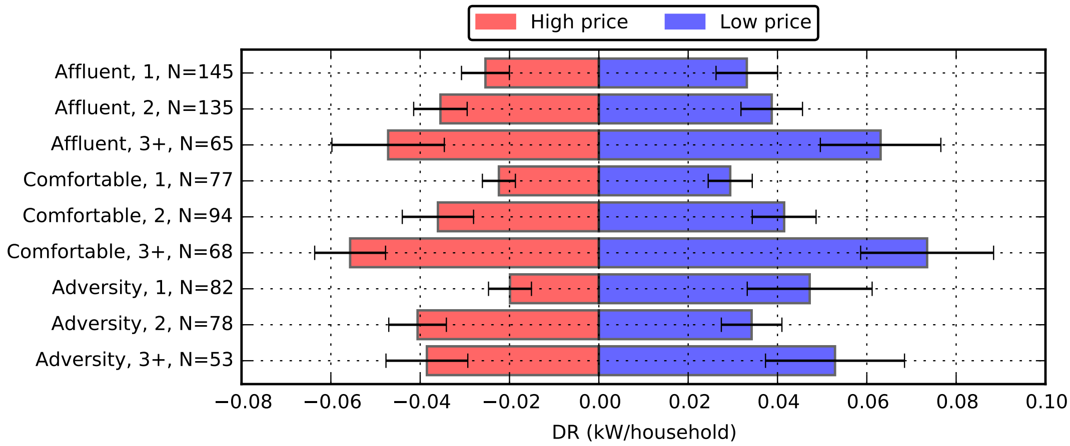

Figure 7.25: Mean DR by LCL Acorn and occupancy class. SF events only. N is the number of households in each group, and the SE of the mean is indicated by black bars.

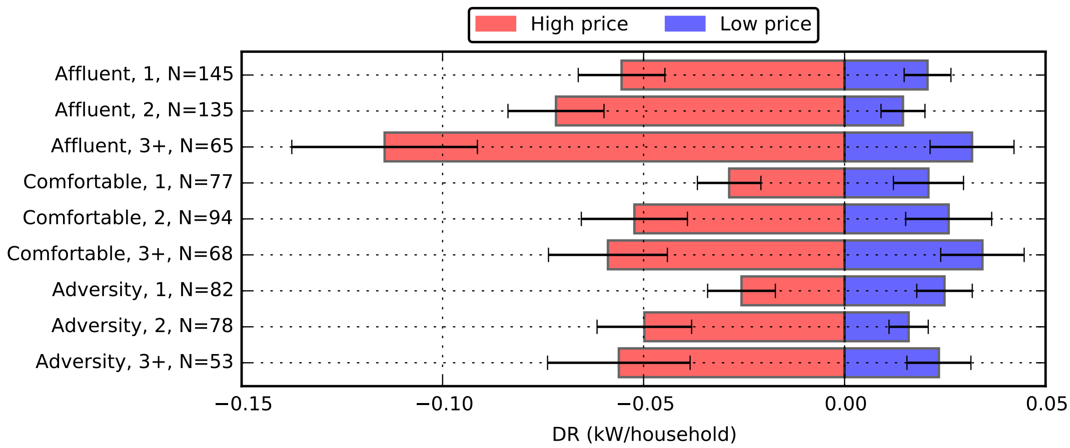

Figure 7.26: Mean DR by LCL Acorn and occupancy class. CM events only. N is the number of households in each group, and the SE of the mean is indicated by black bars.

Load profiles

To provide context for the social breakdowns of DR, average day profiles of the nonToU group are shown in Fig. 7.27 for the LCL Acorn and occupancy classes. It can be seen that there is a significant difference in load profiles between social groups and occupancies, with wealthier households showing increased consumption levels and more pronounced load peaks.

This underlines the value of smart meter data for increasing visibility of demand profiles for individual households. This data can be used to enable more accurate load forecasting for system balancing, load growth analysis for network planning, improved network visibility for operation, and to target tariff and energy saving initiatives more effectively.

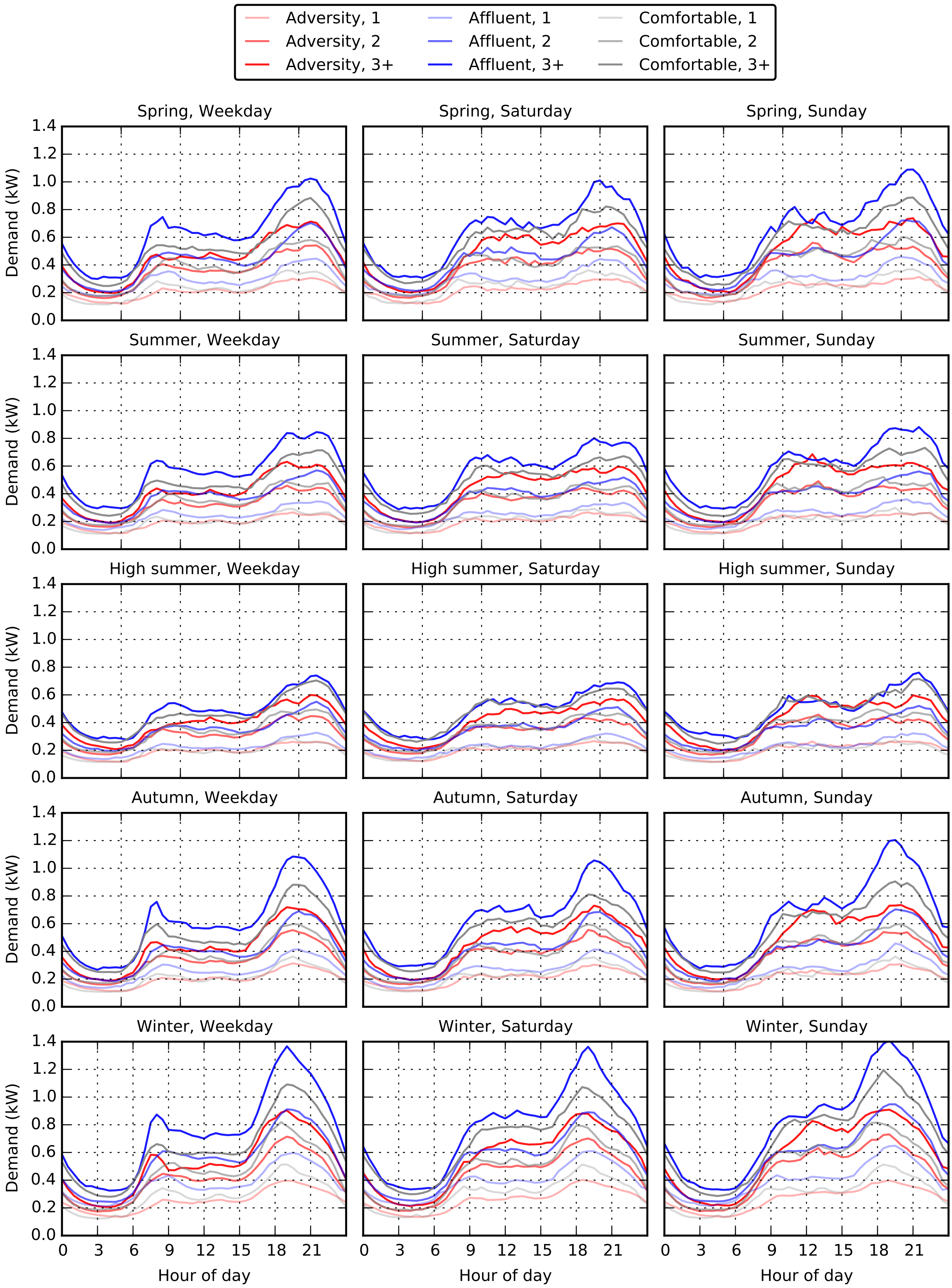

Figure 7.27: Average daily profiles of nonToU group, plotted for the LCL Acorn and occupancy classes.

7.6 Low Carbon London in context

Six key time-of-use (ToU) trials were reviewed in Section 3.3.7. These trials took place within the last 4 decades and were situated within the United Kingdom (UK) and the Republic of Ireland. These results are reproduced here, in Table 7.1 and Fig. 7.28 with the addition of the result of the LCL trial.

Table 7.1: Update of Table 3.11 in Section 3.3.7: Summary of ToU trial results from those considered closely related to the UK. For each trial, the table gives the peak to standard price ratio, the reduction in peak consumption achieved and the number of households in both the treatment and control groups (when available).

Index	Trial	Year	Location	Price ratio	Reduction (%)	N_tou	N_control
1	DTE [70]	1973	GB	3.00	25.0	-	-
2	EMUP [71]	1990	GB	6.59	16.0	250	75
3	NIKP [72]	2004	N. Ireland	1.60	12.0	100	100
4	EDRP EDF [73]	2010	GB	1.69	8.0	170	135
5	IESMT A [67]	2011	Ireland	1.42	7.2	1,368	1,170
6	IESMT B [67]	2011	Ireland	1.84	9.8	511	1,170
7	IESMT C [67]	2011	Ireland	2.27	9.0	1,370	1,170
8	IESMT D [67]	2011	Ireland	2.70	10.9	509	1,170
9	IESMT W [67]	2011	Ireland	2.70	11.6	100	1,170
10	CLNR [76]	2015	GB	1.99	6.4	600	9,000
11	LCL CM	2013	GB	4.72	8.0	988	3,768
12	LCL SF	2013	GB	4.72	7.9	988	3,768

Values for the LCL trial were calculated for both CM and SF events separately by averaging the DR over all time indices at high price and then dividing this by the average of baseline demand over all time indices at high price so that the percentage reduction in demand equals:

$$R_{\%} = 100\left(\frac{a}{b} - 1\right) \tag{7.1}$$

where

$$a = \frac{1}{N_a}\sum_{m\in\mathcal{M}}\sum_{h\in\mathcal{H}}\left(A_{m,h} - B_{m,h}\right) \tag{7.2}$$

$$b = \frac{1}{N_b}\sum_{m\in\mathcal{M}}\sum_{h\in\mathcal{H}} B_{m,h} \tag{7.3}$$

$A_{m,h}$ and $B_{m,h}$ are the actual and baseline demand at measurement index m and household h, $\mathcal{M}$ is the set of all high price time indices for the respective event type, $\mathcal{H}$ is the set of all households in the dToU group, and N_a and N_b are the counts of non-null values in a and b respectively.

Figure 7.28 is a reproduction of Fig. 3.1. The x-axis represents the ratio of the peak to standard price (at the time of the trial) and the y-axis the reported relative reduction in demand. The LCL result for CM events has been added and is indicated by the red circle.

The number of households in each trial group was used to determine a statistical weighting of each result. Assuming that each trial group had the same standard deviation of demand—this number was not provided in most trial reports—and that the control group was proper, the variance of the response measurement, S_R^2, can be shown (proof omitted) to be proportional to the below function of the sample numbers:

$$S_R^2 \propto \frac{1}{N_{\text{tou}}} + \frac{1}{N_{\text{control}}}$$

The data point weightings were set so as to be proportional to $1/S_R^2$, as is the standard in weighted least-squares regression (WLS). These were depicted in the scatter plot via the size of each point's

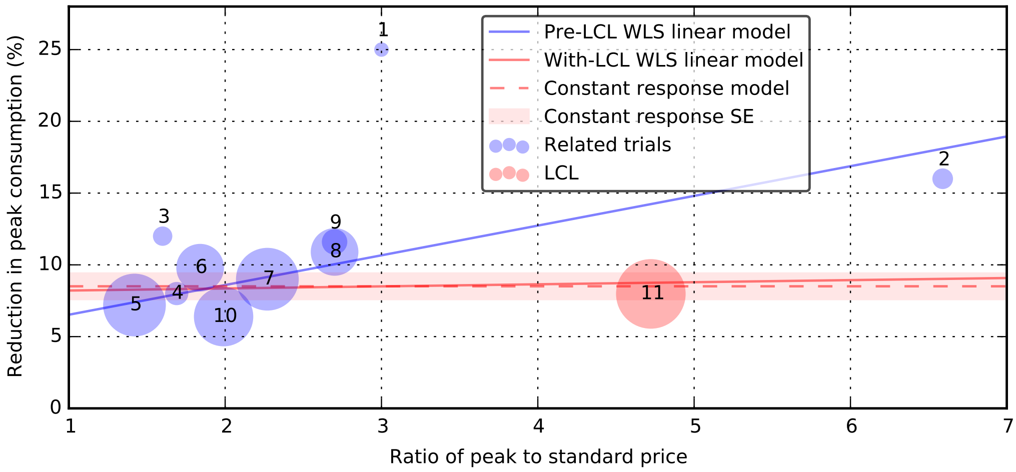

Figure 7.28: Update of Fig. 3.1 in Section 3.3.7: Summary of ToU trial results from those considered closely related to the UK in context with the LCL equivalent result. For each trial, the size of the point is proportional to its weighting, which is a function of the sample numbers. Each trial marker is referenced to the corresponding trial in Table 7.1 via its index number.

marker—larger markers having greater statistical weight. In the case of trial index 1, each group was assumed to contain 50 households. This was considered a conservative estimate (below the minimum reported group population of 75 for the other trials) and allowed the point to feature in the graphic with a small weighting. In addition, a best WLS regression model consisting of a constant and a proportional parameter was fitted to the points to give a sense of whether a trend might exist.

It can be seen that related trials clustered between peak-to-standard price ratios of 1.4–2.7. In this context the LCL trial has entered new territory by using the much higher price ratio of 4.7. This, combined with the statistically robust sample numbers used in the trial, has allowed the LCL result to have significant impact on the interpretation of the data points.

In Fig. 7.28, the pre-LCL linear model is indicated by the solid blue line and the with-LCL linear model by the solid red line. It can be seen that the addition of the LCL data point has changed the trend from one that suggests peak reduction might be proportional to peak-to-standard price ratio, to one that suggests a constant relative reduction in peak consumption in response to a high price signal. For illustration, a constant reduction model has been added to the graph, indicated by the dashed red line, where the red shaded area indicates the 95% confidence interval on the model mean (calculated by WLS regression). The implications of this model are that a constant reduction in demand of 6.8–10.2% might be expected from future ToU or dToU trials at a 95% confidence level. It should be noted that sampling error would have to be added to this in order to calculate a prediction interval.

7.6.1 Evaluation

This cursory comparison with related ToU trials does not control for price event structure such as time-of-day or duration. It should also be remembered that the comparison trials used ToU tariffs while LCL used a dToU tariff where peak price periods could occur at any time of day and over different durations.

The analysis conducted in this section places LCL into context against previous UK related trial results and informs us that we should not necessarily expect an increase in the peak to standard price ratio to result in enhanced peak price consumption reductions.

This section concludes that the current UK related ToU and dToU evidence is not yet sufficient to state the extent to which peak reduction levels are affected by price levels. This is compounded by mixed messages in the existing literature: A recent paper by Faruqui [59] concluded that consumers are responsive to price and that increasing the peak to off-peak price ratio resulted in enhanced peak reductions, but the trend was not linear. Conversely, the recent the Department of Energy and Climate Change, UK (DECC) literature review of residential DR [44, Annex A; figures 4 and 5] formed the view that there is not a strong relationship between peak to off-peak price ratios and peak demand reduction. However, it is worth noting that most of the trials cited in these studies were from North America and therefore the same conclusions may not be applicable in the UK context.

7.7 Summary and conclusions

This chapter provides an overview of the primary results of the Low Carbon London (LCL) dynamic Time-of-Use (dToU) trial. The demand response (DR) signal, calculated as described in Section 5.4, is examined over a number of different stratifications of the response signal.

Constraint Management events. The Constraint Management (CM) events consisted of high price periods that targeted peak hours, flanked by extended low price periods. As such, they are intrinsically asymmetric with short high price periods surrounded by long low price periods. This is reflected in the observed response numbers, with an average demand reduction of 50 W/household and demand increase of 20 W/household. As expected, the 25% most engaged households delivered a larger response with an average reduction of 150 W/household and average increase of 55 W/household. This illustrates the potential of DR to reduce peaks and enhance utilisation of network assets.

The reduction in load during high price periods was always accompanied by an increase in load during the adjacent low price periods. Extended events that targeted peaks on up to three consecutive days were trialled and no significant difference in measured response was observed between days.

The decrease of demand during peak hours and increase during low priced hours is consistent with load shifting. However, such a signature response was not present in the high-price-only Supply Following (SF) events in which the peak was flanked by mid-price periods. This suggests that the apparent load shifting may be caused by opportunistic usage of the lower price electricity. Further investigation is required to identify the reason for this difference.

Supply Following events. SF events targeted high and low price events were used to establish the potential for consumers to respond to dToU signals at different times of the day and throughout the year.

Overall, households responded to high price signals with decreases in consumption levels that were much larger during the colder and darker winter months than in the peak of summer. A similar pattern is observed when the measured responses were analysed by the hour of the day. The demand reduction potential reached its maximum magnitude around the morning and evening peaks (on weekdays). The most engaged quarter of households achieved a mean demand reduction over 150 W/household during these periods, compared to 50 W/household for the average household. The strong correlation between demand reduction potential and absolute demand levels is a positive finding for the CM use case, as the reduction potential during peak demand periods will be higher than suggested by average response numbers.

Households responded to low price signals by increasing their demand levels. This increase was seen to be fairly constant during the waking hours of the day, at a level of 50 W/household across all households and exceeding 150 W/household for the most engaged households. During the night time even the best responders did not achieve an increase of 50 W/household. However, the ability of households to increase power consumption was only very slightly affected by the time of year. During the summer months in particular this led to an asymmetric response to high and low price signals.

These figures suggest an ability of households to assist in supply demand balancing, but this potential is currently limited to waking hours and is significantly larger during winter months. The proliferation of "smart appliances" that can autonomously respond to price signals may provide a more consistent response as human intervention will no longer be needed to activate it.

The dual objectives of Constraint Management and Supply Following may lead to conflicts. For example, an abundance of available wind power or the availability of large amounts of inflexible nuclear plant during low load conditions may result in very low electricity prices. From the system perspective it would be beneficial to use dToU pricing to incentivise customers to increase their consumption levels. However, doing so might cause unanticipated stress on the distribution network. Evidence of such situations was seen during the trials: the 25% most engaged households occasionally responded so strongly to low price signals that a new after diversity demand peak was created. On the other hand it is also possible that the two objectives align leading to synergies between system and network management. This is a common situation when high load conditions coincide with high marginal costs of supply (e.g. during the winter peak).

Socio-economic factors in DR. The responses of the targeted SF trials were analysed against two principal parameters that are known to be strong indicators of energy consumption: household occupancy (1, 2, 3+) and a socio-economic classifier based on the Acorn system. The three socio-economic groups—Affluent, Comfortable and Adversity—can be interpreted as a rough indicator of wealth.

Perhaps surprisingly, the socio-economic class did not have a significant effect on the observed DR for these single events, although results on CM events suggest that households in the Affluent class may respond more strongly to signals that specifically target peak hours. The measured response does depend strongly on occupancy levels, with larger households providing responses of larger magnitude.

Low Carbon London in context. An attempt was made to compare LCL trial results with past time-of-use (ToU) trials that are closely comparable to the United Kingdom (UK) context. Results were compared on two axes; peak price to standard price ratio, and relative reduction in peak price consumption. While the related trials had price ratios in the range 1.4–2.7, LCL had a price ratio of 4.7. This departure from the previous trial cluster, combined with statistically robust sample numbers, meant that the LCL trial contributed significantly to the interpretation of the overall landscape of trial results plotted on these axes. Before LCL the trend might have suggested the possibility of increased response with increased price ratio, however, the addition of the LCL data point makes a relative response look like a better model. A constant response model was fitted to the data points while taking into account the statistical weight of each trail. Using this model, it was estimated that the population peak demand reduction lay in the range 6.8–10.2% at a 95% confidence level.

Though the data was not sufficient to make a conclusive statement regarding the relationship between response level and price ratio, this analysis did place LCL into context against previous UK related trial results and informs us that we should not necessarily expect an increase in the peak price ratio to result in enhanced peak price consumption reductions. More data is required in order to investigate this further.

Chapter 8

Reliability and risk

This chapter builds on the findings of Chapter 7 and proceeds to investigate in detail how the implementation of dynamic Time-of-Use (dToU) tariffs may impact network reliability. Two main questions are addressed: first, the extent to which the distribution network operator (DNO) can count on dToU tariffs to reliably alleviate network constraints; second, how the use of dToU tariffs by suppliers may cause demand to violate network constraints. First, the predictability of Constraint Management (CM) events is examined and two simple response predictor models are proposed. This is followed by an analysis of the network capacity contribution of residential demand response (DR). The first part of this introduces the theory, where it is shown that capacity contribution consists of two components; mean and variance response. This more generally applicable theory is then applied in the specific context of CM events, where its application is shown to be significant under certain conditions. Second, the potential risk to the network from low price event induced demand spikes are empirically examined through use of the dToU trial data set, leading to the times of high network risk being identified.

8.1 Predictability of constraint management event response

The first question that is addressed is how a DNO can use dToU tariffs to alleviate network constraints. This analysis makes use of the CM trials, which specifically targeted load peaks. This section is concerned with quantifying the magnitude and variability of the observed DR for these events. It begins with the introduction of the baseline model that is used to calculate the mean DR values. The distribution of the CM events throughout the year is discussed, and correlations with baseline demand and weather variables examined. A simple model for CM event response magnitude is proposed. Confidence intervals for both this model, and for the response of the dToU group to a future event, are calculated.

8.1.1 The per-household baseline demand model

DR, in the context of this trial, is a reduction in demand relative to what would have been consumed without the price intervention, where price interventions are considered to be deviations from the default price. Quantification of DR thus requires the establishment of a hypothetical baseline demand of the dToU group in the absence of the price event. A linear model was used to establish a baseline demand profile for each household in the dToU group. While the details of its implementation are more thoroughly described in Section 5.4, the salient points are described below.

A key feature of this model is that it relates a household's baseline power consumption to the mean power consumption of the non-time-of-use (nonToU) group. This approach guarantees that the model captures events that could cause an overall bias such as bank holidays, extreme weather and sports events because they are expected to affect the dToU and nonToU groups equally. The predictor variables include the mean demand of the nonToU group and binary variables to allow for some time dependent structure. The baseline demand is calculated as:

$$B_{m,h} = \sum_{w=1}^{W} (\alpha_{w,h} d_w + \beta_{w,h} \overline{A'}_m d_w) + \gamma_h m \tag{8.1}$$

where $B_{m,h}$ is the baseline demand of household h at half-hour measurement index m and household h; $\overline{A'}_m$ is the mean actual demand of the nonToU group at measurement index m; d_w are binary dummy variables, one for each hour w in the $W(= 168)$ hours of the week; and the Greek letters $\{\alpha_{1,h}, \cdots, \alpha_{W,h}\}$, $\{\beta_{1,h}, \cdots, \beta_{W,h}\}$ and γ_h are parameters to be determined by the regression solver. This model effectively relates each half-hour measurement index linearly to the nonToU group mean demand, for each hour of the week, with an overall trend line.

This model was used to calculate a baseline mean demand for each household in the dToU group. Demand response was calculated as:

$$R_{h,m} = A_{h,m} - B_{h,m} \tag{8.2}$$

where $R_{h,m}$ is the inferred DR for household h at measurement ; $A_{h,m}$ and $B_{h,m}$ being the actual measured demand and the calculated baseline of household h respectively. Note that the baseline model outlined above describes the average expected consumption of a household on a given day and hour. In contrast to baseline models that are used for DR contracts [14], it does not attempt to predict random fluctuation in the household's power consumption, including the occasional absence of the inhabitants. As a result, the DR estimates have an intrinsic variability related to the natural variability (diversity) of demand. It can be argued that this approach is most applicable to the constraint management context where a DNO must take a decision on the basis of a load forecast. The baseline model may be considered an approximation of an optimal day-ahead load forecast.

8.1.2 High price demand response by CM event

This section quantifies the observed DR during the high price period of each event. The duration of the high periods, and therefore the number of half-hour measurements taken, depends on the peak that was targeted (defined by time-of-day and day-of-week). For example, in the case of weekday evening peaks, this was 6 hours, from 17:00 to 23:00. Some CM events targeted peaks on consecutive days, and the results for each day are treated as independent observations. We aggregate the measurements into single value measure of DR for each CM event day: mean power reduction over the high price period, calculated as:

$$\overline{R}_h = \frac{1}{N_{\mathcal{M}}} \sum_{m \in \mathcal{M}} R_{h,m} \tag{8.3}$$

where $\overline{R}_h$ is the mean response at high price of household h on a given CM event day, $\mathcal{M}$ is the set of measurement indices for the high price period that the event day, and $N_{\mathcal{M}}$ is the total number of non-null measurements in response matrix $R_{m,h}$ (as defined in Eq. (8.2)) for household h over time indices $\mathcal{M}$. For each of the 21 event days we obtain one mean DR observation, $\overline{R}_h$, from each household in the dToU group.

It is reasonable to assume significant independence and absence of bias of the errors in the inferred DR values for each household, $\overline{R}_h$. Random fluctuations in the nonToU response that affect all baselines are assumed to be small, because of the large number of households in the

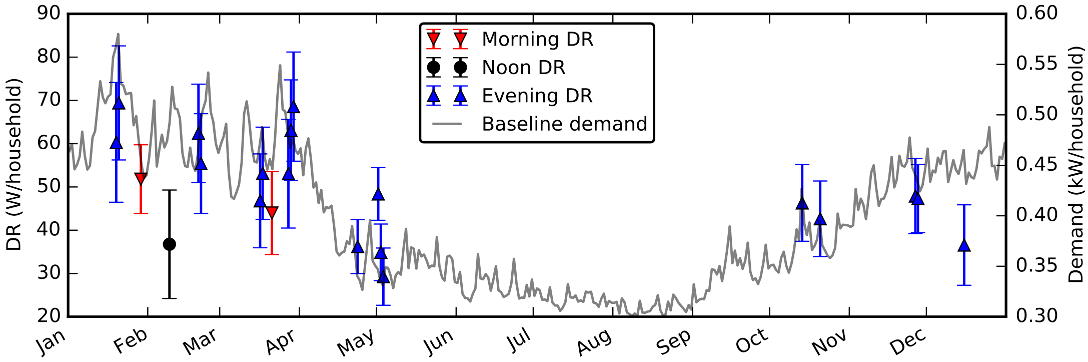

Figure 8.1: Mean high price CM event response against time in the trial year. Error bars indicate the standard error (SE) of the mean measurement. Baseline demand is calculated as the mean demand of each trial day.

nonToU group. As the per-event DR measurements are themselves the mean of the DR estimates from all households in the dToU group, the central limit theorem applies and the SE of the mean is used to estimate the measurement error. The results of this process and the times of the CM events are plotted in Fig. 8.1 and the numerical values are given in Table 8.1.

Table 8.1: CM events and mean DR and SE on the mean for the high price period of each CM event day. "Wall clock" time is used.

Event index	Event name	From	To	Duration (hours)	DR mean (W)	DR standard error (W)
1	P9_2D	19/01/13 17:30	19/01/13 23:00	6	60	14
2	P9_2D	20/01/13 17:30	20/01/13 23:00	6	69	13
3	P3_1D	29/01/13 07:30	29/01/13 10:00	3	52	8
4	P8_1D	09/02/13 10:30	09/02/13 14:00	4	37	13
5	P4_2D	20/02/13 17:30	20/02/13 23:00	6	62	11
6	P4_2D	21/02/13 17:30	21/02/13 23:00	6	55	12
7	P9_2D	16/03/13 17:30	16/03/13 23:00	6	47	11
8	P9_2D	17/03/13 17:30	17/03/13 23:00	6	53	11
9	P3_1D	21/03/13 07:30	21/03/13 10:00	3	44	10
10	P4_3D	27/03/13 17:30	27/03/13 23:00	6	53	13
11	P4_3D	28/03/13 17:30	28/03/13 23:00	6	63	12
12	P4_3D	29/03/13 17:30	29/03/13 23:00	6	69	13
13	P1_1D	23/04/13 17:30	23/04/13 23:00	6	36	6
14	P1_3D	01/05/13 17:30	01/05/13 23:00	6	48	6
15	P1_3D	02/05/13 17:30	02/05/13 23:00	6	35	7
16	P1_3D	03/05/13 17:30	03/05/13 23:00	6	29	7
17	P6_1D	13/10/13 17:30	13/10/13 23:00	6	46	9
18	P6_1D	20/10/13 17:30	20/10/13 23:00	6	43	9
19	P1_2D	26/11/13 17:30	26/11/13 23:00	6	47	9
20	P1_2D	27/11/13 17:30	27/11/13 23:00	6	47	8
21	P9_1D	15/12/13 17:30	15/12/13 23:00	6	36	9

It can be seen in Fig. 8.1 that the CM event DR reduces towards the summer months. As residential demand during summer is significantly less than during winter, this suggests that either weather or the baseline demand may be influencing factors on response magnitude. Furthermore, one may expect dToU DR to change over time as people gain experience with the programme, resulting in either decreasing (novelty wearing off) or increasing trends. To investigate these relations further, we examine the DR correlation with the baseline demand level, three readily available macroscopic weather measurements (temperature, wind speed and solar elevation angle) and the numerical event index. In the case of solar elevation, 0° is considered to be the when the sun is at the horizon, and 90° when the sun is directly overhead.

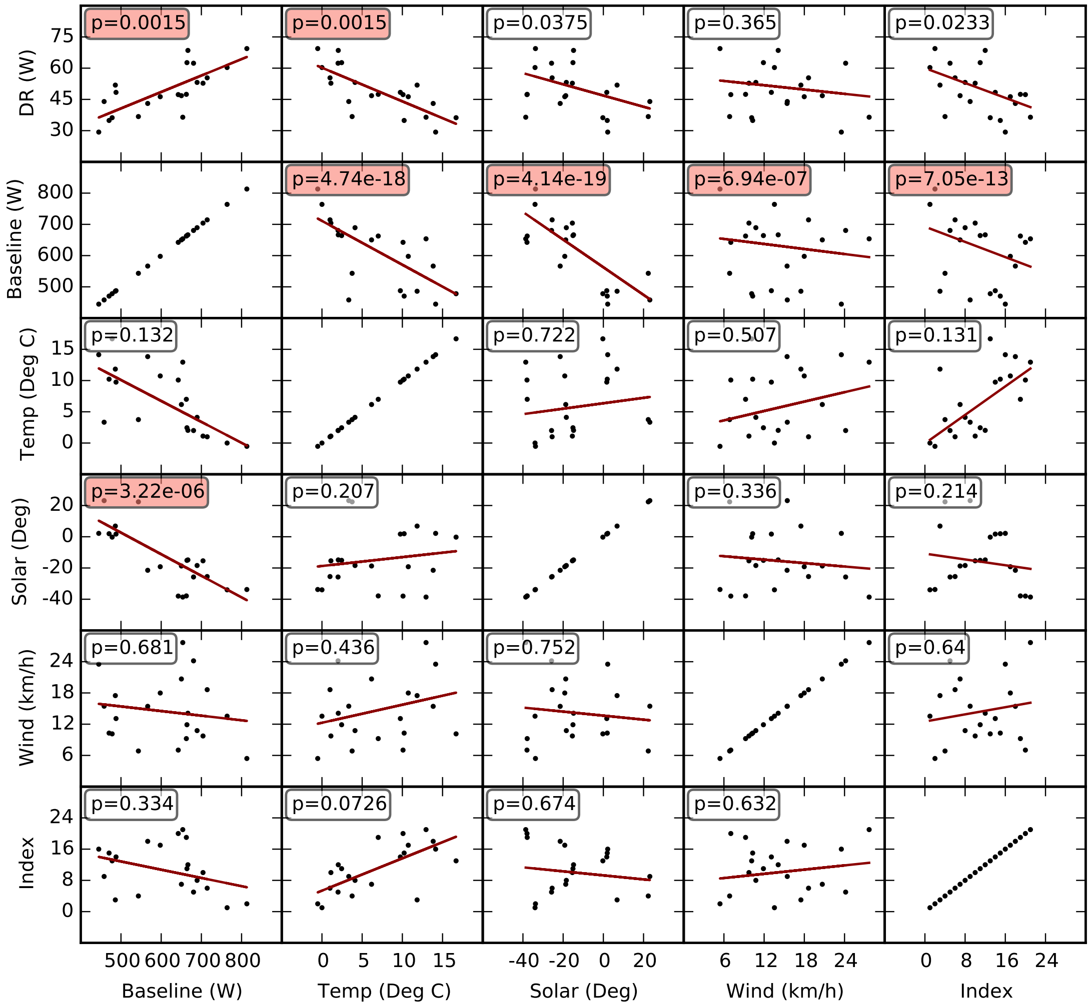

Figure 8.2: Best fit trend lines for CM event DR during the high price period. Trends are examined for potential predictor variables: baseline demand, mean temperature, solar elevation angle, wind speed and chronological event index. Indicated p-values are for the gradient parameters only. Those lower than 0.01 are highlighted as being potentially significant. Temperature and wind data from [99], solar data from [100].

The top row of Fig. 8.2 shows the five listed variables plotted against the demand reductions of the CM events. As measurement errors were estimated independently, weighted least square regression was used to select the best-fit line through the data. The sample weights were set to $1/\epsilon_i^2$, the best linear unbiased estimator, where ϵ_i is the SE of measurement i. The remaining rows show dependencies between pairs of possible explanatory variables; the lines were fitted using ordinary (unweighted) least squares. The p-value for the gradient parameter is shown above each plot. This indicates the computed probability of the null hypothesis that the non-zero value of the fitted parameter was observed as a result of random noise.

8.1.3 A response model for CM events

We proceeded to identify models for the observed responses to CM events. We considered the class of all linear models with a single explanatory variable, either with or without intercept (a constant term). The weighted least squares method was used to fit the linear models to the data using specified SEs for the individual observations. Fitting was performed using a custom made software based on the standard weighted least-squares regression (WLS) fitting algorithm [74]. Suitability of the resulting models was evaluated according to the following criteria:

- *Goodness of fit*: preference was given to models that provided a good fit to the data. This was evaluated using the fraction of sum-of-squares of DR values that was explained by the model. This is similar to the *coefficient of determination* (R^2), though it does not discard the constant (mean response) contribution of the model.
- *Significance of included terms*: p-values were computed for the significance of each of the parameters, quantifying the probability that a non-zero value was obtained by chance. The p-values thus indicate whether the model provides a significantly better fit with the related term than without it. P-values of 0.01 or lower for all parameters, including the constant offset, were required to be classified as a significant model, in order to avoid overfitting.
- *Compatibility of residuals with input errors*: the SE estimates for the individual events provide an independent estimate of the quality of fit that is obtained. For each linear model, standardised residuals were computed. For an accurate model, these residuals should be compatible with a standard normal distribution with mean 0 and variance 1.

We emphasise that this procedure identifies empirical models that provide a sufficient description of the features present in the data. Although they might suggest an underlying mechanism for DR, these models are not postulated with reference to the causes of DR. The *true* population models are almost certainly more complex than the models identified in this manner, but the number of CM events and the accuracy of the associated measurements limits the ability to identify more complex dependencies.

Two simple models were selected according to the criteria described above:

$$R_{\mathrm{CM}}^{\mathrm{demand}} = 0.080 \times [\text{baseline demand}] + [\text{random variation}] \tag{8.4}$$

and

$$R_{\mathrm{CM}}^{\mathrm{temp}} = 59.3 - 1.53 \times [\text{temperature in }^\circ\mathrm{C}] + [\text{random variation}] \tag{8.5}$$

The demand model, Eq. (8.4), can be interpreted as an ability to reduce demand by approximately 8% with respect to the baseline. This simple model accounts for 97.8% of the sum-of-square DR values, the parameter has a p-value of 5×10^{-16}. The normality of the residuals is tested using the "normtest" function in the Python language SciPy library [101]. This test checks if skew and kurtosis of the residuals sample differ significantly from those of a normal distribution using a method proposed by D'Agostino and Pearson [102]. The p-value of this test was 0.83, indicating a high probability that the residuals were normal.

The alternative temperature-based model, Eq. (8.5), suggests an ability to reduce demand that decreases with temperature. This model accounts for 98.0% of the sum-of-square DR values, the maximum p-value of its parameters is 0.001, and the normality test for the standardised residuals results in a p-value of 0.60 [101].

Clearly, both models are consistent with the data. Furthermore, the ability to use either baseline demand or temperature as a dependent variable is rooted in the correlation of baseline load and temperature (see Fig. 8.2). We note that the use of other single predictors resulted in less accurate models, but this does not imply that these factors do not affect DR. The negative correlation between DR and the event index in Fig. 8.2 suggests a reduction in responsiveness over time, but the evidence from the trial is not sufficient to disentangle this from other temporal effects, such as the dominant temperature changes.

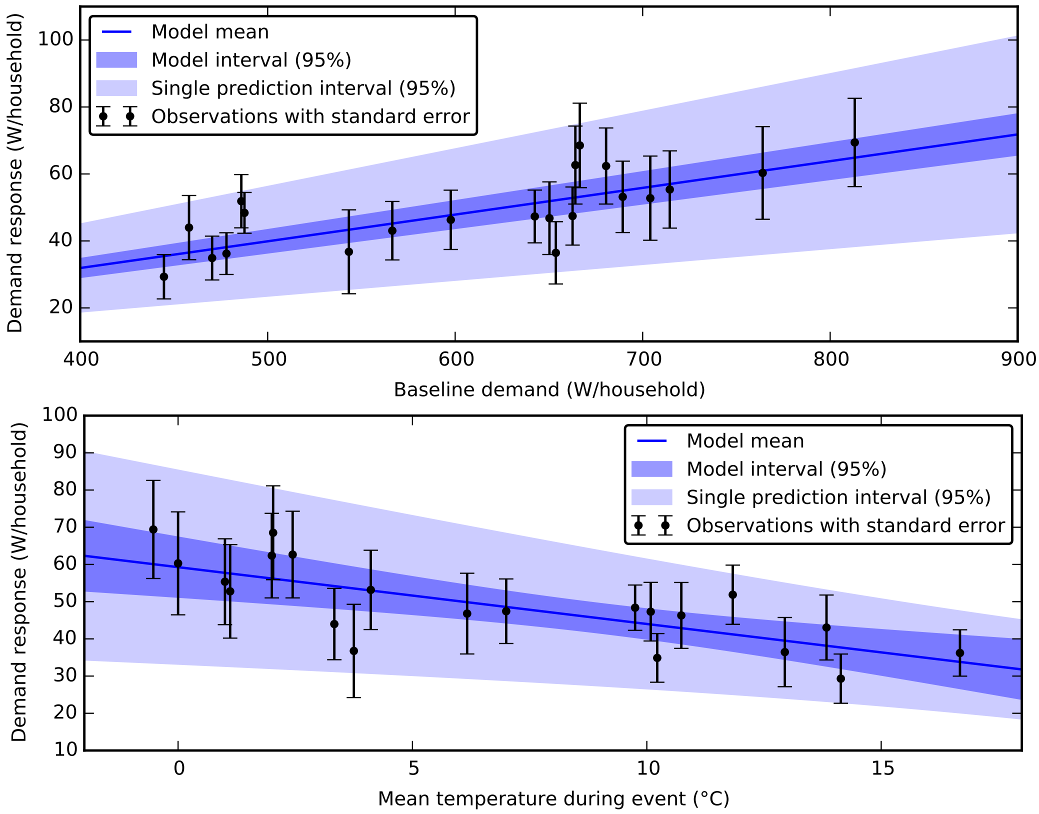

Figure 8.3: Observed DR for CM events (black dots; SEs indicated), alongside fitted empirical models. Demand-proportional (top) and temperature-linear (bottom) models are shown. Dark green bands indicate 95% mean prediction bands (range of models); light green bands indicate 95% single prediction intervals.

Figure 8.3 shows the measured DR values plotted against the baseline demand (top) and average outside temperature during the high-price period (bottom). Standard errors are indicated using vertical lines. The dark blue bands indicate 95% mean prediction intervals for the simple DR models, Eq. (8.4) and Eq. (8.5), derived from confidence intervals on their parameters. These reflect the range of likely models given the observed data. For the demand-based model, Eq. (8.4), (top), the 95% confidence interval corresponds to a load reduction of 7.3%–8.7% of baseline demand.

A question of considerable importance is how this model may be used to predict the magnitude of responses for future events. Note that at this stage we restrict ourselves to predicting future events of the same dToU population, i.e. a hypothetical continuation of the Low Carbon London (LCL) trial. There are two distinct sources of uncertainty associated with future observations. The first contribution is the model uncertainty, represented by the darker blue shaded areas (model mean) in Fig. 8.3. In addition, there is a second contribution related to the households' realised performance compared to their respective baselines (analogous to a measurement error). The two contributions are independent and both are assumed to be normally distributed. The total variance is therefore equal to the sum of variances associated with each contribution:

$$\sigma_{\text{prediction}} = \sqrt{\sigma^2_{\text{model}} + \sigma^2_{\text{sampling}}} \tag{8.6}$$

For each of the linear DR models, a model of the same type was fitted to the measured SEs, resulting in linear noise models $\sigma^{\text{demand}}_{\text{sampling}}(\text{demand})$ and $\sigma^{\text{temp}}_{\text{sampling}}(\text{temp})$. The noise is assumed to be normally distributed according to the fitted standard deviation and 95% confidence intervals for single event predictions were computed by combining both sources of variance. The resulting intervals are indicated by the lighter shaded areas (single prediction interval) in Fig. 8.3. In the case of the proportional demand model, the prediction interval includes load reductions of 4.7–11.2%. The temperature-linear model results in fluctuations of similar magnitude, but with less compact expressions.

8.2 Contribution to network capacity

The analysis up to this point (Section 8.1 and previously in Chapter 7) has considered the DR observed within the LCL trials and determined what information can be extracted regarding the behaviour of the households in the dToU group. In this section an attempt is made to extrapolate these findings to future constraint management scenarios, where the DNO arranges for a high-price signal to be broadcast in order to alleviate network constraints.

In this section, the probabilistic contribution of residential DR to network capacity is defined, and its value is estimated from the dToU trial data. The computed contributions are compared with the naive estimate of the mean DR. For this analysis we shall assume that the measured consumption levels of the dToU group are representative of those of the population as a whole, and that a sample of N households is a selected randomly from the dToU population. Computations will be performed on a per-event basis.

The selection of N random households from the dToU population is appropriate to illustrate the overall range of responses that may be encountered. Conceptually, this reflects a situation where the households are unknown. However, in a situation where dToU signals are regularly used for DNO constraint management, the DNO will be in a position to learn about the response of households connected to specific substations or feeders. This knowledge should then be used to constructed site-specific response profiles, which reduce the magnitude of uncertainty for future events.

8.2.1 Definition of capacity contribution

In probabilistic terms, the required network capacity C may be defined as the capacity that is needed in order to satisfy the expected maximum demand plus a safety margin to cover random load fluctuations with a stated level of confidence (i.e. after-diversity maximum demand). The capacity contribution R of DR is then defined as the change in required network capacity that results from the use of the dToU signal:

$$R = C - C_{\mathrm{CM}} \tag{8.7}$$

where C_{CM} is the required physical network capacity when the dToU signal is used. It is defined such that the probability of reaching the capacity constraint is kept constant with respect to the reference scenario without DR. The value of C_{CM} is thus implicitly defined by the probability equality:

$$P(D_{CM} \geq C_{CM}) = P(D \geq C) \tag{8.8}$$

Here D is the total demand in the absence of a price event; D_{CM} is the total demand if there is a CM event.

For the purpose of this analysis, the total demand D is considered to be the sum of N randomly selected households. When N is sufficiently large, the central limit theorem applies, which allows us to describe the total (or mean) demand distribution of a group of households with the mean and standard deviation of the raw demand measurements. The dToU signal may affect both these distribution parameters. We have already seen, in Section 8.1, that the mean demand is typically reduced by 7% to 9% as a result of a CM event. In a plot of the probability density function (PDF) of total demand, the curve is shifted to the left by this amount. We will see that, in general, the CM event also reduces the variance, so that a narrower, more concentrated distribution will result.

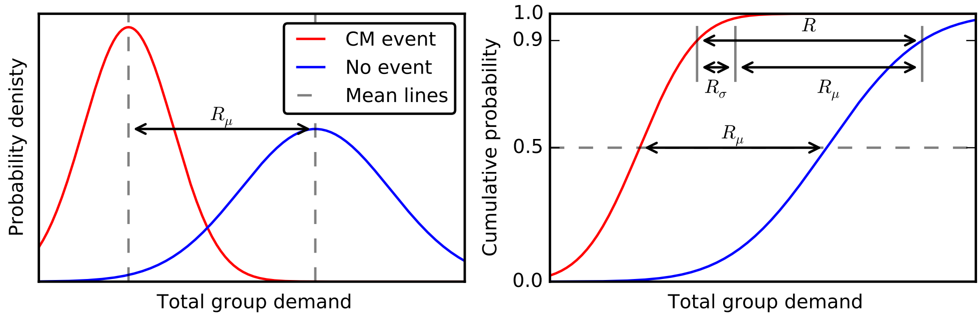

Figure 8.4: Illustrative probability density curves for demand during a CM event and, the counterfactual, what it would have been had there been no event.

These changes are illustrated in Fig. 8.4 (left), where R_μ indicates the difference in mean as a result of the CM event. For ease of annotating the graphic, let us imagine the normal capacity limit is enforced to a probability of 0.9. The network capacity contribution is therefore the difference between the cumulative density functions (CDFs) when both are equal to 0.9. This is illustrated in Fig. 8.4 (right), where R indicates the total network capacity contribution of the CM event. We may think of the capacity contribution (R) of DR as comprising two components: mean shift in demand, which we shall call the mean response R_μ, and change in the confidence interval on this mean which we shall call the variance response R_σ:

$$R = R_\mu + R_\sigma \tag{8.9}$$

The variance response ($R_\sigma = R - R_\mu$) is a result of a decreased dispersion of the group demand relative to the counterfactual no-event situation. This manifests as a steeper CDF curve, as illustrated in Fig. 8.4 (right). Conversely, if the CM event had the opposite effect, increasing the variance of the demand distribution, then the variance response would detract from the capacity contribution of DR.

The mean response (R_μ) is, in the context of this discussion, considered to be a fixed quantity. However, the variance response will be enhanced under the following conditions:

- If the number of households in the CM event is fewer: Because uncertainty in the magnitude of the group demand is proportional to $1/\sqrt{N}$, a reduction in uncertainty (decrease in variance) will have a proportionally greater effect at lower N. As N increases, the total capacity contribution (R) tends towards the mean response (R_μ).
- If the certainty that the capacity limit will not be breached is made more stringent: It can be seen in Fig. 8.4 (right) that the greater the certainty (cumulative probability) that the group demand will not exceed a certain limit, the greater the difference in demand between the no-event and CM event CDF curves.

8.2.2 A group baseline standard-deviation model

The analysis approach outlined in the previous section requires the mean and standard deviation of demand for the dToU group, both during the CM event, and for the hypothetical situation in which the CM event did not occur. A per-household baseline demand model was introduced in Section 8.1.1, and the mean of the household baselines establishes a baseline model for the mean. We now introduce a second baseline model to predict the standard deviation of the dToU group demand during the CM event.

The dToU and nonToU groups are drawn from similar but not identical populations. Therefore a basic linear model is proposed to predict the standard deviation of the dToU group S_m in terms of the nonToU group's mean demand $\overline{A}'_m$ and the group's standard deviation S'_m of demand:

$$S_m = \alpha\overline{A}'_m + \beta S'_m + \gamma \tag{8.10}$$

Here m is the index of the 30-minute measurement block; Greek letters α, β and γ are unknown parameters that were determined by the linear regression solver. The model was fitted on all available data for the dToU and nonToU groups for July-December 2012 and the non-event days of 2013. Cross validation showed the model to have a coefficient of determination (R^2) of 0.94.

8.2.3 Effect on demand mean and standard deviation

Using the two baseline models, one for the dToU group demand mean and the other for its standard deviation, we may observe the effect of the CM event high price on these statistics. The difference between the actual and baseline is calculated for both the demand mean and standard deviation, for each CM event. The reduction in each is depicted in the bar chart in Fig. 8.5.

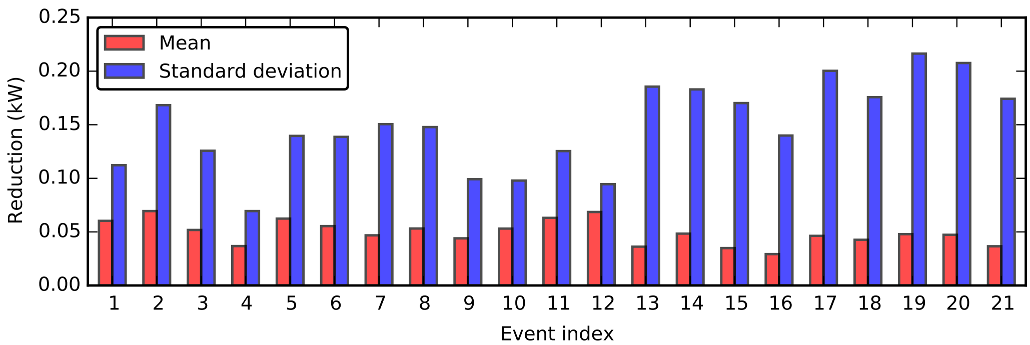

Figure 8.5: Reduction in demand mean and standard deviation as a result of the CM event.

While the reductions in the mean demand (which we call DR) have been discussed in detail, this figure shows that the CM events also have an impact on the standard deviation of demand. In all cases the CM event resulted in a reduced standard deviation of the households' consumption levels. From a network capacity perspective this results in a greater certainty regarding the prediction of future aggregate demand, which may be converted into an effective network capacity contribution using the approach explained in Section 8.2.1.

8.2.4 Mean capacity contribution by number of households

Using the analysis approach described in Section 8.2.1, we calculate the effective network capacity contribution per household, plotted against the number of households in the group. This is repeated for each CM event and shown in Fig. 8.6. Note that these results represent the ensemble of all possible selections of N households (with duplicates) from the group of dToU observations. Analysis for specific customer groups (e.g. those on a particular substation) should be performed using location-specific probability distributions.

On the left of the graph, where the number of households equals 50, the network capacity contribution per household, depicted by the blue lines, substantially exceeds the mean contribution, at between 70 W and 130 W relative to a mean contribution of between 30 W and 70 W. Due to the rapid decline in capacity contribution with group size, a logarithmic scale is used for the number of households in order to increase clarity. For a group size of 1,000 households, the capacity contribution is approaching the mean response contribution—the difference between the actual

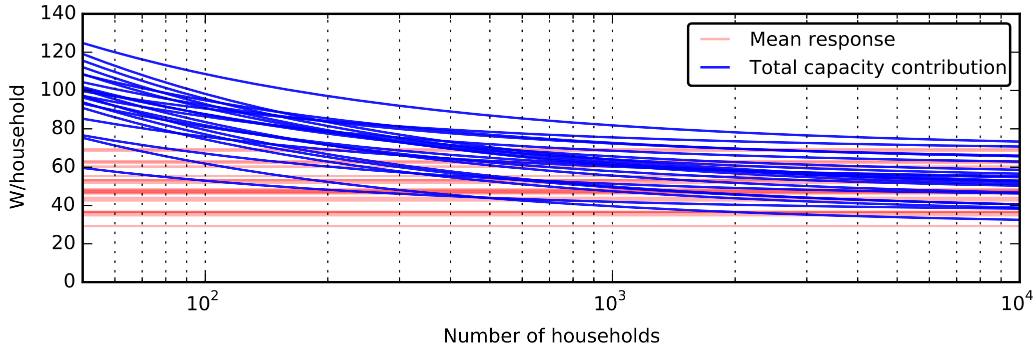

Figure 8.6: Illustrative probability density curves for demand during a CM event and, the counterfactual, what it would have been had there been no event.

and baseline mean demand, depicted in red in the figure. When the group size reaches 10,000 households, little difference can be seen between the total network capacity contribution and the mean response.

8.2.5 Provisioning factor by desired network capacity contribution

By dividing the total mean response of N households by the total network capacity provided by N households, we create a ratio, which we call provisioning factor. This may be thought of as an indicator of the fractional change in the group size necessary to deliver a particular network capacity, as a result of the incorporation of variance response into our DR model. Figure 8.7 shows the provisioning factor plotted against the desired total network capacity for each of the CM events in the trial.

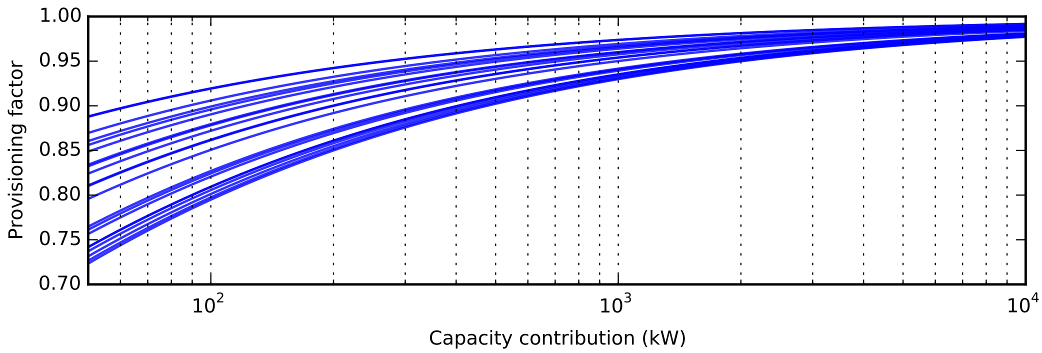

Figure 8.7: Provisioning factor plotted against the desired total network capacity (aggregate mean response) for each of the CM event in the trial.

This figure shows that the reduction in the necessary group size when variance response is considered can be significant when small network capacity contributions are required. For example, if a network capacity contribution of 50 kW was desired, the group size necessary to deliver this may be as much as 25% smaller than that which would be required if only mean response was considered. These savings are quickly lost as the desired capacity contribution is increased. Much past 1 MW, and the mean response alone is almost sufficient to describe the network capacity contribution per household. To place these numbers in perspective, a capacity contribution of 50 kW corresponds to the mean CM contribution of approximately 1,000 households, but the variance response reduces the number of required households to approximately 800—equivalent to two or more distribution substations. A response on the order of 1 MW may be delivered by the customers connected to a primary substation (approximately 10,000-25,000 in domestic areas).

For smaller contributions there is considerable variation in the provisioning factors between each CM event, without a pronounced pattern. Data from a future large-scale rollout of dToU tariffs may be used to try to identify explanatory variables for these differences. Nevertheless, the LCL trial observations give confidence in the *sign* of the deviation: variance reduction results in a contribution that consistently outperforms the mean (provisioning factor less than one).

The maximum relative magnitude of the variance response is only about 25% for a DR event that involves 50 kW of capacity (aggregate mean response), dropping off to 10% for a 1 MW capacity. This may be compared to the uncertainty in response due to inter-event variation (discussed in Section 8.1). For the demand-proportional model, the prediction interval at 95% confidence was shown to be between a 4.7% and 11.2% demand reduction. Given an expectation value of 8.0%, the model prediction could vary more than 40% in either direction. We may therefore conclude that inter-event variation is by far the biggest contributor to uncertainty in residential DR and, in most cases, variance response may be ignored without material consequence.

8.3 The risk of low price signals to network operations

Up to this point, this chapter has been concerned with analysing the potential of dToU signals that aim to strategically alleviate network constraints on behalf of the DNO. This section regards the alternative and arguably more likely case where prices are set by the suppliers in accordance with day-ahead or real-time market conditions. For example, an excess of available wind energy may result in low prices being broadcast to consumers in order to incentivise demand shifting and so avoid wind curtailment. However, such an intervention may boost demand far above previously anticipated levels and thus pose a risk to network operations.

This section aims to quantify the extent to which demand may be boosted by low prices, using data from the LCL Supply Following dToU trials. The Supply Following (SF) events in this trial were designed to sweep through all times of day with a variety of durations of both high and low price events so as to create a general overview of people's willingness to trade flexibility for savings on their energy bill. These events were randomly distributed throughout the year in a randomised-block design. More salient details are given in Chapter 4.

The network capacity required to supply a collection of residential loads is determined by the largest credible peak in the aggregate load (after-diversity maximum demand). For this reason, the remainder of this section will focus on the analysis of peak load levels, comparing observed peak loads with predicted daily peak loads according to the baseline model. It is tempting to consider only annual load peaks (i.e. 'winter peak' scenarios), but this would ignore operational decisions that cause temporary capacity constraints. For example, maintenance work may be scheduled during summer months when the expected peak load levels are lower, or network flows may be rerouted after faults. We therefore consider the occasions when the peak load exceeds its normal daily level as situations that are potentially relevant to the DNO, because they may affect operational decisions.

8.3.1 Day peak compared to event peak

For each SF low price event in the trial, the settlement block with the highest mean consumption was identified (within the low price interval). This block was designated the "event peak", and the corresponding consumption level was recorded. 95% confidence intervals were constructed using the SEs estimated using the individual household measurements for each event peak.

For each event peak, the corresponding baseline demand peak of that day was determined by averaging the baseline demand over all households for each 30-minute settlement block and selecting the maximum value. This was designated the "day peak". The 95% confidence intervals were constructed using the value of baseline standard deviation model (introduced in Section 8.2.2) at the time of the *day peak*.

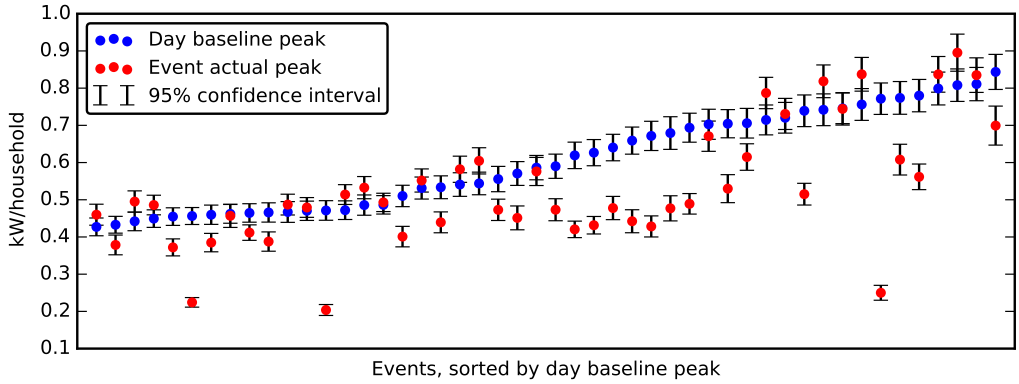

Figure 8.8: Day peak (blue) and event peak (green) demand with respective 95% confidence intervals ($N = 988$).

Figure 8.8 shows the values of the event peak and day peak demand for each low price event in the trial. Events are sorted in order of the magnitude of the day peak. While many of the low price events do not cause demand to exceed the expected peak of the day, it can be seen that a little more than half of the events observed lie close to or exceed the expected day peak.

The distribution of peak breaches shows no obvious sign of depending on the expected peak demand of the day. As the expected peak demand is highly dependent on season of year and day-of-week, this suggests that peak increases may be possible during all season and day types an interpretation which is investigated further in the next section.

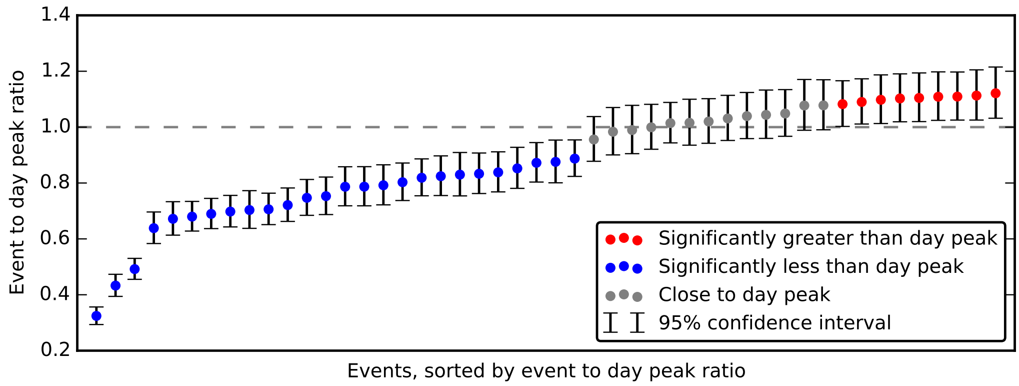

Figure 8.9: Event peak to day peak ratio with 95% confidence intervals ($N = 988$).

To more clearly observe the potential impact of low prices on daily peak loads, we plot the ratio of event peak to day peak consumption, shown sorted by this ratio in Fig. 8.9. Due to the uncertainty in the distribution of the ratios (normality cannot be assumed here), 95% confidence intervals were constructed numerically. For each low price event this was computed as follows: The demand of the day and event peaks were modelled as normally distributed random variables with parameters of the mean and SE of their respective household demand measurements. The ratio of these random variables was sampled 10^6 times and the interval that contained 95% of the sorted values, centred about their mean, was used as the confidence interval.

Ten events (shown in red) resulted in peak load levels that significantly exceeded the baseline peak, and twelve further events (yellow) are not incompatible with an increase in peak load at the 95% confidence level. Furthermore, the measured peak load during the red events exceeded the baseline load by 10% on average. This suggests that the broadcast of low price signals may cause significant increases in peak load, in the order of 10% above the baseline peak level.

8.3.2 Relationship of peak increases with time of day

The lack of correlation between peak increases and the magnitude of the expected day peak (shown in Fig. 8.8 and discussed in the previous section) indicates that peak-increasing events may occur across all seasons, and possibly days of the week. Here, the relationship with type of day (weekday, weekend) and time-of-day is investigated.

We compute the demand to day peak ratio for each actual demand measurement during the low price events. This provides a clearer and more complete overview than the analysis in Fig. 8.9, which only selected the peak level. This approach uses the demand measurements taken during the low price events to form a visual map of the ratio of low price demand to the expected peak demand during the respective day.

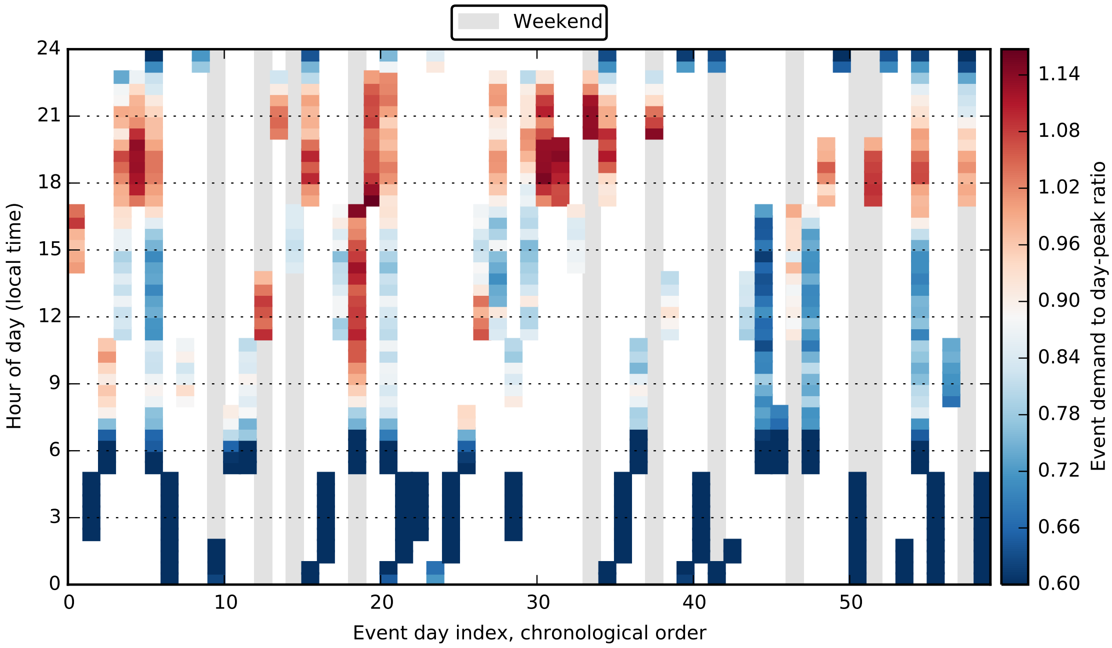

Figure 8.10: Map of the event-peak to day-peak ratio in chronological order, plotted against hour-of-day. Grey backgrounds indicate events that took place during the weekend.

Figure 8.10 shows the image created by the above process. The low price event days are listed along the x-axis in chronological order. For each demand measurement in each low price event, the ratio to day-peak demand is plotted at the corresponding time-of-day on the y-axis. The colour of the measurement point is used to indicate its value. Weekend events are indicated by grey shaded background. Because trial days begin and end at 5am, some events straddled two days. As such, the 48 low price SF events appear across 58 discrete days.

It is immediately evident that most of the peak increases occurred during weekday evenings or weekends. An exception to this trend was the event at index number 26, on 12 July 2013, where the demand on a Friday afternoon exceeded the anticipated daily peak. This event, which may be caused by summer holidays, suggests that care must be taken to anticipate peak load increases even on weekday afternoons. Interestingly, weekday morning events were not observed to pose significant risk of peak increase. As was suggested in the previous section, there does not appear to be a seasonal trend. The low price events were spread, approximately evenly, throughout the year, and yet the same general pattern can be seen in all weekday and weekend events, respectively.

These findings should be viewed in the context of the trial. A number of factors have the potential to greatly change consumer's response to low price signals. Changes to the value of the low price signal, or to influencing factors such as the default or high price, may increase or decrease motivation to respond. Furthermore, increased penetration of home automation technology may enable households to respond at times that they currently find inconvenient, such as while they are out of the home or while they are sleeping.

8.4 Summary and conclusions

This chapter has considered the effects of dynamic time of use tariffs on network constraint management. A statistical analysis of the trial data was performed in order to quantify opportunities and risks from a distribution network operator (DNO) perspective.

Predictability of Constraint Management events. First, the performance of the dynamic Time-of-Use (dToU) trial group over the Constraint Management (CM) trials was analysed with the aim to identify predictive models for the tariff-induced load reduction. Two linear models were identified that match the observed demand response (DR) values: a demand-proportional model and a model where the DR depends linearly on temperature. The simplest model identifies the magnitude of DR as 8.0% of the baseline demand during the peak period (95% confidence range: 7.3%-8.7%). In addition to this descriptive model, a predictive model was derived suggesting that future constraint management events for the same trial population would result in a demand reduction between 4.7% and 11.2% of baseline demand (95% confidence).

It should be noted that the two derived models are heuristic models that relate the observed DR to the most descriptive observables. These do not necessarily imply a causal relation, and relevant factors may be omitted if they are not strictly necessary to explain the data with the observed accuracy. Data from future trials and commercial rollout of dynamic time of use tariffs will provide opportunities to refine these models.

Network capacity contribution. The next step in the analysis was the extrapolation beyond the trial setup, considering an arbitrary number of households of unknown composition. This reflects the situation where the DNO arranges for high-price signals to be broadcast to a set of households in order to alleviate network constraints. To quantify the extent to which DR can alleviate network constraints, the capacity contribution of DR was defined as the change in required network capacity that results from the use of the dToU signal. Here the required capacity is defined in probabilistic terms as the capacity that is needed in order to satisfy the expected maximum demand plus a safety margin to cover random load fluctuations to within a stated prediction interval (i.e. after-diversity maximum demand).

It was shown that the capacity contribution of DR can be decomposed into two components: mean response and variance response. The variance response results from changes in the variance of consumption levels between households. In the case of the constraint management events, the high-price signal was always found to reduce the variance of household consumption levels, even more than suggested by the mean load reduction. This is consistent with trial participants opting to switch off or postpone the use of discretionary large loads, thus reducing the propensity of large load peaks. The variance response thus has the effect of boosting the capacity contribution of DR, as a lower capacity margin is required to anticipate peak load fluctuations.

Capacity contribution of Constraint Management events. To get an impression of the impact that the variance reduction effect has on the capacity contribution, its value was computed across a range of aggregation levels. Furthermore, the consumption distribution of the dToU trial group for each of the events was used as a set of hypothetical collective responses from which the households were sampled, effectively providing a sensitivity regarding response variability. In all cases, the variance contribution boosted the capacity contribution, but by an amount that decreases with the aggregation level. A boost of 25% compared to the mean response was observed at a mean DR capacity contribution of 50 kW, decreasing to 10% at 1 MW and 5% at 10 MW. These are significant figures, but they are outweighed by the observed variability in the mean response itself, with fluctuations of 40% or more around the expected value. Therefore, in most cases, the additional contribution of variance response may be ignored without material consequence.

Risk to the network from low price induced demand spikes. Finally, the focus shifted to potential conflicts between the DNO's local network management aims and the supplier's incentive to respond to wholesale electricity markets. At times of abundant wind power availability, the suppliers may broadcast low prices to consumers in order to incentivise demand shifting. However, the resulting additional demand may boost local demand far above previously anticipated levels and thus interfere with network operations.

The extent to which demand may be boosted by low prices was analysed using data from the Low Carbon London (LCL) Supply Following dToU trials. It was confirmed that there is a considerable risk of increasing the load on distribution networks, with 22 out of 48 low price events achieving maximum loads that are consistent with or higher than the daily peak load, and 10 of those showing load levels that are significantly higher than the baseline (95% confidence). The enhanced load peaks all occurred on weekday evenings and weekend afternoons, but their occurrence does not appear to depend on the magnitude of the expected peak demand of the day. We note that these findings must be taken in the context of the trial: changes to the price signals may increase or reduce motivation to respond, while increased penetration of home automation may make it easier for consumer to respond at hitherto inconvenient times (e.g. sleeping or working hours).

Chapter 9

Metadata analysis

As part of the Low Carbon London (LCL) residential dynamic Time-of-Use (dToU) tariff trial, substantial metadata was collected on participating households. This chapter presents a correlation analysis of these metadata variables. Spearman's rank correlation coefficient was use to analyse the relationships between some 200 variables. The significant correlation coefficients were then used to construct weighted correlation network graphs in order to display and further analyse the relationships. The chapter begins with a description of the data that was analysed, moving on to the details of the analysis approach and a results section where the findings are discussed.

9.1 Data

This section describes the data set to be analysed, including the data cleansing and encoding that was performed in order to create the array of numerical variables necessary for input into the correlation analysis.

9.1.1 Sources

Data sources are divided into two high level categories; *primary* and *metadata*. Primary variables are all those that were calculated using smart meter (SM) consumption data and include demand response (DR), consumption and consumer engagement variables—the analysis of which has been described in previous chapters. All other variables are considered metadata in that they give context to the primary variables that are the focus of this investigation.

Raw metadata was gathered from the below sources:

- CACI's Acorn consumer classification data [87]: This consisted of a socioeconomic classification for each household in the trial. Acorn Group consisted of 17 classes, $\{A, \cdots, Q\}$, which roughly correspond with decreasing household wealth.
- Appliance survey: Before the commencement of the LCL dToU trial, an appliance ownership survey was sent to every household. Survey questions focused on the physical properties of the accommodation and the electric appliances contained within. 1,870 submissions were received from the non-time-of-use (nonToU) group and 990 from the dToU group.
- Attitudes survey: After the completion of the trial, a survey that focused on consumer's attitudes to the dToU tariff was sent to each household in the dToU group. 714 submissions were received in total.

9.1.2 Cleansing and encoding

Administering and collating survey response data was performed by LCL programme partner EDF Energy. High level details of the process are recounted here in order to provide context.

Trial participants had the option to respond to surveys either by returning hand written response forms or using an online web form. In the case of the hand written forms, these were transcribed into a database by hand.

Survey results were made available from the LCL Operational Data Store (ODS) (see Section 4.9 for more details) in raw (as transcribed) string format. Results were stored in one table with a column for each question and a row for each response, keyed to respective participants via a unique and anonymous code.

During data checking it was noted that multiple responses existed for 15 trial participants. This may be due to some responders replying in both paper and online format, or transcription error. The actual cause was unknown but the low number of such cases meant that it was not a major concern for the quality of the overall dataset. So as not to unnecessarily discard data, duplicate survey responses were combined so that, for questions where only one response existed, that response was taken, and for questions where multiple responses existed, the first response was arbitrarily taken.

Our household would prefer to remain on a tariff like Economy Alert if possible.
(Please select one of the following)

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Figure 9.1: Example of a question from the End of Trial survey. “Economy Alert” was the name given to the dToU tariff.

While primary data was already in numerical form, metadata had to first be encoded into a numerical format. Survey answers were encoded into three different numerical types:

- Nominal: Used to convert categorical fields and true/false statements into binary form; 1 if the item was present or true, 0 if not. For example, 1 if the household was in the dToU group and 0 if not.
- Ordinal: Used where there were multiple responses and it was possible to place these responses into a qualitative rank order, but not quantitatively state the extent to which one response was greater or less than another. For example, the degree to which a responder agrees with a given statement in the survey. An example of such a question is shown in Fig. 9.1.
- Interval scale: Used where there were multiple responses and the responses were quantifiable in such a way that the relative differences between one response and another were definable. For example, the number of people in a household.

Not all survey responses had a direct one-to-one mapping with the resulting encoded variables. Some responses were combined when it was believed that too fine detail would result in the variables being under populated with responses. For example, the details of the ages and genders of the household occupants were dealt with on a person by person basis in the survey, but were encoded as the total number of occupants within a chosen age range in one variable, and the total number of occupants of a given gender in another. There were many different ways variables could be combined in such a manner; our choices were based on intuition and preliminary data analysis.

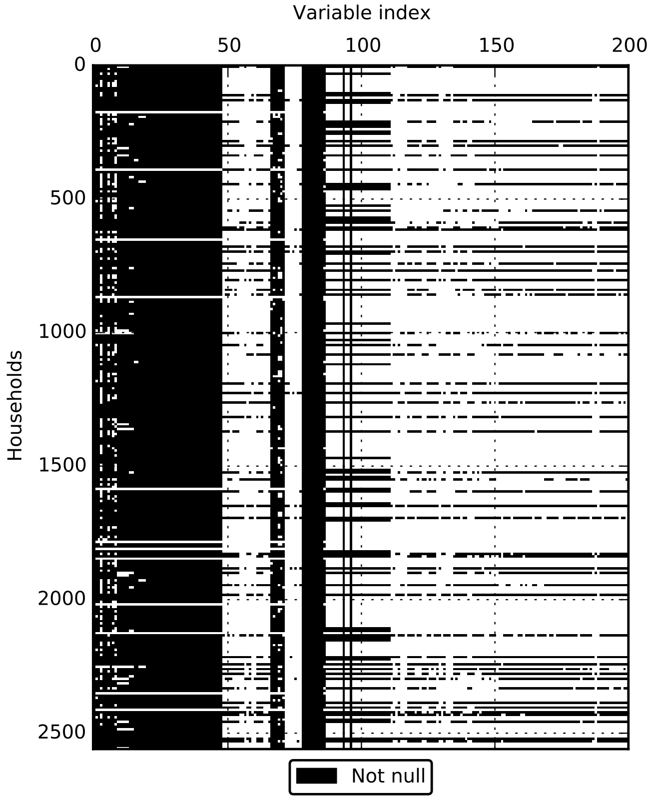

Figure 9.2: Depiction of the sparsity of the design matrix. Variable index corresponds to those listed in Appendix B.1.

Some 200 encoded variables were created by this process. The analysis ready variables were represented by a floating point array, the design matrix, in the shape $2,560 \times 200$. Here, columns were used to represent each variable and each row represented a household in the trial. A full list of the 200 variables, including number of responses for each, is given in Appendix B.1. That the number of rows was significantly less than the total 4,756 households chosen for analysis was due to some households not responding to the surveys. A sense of the sparsity of the design matrix is given by Fig. 9.2. The array was stored using the Python language NumPy [103] *float64* data type. This allowed missing values to be encoded as NaNs (not a number) for efficient computational analysis.

9.2 Analysis

9.2.1 Correlation calculation

Correlation coefficients can be calculated in a number of different ways. Two of the most common methods are "Pearson's product moment" and "Spearman's rank". The principal difference is that the former is a measure of the linear dependence between two variables and the latter a measure of a more general monotonic relationship.

Pearson's correlation coefficient is defined as:

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} \tag{9.1}$$

where X and Y are the variables, Cov is the covariance function and σ_X and σ_Y represent the standard deviation of each variable. Spearman's rank correlation coefficient is calculated in the same way as Pearson's with the addition that both variables are converted to rank indices before computing Pearson's correlation coefficient.

The ordinals used in Spearman's method begin at 1 with an increment of one for each subsequent ordinal. Equal values within variables are assigned the same rank ordinal where that ordinal is equal to the average of their positions in the ascending order of the values. For example, $[1, 2, 2, 3]$ becomes $[1, 2.5, 2.5, 4]$ when converted this way.

The objective of this metadata analysis was to discover any relationships that might exist between pairs of variables. With linear relationships being considered a subset of the more general set of monotonic relationships, Spearman's rank correlation coefficient was the more appropriate in this respect. In addition, as some variables were encoded as ordinals—hence would effectively be equal to Spearman's method even if processed using Pearson's method—and others as interval scales, use of Spearman's rank correlation method ensured a consistent treatment across all variable types.

9.2.2 Variable groups

In order to aid the analysis of the results involving hundreds of variables, some high level variable groups were defined. Each group had a single common theme and one group was assigned to each variable. Grouping was performed manually with judgement used in both selecting the group themes and assigning variables to groups. Table 9.1 lists the defined variable groups. A comprehensive list of all variables, including their assigned groups, can be found in Appendix B.1.

These variable groups were particularly useful for filtering results and in visually representing the variable relationships in the weighted correlation network graphs described later.

9.2.3 Root variables

A number of variables were chosen against which all other variables would be correlation tested. These variables, so called because they formed the root nodes in the later described correlation network analysis, are listed in Table 9.2. The choice of root variables mainly focussed on DR and various stratifications thereof, with engagement rank and annual consumption included as they were shown to be strong determinant DR in chapters Chapter 6 and Chapter 8 respectively. The binary variable Is_dToU was included in order to test whether there were any significant differences between the dToU and nonToU groups. In total, 25 root variables were chosen from the full set of 200 variables listed in Appendix B.1.

The motivation for choosing a small set of root variables, as opposed to performing the correlation analysis on all variables, was related to the statistical significance of results. As will be explained in Section 9.2.4, the stringency of the p-value threshold for significance is dependant on the total number of correlation tests performed. This is because each test increases the probability that a false positive may be returned (known as type I error). As such, it is desirable to limit the number of tests to those that are of interest lest the p-value threshold become so stringent that genuine correlations are rejected as insignificant (known as type II error). While there are workarounds for analysis where large numbers of tests are performed [104], the tradeoff between simplicity of design and the value of learning that may result from a more complex approach were considered unjustified for this analysis.

Table 9.1: List and description of the high level metadata variable groups. The full list of 200 variables is given in Appendix B.1.

Group	Description
Accommodation	Physical parameters of the accommodation
Appliances	Numbers of appliances in the premises
Behaviour_change	Behaviour that has changed as a result of the dToU tariff
Behaviour_inflexible	Behaviour that was not changeable despite the tariff
Behaviour_normal	Behaviour before the trial
Behaviour_timer_use	Use of timers to enable load shifting
People	Ages, genders and numbers of people living in the accommodation
Primary	Variables that were derived from the SM consumption data
Report_high	Self reported ability to shift load away from high price periods, for appliances and times
Report_low	Self reported ability to shift load into low price periods, for appliances and times
Response_helper	Behavioural and appliance aspects that helped in responding to the tariff
Response_limiter	Behavioural and appliance aspects that hindered in responding to the tariff
Trial_impressions	Attitudes and sentiments towards being on the dToU tariff

Table 9.2: List of root variables against which all other variables were tested. This is a subset of the 200 variables listed in Appendix B.1.

Index	Group	Alias	Description
1	Accommodation	Acorn_group	CACI socioeconomic classification
100	Primary	All_default_price	DR signal at default price
101	Primary	All_high_price	DR signal at high price
102	Primary	All_high_price_relative	DR/baseline signal at high price
103	Primary	All_low_price	DR signal at low price
104	Primary	All_low_price_relative	DR/baseline at low price
105	Primary	CM_high_price	DR signal at high price for CM events
106	Primary	Consumption_annual	Total annual consumption
107	Primary	DR_training_index	DR signal calculated over the baseline model training index
108	Primary	Engagement_rank	Non-parametric consumer engagement rank (lower is better)
109	Primary	Is_dToU	Binary 1 if in dToU group, 0 if in nonToU group
110	Primary	SF_high_price	DR signal at high price for SF events
111	Primary	SF_high_wd	DR signal at high price for SF weekday events
112	Primary	SF_high_wd_00_07	DR signal at high price for SF weekdays events between 00h and 07h
113	Primary	SF_high_wd_07_10	DR signal at high price for SF weekdays events between 07h and 10h
114	Primary	SF_high_wd_10_17	DR signal at high price for SF weekdays events between 10h and 17h
115	Primary	SF_high_wd_18_21	DR signal at high price for SF weekdays events between 18h and 21h
116	Primary	SF_high_we	DR signal at high price for SF weekend events
117	Primary	SF_low_price	DR signal at low price for SF events
118	Primary	SF_low_wd	DR signal at low price for SF weekday events
119	Primary	SF_low_wd_00_07	DR signal at low price for SF weekdays events between 00h and 07h
120	Primary	SF_low_wd_07_10	DR signal at low price for SF weekdays events between 07h and 10h
121	Primary	SF_low_wd_10_17	DR signal at low price for SF weekdays events between 10h and 17h
122	Primary	SF_low_wd_18_21	DR signal at low price for SF weekdays events between 18h and 21h
123	Primary	SF_low_we	DR signal at low price for SF weekend events

9.2.4 Threshold p-value

The significance of correlation results was determined by calculating their associated p-values, a form of null-hypothesis test. This section describes the approach taken to determining threshold p-value (the value below which results were considered significant), the Fisher transformation method used to calculate each p-value, and a validation of this method against an alternative non-parametric method using a bootstrap approach.

The significance of the calculated sample correlation coefficients r were measured by their respective p-values p. P-values are an estimate of the probability of measuring a value of r under the assumption of a null hypothesis; that the true population correlation coefficient, ρ, equals zero. An observed value of r was considered significant if $p \leq \alpha$, where α is a chosen threshold value.

For this analysis, α is chosen so that the probability of a single false positive result is less than 0.05, a value which is customary in statistical inference [74]. The Bonferroni inequality [104] is used to determine the threshold p-value.

For a single test, α indicates the threshold probability of receiving a false positive, so $(1-\alpha)$ is the probability of not receiving a false positive. With 200 variables in total and 25 of these chosen for full analysis against all the others, 200×25 correlation tests were conducted. α can therefore be found by solving:

$$(1-\alpha)^{200\times 25} \leq 1 - 0.05$$

Using the approximation that $\log(1-x) \approx x$ when x is small,

$$\alpha \approx \frac{0.05}{200 \times 25} = 10^{-5} \tag{9.2}$$

Correlation coefficients were considered to be equally significant if their p-values were less than 10^{-5}. Correlation coefficients with p-values above the threshold were disregarded as insignificant. The next sections discuss how the p-values were calculated.

9.2.5 Calculating the p-value

Two approaches were looked at for the calculation of correlation coefficient p-values; the Fisher transformation and bootstrapping. Of these, the Fisher transformation was chosen. This section first introduces the Fisher transformation, followed by a justification of its use through comparison with the alternative bootstrap method.

Fisher transformation

The "Fisher transformation" can be used to calculate the p-value of a sample correlation coefficient r. The method is a standard approach and it is used in major commercial statistics packages such as SAS [105].

The Fisher transformation transforms the measured sample correlation coefficient r into Z_r, an approximately normally distributed variable with expectation value $E(Z_r)$ and variance $V(Z_r)$, as defined below:

$$Z_r = \tanh^{-1}(r) = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right) \tag{9.3}$$

$$E(Z_r) = \zeta + \frac{\rho}{2(N-1)} \tag{9.4}$$

$$V(Z_r) = \frac{1}{N-3} \tag{9.5}$$

where

$$\zeta = \tanh^{-1}(\rho) \tag{9.6}$$

ρ is the population correlation coefficient and N indicates the number of samples that r was calculated with. From Z_r, a standard Z-test may be used to determine the p-value of r, which is the probability of the null hypothesis H_0; that the true population correlation coefficient ρ equals zero and therefore the calculated sample correlation coefficient r is the result of chance alone.

As can be seen from the above equations, if $\rho_0 = 0$ then $E(Z_0) = 0$. The *two tailed* p-value, p, is then given by the fractional area in the tails of the normal distribution function $\mathcal{N}$, parameterised with mean equal to 0 and variance equal to $V(Z_r)$, such that:

$$p = 1 - \int_{-Z_r}^{Z_r} \mathcal{N}(Z)\,\mathrm{d}Z \tag{9.7}$$

Comparison with bootstrap

The use of the Fisher transformation in conjunction with a Z-test requires the assumption that the random error is from a bivariate normal distribution. As a sense check on the appropriateness of the Fisher transformation method, a selection of calculated p-values were compared to an alternative approach known as bootstrapping. Bootstrapping is a non-parametric method and therefore make weaker assumptions regarding the distribution of errors. Major discrepancies between p-values calculated by the two different approaches would indicate that one of the methods is a poor approximation of the true significance of results. Conversely, if the two approaches produced similar results, it may be reasonable to assume they are both equally good methods for calculating significance.

For any two given measurement vectors of length N, calculating the p-value of their correlation coefficient r using the bootstrap method involved the following steps:

1. N pairs of measurements were randomly picked, with replacement, from the measurement vectors
2. The correlation coefficient of the resulting vectors is calculated
3. Steps 1 and 2 are repeated 20,000 times so as to build a set of hypothetical correlation coefficients
4. The actual correlation coefficient of the measurement vectors is calculated
5. The percentile of the actual correlation coefficient within the sorted array of hypothetical correlation coefficients is calculated
6. This value is doubled so as to provide an estimate of the two tailed p-value

Using the above described process, the Fisher transform calculated p-values were compared to bootstrap calculated p-values for a number of different variable combinations. For this illustration, Acorn group was chosen as it is one of the few variables that had a value for every household in the trial. This was tested agains all 200 variables in the data set, thus ensuring a representative range of sample numbers and variable types. A full list of variables with associated sample numbers can be found in Appendix B.1.

Figure 9.3 shows the results of this process. It can be seen that both the Fisher transformation and bootstrap methods span a similarly wide range of p-values, as would be expected given the 200 different variables tested.

Though the histogram of the differences between the p-values shows that this difference is generally small—for 90% of the 200 variables tested, the difference between the correlation coefficient p-values was less than 0.007—it is still large in relation to the chosen value of $\alpha (= 10^{-5})$. This is due to limits in the resolution of the bootstrap method, which depends on the number of samples, their distribution and the number of bootstrap iterations performed. Nevertheless, the similarity

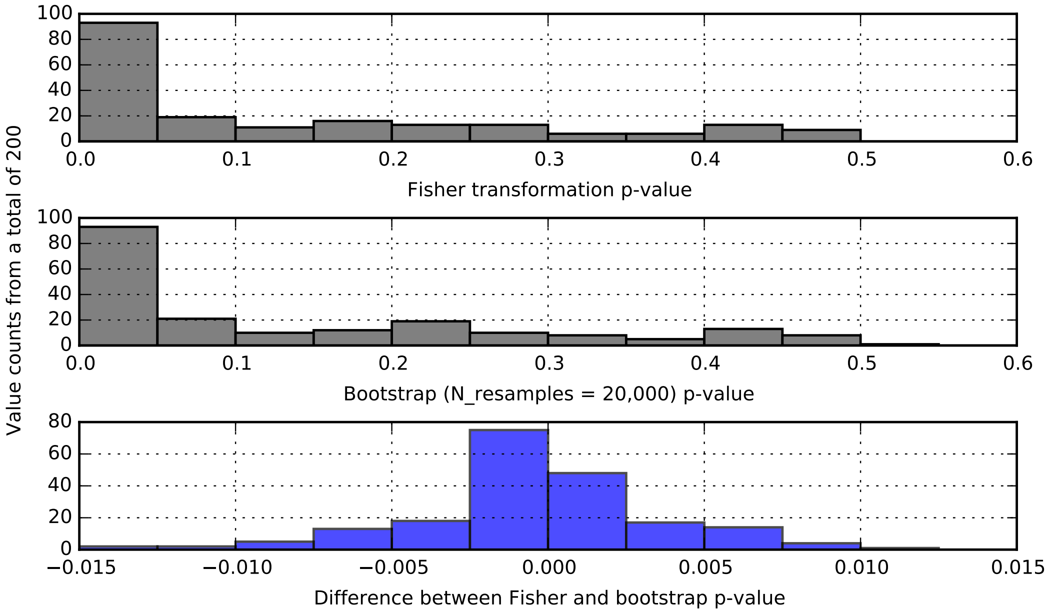

Figure 9.3: Comparison of calculated p-value histograms using the both the Fisher transformation and bootstrap methods.

of the bootstrap method output suggests that the Fisher transformation is an appropriate method for determining the p-value of the correlation coefficients.

Being a purely analytical method, the Fisher transformation is only limited by the accuracy of its founding assumption that the sample vectors follow a bivariate normal distribution. A further advantage of the Fisher transformation over the bootstrap is that, with it being one calculation, it is computationally cheap to perform, while the bootstrap approach requires tens of thousands of repeated calculations for each returned p-value. The Fisher transformation was chosen and all p-values henceforth reported were calculated using this method.

9.2.6 Weighted correlation network analysis

Weighted correlation networks can be applied to most high dimensional data sets and are often used in biology for the analysis of gene expression data [106]. Such networks allow relationships between large numbers of variables to be easily visualised in the form of a graph, and provide a platform for analysis techniques, such as clustering and data reduction, that are based within graph theory.

For this analysis, we used the calculated pairwise correlation coefficients to construct a weighted network in order to visualise the relationships between variables in graph form. Network graphs were constructed so that each node represented a variable and connecting edges between these nodes represented the correlation between the two variables. Edges were not directed and were weighted according to the magnitude of their corresponding Spearman's rank correlation coefficient, normalised so that maximum edge weight in the resulting graph was 1 for the maximum correlation coefficient and 0.1 for the minimum correlation coefficient—for visualisation purposes that will become apparent later in this section, it was necessary to ensure a minimum level of edge weighting. Edges were only added when the associated correlation coefficient p-values were below the threshold level of 10^{-5}. Nodes with no edges (with no significant correlations) were removed from the graph. Self looping edges (self correlations; one by definition) were also removed.

A number of graphs were created from individual subsets of the root variables, shown in Table 9.2, in order to investigate different relationships. For each root variable, correlations were calculated for all other variables in the data set, listed in Appendix B.1. As already described, only variables that had significant correlations with other variables were included in the resulting graphs.

Graph were displayed using the FruchtermanReingold force directed placement algorithm [107] as implemented by the NetworkX [108] Python language library. The FruchtermanReingold placement algorithm takes its inspiration from physical systems, modelling edges as springs and nodes as metal spheres with equal electrical charge so that edges attract according to Hook's law and nodes repel according to the Coulomb force law in two dimensional space. The algorithm is initialised by placing the graph nodes randomly within the graph area, then iteratively updating their positions according to the forces acting upon them. Making a departure from the physical analogy, instead of setting acceleration proportional to the net force on each node, it is the velocity which is instead set. This ensures that, instead of the dynamic equilibria that might be reached in the physical world, the node positions converge towards a static equilibria. The algorithm takes a parameter, k, which moderates both the attractive and repulsive forces in such a way that it defines the optimal distance between nodes.

In this analysis, the value of parameter k was determined manually for each graph through experimentation in order to achieve an aesthetically pleasing layout. For all graphs, 10,000 updates of the FruchtermanReingold node layout algorithm were performed in order to calculate final node positions; a number that was more than enough to ensure the observed clustering structures were broadly replicable over multiple runs.

For graphical representation, edge weights (which indicate strength of correlation) were depicted by line thickness and variable groups were depicted via the shape and colour of the nodes. Nodes were also labeled by variable index number, according to the table given in Appendix B.1, so that the individual variables are identifiable if required.

The relative positions of the nodes joined by edges can then be used to infer the relative strength of the relationships between them; nodes that are closer have a stronger relationship than nodes that are further away. When interpreting the graphs it should be remembered that the attractive (and repulsive) forces acting on nodes are additive, so it is possible for multiple weak relationships pulling in the same direction to have the same effect as a single strong relationship.

9.3 Results

This section presents the results of the correlation analysis and subsequent weighted correlation network analysis. Correlation results were calculated for each of the root variables, listed in Table 9.2, against all other variables, listed in Appendix B.1. As the results of the correlation analysis consisted of large volumes of tabled data, they will not be reproduced in this section save for the single example, Table 9.3 that lists significant correlations with Acorn group. This illustrates the nature of the raw results; for reference, a full set of correlation results are provided in Appendix B.2.

Results presented in this section are sectioned according to the root variables under investigation. Weighted correlation network graph figures presented in this section were built according to the method described in Section 9.2.6, and readers should refer here for the details of its implementation and interpretation.

9.3.1 Experimental group

Other than a low correlation with Acorn group of -0.09, there were no significant correlations with the binary variable Is_dToU. The weak, negative correlation with socio-economic group suggest that there may be a slight skew towards more wealthy households in the dToU group.

Table 9.3: Example results table for root variable "Acorn_group" group (where lower is wealthier) showing all significant correlations. Significance was defined as being when the Spearman correlation coefficient, r, has a two tailed p-value, p, of less than 10^{-5}. N indicates the number of sample pairs used to calculate r.

Index	Group	Alias	N	r	p
29	Appliances	N_dishwashers	2,425	-0.27	3.48E-43
24	Appliances	Lighting_N_halogen	2,425	-0.24	8.96E-34
69	Behaviour_normal	Work_from_home	2,333	-0.23	1.77E-30
19	Accommodation	N_rooms	2,396	-0.22	9.06E-28
41	Appliances	N_printers	2,425	-0.20	2.63E-23
18	Accommodation	N_bedrooms	2,395	-0.17	1.43E-17
106	Primary	Consumption_annual	2,560	-0.17	1.26E-18
11	Accommodation	Heating_interface_set_times	2,256	-0.17	1.55E-15
38	Appliances	N_ovens_electric	2,425	-0.16	3.20E-16
42	Appliances	N_routers	2,425	-0.16	1.99E-15
27	Appliances	Lighting_N_traditional	2,425	-0.16	4.85E-15
36	Appliances	N_laptop_PCs	2,425	-0.13	4.38E-11
40	Appliances	N_portable_electric_heaters	2,425	-0.12	6.35E-10
28	Appliances	N_desktop_PCs	2,425	-0.12	2.86E-09
25	Appliances	Lighting_N_LED	2,425	-0.12	1.05E-08
71	Behaviour_normal	Feeling_about_lifestyle_and_environment	2,388	-0.11	3.69E-08
5	Accommodation	Has_gas_heating	2,402	-0.10	1.04E-06
15	Accommodation	Is_house	2,349	-0.10	1.60E-06
109	Primary	Is_dToU	2,560	-0.09	9.52E-06
14	Accommodation	Is_flat	2,349	0.10	1.60E-06
9	Accommodation	Has_wall_insulation	1,555	0.13	1.48E-07
2	Accommodation	Has_double_glazing	2,253	0.16	9.81E-14
105	Primary	CM_high_price	887	0.16	1.92E-06
132	Report_high	Lighting	610	0.18	6.38E-06
148	Report_low	Kettle	568	0.20	2.16E-06

The overall lack of any other correlations with Is_dToU indicates that both groups contain similar mixes of households with respect to the variables analysed; this provides a post-hoc validation of the objectives of the trial recruitment process (Section 4.4).

9.3.2 Annual consumption

We continue with annual consumption as it is a well understood variable [109, 110] and therefore corroboration of expectations here may serve as a form of validation of the analysis method.

The annual consumption variable can be seen visually represented by an orange hexagon (index 94) in Fig. 9.4. Also, grey circles, grey squares and brown circles represent the accommodation, appliance and people related variables respectively.

As expected, the strongest correlations with annual consumption were with variables representing the numbers of appliances, the physical characteristics of the accommodation (number of rooms etc.) and the numbers of people in the household. In addition to these these variables groups, the annual consumption was seen to be correlated with "Trial impressions" and one "Behaviour normal" variable, represented by the yellow and green circles respectively. At index number 72, the green circle represents the behaviour of working from home, so it is unsurprising that this has a positive correlation with annual consumption. The yellow circle at index number 189 represents the amount of money that people feel they would need to save in order for being on a dToU like tariff to be worthwhile. For this too, it is unsurprising that people with larger annual electricity bills wish for greater savings.

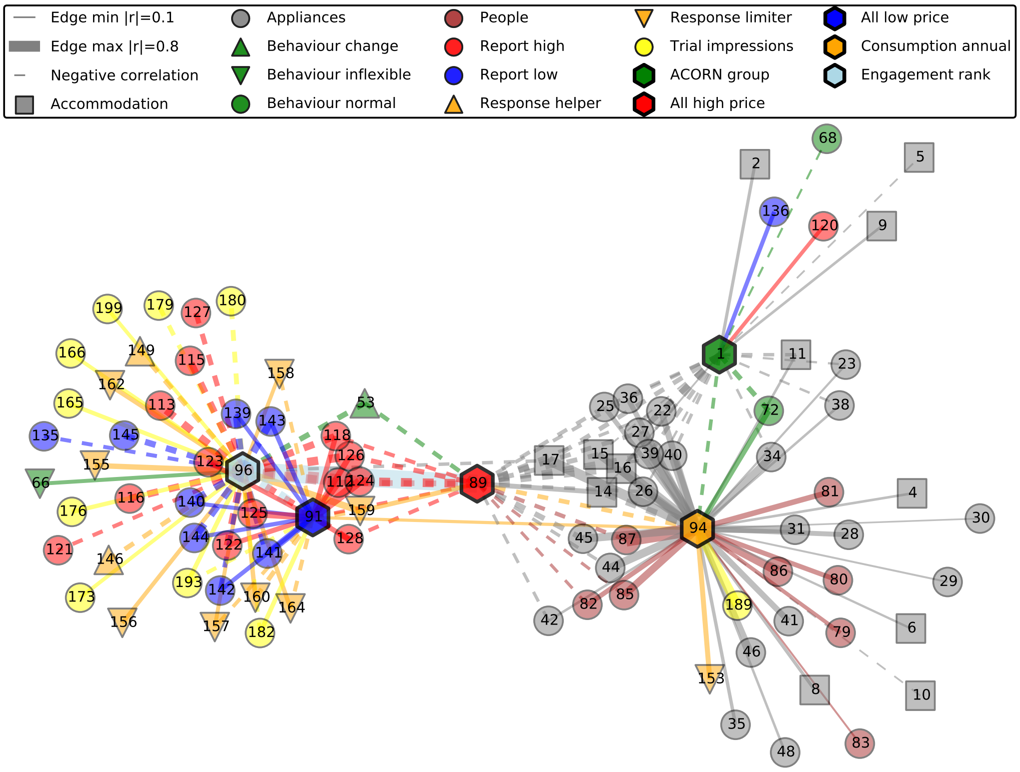

Figure 9.4: Weighted correlation network graph where DR representing nodes are measured in units of absolute DR (kW).

9.3.3 Engagement rank

Engagement rank, first introduced in Chapter 6, is pictured as the light blue hexagon in Fig. 9.4, correlated most strongly with the average high and low price DR metrics (red and dark blue), as expected.

The next strongest correlations were self reporting of response to price signals, indicated by the red and blue circles for high and low price respectively. These variables indicate the self reported use of appliances for shifting and recognition of specific times that were good for shifting. At high price, for example, these variables included behaviours such as being able to respond easily using various appliances; electric oven, electric shower, ironing and washing machine. The corresponding indices, 115, 116, 118, and 125, can be seen close to the engagement rank node. A similar pattern is observed for low price self reported response behaviour.

Response limiters (indicated by orange downwards facing triangles in the graph) of tariff complexity, not knowing when price changes were, and savings being too small featured strongly in association with low trial engagement. Variables indicating fixed appliance usage routines, including one specifically for inflexible usage of the washing machine, also correlated here.

The engagement rank was calculated using a data driven algorithm (Section 6.3) and was designed to detect the extent to which a response was a result of the deliberate choices of the consumer. Another way to measure deliberate choices is through survey responses to behavioural questions. Therefore, that the survey responses pertaining to behaviour are amongst the strongest correlations with the engagement ranking metric, is a positive validation of its functionality and a validation of the value of self-report methods too.

These findings allude to the possibility of conscious behaviour change as people become more experienced at responding to dToU like tariffs. This in turn suggests that the response level from the more engaged household in this trial may be representative of a larger fraction of the population in a future where the populace has gained more experience at responding to such tariffs.

9.3.4 Acorn group

Acorn group was encoded according to the alphabetical ordinals of their group labels so that group $\{A, \cdots, Q\}$ became $\{1, \cdots, 17\}$, where lower numbers loosely correspond to increasing household wealth. The Acorn variable is depicted in Figs. 9.4 and 9.6 as a green hexagon.

Correlations with Acorn group were nearly all appliance ownership or accommodation related, and the number of appliances owned and rooms in the accommodation correlated positively with increasing wealth. The only significant Primary DR metric was that of Constraint Management (CM), where high price response correlated negatively with increasing wealth. As expected, this suggests that lower income are more price sensitive.

It was interesting to note that the number of halogen light bulbs and working from home both correlated with increasing wealth, while energy efficiency measures, such as double glazing and wall insulation, had a negative correlation with increasing wealth.

9.3.5 High price demand response

High price absolute DR is represented by the red hexagon in Fig. 9.4, and also in Fig. 9.5, where the two DR root variables (high and low price DR) are isolated for clarity. Absolute high price DR is calculated over all events in the trial year as defined in Eq. (5.7). It should be noted that, as high price events incentivise a reduction in demand, a negative correlations correspond to increasing response magnitude.

The closest relationships were found to exist with accommodation parameters and appliance number variables (excluding other high price DR metrics). The strongest of these correlations were variables 16 and 17, indicate the number of bedrooms and total number of rooms respectively, and 27 and 45 indicate the number of dishwashers and tumble driers respectively. Three variables relating to the numbers of people were also present, the strongest correlating of these being index 87; the total number of people in the household.

It is concievable that the above mentioned variables are all effectively proxies for overall consumption level and that this is the underlying driver of high price DR magnitude. This hypothesis is investigated later.

Self reported appliance responsiveness showed strong correlations with the ability to shift the use of tumble driers, dishwashers, ironing and the use of the washing machine, here listed in order of strength of relationship. Reported flexibility in when consumers could respond was also seen to be important. In order of strength, these were; weekday afternoons, weekday mornings and Saturdays.

Of the response helper variables, only the use of substitute fuel correlated with high price DR, but this correlated strongly. The substitute fuel was assumed to be gas as the use of smoke emitting fuels has been banned in London [111]. This alludes to synergies between the use of gas for heating and cooking, and reducing stress on the electricity system.

Like low price DR response (discussed next), forgetting price changes was seen to be a response limiter. However, unlike low price DR, this was found to be the only significant response limiter.

9.3.6 Low price demand response

Low price absolute DR is represented by the blue hexagon in Fig. 9.4, and also in Fig. 9.5, where the two DR root variables (high and low price DR) are isolated for clarity. Absolute low price DR is calculated over all events in the trial year as defined in Eq. (5.7). It should be noted that, as

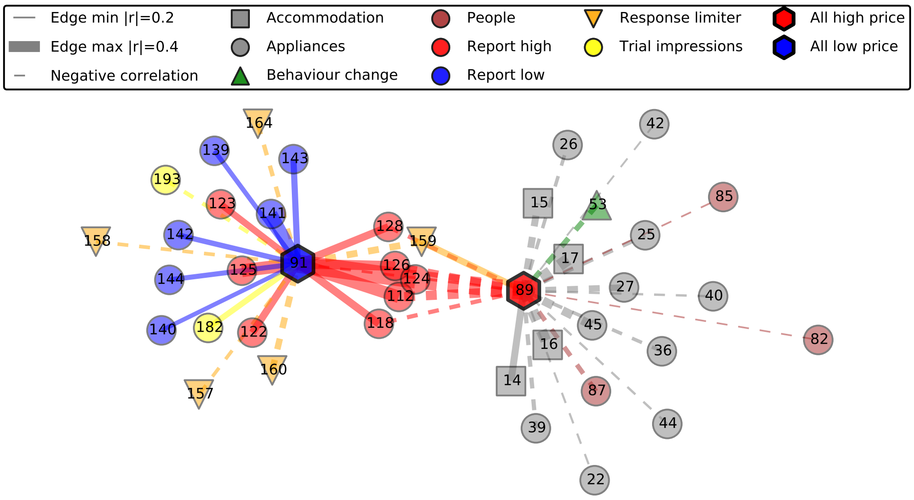

Figure 9.5: Weighted correlation network graph with root nodes representing absolute DR (kW) only. Only variables correlated with high and low price DR are shown.

low price events incentivise an increase in demand, a positive correlations correspond to increasing response magnitude.

In contrast to high price response, there were no appliance ownership correlations and almost no significant accommodation factors; only the number of rooms in the accommodation were significant. This suggests that many households can respond to low price signals if they are motivated.

Low price DR is strongly correlated with self reported response variables for both the low and high price events. This is in contrast to the high price DR variable that only correlates with self reported response at high price. This suggests the hypothesis that good responders to low price signals may be a subset of good responders to high price signals.

Indeed, the correlation coefficient magnitudes for the high price self reported responses were generally greater than for low, at around 0.3 and 0.2 respectively. However, this may be a reflection of a higher price ratio, relative to the default price (Section 4.8.1), at high price than at low price.

The response limiter shared by both high and low price DR was index 159; the variable indicating the extent to which people agreed that forgetting price changes hindered their ability to respond. Low price DR also correlated with four additional response limiters that were not shared with high price DR: index 157 and 158 pertained to the specific users of appliances being fixed (e.g. only some people in the household could/would use the appliance) and to fixed appliances usage routines, while 160 and 164 were related to the difficulty of increasing electricity consumption and the savings from doing so being too small. This tallies with the trial impression variable at index 193, that users may consider it worthwhile to sign up to a future dToU tariff if the price events (low price in this case) were longer and simpler.

A correlation with trial impression variable 182, indicating the extent to which participants talked about the tariff, was also present for low price DR but not for high.

9.3.7 Relative demand response

Chapter 8 offered two simple models for CM event response level. One of these, Eq. (8.4), was a one parameter linear model using the baseline demand variable as an input. The implication of this model is that DR magnitude is proportional to the magnitude of the baseline demand at the time of the event. In the same vein, in Section 9.3.5, it was noted that high price DR correlated strongly with variables relating to appliance ownership and the physical parameters of the accommodation,

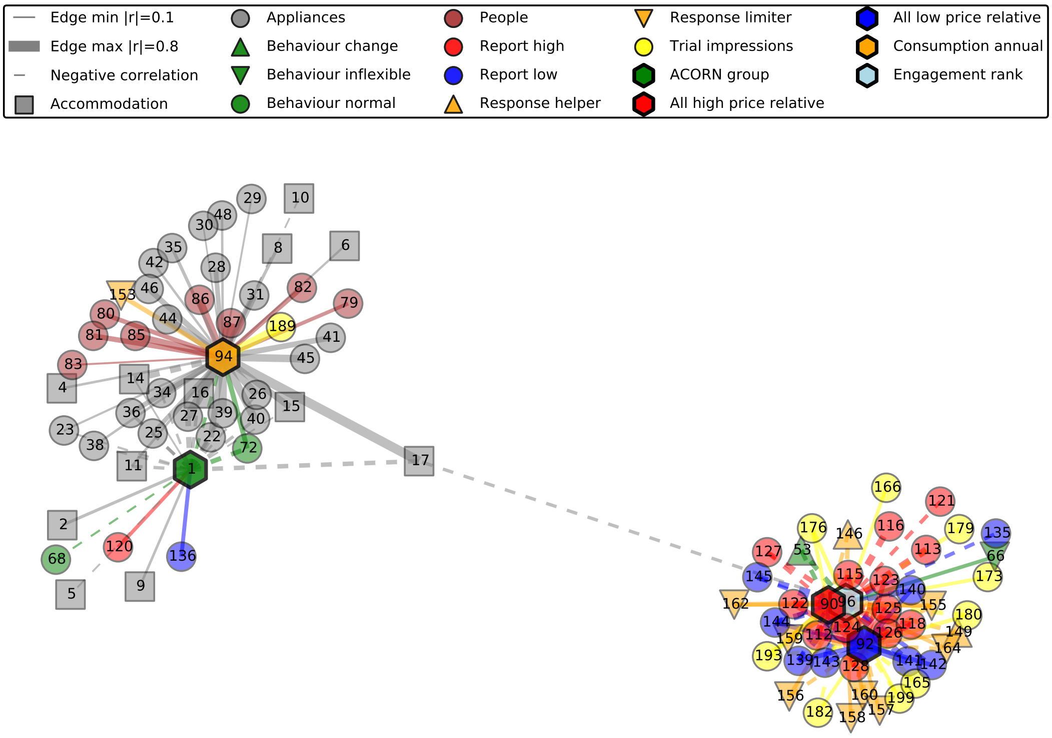

Figure 9.6: Weighted correlation network graph where DR representing nodes are measured in units of DR relative to their annual consumption (kW/kWh).

and that these parameters may be a proxy for overall consumption levels. To test this hypothesis, DR is transformed so as to be measured as ratio with respect to annual consumption.

Relative DR for household h is calculated as:

$$R_{\mathrm{rel},h} = \frac{1}{N_1 \cdot a_h} \sum_{m \in \mathcal{M}_1} R_{m,h} \tag{9.8}$$

where the annual baseline consumption a_h for household h is calculated as

$$a_h = \sum_{m \in \mathcal{M}_{2013}} B_{m,h} \tag{9.9}$$

$R_{m,h}$ and $B_{m,h}$ are the DR and baseline demand variables respectively, calculated for household h at measurement index m, as defined in Eq. (5.5) and Eq. (5.8). $\mathcal{M}_1$ is the set of measurement indices that defined the desired time stratification, and N_1 is the cardinality of that set. For example, for the average relative response at high price over all events, this was the set of measurement indices during which the high price signal was in effect. $\mathcal{M}_{2013}$ is the set of measurement indices in the trial year of 2013.

Figure 9.6 shows weighted correlation network graph with the same root variables as Fig. 9.4, but with the two representing high and low price DR (still represented by the red and blue hexagons respectively) transformed into annual-consumption relative units.

The most immediate striking observation is that the graph is now clearly divided into two almost unrelated clusters; one surrounding the DR, and engagement rank variables, and the other, looser cluster surrounding the annual consumption and Acorn group variables. That the link between the accommodation and appliance variables and high price DR that existed in Fig. 9.4 has dropped away supports the view that accommodation and appliance variables are proxies for annual consumption. This suggests that there is little in the way of objective/external data that can predict annual-consumption relative responsiveness levels.

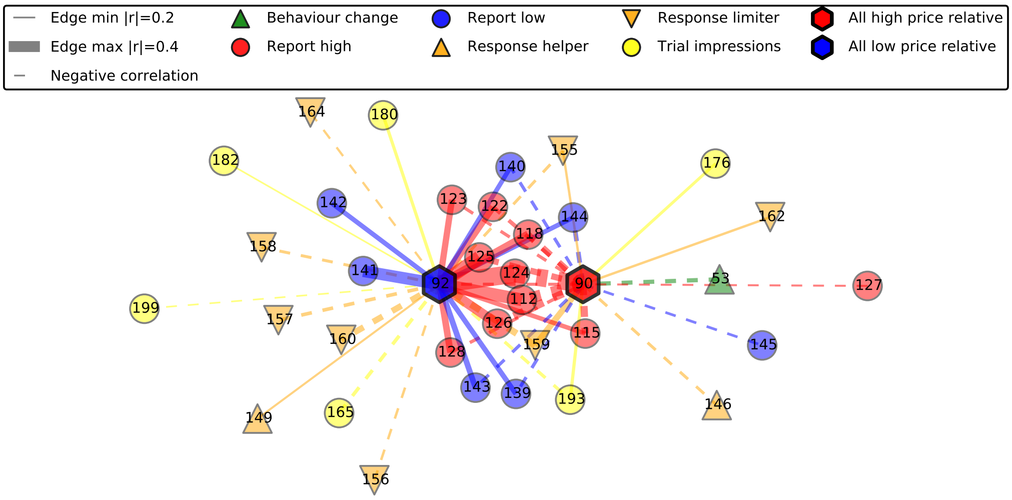

Figure 9.7: Weighted correlation network graph with root nodes representing relative DR (kW/kWh) only.

The only remaining link between the two clusters is a relatively weak correlation between number of rooms in the accommodation, variable 17, and engagement rank. The sign of this correlation suggest that more engaged consumers tended to have more rooms in their accommodation.

9.3.8 Event type

Stratification of DR by Constraint Management (CM) and Supply Following (SF) event types were examined using the absolute measure of DR (measured in kW). Figure 9.8 shows the weighted correlation network graph for root nodes of CM high price response and both low and high price response for Supply Following (SF) events.

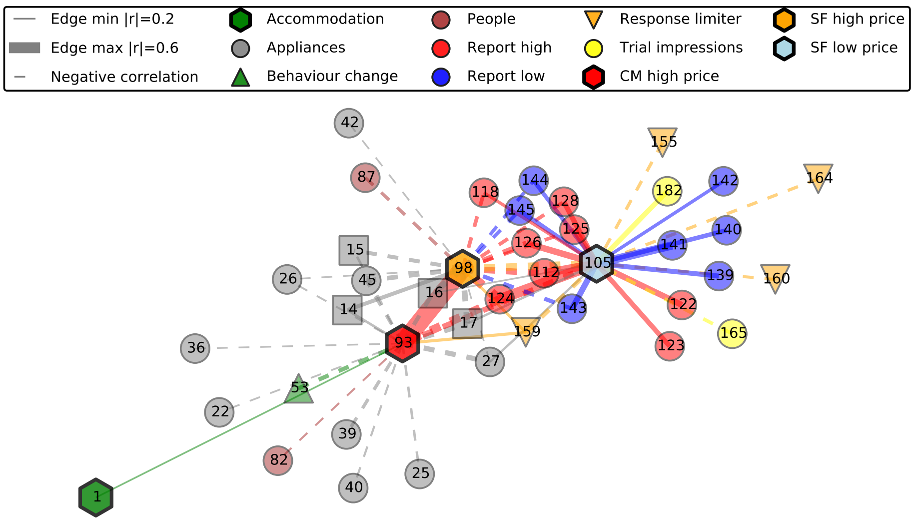

Figure 9.8: Weighted correlation network graph with root nodes representing Constraint Management (CM) and Supply Following (SF) DR (kW) high price CM and SF events, and low price in SF events only.

High price CM event response correlates more strongly with the accommodation and appliance ownership variable groups (grey circles and squares) than high price SF event response. These additional variables represent computer equipment, 39 and 40, the numbers of traditional and halogen bulbs, 25 and 22, and electric ovens, 36. The signs of the correlation and the variables (high price DR is a reduction in demand and therefore negative) indicates that increases in their numbers corresponds with an increase in CM response magnitude. This makes sense in the context of discretionary loads being switched off during a high price event. For example, households with lower efficiency lighting and electric ovens have more demand to reduce.

High price SF response correlates more strongly with self reported responsiveness variables (red and blue circles) than CM type events. This may indicate a greater conscious effort to respond. It may also be due to perceptions: With many more SF event days than CM, it may be that, when estimating self response level, survey responders considered the SF element of the trial to be the most significant element in their own response and therefore answered with respect to this.

Low price SF response showed similar trends as low price DR over all event types.

9.3.9 Time

Stratifications of DR by a selection of times-of-day and days-of-week were investigated. Average DR was calculated according to Eq. (5.7) where the sets of measurement indices ($\mathcal{M}$) were limited to SF event type at the defined times and days. The following time-of-day stratifications were tested for high and low prices:

- Night: 00:00 to 07:00
- Morning: 07:00 to 10:00
- Afternoon: 10:00 to 17:00
- Evening: 18:00 to 21:00

These were represented for high price in variables 100–103 and for low price in variables 107–110. Time stratifications were only calculated for weekdays. Stratifications of average DR over all times of day were made for for weekdays; variables 99 and 106; and weekend days; variables 104 and 111.

In general there were fewer correlations with time stratified DR variables than with the overall DR variables. This is probably due to statistical noise introduced by averaging over fewer events. It was therefore difficult to determine whether differences in correlations between the time stratifications were the result of statistical noise or a significant feature of the stratification itself.

There were almost no significant correlations outside the primary variable group for both high and low price DR at night. The only exception to this was washing machine use (variable 125) for low price events. The number of correlations increased in the morning period with 6 and 8 non-primary correlations for high and low price events respectively. The afternoon period showed the most correlations for both high and low price events, with counts of 15 and 19 respectively. Correlation numbers then reduced again in the evening, with a count of 6 for both low and high price events.

The types of correlations observed generally followed the same trends already noted in Section 9.3.5 and Section 9.3.6: High price DR tended to correlate with the physical variables of accommodation and appliance ownership, while low price DR correlated with the self reported response variables for both low and high price. A full list of significant correlations can be found Appendix B.1.

There were no major differences in the relationship trends between weekday and weekend DR at both low and high price. Though it is worth noting that the self reported response variables for the period in question (weekdays or weekends) were always significantly correlated for that period. This may be taken as validation of the self-report survey method for assessing engagement.

9.4 Summary and conclusions

This chapter presented a correlation analysis of the metadata variables collected in the Low Carbon London (LCL) residential dynamic Time-of-Use (dToU) trial. Spearman's rank correlation coefficient was used to analyse the relationships between metadata and primary variables that were derived from the smart meter (SM) collected consumption data, such as demand response (DR). Some 200 variables were analysed in total. Statistically significant correlations were also used to produce weighted correlation network graphs in order to display and further analyse the relationships. Only correlation coefficients with a p-value of less than 10^{-5} were considered significant. Key findings were:

Annual consumption. Annual consumption was seen to correlate strongly with variables representing physical parameters of the accommodation and appliance ownership. In addition, working from home was found to increase annual consumption.

Engagement rank. Engagement rank was seen to correlate strongly with both high and low price DR metrics and the self reported responsiveness data that was collected via survey. This correlation alludes to the possibility of increasing consumer responsiveness through behaviour change and also acts as a validation of the power of the data driven engagement ranking technique introduced in Chapter 6.

Acorn group. Acorn group correlated strongly with physical accommodation and appliance ownership related variables. Furthermore, increasing wealth was seen to correspond to increased numbers of halogen light bulbs and decreased uptake of energy efficiency measures such as wall insulation. Correlations with DR related variables were weak, with only one correlating significantly; Constraint Management (CM) event response at high price.

Absolute high price DR. High price DR correlated strongest with physical accommodation or appliance ownership related variables. It was suggested that these may be a proxy for the overall consumption level, and that this is the underlying driver of high price DR magnitude. Strong correlations were seen between the absolute DR variable and self reported responsiveness with appliances; tumble drier, dishwasher, ironing and washing machine use, in order of decreasing strength of relationship. The use of a substitute fuel (assumed to be gas as smoke emitting fuels are banned in London) also correlated strongly with enhanced DR. This alludes to synergies between the use of gas for heating and cooking, and the electricity system.

Absolute low price DR. Most striking for low price DR was that, in contrast to high price DR, there were no appliance ownership and almost no physical accommodation related variable correlations. Only the number of rooms in the accommodation was significant, but the relationship here was weak. Instead, strong correlations were found for variables relating to self reported responsiveness. Interestingly, while both were strong, correlation with high price responsiveness variables were even stronger than for low price responsiveness. This suggests the hypothesis that good low price responders may be a subgroup of good high price responders—a point that warrants further investigation. Response limiters also correlated with reduced DR, with sentiments of inflexible appliance use cycles and savings being too small both correlating with reduced DR levels.

Consumption relative DR. Consumption relative DR variables were created to test the premise that DR response levels were strongly dependent on normal consumption levels. In making the switch from high price DR measured absolutely (kW), to an annual-consumption relative metric, all appliance and physical accommodation variable correlations were seen to drop away. This is evidence that DR is strongly related to normal consumption levels and that there may be little in the way of objective/external data that can predict annual-consumption relative responsiveness levels.

DR by event type. Average DR variables for both CM and Supply Following (SF) event types were created to examine differences between the drivers for good response in each type. It was observed that high price CM response correlated much more with physical accommodation and appliance variables than SF event response.

Time stratifications. Stratifications of DR over both times-of-day and weekend/weekday yielded no new findings.

Chapter 10

Summary and conclusions

This section presents an overall summary of the work described in this thesis. To provide context, it begins with a summary of the Low Carbon London (LCL) dynamic Time-of-Use (dToU) trial design and analysis basis, after which key findings form each of the results chapters are presented. The chapter ends with suggestions for the further development of this work.

10.1 Trial design and analysis basis

Two trial groups were formed, one to receive the experimental dToU tariff, the other to act as a reference point for consumption on a standard flat-rate tariff, which we call the non-time-of-use (nonToU) group. Households were recruited onto the trial by programme partner EDF Energy from their existing customer base in the London area. For all households, smart meters (SMs) were installed and consumption was measured at 30 minute intervals. 5,533 households opted into the trial and, of these, 1,119 opted into the experimental dToU tariff group. Throughout the recruitment process, Acorn socioeconomic groups [87] were used to guide recruitment and ensure that both trial groups had socioeconomic class ratios similar to that of London. After trial attrition, data cleansing and validation, consumption data sets for the trial year of 2013 were available from 988 households in the dToU group and 3,768 households in the nonToU group.

The dToU tariff utilised three price bands at 3.99, 11.76 and 67.2 pence per kWh, which are referred to as the low, default and high price. The tariff was designed to examine the potential of dToU tariffs to deliver residential demand response (DR) to the Supplier, where it may contribute to system balancing through Supply Following (SF) actions, and to the distribution network operator (DNO), where it may be used for network Constraint Management (CM). The dToU tariff was hence designed with two high level classes of event:

- CM events were designed to reduce demand at the typical peak load times of the day. Elexon's load Profile Class 1 archetype was used to identify the times of peak load occurrence during the year. As such, CM events typically covered late autumn, winter and early spring seasons, and occurred during weekday mornings and evenings, Sunday afternoons and Saturday evenings. In order to stimulate the maximum possible demand reduction, the peak reducing high price signal was flanked on both sides by low price periods for the remainder of the respective trial day. CM events were scheduled to cover one, two and three consecutive days. 13 CM events were scheduled during the trial, which, when disaggregating consecutive event days, covered 21 separate event days.
- SF events were designed to inform the potential use of dToU for supply balancing and, as such, were designed to provide data on the availability of DR at different times of day, seasons

of year, and for a range of durations. Event durations of 3, 6, 12 and 24 hours were used for both high and low prices, uniformly subtending the day via staggered start times. Each unique event, defined by price, start-time and duration, was repeated 3 times during the trial year. Events were placed throughout the year in a randomised-block design such that noise from time-of-day would approximately cancel upon analysis.

Demand response, measured in kW, was defined as the change in demand induced by either a high or low price signal, relative to a counterfactual baseline demand—that which would have been consumed in the absence of the price signal. A linear regression model was used to compute per-household baseline demand. Dummy variables were included to modulate for temporal factors, one binary variable for each hour of the week and an index variable to account for gradual load growth. By coupling the baseline to the mean consumption level of the nonToU group, it correctly accounts for non-standard days (e.g. bank holidays) and special events. While each baseline demand model reflects only the average behaviour of that household, random deviations from the model will tend to cancel in aggregate operations (e.g. the mean demand of a group of households). This approach therefore coupled the benefits of a mean response model with the ability to arbitrarily stratify household groupings.

10.2 Consumer engagement

Chapter 6 presented three approaches to measuring the level of engagement of residential consumers with a dToU tariff. The first two were aggregate measures and examined engagement from prospectives of the energy supplier or network operator, and the residential consumer respectively. The third method provided a means to quantify the level of engagement of the individual household. This was developed into an engagement ranking index, which may be used for the stratification of the DR response signal.

Consumption shift. From the perspective of the DNO and retail energy supplier, the function of a dToU tariff is to incentivise a reduction in consumption during the high price periods and an increase in consumption during the low price periods, relative to the normal consumption levels that would be present at those times. In this context, the most basic indicator of consumer engagement with the tariff would be the observation of a change in the proportion of energy consumed at each price level in the dToU group. To this end, a simple approach was developed that used the nonToU group consumption as a reference point from which to measure change in consumption at the high and low price levels. Natural differences between average group load profile were accounted for by focusing only on SF events. These events had the desirable features of being approximately uniformly distributed across high and low prices, and the times of day in which they took place; then were also randomly distributed throughout the days of the trial. These features mean that natural differences between group load profiles should average towards zero over increasing time scales. Overall it was found that the dToU group had reduced average consumption during high price periods by 7–9%, and increased average consumption during low price periods by 11–14% (95% confidence ranges).

Change in annual bills. From the perspective of the residential energy consumer, the main motivation for switching to a dToU tariff is the opportunity to make savings on their annual energy bills. Examination of the changes in annual bills can therefore provided a first indicator of the level of consumer engagement with the tariff. Over the trial year of 2013, it was calculated that 85% of households on the dToU tariff received lower annual bills than they would have had on the standard flat tariff of the nonToU group, with the mean reduction in bill being 4.9%.

Engagement rank index. To classify the engagement of individual households with the trials a measure of responsiveness to dToU signals was developed to determine the likelihood that the realised annual bill came about by chance; i.e. that the household had paid no attention to the dToU signal. If this likelihood is very low, it is assumed that the household has actively responded to the signal, whereas a high likelihood is consistent with a lack of engagement. The likelihood measures were used to rank all households according to their perceived responsiveness to dToU signals.

As expected, highly engaged households (low rank index) tend to decrease their consumption in response to high price signals and increase their consumption in response to low price signals, and the magnitude of the response generally decreases with increasing rank index. An important characteristic of the responsiveness ranking is that it does not strictly select for those households with the largest absolute DR, which tend to be the largest consumers of electricity. By measuring the statistical properties of a household's energy consumption the method also identifies consumers that deliver small but consistent DR contributions.

Averaged across all trials and households, the high price signal resulted in a decrease in demand of 0.04 kW/household relative to the default price signal, and the low price signal resulted in an increase of 0.03 kW/household. The most responsive 25% of households outperformed these values by a factor of three, with a decrease of 0.11 kW/household and an increase of 0.08 kW/household, respectively.

The responsiveness ranking also played a key role in the extrapolation of results in subsequent chapters, where it was proposed that highly ranked households may be indicative of future, "business as usual" consumers who are increasingly responsive to dToU signals, either manually or mediated by home automation devices and services. As well as use as an analytical tool, the engagement ranking index may also have operational application as a means of targeting rewards for participating in DR schemes, or to select households that would be well suited for particular tariff types.

10.3 Response stratification

Chapter 7 provided an overview of the primary results of the LCL dToU trial. The DR signal, calculated as described in Section 5.4, was examined over a number of different stratifications of the response signal.

Constraint Management events. The CM events were intrinsically asymmetric with short high price periods surrounded by long low price periods. This was reflected in the observed response numbers, with an average demand reduction of 50 W/household and demand increase of 20 W/household. As expected, the 25% most engaged households delivered a larger response with an average reduction of 150 W/household and average increase of 55 W/household. This illustrates the potential of DR to reduce peaks and enhance utilisation of network assets.

The reduction in load during high price periods was always accompanied by an increase in load during the adjacent low price periods. Extended events that targeted peaks on up to three consecutive days were trialled and no significant difference in measured response was observed between days.

The decrease of demand during peak hours and increase during low priced hours is consistent with load shifting. However, such a signature response was not present in the high price only SF events in which the peak was flanked by default price periods. This suggests that the apparent load shifting may be caused by opportunistic usage of the lower price electricity. Further investigation is required to identify the reason for this difference.

Supply Following events. SF events were used to establish the potential for consumers to respond to dToU signals at different times of the day and throughout the year. Overall, households responded to high price signals with decreases in consumption levels that were much larger during the colder and darker winter months than in the peak of summer. A similar pattern is observed when the measured responses were analysed by the hour of the day. The demand reduction potential reached its maximum magnitude around the morning and evening peaks (on weekdays). The most engaged quarter of households achieved a mean demand reduction over 150 W/household during these periods, compared to 50 W/household for the average household. The strong correlation between demand reduction potential and absolute demand levels is a positive finding for the CM use case, as the reduction potential during peak demand periods will be higher than suggested by average response numbers.

Households responded to low price signals by increasing their demand levels. This increase was seen to be fairly constant during the waking hours of the day, at a level of 50 W/household across all households and exceeding 150 W/household for the most engaged households. During the night time even the best responders did not achieve an increase of 50 W/household. However, the ability of households to increase power consumption was only very slightly affected by the time of year. During the summer months in particular this led to an asymmetric response to high and low price signals.

These figures suggest an ability of households to assist in supply demand balancing, but this potential is currently limited to waking hours and is significantly larger during winter months. The proliferation of "smart appliances" that can autonomously respond to price signals may provide a more consistent response as human intervention will no longer be needed to activate it.

The dual objectives of Constraint Management and Supply Following may lead to conflicts. For example, an abundance of available wind power or the availability of large amounts of inflexible nuclear plant during low load conditions may result in very low electricity prices. From the system perspective it would be beneficial to use dToU pricing to incentivise customers to increase their consumption levels. However, doing so might cause unanticipated stress on the distribution network. Evidence of such situations was seen during the trials: the 25% most engaged households occasionally responded so strongly to low price signals that a new after diversity demand peak was created (for their subpopulation only). On the other hand it is also possible that the two objectives align leading to synergies between system and network management. This is a common situation when high load conditions coincide with high marginal costs of supply (e.g. during the winter peak).

Socio-economic factors in DR. The responses of the targeted SF trials were analysed against two principal parameters that are known to be strong indicators of energy consumption: household occupancy (1, 2, 3+) and a socio-economic classifier based on the Acorn system. The three socio-economic groups—Affluent, Comfortable and Adversity—can be interpreted as a rough indicator of wealth.

Perhaps surprisingly, the socio-economic class did not have a significant effect on the observed DR for these single events, although results on CM events suggest that households in the Affluent class may respond more strongly to signals that specifically target peak hours. The measured response does depend strongly on occupancy levels, with larger households providing responses of larger magnitude.

Low Carbon London in context. An attempt was made to compare LCL trial results with past time-of-use (ToU) trials that are comparable. Results were compared on two axes; peak price to standard price ratio, and relative reduction in peak price consumption. While the related trials had price ratios in the range 1.4–2.7, LCL had a price ratio of 4.7. This departure from the previous trial cluster, combined with statistically robust sample numbers, meant that the LCL trial contributed significantly to the interpretation of the overall landscape of trial results plotted on these axes. Before LCL the trend might have suggested the possibility of increased response with increased price ratio, however, the addition of the LCL data point makes a relative response look like a better model. A constant response model was fitted to the data points while taking into account the statistical weight of each trail. Using this model, it was estimated that the population peak demand reduction lay in the range 6.8–10.2% at a 95% confidence level.

Though the data was not sufficient to make a conclusive statement regarding the relationship between response level and price ratio, this analysis did place LCL into context against previous United Kingdom (UK) related trial results and informs us that we should not necessarily expect an increase in the peak price ratio to result in enhanced peak price consumption reductions. More data is required in order to investigate this further.

10.4 Reliability and risk

Chapter 8 considered the effects of dynamic time of use tariffs on network constraint management. A statistical analysis of the trial data was performed in order to quantify opportunities and risks from a DNO perspective.

Predictability of Constraint Management events. The performance of the dToU trial group over the CM trials was analysed with the aim to identify predictive models for the tariff-induced load reduction. Two linear models were identified that match the observed DR values: a demand-proportional model and a model where the DR depends linearly on temperature. The simplest model identifies the magnitude of DR as 8.0% of the baseline demand during the peak period (95% confidence range: 7.3%-8.7%). In addition to this descriptive model, a predictive model was derived suggesting that future constraint management events for the same trial population would result in a demand reduction between 4.7% and 11.2% of baseline demand (95% confidence).

It should be noted that the two derived models are heuristic models that relate the observed DR to the most descriptive observables. These do not necessarily imply a causal relation, and relevant factors may be omitted if they are not strictly necessary to explain the data with the observed accuracy. Data from future trials and commercial rollout of dynamic time of use tariffs will provide opportunities to refine these models.

Network capacity contribution. The next step in the analysis was the extrapolation beyond the trial setup, considering an arbitrary number of households of unknown composition. This reflects the situation where the DNO arranges for high-price signals to be broadcast to a set of households in order to alleviate network constraints. To quantify the extent to which DR can alleviate network constraints, the capacity contribution of DR was defined as the change in required network capacity that results from the use of the dToU signal. Here the required capacity is defined in probabilistic terms as the capacity that is needed in order to satisfy the expected maximum demand plus a safety margin to cover random load fluctuations to within a stated prediction interval (i.e. after-diversity maximum demand).

It was shown that the capacity contribution of DR can be decomposed into two components: mean response and variance response. The variance response results from changes in the variance of consumption levels between households. In the case of the constraint management events, the high-price signal was always found to reduce the variance of household consumption levels, even more than suggested by the mean load reduction. This is consistent with trial participants opting to switch off or postpone the use of discretionary large loads, thus reducing the propensity of large load peaks. The variance response thus has the effect of boosting the capacity contribution of DR, as a lower capacity margin is required to anticipate peak load fluctuations.

Capacity contribution of Constraint Management events. To get an impression of the impact that the variance reduction effect has on the capacity contribution, its value was computed across a range of aggregation levels. Furthermore, the consumption distribution of the dToU trial group for each of the events was used as a set of hypothetical collective responses from which the households were sampled, effectively providing a sensitivity regarding response variability. In all cases, the variance contribution boosted the capacity contribution, but by an amount that decreases with the aggregation level. A boost of 25% compared to the mean response was observed at a mean DR capacity contribution of 50 kW, decreasing to 10% at 1 MW and 5% at 10 MW. These are significant figures, but they are outweighed by the observed variability in the mean response itself, with fluctuations of 40% or more around the expected value. Therefore, in most cases, the additional contribution of variance response may be ignored without material consequence.

Risk to the network from low price induced demand spikes. Finally, the focus shifted to potential conflicts between the DNO's local network management aims and the supplier's incentive to respond to wholesale electricity markets. At times of abundant wind power availability, the suppliers may broadcast low prices to consumers in order to incentivise demand shifting. However, the resulting additional demand may boost local demand far above previously anticipated levels and thus interfere with network operations.

The extent to which demand may be boosted by low prices was analysed using data from the LCL Supply Following dToU trials. It was confirmed that there is a considerable risk of increasing the load on distribution networks, with 22 out of 48 low price events achieving maximum loads that are consistent with or higher than the daily peak load, and 10 of those showing load levels that are significantly higher than the baseline (95% confidence). The enhanced load peaks all occurred on weekday evenings and weekend afternoons, but their occurrence does not appear to depend on the magnitude of the expected peak demand of the day. We note that these findings must be taken in the context of the trial: changes to the price signals may increase or reduce motivation to respond, while increased penetration of home automation may make it easier for consumer to respond at hitherto inconvenient times (e.g. sleeping or working hours).

10.5 Metadata analysis

This chapter presented a correlation analysis of the metadata variables collected in the LCL residential dToU trial. Spearman's rank correlation coefficient was used to analyse the relationships between metadata and primary variables that were derived from the SM collected consumption data, such as DR. Some 200 variables were analysed in total. Statistically significant correlations were also used to produce weighted correlation network graphs in order to display and further analyse the relationships. Only correlation coefficients with a p-value of less than 10^{-5} were considered significant. Key findings were:

Annual consumption. Annual consumption was seen to correlate strongly with variables representing physical parameters of the accommodation and appliance ownership. In addition, working from home was found to increase annual consumption.

Engagement rank. Engagement rank was seen to correlate strongly with both high and low price DR metrics and the self reported responsiveness data that was collected via survey. This correlation alludes to the possibility of increasing consumer responsiveness through behaviour change and also acts as a validation of the power of the data driven engagement ranking technique introduced in Chapter 6.

Acorn group. Acorn group correlated strongly with physical accommodation and appliance ownership related variables. Furthermore, increasing wealth was seen to correspond to increased numbers of halogen light bulbs and decreased uptake of energy efficiency measures such as wall insulation. Correlations with DR related variables were weak, with only one correlating significantly; CM event response at high price, which was larger for wealthier households.

Absolute high price DR. High price DR correlated strongest with physical accommodation or appliance ownership related variables. It was suggested that these may be a proxy for the overall consumption level, and that this is the underlying driver of high price DR magnitude. Strong correlations were seen between the absolute DR variable and self reported responsiveness with appliances; tumble drier, dishwasher, ironing and washing machine use, in order of decreasing strength of relationship. The use of a substitute fuel (assumed to be gas as smoke emitting fuels are banned in London) also correlated strongly with enhanced DR. This alludes to synergies between the use of gas for heating and cooking, and the electricity system.

Absolute low price DR. Most striking for low price DR was that, in contrast to high price DR, there were no appliance ownership and almost no physical accommodation related variable correlations. Only the number of rooms in the accommodation was significant, but the relationship here was weak. Instead, strong correlations were found for variables relating to self reported responsiveness. Interestingly, while both were strong, correlation with high price responsiveness variables were even stronger than for low price responsiveness. This suggests the hypothesis that good low price responders may be a subgroup of good high price responders—a point that warrants further investigation. Response limiters also correlated with reduced DR, with sentiments of inflexible appliance use cycles and potential savings being too small both correlating with reduced DR levels.

Consumption relative DR. Consumption relative DR variables were created to test the premise that DR response levels were strongly dependent on normal consumption levels. In making the switch from high price DR measured absolutely (kW), to an annual-consumption relative metric, all appliance and physical accommodation variable correlations were seen to drop away. This is evidence that DR is strongly related to normal consumption levels and that there may be little in the way of objective/external data that can predict annual-consumption relative responsiveness levels.

DR by event type. Average DR variables for both CM and SF event types were created to examine differences between the drivers for good response in each type. It was observed that high price CM response correlated much more with physical accommodation and appliance variables than SF event response.

Time stratifications. Stratifications of DR over both times-of-day and weekend/weekday yielded no new findings.

10.6 Future work

The trial data and result of the analysis presented in this report provide a solid basis for future research and development regarding the role of dToU-mediated DR in network operation and planning. This report has focused on the magnitude, predictability, risk to the network, and determinants of dToU mediated residential DR. Potential extensions of these themes are listed below:

Do dToU tariffs inspire an overall reduction in energy consumption? From the perspective of energy conservation, one of the more frequently asked question of dToU tariffs is whether they will inspire consumers to an overall reduction in energy consumption. This may be answered by designing future dToU tariff trials with either sufficient benchmarking periods before the application of experimental tariffs, and/or by ensuring that control groups are sampled from the same populations (i.e. avoid a secondary round of of recruitment from the control group to create the treatment group) so that direct comparison is valid.

Is there evidence of learning/novelty effects with regards to response persistence? This question effectively asks whether the magnitude of the DR signal increases (learning effect) or decreases (novelty effect) over the duration of the trial year. With only one year of data and many external variables that affect response level, disambiguating a change in response magnitude from confounding variables is an obstacle to answering this research question. Future trials may overcome this by running trials for increased durations, ideally covering a minimum of two years.

Are good low price responders a subset of good high price responders? In Chapter 9 it was observed that the low price DR signal correlated with the self reported responsiveness variables, to both low and high price signals, in roughly even numbers. In contrast, the high price DR signal correlated strongly with high price responsiveness variables, but only weakly with low price responsiveness variables. This observation suggests the hypothesis that good responders to low price signals are a subset of good responders to high price signals. This may be tested by constructing two new engagement ranking indices (according to the method described in Section 6.3), one trained on engagement with the low price signal only, index A, and the other trained on the high price signal only, index B. Using these, the mean ranks of a given fraction of the best responding households in index A, may be compared to the mean rank of these same households in index B, and *vice versa*. If the mean rank of the best responding fraction of households in index A have a significantly lower mean rank in index B, but this same fraction of the best responding households of index B does not have a significantly lower mean rank in index A, then we may conclude that good low price responders (index A) are a subset of good high price responders (index B).

Is load shifting really occurring? For CM events, as reported in Chapter 7, decrease of demand during peak hours and increase during low priced hours is consistent with load shifting. However, such a signature response was not present in the high price only SF events in which the peak was flanked by default price periods. This suggests that the apparent load shifting may be caused by opportunistic usage of the lower price electricity. This question is therefore one of disambiguating the motivations and/or appliances used to respond to price signals. Further understanding may be possible via a deeper analysis of the LCL survey responses, with a focus on those that were designed to determine if consumers found using specific appliances to respond easy or difficult. Alternatively, a more direct approach for use in a future trial might be the use of appliance sub-metering, or high resolution metering combined with load-disaggregation algorithms.

Increased diversity of price level and notification lead times. The Supply Following trials have established that dToU signals can be used to increase or decrease average household demand by varying amounts depending on the timing and duration of the price signal. Future experiments could refine the results by introducing more price points to determine price elasticity in a more granular fashion than is possible with ‘high’ and ‘low’ tariffs used here. If the technical and regulatory setup permits, shorter lead times than the current day-ahead signal may also be investigated, which would permit more accurate matching of demand with forecast renewable generation patterns.

Commercially realistic dToU tariff trials. The value of DR to both the supplier and the DNO is largely determined by the predictability of response to price signals. The development of better response predictors is, to a large extent, driven by the availability of historic response data (with which new models may be trained), hence the value of a given dToU tariff offering is strongly dependent on the volume of data available on its real application. This feature of dToU tariffs may create a standoff between suppliers as they may fear that the first to offer a dToU tariff would have to accept a loss until sufficient data was collected. In contrast, subsequent dToU tariff offerings may be able to free-ride from the learning of the loss making first. This standoff may be broken if a sufficiently large publicly available data set were available, and indeed, if the commercial viability of residential DR had been demonstrated. LCL has made a start here, providing the first publicly available dToU data set for the UK, but more is needed. The LCL dToU tariff was developed primarily to collect data for research, without consideration of commercial viability. Future residential DR trials should therefore focus on dToU tariffs that are more realistic representations of potential future “business as usual” offerings.

Incorporating DR response models into existing power system models. Development of DR predictor models and their integration into existing generator dispatch and network models may help to determine both the value of DR and the situations for which it is most useful. Potential risks to the system may be explored with the same approach.

Bottom-up load modelling of future consumers. By definition the LCL dToU trial has measured the response of *today’s* households outfitted with current appliances. Because network planning procedures and security standards have long-term implications they should anticipate the response of future households. In this report a start was made by identifying the contribution of the most responsive households, but this should be complemented by a bottom-up approach to modelling the flexibility of future consumers and responsive appliances.

References

[1] UK Parliament. *Climate Change Act 2008*. London, Nov. 2008. URL: www.legislation.gov.uk.

[2] Department of Energy and Climate. *UK Renewable Energy Roadmap*. White Paper. London, 2011. URL: www.decc.gov.uk.

[3] Department for Energy and Climate Change. *2050 Pathways Analysis*. London, 2010. URL: www.decc.gov.uk.

[4] National Grid. *Operating the Electricity Transmission Networks in 2020*. National Grid, June 2011. URL: www.nationalgrid.com.

[5] G. Strbac, C. K. Gan, M. Aunedi, V. Stanojevic, P. Djapic, J. Dejvises, P. Mancarella, A. Hawkes, and D. Pudjianto. *Benefits of advanced smart metering for demand response based control of distribution networks*. London: Imperial College London, 2010. URL: www.cts.cv.ic.ac.uk.

[6] G. Strbac, M. Aunedi, D. Pudjianto, P. Djapic, S. Gammons, and R. Druce. *Understanding the Balancing Challenge*. London: Imperial College London and NERA Economic Consulting, 2012. URL: www.gov.uk.

[7] Ofgem. *Smart metering implementation programme: Response to prospectus consultation*. White Paper. London, UK: Department of Energy and Climate Change, 2011. URL: www.gov.uk.

[8] Department of Energy and Climate Change. *Smart Metering Implementation Programme: Smart Metering Equipment Technical Specifications Version 2*. London, 2013. URL: www.decc.gov.uk.

[9] D. S. Kirschen. "Demand-side view of electricity markets". In: *IEEE Transactions on Power Systems* 18.2 (2003), pp. 520–527. ISSN: 08858950. DOI: 10.1109/TPWRS.2003.810692. URL: ieeexplore.ieee.org.

[10] D. Kirschen and G. Strbac. *Fundamentals of power system economics*. Reprinted with corr. 2005. Chichester: Wiley, 2005. ISBN: 9780470845721.

[11] US Department of Energy. *Benefits of demand response in electricity markets and recommendations for achieving them*. 2006. URL: www.energy.gov.

[12] D. Hurley, P. Peterson, and M. Whited. *Demand Response as a Power System Resource*. Synapse Energy Economics, 2013. URL: www.synapse-energy.com.

[13] M. L. Goldberg and G. K. Agnew. *Measurement and Verification for Demand Response*. Department of Energy, USA, 2013. URL: emp.lbl.gov.

[14] EnerNOC. *The demand response baseline*. 2011. URL: www.enernoc.com.

[15] Association of Edison Illuminating Companies. *Demand response measurement and verification*. 2009. URL: www.smartgrid.gov.

[16] Elexon. *What is the balancing mechanism?* 2015. URL: www.elexon.co.uk.

[17] National Grid. *Balancing services*. 2015. URL: www.nationalgrid.com.

[18] M. Aunedi, J. E. O. Calderon, V. Silva, P. Mitcheson, and G. Strbac. *Economic and Environmental Impact of Dynamic Demand*. Centre for Sustainable Electricity and Distributed Generation: Imperial College London, 2008.

[19] National Grid. *Reserve services*. 2015. URL: www.nationalgrid.com.

[20] Elexon. *BM Units*. 2015. URL: www.elexon.co.uk.

[21] *The Balancing and Settlement Code*. Version 7. Elexon, Nov. 2013. URL: www.elexon.co.uk.

[22] J. Yang, X. Bai, D. Strickland, L. Jenkins, and A. Cross. "Dynamic Network Rating for Low Carbon Distribution Network Operation – A U.K. Application". In: *IEEE Transactions on Smart Grid* 6.2 (Mar. 2015), pp. 988–998. ISSN: 1949-3053. DOI: 10.1109/TSG.2015.2389711. URL: ieeexplore.ieee.org.

[23] *Requirements for electrical installations (BS 7671)*. London: The Institution of Engineering and Technology, UK, 2015. ISBN: 1849197695 9781849197694. URL: electrical.theiet.org.

[24] Y. Agalgaonkar, B. Pal, and R. Jabr. "Distribution Voltage Control Considering the Impact of PV Generation on Tap Changers and Autonomous Regulators". In: *IEEE Transactions on Power Systems* 29.1 (Jan. 2014), pp. 182–192. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2013.2279721. URL: ieeexplore.ieee.org.

[25] R. Walling, R. Saint, R. Dugan, J. Burke, and L. Kojovic. "Summary of Distributed Resources Impact on Power Delivery Systems". In: *IEEE Transactions on Power Delivery* 23.3 (July 2008), pp. 1636–1644. ISSN: 0885-8977. DOI: 10.1109/TPWRD.2007.909115. URL: ieeexplore.ieee.org.

[26] National Grid. *NETS Security and Quality of Supply Standard*. 2012. URL: www.nationalgrid.com.

[27] G. Strbac. "Demand side management: Benefits and challenges". In: *Energy Policy* 36.12 (Dec. 2008), pp. 4419–4426. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2008.09.030. URL: www.sciencedirect.com.

[28] Energy Networks Association. *Engineering Recommendation P2/6: Security of supply*. Engineering Recommendation. July 2006. URL: www.energynetworks.org.

[29] Energy Networks Association. *Engineering technical report 130-2: Application guide for assessing the capacity of networks containing distributed generation*. London: Engineering Directorate, 2014, pp. 1–29. URL: www.energynetworks.org.

[30] Energy Networks Association. *Amendment to ETR130 to Account for Demand Side Response*. Engineering Technical Report. GB Distribution Code Review Panel, 2014. URL: www.energynetworks.org.

[31] Department of Energy and Climate Change. *Consultation on the draft Electricity Market Reform Delivery Plan*. URN:13D/139. London, 2013. URL: www.gov.uk.

[32] National Grid. *Winter outlook*. 2014. URL: www.nationalgrid.com.

[33] Ofgem. *Electricity Capacity Assessment Report*. London, 2013. URL: www.ofgem.gov.uk.

[34] Department of Energy and Climate Change. *Digest of United Kingdom energy statistics 2013*. 2013. URL: www.decc.gov.uk.

[35] J. Perkins. *National Grid's Demand Side Opportunities*. Presentation. 2007. URL: www.nationalgrid.com (visited on 2015).

[36] National Grid. *Demand Side Balancing Reserve and Supplemental Balancing Reserve*. UK, 2013. URL: www.nationalgrid.com.

[37] Department of Energy and Climate Change. *Implementing Electricity Market Reform*. UK, June 2014. URL: www.gov.uk.

[38] National Grid. *Provisional Auction Results: T-4 Capacity Market Auction 2014*. 2014. URL: www.gov.uk.

[39] A. Nourai, V. I. Kogan, and C. M. Schafer. "Load leveling reduces T&D line losses". In: *IEEE Transactions on Power Delivery* 23.4 (2008), pp. 2168–2173. ISSN: 08858977. DOI: 10.1109/TPWRD.2008.921128. URL: ieeexplore.ieee.org.

[40] Sohn Associates Limited. *Electricity distribution system losses*. Ofgem, 2009. URL: www.ofgem.gov.uk.

[41] Y. T. Tan and D. Kirschen. *Classification of control for demand side participation*. University of Manchester, 2007. URL: www.ee.washington.edu.

[42] Energy Networks Association. *Radio teleswitching service*. 2014. URL: www.energynetworks.org.

[43] E.On. *Distribution Use of System (DUoS) charges for half-hourly metered customers*. 2015. URL: www.eon-uk.com.

[44] Department for Energy and Climate Change. *Demand Side Response in the domestic sector: a literature review of major trials*. London: Frontier Economics and Sustainability First, 2012. URL: www.decc.gov.uk.

[45] National Grid. *2011 National Electricity Transmission System (NETS) Seven Year Statement*. 2011. URL: www.nationalgrid.com.

[46] K. Kariuki and R. Allan. "Evaluation of reliability worth and value of lost load". In: *Generation, Transmission and Distribution, IEE Proceedings-* 143.2 (Mar. 1996). DOI: 10.1049/ip-gtd:19960191. URL: ieeexplore.ieee.org.

[47] D. Kirschen, K. Bell, D. Nedic, D. Jayaweera, and R. Allan. "Computing the value of security". In: *IEE Proceedings - Generation, Transmission and Distribution* 150.6 (2003), p. 673. DOI: 10.1049/ip-gtd:20030796. URL: ieeexplore.ieee.org.

[48] V. Trovato, S. H. Tindemans, and G. Strbac. "Security Constrained Economic Dispatch with Flexible Thermostatically Controlled Loads". In: *5th IEEE PES Innovative Smart Grid Technologies Europe*. Istanbul, 2014, pp. 1–6. ISBN: 9781479977208. URL: ieeexplore.ieee.org.

[49] D. S. Kirschen, G. Strbac, P. Cumperayot, and D. de Paiva Mendes. "Factoring the elasticity of demand in electricity prices". In: *IEEE Transactions on Power Systems* 15.2 (May 2000), pp. 612–617. DOI: 10.1109/59.867149. URL: ieeexplore.ieee.org.

[50] R. Yu, W. Yang, and S. Rahardja. "Optimal real-time price based on a statistical demand elasticity model of electricity". In: *2011 IEEE First International Workshop on Smart Grid Modeling and Simulation (SGMS)*. IEEE, Oct. 2011, 90â“95. ISBN: 978-1-4673-0195-4. DOI: 10.1109/SGMS.2011.6089205. URL: ieeexplore.ieee.org.

[51] G. Strbac, M. Aunedi, D. Pudjianto, P. Djapic, F. Teng, A. Sturt, D. Jackravut, R. Sansom, V. Yufit, and N. Brandon. *Strategic Assessment of the Role and Value of Energy Storage Systems in the UK Low Carbon Energy Future*. Imperial College London, June 2012. URL: www.carbontrust.com.

[52] V. Stanojevid. “Enhancing Performance of Electricity Networks through Application of Demand Side Response and Storage Technologies”. Imperial College London, 2012. URL: spiral.imperial.ac.uk.

[53] D. Papadaskalopoulos. “A Mechanism for Decentralized Participation of Flexible Demand in Electricity Markets”. Imperial College London, 2013. URL: spiral.imperial.ac.uk.

[54] T. Gönen. *Electric power distribution engineering.* 3rd edition. Boca Raton: CRC Press, 2014. ISBN: 9781482207002.

[55] C. Bary. “Coincidence-Factor Reationships of Electric-Service-Load Characteristics”. In: *Transactions of the American Institute of Electrical Engineers* 64.9 (Sept. 1945), pp. 623–629. ISSN: 0096-3860. DOI: 10.1109/T-AIEE.1945.5059190. URL: ieeexplore.ieee.org.

[56] UK Power Networks. *Low Carbon London programme website.* 2015. URL: innovation.ukpowernetworks.co.uk.

[57] J. Schofield, R. Carmichael, S. Tindemans, M. Woolf, M. Bilton, and G. Strbac. *Residential consumer responsiveness to time-varying pricing.* Report A3 for the “Low Carbon London” LCNF project: Imperial College London, 2014. URL: innovation.ukpowernetworks.co.uk.

[58] R. Carmichael, J. Schofield, M. Woolf, M. Bilton, R. Ozaki, and G. Strbac. *Residential consumer attitudes to time-varying pricing.* Report A2 for the "Low Carbon London" LCNF project: Imperial College London, 2014. URL: innovation.ukpowernetworks.co.uk.

[59] A. Faruqui and J. Palmer. “The Discovery of Price Responsiveness–A Survey of Experiments Involving Dynamic Pricing of Electricity”. In: *Social Science Research Network* (2012). URL: papers.ssrn.com.

[60] K. Ehrhardt-Martinez, K. A. Donnelly, and J. A. Laitner. *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities.* Washington DC: American Council for an Energy-Efficient Economy, 2010. URL: www.aceee.org.

[61] A. Faruqui and S. Sergici. *Household response to dynamic pricing of electricity: A survey of 15 experiments.* Vol. 38. 2. 2010, pp. 193–225. ISBN: 1114901091. DOI: 10.1007/s11149-010-9127-y. URL: link.springer.com.

[62] H. Bulkeley, P. Matthews, G. Whitaker, S. Bell, R. Wardle, S. Lyon, and G. Powells. *High level summary of learning: Domestic Smart Meter Customers on Time of Use Tariffs.* Customer Lead Network Revolution project, Northern Power Grid, 2015. URL: www.networkrevolution.co.uk.

[63] Summit Blue Consulting. *Evaluation of the 2006 Energy-Smart Pricing Plan.* Chicago: CNT Energy, 2007. URL: www.smartgridnews.com.

[64] D. J. Hammerstrom, R. Ambrosio, T. A. Carlon, J. G. DeSteese, G. R. Horst, R. Kajfasz, L. Kiesling, P. Michie, R. G. Pratt, M. Yao, J. Brous, D. P. Chassin, R. T. Guttromson, O. M. Järvegren, S. Katipamula, N. T. Le, T. V. Oliver, and S. Thompson. *Pacific Northwest GridWise testbed demonstration projects: Part I . Olympic Peninsula Project.* Washington: Pacific Northwest National Laboratory, 2007. URL: eioc.pnnl.gov.

[65] D. Crossley. *Worldwide survey of network driven demand-side management projects.* 1. Energy Futures Australia Pty Ltd, 2008. URL: www.ieadsm.org.

[66] J. Stromback, C. Dromacque, and M. H. Yassin. *The potential of smart meter enabled programs to increase energy and systems efficiency: a mass pilot comparison.* VaasaETT, 2011. URL: www.vaasaett.com.

[67] Comission for Energy Regulation, Ireland. *Electricity Smart Metering Customer Behaviour Trials (CBT) Findings Report*. Dublin, 2011. URL: www.cer.ie.

[68] A. Faruqui, S. Sergici, and J. Palmer. *The Impact of Dynamic Pricing on Low Income Customers*. The Brattle Group, 2010. URL: www.edisonfoundation.net.

[69] Rocky Mountain Institute. *Automated demand response system pilot: Final report, volume 1*. 2006. URL: www.smartgridinformation.info.

[70] The Electricity Council. *The Domestic Tariffs Experiment*. London, 1974.

[71] S. Allera and A. Cook. “The energy management unit project-domestic customer response to a multi-rate tariff”. In: *Seventh International Conference on Metering Apparatus and Tariffs for Electricity Supply*. 1992, pp. 47–52. URL: ieeexplore.ieee.org.

[72] G. Owen and J. Ward. *Smart meters in Great Britain: the next steps?* Paper 6: Case studies. Sustainability First, 2007. URL: www.sustainabilityfirst.org.uk.

[73] AECOM. *Energy Demand Research Project: Final Analysis*. For the Office for Gas and Electricity Ombudsman, UK, 2011. URL: www.ofgem.gov.uk.

[74] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye. *Probability and statistics for engineers and scientists*. 9th ed. Boston: Pearson Education, 2011. ISBN: 9780321629111.

[75] Experian. *Mosaic consumer classifications*. 2015. URL: www.experian.co.uk.

[76] G. Whitaker, R. Wardle, C. Barteczko-Hibbert, P. Matthews, H. Bulkeley, and G. Powells. *Insight Report: Domestic Time of Use Tariff: A comparison of the time of use tariff trial to the baseline domestic profiles*. Customer Lead Network Revolution project, Northern Power Grid, 2013. URL: www.networkrevolution.co.uk.

[77] C. Barteczko-Hibbert, G. Whitaker, S. Slater, and J. Groizard. *Insight Report: Baseline Domestic Profile: Test Cell 1a Customer Subgroup Analysis*. Customer Lead Network Revolution project, Northern Power Grid, 2015. URL: www.networkrevolution.co.uk.

[78] J. Jorgensen, S. Sorensen, K. Behnke, and P. Eriksen. “EcoGrid EU - A prototype for European Smart Grids”. In: *2011 IEEE Power and Energy Society General Meeting*. 2011, pp. 1–7. DOI: 10.1109/PES.2011.6038981. URL: ieeexplore.ieee.org.

[79] D. Gantenbein, C. Binding, B. Jansen, A. Mishra, and O. Sundstrom. “EcoGrid EU: An efficient ICT approach for a sustainable power system”. In: *Sustainable Internet and ICT for Sustainability (SustainIT), 2012*. 2012, pp. 1–6. URL: ieeexplore.ieee.org.

[80] University College London, Energy Institute. *Smart Energy GB collaboration on consumer demand for time-of-use tariffs*. 2015. URL: www.bartlett.ucl.ac.uk.

[81] Project SOLA BRISTOL. *Six monthly progress report Reporting period: June 2014 to November 2014*. Western Power Distribution, 2014. URL: westernpowerinnovation.co.uk.

[82] UK Power Networks. *Low Carbon Networks Fund Full Submission Pro-forma: Vulnerable Customers and Energy Efficiency (VCEE)*. 2013. URL: innovation.ukpowernetworks.co.uk.

[83] W. K. Lee, G. S. K. Fung, H. Y. Lam, F. H. Y. Chan, and M. Lucente. “Exploration on load signatures”. In: *International conference on electrical Engineering (ICEE)*. Vol. 2. 2004, pp. 690–694. URL: www.lucente.us.

[84] Green eMotion project. *Project website*. 2013. URL: www.greenemotion-project.eu.

[85] M. Aunedi, M. Woolf, M. Bilton, and G. Strbac. *Impact & opportunities for wide-scale EV deployment*. Report B1 for the "Low Carbon London" LCNF project: Imperial College London, 2014. URL: innovation.ukpowernetworks.co.uk.

[86] S. M. Ross. *Introduction to Probability and Statistics for Engineers and Scientists*. 4th ed. Elsevier Science, 2009. ISBN: 9780123704832.

[87] CACI. *Acorn socioeconomic classification data*. 2012. URL: acorn.caci.co.uk.

[88] Greater London Authority. *London Datastore*. 2014. URL: data.london.gov.uk.

[89] Ordnance Survey. *Map data*. 2014. URL: www.ordnancesurvey.co.uk.

[90] G. Sinden. *Wind power and the UK wind resource*. Oxford: University of Oxford, Environmental Change Institute, 2005. URL: www.eci.ox.ac.uk.

[91] Elexon. *Elexon Portal*. 2014. URL: elexonportal.co.uk.

[92] Elexon. *What are the Profile Classes?* 2014. URL: www.elexon.co.uk.

[93] OSIsoft. *PI Server [database]*. 2015. URL: www.osisoft.com.

[94] Microsoft. *SQL Server 2014*. 2015. URL: www.microsoft.com (visited on 03/21/2015).

[95] Pandas. *Pandas Python Data Analysis Library*. Version 16.0. 2015. URL: pandas.pydata.org.

[96] PyData. *A community for developers and users of Python data tools*. 2015. URL: pydata.org.

[97] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830. URL: scikit-learn.org.

[98] I. Konstantelos, M. Sun, and G. Strbac. *Quantifying Demand Diversity of Households*. Report for the "Low Carbon London" LCNF project: Imperial College London, 2014. URL: innovation.ukpowernetworks.co.uk.

[99] Weather Underground. *Historic weather data*. 2014. URL: www.wunderground.com.

[100] B. Stafford. *Pysolar v0.4.3*. 2011. URL: pysolar.org.

[101] E. Jones, T. Oliphant, P. Peterson, et al. *SciPy: Open source scientific tools for Python, Version 15.1*. 2001. URL: www.scipy.org.

[102] R. D'Agostino and E. S. Pearson. "Tests for departure from normality". In: *Biometrika* 60.3 (1973), pp. 613–622. URL: www.jstor.org.

[103] NumPy. *Package for scientific computing with Python*. 2015. URL: www.numpy.org.

[104] R. J. Simes. "An improved bonferroni procedure for multiple tests of significance". In: *Biometrika* 73.3 (1986), pp. 751–754. ISSN: 00063444. DOI: 10.1093/biomet/73.3.751. URL: oxfordjournals.org.

[105] SAS Institute Inc. *Base SAS 9.2 Procedures Guide: Statistical Procedures*. Third Editition. Cary, NC: SAS Publishing, 2010. ISBN: 9781607644514. URL: sas.com.

[106] B. Zhang and S. Horvath. "A general framework for weighted gene co-expression network analysis." In: *Statistical applications in genetics and molecular biology* 4.1 (2005). DOI: 10.2202/1544-6115.1128. URL: www.degruyter.com.

[107] T. M. J. Fruchterman and E. M. Reingold. "Graph Drawing by Force-directed Placement". In: *Software-Practice and Experience* 21 (1991), pp. 1129–1164. DOI: 10.1002/spe.4380211102. URL: wiley.com.

[108] A. A. Hagberg, D. A. Schult, and P. J. Swart. "Exploring network structure, dynamics, and function using NetworkX". In: *Proceedings of the 7th Python in Science Conference (SciPy2008)*. Pasadena, CA USA, Aug. 2008, pp. 11–15. URL: www.lanl.gov.

[109] Y. G. Yohanis, J. D. Mondol, A. Wright, and B. Norton. “Real-life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use”. In: *Energy and Buildings* 40.6 (2008), pp. 1053–1059. DOI: `10.1016/j.enbuild.2007.09.001`. URL: `www.sciencedirect.com`.

[110] J.-P. Zimmermann, M. Evans, J. Griggs, N. King, L. Harding, P. Roberts, and C. Evans. *Household Electricity Survey: A study of domestic electrical product usage*. Tech. rep. Didcot, Oxfordshire: Department for Environment Food and Rural Affiars, May 2013. URL: `www.gov.uk`.

[111] UK Parliament. *Clean Air Act, 1956*. Act of Parliament. London, UK: UK Government, 1956. URL: `legislation.gov.uk`.

Appendix A

Experimental group populations

A.1 Targeting the correct number of experimental units

This section is intended to, first, describe how the target number of experimental units were chosen, and second, serve as a guide to future trial planners. Suppose that the total number of experimental units in the trial is given by:

$$N = n_c + Tn_t \tag{A.1}$$

where n_c is the number of experimental units in the control group, n_t is the number of experimental units per treatment group and T is the number of treatment groups. Where a treatment represents any controlled difference between the groups for which the implications should be quantified.

Two approaches for determining the correct number for each of the groups are considered:

1. Assuming a fixed number of experimental units, their optimal allocation is considered.
2. Given a minimum resolution and confidence, minimum numbers of experimental units are considered.

For the purposes of this analysis, let treatment group t have a population mean of μ_t and the control group one of μ_c. We will be comparing sample mean $\bar{x}_t$, from treatment group t, to sample mean $\bar{x}_c$ from the control group. It is assumed that the variance of the control group is the same as the variance of the treatment group. This may not be the case, but without additional information, this is a good starting assumption. The variance of the difference between the means is therefore given by:

$$\sigma^2_{\bar{x}_t - \bar{x}_c} = \frac{\sigma^2}{n_c} + \frac{\sigma^2}{n_t} \tag{A.2}$$

where σ^2 is the population variance of both the control group and the treatment group, and n_c and n_t are the numbers of experimental units assigned to the control group and to each of the treatment groups respectively.

A.1.1 Allocating a fixed number of experimental units

Given a fixed number of experimental units, their optimum allocation is determined. The following assumptions are made:

1. We have a resource of N experimental units and want to distribute them to achieve the minimum standard error (SE) for comparisons between means of samples from a control group and any one of T treatment groups.
2. The treatment groups will be compared to the control group only.

3. Due to symmetry, each treatment group will consist of the same number of experimental units.

The optimal allocation of experimental units occurs when the variance of the difference (Eq. (A.2)) is minimised. Substituting for n_c from Eq. (A.1) into Eq. (A.2), the minimisation becomes:

$$\min\left[\sigma^2_{\bar{x}_t-\bar{x}_c} = \frac{\sigma^2}{N - Tn_t} + \frac{\sigma^2}{n_t}\right] \tag{A.3}$$

which will be at a minimum when:

$$\frac{dN(n_t)}{dn_t} = \frac{T\sigma^2}{(N - Tn_t)^2} - \frac{\sigma^2}{n_t^2} = 0 \tag{A.4}$$

There are two solutions to the above, but only one is physically relevant:

$$n_t = \begin{cases} \dfrac{N\left(T - \sqrt{T}\right)}{T\,(T-1)}, & \text{if } T > 1 \\ \dfrac{N}{2}, & \text{if } T = 1 \end{cases} \tag{A.5}$$

and therefore

$$n_c = \begin{cases} N - \dfrac{N\left(T - \sqrt{T}\right)}{(T-1)}, & \text{if } T > 1 \\ \dfrac{N}{2}, & \text{if } T = 1 \end{cases} \tag{A.6}$$

For example, if there were 5,000 experimental units and three treatment groups, the optimal assignment of experimental units would be 1,830 for the control group and 1,056 for each of the three treatment groups.

A.1.2 Targeting numbers to achieve a minimum resolution and confidence

Given a minimum resolution and confidence, minimum numbers of experimental units are considered. The following assumptions are made:

1. We would like to minimise N while ensuring that, for all T comparisons between a treatment group and the control group means, a difference of $E\mu_c$ can be observed to at least a confidence of C, where C and E are numbers between 0 and 1.
2. The treatment groups will be compared to the control group only.
3. Due to symmetry, each treatment group will consist of the same number of experimental units.

The approach taken is effectively that of the Z-test[1] where the free variable is the total number of samples, N. Using this definition, the Z value of the difference is:

$$Z = \frac{E\mu_c}{\sigma_{\bar{x}_i-\bar{x}_c}} \tag{A.7}$$

where the variance of the difference between the treatment group mean and control group mean is given by Eq. (A.2). The value of Z is determined as for the Z-test, where is a function of C;

[1] For a quick description of the test, see `http://en.wikipedia.org/wiki/Z-test`

the fractional probability in either 1-tail or 2-tails of the gaussian distribution. A discussion when each should be used can be found in most statistical textbooks [74] and will not be given here.

Starting from Eq. (A.1), the minimisation to be performed is:

$$\min\left[N = n_c + T n_t\right] \tag{A.8}$$

Substituting Eqs. (A.2) and (A.7) into Eq. (A.8) allows us to eliminate the variables $\sigma_{\bar{x}_i - \bar{x}_c}$ and n_t so that we are left with:

$$\min\left[N = \frac{n_c\left(E^2 {\mu_c}^2 n_c - \sigma^2 Z^2 + T\sigma^2 Z^2\right)}{E^2 {\mu_c}^2 n_c - \sigma^2 Z^2}\right] \tag{A.9}$$

Which will be at a minimum when:

$$\frac{dN(n_c)}{dn_c} = 0 \tag{A.10}$$

This has two solutions, but only one is physically relevant. Taking this solution for n_c, and with a little algebra, the following can be found:

$$n_c = \left(\frac{\sigma Z}{\mu_c E}\right)^2 \cdot \left(1 + \sqrt{T}\right) \tag{A.11}$$

$$n_t = \left(\frac{\sigma Z}{\mu_c E}\right)^2 \cdot \frac{\left(1 + \sqrt{T}\right)}{\sqrt{T}} \tag{A.12}$$

$$N = \left(\frac{\sigma Z}{\mu_c E}\right)^2 \cdot \left(1 + \sqrt{T}\right)^2 \tag{A.13}$$

It is worth noting that the control to treatment group ratio is only a function of T:

$$\frac{n_c}{n_t} = \sqrt{T} \tag{A.14}$$

For example, with 3 treatment groups, a standard deviation to mean ratio of 0.75, a change of 5% ($E = 0.05$) to be measured at 95% confidence, and taking the 1-tail definition of Z, one would require 1,663 experimental units in the control group and 960 experimental units in each of the three treatment groups, giving a total of 4,544 experimental units. In contrast, a total of 6,451 would be required if the 2-tails definition of Z were used.

A.2 Low Carbon London target population estimates

This section the process used to estimate the number of samples (households) required in each experimental group, and by extension, the number of treatment groups to include in the trial. The number of samples required for statistical robustness is dependent on the number of treatment groups and the choice of *confidence* and *resolution*. Here, resolution is defined as the difference between the two group means which is desired to be measurable. It is important to note that the values of confidence and resolution are tradable, which is to say, increasing one will will have a detrimental effect on the other. The trade-off relationship between the two is dependant on a number of other parameters; the respective sample means, their standard deviations and the number of samples in each group.

As the most common analysis to be conducted on such panel data is the comparison of two group means, the confidence and resolution were defined according to the Z-test, a standard statistical test used to determine whether two means can be said to be from different populations.

A.2.1 Parameters

The key parameters in the choice of group populations were:

Target confidence: The choice of a satisfactory level of confidence is subjective, though by convention, for social science experiments (this trial may be considered one), levels of 90% or 95% are typically chosen. 95% confidence was aimed for. We use the 2-tailed definition of confidence in this section.

Target resolution: The expected change was informed by the literature on past trials as described in Chapter 3. Table 3.1 shows that the peak reduction reported from past trials ranged from 0% to 12%. As this trial is most similar in terms of its objectives, cultural background and climate situation to that of the Ireland Electricity Smart Metering Trials (IESMT), designing to trial to measure at least a 5% difference (between group means) was considered reasonable.

Estimated mean and standard deviation: In order to obtain estimates of the likely means and standard deviations of our sample group parameters, programme partner EDF Energy provided 10,000 randomly selected estimated annual consumption (EAC) figures from its customers within the London area. It was assumed that the variance of this data could be taken as a proxy for the variance of other group metrics. Statistics from this data set are given in Table A.1.

Data set	Mean (kWh)	St. Dev. (kWh)
Dual fuel (2010)	3,327	2,280
Dual fuel (2011)	3,271	2,270
Electricity only (2010)	4,116	3,160
Electricity only (2011)	4,162	3,220

Table A.1: Statistics for EAC data from the London area. 10,000 households in each data set. Data courtesy of EDF Energy

Attrition rate: Attrition rate is the fraction of the original group population that are estimated to leave the trial before completion. Sample numbers must be made sufficiently higher than the statistically defined target in order to account for attrition losses. EDF Energy provided a conservative estimate of attrition of 20%.

Total target sample numbers, accounting for attrition, are calculated as:

$$N_{\text{target}} = \frac{N_{\text{stat}}}{1-a} \tag{A.15}$$

where a is the attrition rate and N_{stat} is the number of samples necessary for statistically valid results, given by Eqs. (A.11) to (A.13) for respective groups.

A.2.2 Group target populations

Informed by the results of previous trials and the EAC data, the following assumptions were made:

- Up to 6000 households available for the trial.
- An attrition rate of no more than 20%.
- The trial would be designed, as far as possible, to detect 5% difference between group means.
- The mean-to-standard-deviation ratio for consumption should be similar to the same ratio for peak demand.
- Mean consumption was expected to be approximately 4,000 kWh per year.
- Standard deviation of consumption was expected to be approximately 3,000 kWh.

With two trial groups, the optimal group populations occur when there is an equal split of the available households between both groups. Using Eq. (A.13), it can be shown that, given the maximum of 6,000 samples thought to be available to the trial, and with an attrition rate of 20%, such a split would allow the measurement of a 3.5% difference between group means at the 90% confidence level.

However, it was expected that an even split between the control and experimental groups might not be possible owing to the opt-in nature of the trial. An additional calculation was made to determine the minimum number of samples required in each group in order to be able to measure a 5% group difference at 90% confidence. It was found (working omitted) that, in order to meet this criteria, 1,217 samples were required in each group after attrition. Including attrition, minimum group recruitment targets were set to 1,521 for each group.

Appendix B

Metadata analysis

This appendix holds data tables in association with the work described in Chapter 9.

B.1 Metadata variables

This section provides a list of the metadata variables used in the correlation analysis of Chapter 9. The Index column in Table B.1 corresponds to the numbered items in the weighted correlation network graphs presented in Chapter 9. Variables were first grouped according to their high level categories, then given an alias that briefly describes what they represent. The number of responses from the dToU and nonToU groups are given in the N_dToU and N_nonToU columns respectively.

Table B.1: List of metadata variables tested in Chapter 9.

Index	Group	Alias	N_dToU	N_nonToU
1	Accommodation	ACORN_group	887	1673
2	Accommodation	Has_double_glazing	784	1469
3	Accommodation	Has_electric_heating	828	1574
4	Accommodation	Has_floor_insulation	509	861
5	Accommodation	Has_gas_heating	828	1574
6	Accommodation	Has_hot_water_tank	658	1200
7	Accommodation	Has_immersion_heater	832	1576
8	Accommodation	Has_loft_insulation	706	1317
9	Accommodation	Has_wall_insulation	563	992
10	Accommodation	Heating_interface_manual	767	1489
11	Accommodation	Heating_interface_set_times	767	1489
12	Accommodation	Heating_interface_thermostatic_controller	767	1489
13	Accommodation	Heating_interface_thermostatic_valves	767	1489
14	Accommodation	Is_flat	814	1535
15	Accommodation	Is_house	814	1535
16	Accommodation	N_bedrooms	832	1563
17	Accommodation	N_rooms	832	1564
18	Accommodation	SM_in_Hallway	825	1557
19	Accommodation	SM_in_Kitchen	825	1557
20	Accommodation	SM_in_Living_Room	825	1557
21	Appliances	Lighting_N_LED	837	1588
22	Appliances	Lighting_N_fluorescent	837	1588
23	Appliances	Lighting_N_halogen	837	1588
24	Appliances	Lighting_N_low_energy	837	1588
25	Appliances	Lighting_N_traditional	837	1588
26	Appliances	N_desktop_PCs	837	1588
27	Appliances	N_dishwashers	837	1588
28	Appliances	N_freezers	837	1588

Continued on next page

Continued from previous page

Index	Group	Alias	N_dToU	N_nonToU
29	Appliances	N_fridge_freezers	837	1588
30	Appliances	N_fridges	837	1588
31	Appliances	N_game_consoles	837	1588
32	Appliances	N_hobs_electric	837	1588
33	Appliances	N_hobs_gas	837	1588
34	Appliances	N_laptop_PCs	837	1588
35	Appliances	N_microwaves	837	1588
36	Appliances	N_ovens_electric	837	1588
37	Appliances	N_over_sink_water_heaters	837	1588
38	Appliances	N_portable_electric_heaters	837	1588
39	Appliances	N_printers	837	1588
40	Appliances	N_routers	837	1588
41	Appliances	N_set_top_boxes	837	1588
42	Appliances	N_showers_electric	837	1588
43	Appliances	N_standby_savers	837	1588
44	Appliances	N_televisions	837	1588
45	Appliances	N_tumble_driers	837	1588
46	Appliances	N_video_players	837	1588
47	Appliances	N_washer_driers	837	1588
48	Appliances	N_washing_machines	837	1588
49	Behaviour_change	Cycle_of_appliance	600	0
50	Behaviour_change	Leaving_home	600	0
51	Behaviour_change	Lighting_reduction	628	0
52	Behaviour_change	Substitute_cooking	606	0
53	Behaviour_change	Substitute_fuel	589	0
54	Behaviour_change	Users_of_appliances	601	0
55	Behaviour_change	Working_hours	585	0
56	Behaviour_inflexible	Dishwasher	301	0
57	Behaviour_inflexible	Electric_heater	214	0
58	Behaviour_inflexible	Electric_hob	221	0
59	Behaviour_inflexible	Electric_oven	441	0
60	Behaviour_inflexible	Electric_shower	228	0
61	Behaviour_inflexible	Immersion_heater	204	0
62	Behaviour_inflexible	Ironing	630	0
63	Behaviour_inflexible	Kettle	637	0
64	Behaviour_inflexible	Lighting	660	0
65	Behaviour_inflexible	Tumble_drier	270	0
66	Behaviour_inflexible	Washing_machine	611	0
67	Behaviour_normal	Work_from_home	814	1519
68	Behaviour_normal	Concerned_about_climate_change	799	1461
69	Behaviour_normal	Feeling_about_lifestyle_and_environment	830	1558
70	Behaviour_normal	Interest_in_micro_generation	686	1238
71	Behaviour_normal	Interest_in_renewable_electricity	708	1242
72	Behaviour_normal	Reads_paper	822	1530
73	Behaviour_timer_use	Dishwasher	90	0
74	Behaviour_timer_use	Electric_heating	51	0
75	Behaviour_timer_use	Immersion_heater	66	0
76	Behaviour_timer_use	Tumble_drier	52	0
77	Behaviour_timer_use	Washer_drier	43	0
78	Behaviour_timer_use	Washing_machine	161	0
79	People	Age_N_0_to_11	887	1673
80	People	Age_N_12_to_17	887	1673
81	People	Age_N_18_to_34	887	1673
82	People	Age_N_35_to_54	887	1673
83	People	Age_N_55_to_74	887	1673
84	People	Age_N_75_or_older	887	1673
85	People	Gender_N_females	887	1673
86	People	Gender_N_males	887	1673

Continued on next page

Continued from previous page

Index	Group	Alias	N_dToU	N_nonToU
87	People	N_of	797	1468
88	Primary	All_default_price	887	0
89	Primary	All_high_price	887	0
90	Primary	All_high_price_relative	887	0
91	Primary	All_low_price	887	0
92	Primary	All_low_price_relative	887	0
93	Primary	CM_high_price	887	0
94	Primary	Consumption_annual	887	1673
95	Primary	DR_training_index	887	0
96	Primary	Engagement_rank	887	0
97	Primary	Is_dToU	887	1673
98	Primary	SF_high_price	887	0
99	Primary	SF_high_wd	887	0
100	Primary	SF_high_wd_00_07	887	0
101	Primary	SF_high_wd_07_10	887	0
102	Primary	SF_high_wd_10_17	887	0
103	Primary	SF_high_wd_18_21	887	0
104	Primary	SF_high_we	887	0
105	Primary	SF_low_price	887	0
106	Primary	SF_low_wd	887	0
107	Primary	SF_low_wd_00_07	887	0
108	Primary	SF_low_wd_07_10	887	0
109	Primary	SF_low_wd_10_17	887	0
110	Primary	SF_low_wd_18_21	887	0
111	Primary	SF_low_we	887	0
112	Report_high	Dishwasher	283	0
113	Report_high	Electric_heater	199	0
114	Report_high	Electric_hob	200	0
115	Report_high	Electric_oven	426	0
116	Report_high	Electric_shower	209	0
117	Report_high	Immersion_heater	179	0
118	Report_high	Ironing	588	0
119	Report_high	Kettle	601	0
120	Report_high	Lighting	610	0
121	Report_high	Night	565	0
122	Report_high	Saturdays	600	0
123	Report_high	Sundays	582	0
124	Report_high	Tumble_drier	256	0
125	Report_high	Washing_machine	597	0
126	Report_high	Weekday_afternoons	583	0
127	Report_high	Weekday_evenings	575	0
128	Report_high	Weekday_mornings	579	0
129	Report_low	Dishwasher	275	0
130	Report_low	Electric_heater	186	0
131	Report_low	Electric_hob	188	0
132	Report_low	Electric_oven	409	0
133	Report_low	Electric_shower	191	0
134	Report_low	Immersion_heater	161	0
135	Report_low	Ironing	558	0
136	Report_low	Kettle	568	0
137	Report_low	Lighting	573	0
138	Report_low	Night	544	0
139	Report_low	Saturdays	576	0
140	Report_low	Sundays	564	0
141	Report_low	Tumble_drier	247	0
142	Report_low	Washing_machine	567	0
143	Report_low	Weekday_afternoons	557	0
144	Report_low	Weekday_evenings	558	0

Continued on next page

Continued from previous page

Index	Group	Alias	N_dToU	N_nonToU
145	Report_low	Weekday_mornings	558	0
146	Response_helper	Calendar_or_notepad	578	0
147	Response_helper	Coordinating_with_household_members	523	0
148	Response_helper	Flexible_working_hours	485	0
149	Response_helper	In_home_display	634	0
150	Response_helper	Phone_messages	557	0
151	Response_helper	Working_from_home	496	0
152	Response_limiter	Being_at_home	623	0
153	Response_limiter	Children_routine	447	0
154	Response_limiter	Comfort_and_convenience	606	0
155	Response_limiter	Complexity	613	0
156	Response_limiter	Complicated_changes	619	0
157	Response_limiter	Fixed_appliance_use_routine	608	0
158	Response_limiter	Fixed_appliance_users	604	0
159	Response_limiter	Forget_price_changes	621	0
160	Response_limiter	Increasing_elec_usage	608	0
161	Response_limiter	Inflexible_working	549	0
162	Response_limiter	Knowing_when_rates_change	625	0
163	Response_limiter	Notice_too_short	618	0
164	Response_limiter	Savings_too_small	616	0
165	Trial_impressions	Too_much_hassle	633	0
166	Trial_impressions	Convenience_trumps_saving	634	0
167	Trial_impressions	Curious_of_others_performance	639	0
168	Trial_impressions	Curious_to_talk_to_others	622	0
169	Trial_impressions	Effort_was_sustainable	629	0
170	Trial_impressions	Enjoyable	626	0
171	Trial_impressions	Feedback_letter_clear	634	0
172	Trial_impressions	Feedback_letter_useful	631	0
173	Trial_impressions	Frustrating_lack_of_effort_from_some	591	0
174	Trial_impressions	Frustrating_not_knowing_reason	629	0
175	Trial_impressions	Frustrating_when_cannot_respond	632	0
176	Trial_impressions	Guarantee_meant_less_effort	627	0
177	Trial_impressions	High_easy	639	0
178	Trial_impressions	In_home_display_clear	635	0
179	Trial_impressions	In_home_display_useful	627	0
180	Trial_impressions	Low_easy	636	0
181	Trial_impressions	More_in_control_of_bill	633	0
182	Trial_impressions	Often_talked_about_it	614	0
183	Trial_impressions	Reduced_consumption	606	0
184	Trial_impressions	Reduced_our_comfort	628	0
185	Trial_impressions	Reduced_our_consumption	627	0
186	Trial_impressions	Renewable_link_behaviour_change	603	0
187	Trial_impressions	Savings_estimate	458	0
188	Trial_impressions	Savings_to_be_worthwhile_percent	309	0
189	Trial_impressions	Savings_to_be_worthwhile_pounds	345	0
190	Trial_impressions	Should_be_offered_to_everyone_if_efficient	640	0
191	Trial_impressions	Should_be_offered_to_everyone_if_fairer	635	0
192	Trial_impressions	Signup_Guarantee_needed	635	0
193	Trial_impressions	Signup_If_longer_simpler_events	616	0
194	Trial_impressions	Signup_If_more_predictable	609	0
195	Trial_impressions	Signup_If_no_evening_peak_changes	601	0
196	Trial_impressions	Signup_If_only_weekdays	600	0
197	Trial_impressions	Signup_If_renewable_link	604	0
198	Trial_impressions	Tarif_Signup_If_only_weekends	601	0
199	Trial_impressions	Too_complex	640	0
200	Trial_impressions	Want_to_stay_on	632	0

B.2 Correlation results tables

This section provides the tables of significant correlations for each of the root variables listed in Table 9.2 against all other variables listed in Table B.1. Significance was defined as having a p-value less than 10^{-5}.

Table B.2: Table of significant correlations against root variable ACORN_group.

Index	Group	Alias	N	r	p
136	Report_low	Kettle	568	0.2	2.2e-06
120	Report_high	Lighting	610	0.18	6.4e-06
93	Primary	CM_high_price	887	0.16	1.9e-06
2	Accommodation	Has_double_glazing	2,253	0.16	9.8e-14
9	Accommodation	Has_wall_insulation	1,555	0.13	1.5e-07
14	Accommodation	Is_flat	2,349	0.099	1.6e-06
97	Primary	Is_dToU	2,560	-0.087	9.5e-06
15	Accommodation	Is_house	2,349	-0.099	1.6e-06
5	Accommodation	Has_gas_heating	2,402	-0.099	1e-06
68	Behaviour_normal	Feeling_about_lifestyle_and_environment	2,388	-0.11	3.7e-08
23	Appliances	Lighting_N_LED	2,425	-0.12	1e-08
26	Appliances	N_desktop_PCs	2,425	-0.12	2.9e-09
38	Appliances	N_portable_electric_heaters	2,425	-0.12	6.3e-10
34	Appliances	N_laptop_PCs	2,425	-0.13	4.4e-11
25	Appliances	Lighting_N_traditional	2,425	-0.16	4.9e-15
40	Appliances	N_routers	2,425	-0.16	2e-15
36	Appliances	N_ovens_electric	2,425	-0.16	3.2e-16
11	Accommodation	Heating_interface_set_times	2,256	-0.17	1.6e-15
94	Primary	Consumption_annual	2,560	-0.17	1.3e-18
16	Accommodation	N_bedrooms	2,395	-0.17	1.4e-17
39	Appliances	N_printers	2,425	-0.2	2.6e-23
17	Accommodation	N_rooms	2,396	-0.22	9.1e-28
72	Behaviour_normal	Work_from_home	2,333	-0.23	1.8e-30
22	Appliances	Lighting_N_halogen	2,425	-0.24	9e-34
27	Appliances	N_dishwashers	2,425	-0.27	3.5e-43

Table B.3: Table of significant correlations against root variable All_default_price.

Index	Group	Alias	N	r	p
88	Primary	All_default_price	887	1	0
89	Primary	All_high_price	887	0.49	2.3e-56
98	Primary	SF_high_price	887	0.48	2e-54
99	Primary	SF_high_wd	887	0.48	1.6e-53
90	Primary	All_high_price_relative	887	0.46	4.9e-50
102	Primary	SF_high_wd_10_17	887	0.4	2.8e-36
93	Primary	CM_high_price	887	0.37	8.4e-31
103	Primary	SF_high_wd_18_21	887	0.32	2.2e-23
100	Primary	SF_high_wd_00_07	887	0.32	3.4e-23
101	Primary	SF_high_wd_07_10	887	0.31	5.9e-22
104	Primary	SF_high_we	887	0.27	7.5e-17
91	Primary	All_low_price	887	0.26	3.4e-15
92	Primary	All_low_price_relative	887	0.24	7.5e-13
96	Primary	Engagement_rank	887	0.18	2.8e-08
107	Primary	SF_low_wd_00_07	887	0.16	1.1e-06
106	Primary	SF_low_wd	887	0.15	7.9e-06
95	Primary	DR_training_index	887	-0.15	5.4e-06

Table B.4: Table of significant correlations against root variable All_high_price.

Index	Group	Alias	N	r	p
89	Primary	All_high_price	887	1	0
90	Primary	All_high_price_relative	887	0.91	0
98	Primary	SF_high_price	887	0.89	0
99	Primary	SF_high_wd	887	0.87	0
93	Primary	CM_high_price	887	0.84	1.3e-288
96	Primary	Engagement_rank	887	0.77	9.8e-206
102	Primary	SF_high_wd_10_17	887	0.75	4.6e-180
103	Primary	SF_high_wd_18_21	887	0.66	6.6e-124
101	Primary	SF_high_wd_07_10	887	0.52	2.3e-66
104	Primary	SF_high_we	887	0.51	5.4e-63
88	Primary	All_default_price	887	0.49	2.3e-56
100	Primary	SF_high_wd_00_07	887	0.4	1.2e-35
14	Accommodation	Is_flat	814	0.23	5.6e-11
159	Response_limiter	Forget_price_changes	621	0.22	2.8e-08
82	People	Age_N_35_to_54	887	-0.15	5.4e-06
85	People	Gender_N_females	887	-0.15	4.4e-06
42	Appliances	N_showers_electric	837	-0.16	6.4e-06
22	Appliances	Lighting_N_halogen	837	-0.16	4e-06
44	Appliances	N_televisions	837	-0.16	2.7e-06
40	Appliances	N_routers	837	-0.16	2.6e-06
92	Primary	All_low_price_relative	887	-0.17	6e-07
91	Primary	All_low_price	887	-0.17	5.6e-07
110	Primary	SF_low_wd_18_21	887	-0.18	1.2e-07
36	Appliances	N_ovens_electric	837	-0.18	2.1e-07
26	Appliances	N_desktop_PCs	837	-0.18	2e-07
118	Report_high	Ironing	588	-0.19	4.7e-06
39	Appliances	N_printers	837	-0.19	3.9e-08
128	Report_high	Weekday_mornings	579	-0.19	4.8e-06
25	Appliances	Lighting_N_traditional	837	-0.19	3e-08
87	People	N_of	797	-0.2	2.5e-08
53	Behaviour_change	Substitute_fuel	589	-0.2	5e-07
94	Primary	Consumption_annual	887	-0.21	1.8e-10
27	Appliances	N_dishwashers	837	-0.22	2e-10
15	Accommodation	Is_house	814	-0.23	5.6e-11
126	Report_high	Weekday_afternoons	583	-0.23	1.2e-08
111	Primary	SF_low_we	887	-0.24	7.9e-13
108	Primary	SF_low_wd_07_10	887	-0.25	5e-14
45	Appliances	N_tumble_driers	837	-0.25	2.4e-13
112	Report_high	Dishwasher	283	-0.26	9.1e-06
16	Accommodation	N_bedrooms	832	-0.28	6.6e-17
106	Primary	SF_low_wd	887	-0.29	2.1e-18
17	Accommodation	N_rooms	832	-0.29	2.6e-18
105	Primary	SF_low_price	887	-0.3	5.2e-20
109	Primary	SF_low_wd_10_17	887	-0.32	1.3e-23
124	Report_high	Tumble_drier	256	-0.38	1.9e-10

Table B.5: Table of significant correlations against root variable All_high_price_relative.

Index	Group	Alias	N	r	p
90	Primary	All_high_price_relative	887	1	0
89	Primary	All_high_price	887	0.91	0
96	Primary	Engagement_rank	887	0.83	1.6e-279
98	Primary	SF_high_price	887	0.81	2.1e-240
99	Primary	SF_high_wd	887	0.79	9.6e-227
93	Primary	CM_high_price	887	0.79	1.4e-222
102	Primary	SF_high_wd_10_17	887	0.66	2.2e-121
103	Primary	SF_high_wd_18_21	887	0.61	4e-97
88	Primary	All_default_price	887	0.46	4.9e-50
101	Primary	SF_high_wd_07_10	887	0.45	7.4e-47
104	Primary	SF_high_we	887	0.45	1.9e-46
100	Primary	SF_high_wd_00_07	887	0.37	1.1e-31
159	Response_limiter	Forget_price_changes	621	0.25	4.8e-10
193	Trial_impressions	Signup_If_longer_simpler_events	616	0.2	4.2e-07
176	Trial_impressions	Guarantee_meant_less_effort	627	0.2	6e-07
162	Response_limiter	Knowing_when_rates_change	625	0.19	1.7e-06
155	Response_limiter	Complexity	613	0.19	2.3e-06
91	Primary	All_low_price	887	-0.16	1.4e-06
127	Report_high	Weekday_evenings	575	-0.18	9.1e-06
110	Primary	SF_low_wd_18_21	887	-0.19	2.1e-08
145	Report_low	Weekday_mornings	558	-0.19	4.2e-06
146	Response_helper	Calendar_or_notepad	578	-0.2	1.6e-06
92	Primary	All_low_price_relative	887	-0.2	2.1e-09
123	Report_high	Sundays	582	-0.2	8.8e-07
140	Report_low	Sundays	564	-0.2	1.3e-06
143	Report_low	Weekday_afternoons	557	-0.2	1.5e-06
128	Report_high	Weekday_mornings	579	-0.2	7.4e-07
139	Report_low	Saturdays	576	-0.2	7.4e-07
122	Report_high	Saturdays	600	-0.21	1.2e-07
53	Behaviour_change	Substitute_fuel	589	-0.22	9.4e-08
144	Report_low	Weekday_evenings	558	-0.22	8.4e-08
108	Primary	SF_low_wd_07_10	887	-0.24	9e-13
111	Primary	SF_low_we	887	-0.25	7.3e-14
125	Report_high	Washing_machine	597	-0.25	8.3e-10
118	Report_high	Ironing	588	-0.25	1.1e-09
106	Primary	SF_low_wd	887	-0.26	1.4e-15
115	Report_high	Electric_oven	426	-0.26	3.4e-08
126	Report_high	Weekday_afternoons	583	-0.27	3e-11
105	Primary	SF_low_price	887	-0.28	1.7e-17
109	Primary	SF_low_wd_10_17	887	-0.31	6.2e-21
112	Report_high	Dishwasher	283	-0.32	2.9e-08
124	Report_high	Tumble_drier	256	-0.42	1.8e-12

Table B.6: Table of significant correlations against root variable All__low__price.

Index	Group	Alias	N	r	p
91	Primary	All__low__price	887	1	0
92	Primary	All__low__price__relative	887	0.95	0
105	Primary	SF__low__price	887	0.81	2.1e-241
106	Primary	SF__low__wd	887	0.8	1.6e-233
109	Primary	SF__low__wd__10__17	887	0.71	4e-152
108	Primary	SF__low__wd__07__10	887	0.57	3.7e-84
110	Primary	SF__low__wd__18__21	887	0.56	2.2e-80
107	Primary	SF__low__wd__00__07	887	0.42	6.9e-40
111	Primary	SF__low__we	887	0.4	2e-35
124	Report__high	Tumble__drier	256	0.38	1.8e-10
125	Report__high	Washing__machine	597	0.31	1.1e-14
141	Report__low	Tumble__drier	247	0.3	9.1e-07
112	Report__high	Dishwasher	283	0.3	2.1e-07
126	Report__high	Weekday__afternoons	583	0.26	7.3e-11
88	Primary	All__default__price	887	0.26	3.4e-15
118	Report__high	Ironing	588	0.24	6.1e-09
128	Report__high	Weekday__mornings	579	0.23	1.1e-08
122	Report__high	Saturdays	600	0.23	9.1e-09
123	Report__high	Sundays	582	0.22	7.6e-08
143	Report__low	Weekday__afternoons	557	0.21	2.9e-07
182	Trial__impressions	Often__talked__about__it	614	0.21	1.3e-07
142	Report__low	Washing__machine	567	0.2	1.1e-06
144	Report__low	Weekday__evenings	558	0.2	2e-06
139	Report__low	Saturdays	576	0.19	2.5e-06
140	Report__low	Sundays	564	0.19	5.3e-06
94	Primary	Consumption__annual	887	0.15	3.8e-06
90	Primary	All__high__price__relative	887	-0.16	1.4e-06
103	Primary	SF__high__wd__18__21	887	-0.16	9.2e-07
89	Primary	All__high__price	887	-0.17	5.6e-07
157	Response__limiter	Fixed__appliance__use__routine	608	-0.18	9.9e-06
158	Response__limiter	Fixed__appliance__users	604	-0.18	8.8e-06
193	Trial__impressions	Signup__If__longer__simpler__events	616	-0.19	3.2e-06
99	Primary	SF__high__wd	887	-0.19	1.7e-08
164	Response__limiter	Savings__too__small	616	-0.19	1.9e-06
102	Primary	SF__high__wd__10__17	887	-0.19	8.2e-09
101	Primary	SF__high__wd__07__10	887	-0.2	6.9e-10
98	Primary	SF__high__price	887	-0.22	7.5e-11
160	Response__limiter	Increaseing__elec__usage	608	-0.22	5.3e-08
159	Response__limiter	Forget__price__changes	621	-0.22	1.7e-08
104	Primary	SF__high__we	887	-0.26	8e-16
96	Primary	Engagement__rank	887	-0.45	1.3e-46

Table B.7: Table of significant correlations against root variable All_low_price_relative.

Index	Group	Alias	N	r	p
92	Primary	All_low_price_relative	887	1	0
91	Primary	All_low_price	887	0.95	0
105	Primary	SF_low_price	887	0.79	2.5e-225
106	Primary	SF_low_wd	887	0.78	1.5e-216
109	Primary	SF_low_wd_10_17	887	0.7	1.4e-143
110	Primary	SF_low_wd_18_21	887	0.57	4e-83
108	Primary	SF_low_wd_07_10	887	0.57	1e-81
124	Report_high	Tumble_drier	256	0.43	4e-13
107	Primary	SF_low_wd_00_07	887	0.41	6.3e-39
111	Primary	SF_low_we	887	0.41	1.4e-37
125	Report_high	Washing_machine	597	0.36	1.2e-19
112	Report_high	Dishwasher	283	0.34	2e-09
141	Report_low	Tumble_drier	247	0.33	6.7e-08
126	Report_high	Weekday_afternoons	583	0.3	9.1e-14
118	Report_high	Ironing	588	0.28	6.2e-12
122	Report_high	Saturdays	600	0.26	1.1e-10
128	Report_high	Weekday_mornings	579	0.25	5.9e-10
123	Report_high	Sundays	582	0.25	1.1e-09
143	Report_low	Weekday_afternoons	557	0.24	1e-08
88	Primary	All_default_price	887	0.24	7.5e-13
139	Report_low	Saturdays	576	0.23	2.2e-08
144	Report_low	Weekday_evenings	558	0.22	7.5e-08
142	Report_low	Washing_machine	567	0.22	6e-08
140	Report_low	Sundays	564	0.22	8.6e-08
115	Report_high	Electric_oven	426	0.22	3.9e-06
180	Trial_impressions	Low_easy	636	0.2	2.9e-07
149	Response_helper	In_home_display	634	0.18	2.8e-06
182	Trial_impressions	Often_talked_about_it	614	0.18	8.7e-06
103	Primary	SF_high_wd_18_21	887	-0.15	3.6e-06
89	Primary	All_high_price	887	-0.17	6e-07
199	Trial_impressions	Too_complex	640	-0.18	5.7e-06
101	Primary	SF_high_wd_07_10	887	-0.18	4.8e-08
102	Primary	SF_high_wd_10_17	887	-0.18	3.3e-08
99	Primary	SF_high_wd	887	-0.18	3.1e-08
164	Response_limiter	Savings_too_small	616	-0.19	2.7e-06
156	Response_limiter	Complicated_changes	619	-0.19	1.7e-06
155	Response_limiter	Complexity	613	-0.2	8.3e-07
158	Response_limiter	Fixed_appliance_users	604	-0.2	9.5e-07
90	Primary	All_high_price_relative	887	-0.2	2.1e-09
193	Trial_impressions	Signup_If_longer_simpler_events	616	-0.2	3.7e-07
165	Trial_impressions	Too_much_hassle	633	-0.21	1.4e-07
157	Response_limiter	Fixed_appliance_use_routine	608	-0.21	2.4e-07
98	Primary	SF_high_price	887	-0.21	3e-10
104	Primary	SF_high_we	887	-0.24	4.2e-13
160	Response_limiter	Increaseing_elec_usage	608	-0.24	7.8e-10
159	Response_limiter	Forget_price_changes	621	-0.26	8e-11
96	Primary	Engagement_rank	887	-0.49	1.7e-57

Table B.8: Table of significant correlations against root variable CM_high_price.

Index	Group	Alias	N	r	p
93	Primary	CM_high_price	887	1	0
89	Primary	All_high_price	887	0.84	1.3e-288
90	Primary	All_high_price_relative	887	0.79	1.4e-222
96	Primary	Engagement_rank	887	0.65	5e-119
99	Primary	SF_high_wd	887	0.56	9.1e-80
98	Primary	SF_high_price	887	0.56	1.3e-77
102	Primary	SF_high_wd_10_17	887	0.48	4.5e-54
103	Primary	SF_high_wd_18_21	887	0.46	7.2e-49
88	Primary	All_default_price	887	0.37	8.4e-31
101	Primary	SF_high_wd_07_10	887	0.33	6.2e-24
104	Primary	SF_high_we	887	0.31	9.3e-22
100	Primary	SF_high_wd_00_07	887	0.25	8.5e-14
159	Response_limiter	Forget_price_changes	621	0.19	1.3e-06
14	Accommodation	Is_flat	814	0.18	2.4e-07
1	Accommodation	ACORN_group	887	0.16	1.9e-06
36	Appliances	N_ovens_electric	837	-0.16	3.4e-06
94	Primary	Consumption_annual	887	-0.16	1.6e-06
22	Appliances	Lighting_N_halogen	837	-0.16	2.7e-06
26	Appliances	N_desktop_PCs	837	-0.16	1.8e-06
82	People	Age_N_35_to_54	887	-0.17	6.2e-07
40	Appliances	N_routers	837	-0.17	7.8e-07
25	Appliances	Lighting_N_traditional	837	-0.18	2.1e-07
15	Accommodation	Is_house	814	-0.18	2.4e-07
111	Primary	SF_low_we	887	-0.19	2.4e-08
45	Appliances	N_tumble_driers	837	-0.19	1.4e-08
39	Appliances	N_printers	837	-0.2	6.1e-09
16	Accommodation	N_bedrooms	832	-0.22	2.2e-10
53	Behaviour_change	Substitute_fuel	589	-0.22	7.5e-08
108	Primary	SF_low_wd_07_10	887	-0.22	1.3e-11
27	Appliances	N_dishwashers	837	-0.24	9.9e-13
106	Primary	SF_low_wd	887	-0.25	7.8e-14
17	Accommodation	N_rooms	832	-0.25	1.6e-13
105	Primary	SF_low_price	887	-0.25	1.6e-14
109	Primary	SF_low_wd_10_17	887	-0.3	7.8e-21
124	Report_high	Tumble_drier	256	-0.32	1.2e-07

Table B.9: Table of significant correlations against root variable Consumption_annual.

Index	Group	Alias	N	r	p
94	Primary	Consumption_annual	2,560	1	0
87	People	N_of	2,265	0.56	2.9e-203
16	Accommodation	N_bedrooms	2,395	0.46	4.3e-133
17	Accommodation	N_rooms	2,396	0.44	3.3e-118
27	Appliances	N_dishwashers	2,425	0.4	7.4e-99
189	Trial_impressions	Savings_to_be_worthwhile_pounds	345	0.4	2.7e-15
39	Appliances	N_printers	2,425	0.38	4.7e-87
40	Appliances	N_routers	2,425	0.36	7.9e-79
44	Appliances	N_televisions	2,425	0.36	3.1e-78
86	People	Gender_N_males	2,560	0.36	1.4e-80
26	Appliances	N_desktop_PCs	2,425	0.36	2.9e-75
45	Appliances	N_tumble_driers	2,425	0.36	8.3e-75
34	Appliances	N_laptop_PCs	2,425	0.34	1.3e-67
85	People	Gender_N_females	2,560	0.33	4.6e-67
31	Appliances	N_game_consoles	2,425	0.33	2.4e-63
15	Accommodation	Is_house	2,349	0.31	6.1e-55
22	Appliances	Lighting_N_halogen	2,425	0.29	6.8e-49
46	Appliances	N_video_players	2,425	0.27	2.4e-42
41	Appliances	N_set_top_boxes	2,425	0.27	3e-42
81	People	Age_N_18_to_34	2,560	0.25	1.3e-38
36	Appliances	N_ovens_electric	2,425	0.24	5.3e-34
72	Behaviour_normal	Work_from_home	2,333	0.24	3.8e-32
153	Response_limiter	Childen_routine	447	0.23	7.9e-07
28	Appliances	N_freezers	2,425	0.23	4.5e-30
82	People	Age_N_35_to_54	2,560	0.23	5.2e-31
80	People	Age_N_12_to_17	2,560	0.22	1.3e-30
25	Appliances	Lighting_N_traditional	2,425	0.22	4.1e-28
79	People	Age_N_0_to_11	2,560	0.19	2e-21
35	Appliances	N_microwaves	2,425	0.17	2e-16
106	Primary	SF_low_wd	887	0.16	7.6e-07
105	Primary	SF_low_price	887	0.16	1.3e-06
8	Accommodation	Has_loft_insulation	2,023	0.16	1e-12
91	Primary	All_low_price	887	0.15	3.8e-06
4	Accommodation	Has_floor_insulation	1,370	0.13	1.4e-06
11	Accommodation	Heating_interface_set_times	2,256	0.13	6.4e-10
42	Appliances	N_showers_electric	2,425	0.13	2.9e-10
38	Appliances	N_portable_electric_heaters	2,425	0.13	3e-10
48	Appliances	N_washing_machines	2,425	0.13	3e-10
6	Accommodation	Has_hot_water_tank	1,858	0.12	3.4e-07
23	Appliances	Lighting_N_LED	2,425	0.11	1.7e-08
83	People	Age_N_55_to_74	2,560	0.1	1.1e-07
29	Appliances	N_fridge_freezers	2,425	0.1	2.5e-07
30	Appliances	N_fridges	2,425	0.092	5.2e-06
10	Accommodation	Heating_interface_manual	2,256	-0.097	4e-06
93	Primary	CM_high_price	887	-0.16	1.6e-06
103	Primary	SF_high_wd_18_21	887	-0.17	6.1e-07
1	Accommodation	ACORN_group	2,560	-0.17	1.3e-18
99	Primary	SF_high_wd	887	-0.2	1.3e-09
89	Primary	All_high_price	887	-0.21	1.8e-10
98	Primary	SF_high_price	887	-0.22	7e-11
102	Primary	SF_high_wd_10_17	887	-0.22	1.9e-11
104	Primary	SF_high_we	887	-0.22	1.4e-11
101	Primary	SF_high_wd_07_10	887	-0.25	2.6e-14
14	Accommodation	Is_flat	2,349	-0.31	6.1e-55

Table B.10: Table of significant correlations against root variable DR_training_index.

Index	Group	Alias	N	r	p
95	Primary	DR_training_index	887	1	0
88	Primary	All_default_price	887	-0.15	5.4e-06

Table B.11: Table of significant correlations against root variable Engagement_rank.

Index	Group	Alias	N	r	p
90	Primary	All_high_price_relative	887	0.83	1.6e-279
89	Primary	All_high_price	887	0.77	9.8e-206
98	Primary	SF_high_price	887	0.72	3.4e-159
99	Primary	SF_high_wd	887	0.7	5.6e-147
93	Primary	CM_high_price	887	0.65	5e-119
102	Primary	SF_high_wd_10_17	887	0.6	3.9e-96
103	Primary	SF_high_wd_18_21	887	0.54	3.3e-71
104	Primary	SF_high_we	887	0.43	3.4e-42
101	Primary	SF_high_wd_07_10	887	0.41	3.7e-39
159	Response_limiter	Forget_price_changes	621	0.31	4.3e-15
100	Primary	SF_high_wd_00_07	887	0.3	7.3e-20
155	Response_limiter	Complexity	613	0.25	6.3e-10
162	Response_limiter	Knowing_when_rates_change	625	0.23	4.4e-09
193	Trial_impressions	Signup_If_longer_simpler_events	616	0.21	1.1e-07
160	Response_limiter	Increaseing_elec_usage	608	0.21	2.1e-07
164	Response_limiter	Savings_too_small	616	0.21	2.3e-07
176	Trial_impressions	Guarantee_meant_less_effort	627	0.2	2.8e-07
165	Trial_impressions	Too_much_hassle	633	0.2	3.9e-07
156	Response_limiter	Complicated_changes	619	0.19	1.1e-06
157	Response_limiter	Fixed_appliance_use_routine	608	0.19	2.8e-06
66	Behaviour_inflexible	Washing_machine	611	0.19	3.3e-06
88	Primary	All_default_price	887	0.18	2.8e-08
173	Trial_impressions	Frustrating_lack_of_effort_from_some	591	0.18	8.2e-06
166	Trial_impressions	Convenience_trumps_saving	634	0.18	4e-06
158	Response_limiter	Fixed_appliance_users	604	0.18	6.9e-06
199	Trial_impressions	Too_complex	640	0.18	5.5e-06
17	Accommodation	N_rooms	832	-0.17	1.6e-06
182	Trial_impressions	Often_talked_about_it	614	-0.18	6.8e-06
135	Report_low	Ironing	558	-0.19	8.6e-06
121	Report_high	Night	565	-0.19	5.7e-06
179	Trial_impressions	In_home_display_useful	627	-0.22	4.4e-08
142	Report_low	Washing_machine	567	-0.22	1.4e-07
146	Response_helper	Calendar_or_notepad	578	-0.22	7.6e-08
180	Trial_impressions	Low_easy	636	-0.23	7.7e-09
127	Report_high	Weekday_evenings	575	-0.24	9.7e-09
149	Response_helper	In_home_display	634	-0.24	1.3e-09
53	Behaviour_change	Substitute_fuel	589	-0.25	7.5e-10
107	Primary	SF_low_wd_00_07	887	-0.25	1.4e-14
145	Report_low	Weekday_mornings	558	-0.28	9.8e-12
128	Report_high	Weekday_mornings	579	-0.29	8.5e-13
140	Report_low	Sundays	564	-0.3	4.4e-13
123	Report_high	Sundays	582	-0.3	1.8e-13
115	Report_high	Electric_oven	426	-0.3	1.7e-10
139	Report_low	Saturdays	576	-0.31	3.6e-14
118	Report_high	Ironing	588	-0.31	4.9e-15
144	Report_low	Weekday_evenings	558	-0.32	1.2e-14
116	Report_high	Electric_shower	209	-0.32	2.4e-06
122	Report_high	Saturdays	600	-0.32	4.6e-16
141	Report_low	Tumble_drier	247	-0.32	1.6e-07
143	Report_low	Weekday_afternoons	557	-0.33	7.3e-16
113	Report_high	Electric_heater	199	-0.34	7.3e-07
125	Report_high	Washing_machine	597	-0.36	5.5e-20

Continued on next page

Continued from previous page

Index	Group	Alias	N	r	p
111	Primary	SF_low_we	887	-0.36	2.8e-29
110	Primary	SF_low_wd_18_21	887	-0.37	2.9e-30
126	Report_high	Weekday_afternoons	583	-0.37	1.3e-20
112	Report_high	Dishwasher	283	-0.4	2.2e-12
108	Primary	SF_low_wd_07_10	887	-0.4	3.7e-36
91	Primary	All_low_price	887	-0.45	1.3e-46
124	Report_high	Tumble_drier	256	-0.48	3.9e-17
92	Primary	All_low_price_relative	887	-0.49	1.7e-57
109	Primary	SF_low_wd_10_17	887	-0.5	1.6e-60
106	Primary	SF_low_wd	887	-0.51	2.1e-62
105	Primary	SF_low_price	887	-0.53	4.8e-68

Table B.12: Table of significant correlations against root variable Is_dToU.

Index	Group	Alias	N	r	p
97	Primary	Is_dToU	2,560	1	0
1	Accommodation	ACORN_group	2,560	-0.087	9.5e-06

Table B.13: Table of significant correlations against root variable SF_high_price.

Index	Group	Alias	N	r	p
98	Primary	SF_high_price	887	1	0
99	Primary	SF_high_wd	887	0.96	0
89	Primary	All_high_price	887	0.89	0
102	Primary	SF_high_wd_10_17	887	0.83	1.2e-268
90	Primary	All_high_price_relative	887	0.81	2.1e-240
103	Primary	SF_high_wd_18_21	887	0.72	8.7e-161
96	Primary	Engagement_rank	887	0.72	3.4e-159
101	Primary	SF_high_wd_07_10	887	0.6	6e-96
104	Primary	SF_high_we	887	0.58	8.7e-88
93	Primary	CM_high_price	887	0.56	1.3e-77
88	Primary	All_default_price	887	0.48	2e-54
100	Primary	SF_high_wd_00_07	887	0.47	7.4e-51
14	Accommodation	Is_flat	814	0.21	5.9e-10
159	Response_limiter	Forget_price_changes	621	0.19	2.9e-06
27	Appliances	N_dishwashers	837	-0.16	4.4e-06
26	Appliances	N_desktop_PCs	837	-0.16	3.7e-06
42	Appliances	N_showers_electric	837	-0.16	3.4e-06
110	Primary	SF_low_wd_18_21	887	-0.19	1.1e-08
87	People	N_of	797	-0.2	1.6e-08
144	Report_low	Weekday_evenings	558	-0.2	1.7e-06
145	Report_low	Weekday_mornings	558	-0.2	1.2e-06
125	Report_high	Washing_machine	597	-0.2	4.4e-07
128	Report_high	Weekday_mornings	579	-0.21	5e-07
92	Primary	All_low_price_relative	887	-0.21	3e-10
118	Report_high	Ironing	588	-0.21	2.9e-07
15	Accommodation	Is_house	814	-0.21	5.9e-10
91	Primary	All_low_price	887	-0.22	7.5e-11
94	Primary	Consumption_annual	887	-0.22	7e-11
143	Report_low	Weekday_afternoons	557	-0.22	1.7e-07
108	Primary	SF_low_wd_07_10	887	-0.23	9.4e-12
45	Appliances	N_tumble_driers	837	-0.23	1.3e-11
126	Report_high	Weekday_afternoons	583	-0.25	9.2e-10
111	Primary	SF_low_we	887	-0.25	1.9e-14
17	Accommodation	N_rooms	832	-0.26	3.5e-14
16	Accommodation	N_bedrooms	832	-0.27	4e-15
112	Report_high	Dishwasher	283	-0.27	3.4e-06
106	Primary	SF_low_wd	887	-0.28	4.4e-17
105	Primary	SF_low_price	887	-0.3	1.3e-19
109	Primary	SF_low_wd_10_17	887	-0.3	3.4e-20
124	Report_high	Tumble_drier	256	-0.33	3.1e-08

Table B.14: Table of significant correlations against root variable SF_high_wd.

Index	Group	Alias	N	r	p
99	Primary	SF_high_wd	887	1	0
98	Primary	SF_high_price	887	0.96	0
89	Primary	All_high_price	887	0.87	0
102	Primary	SF_high_wd_10_17	887	0.8	5e-230
90	Primary	All_high_price_relative	887	0.79	9.6e-227
103	Primary	SF_high_wd_18_21	887	0.73	3.2e-167
96	Primary	Engagement_rank	887	0.7	5.6e-147
93	Primary	CM_high_price	887	0.56	9.1e-80
101	Primary	SF_high_wd_07_10	887	0.54	2.8e-73
88	Primary	All_default_price	887	0.48	1.6e-53
100	Primary	SF_high_wd_00_07	887	0.44	2.1e-44
104	Primary	SF_high_we	887	0.4	8.5e-37
14	Accommodation	Is_flat	814	0.21	7.4e-10
27	Appliances	N_dishwashers	837	-0.17	1.3e-06
26	Appliances	N_desktop_PCs	837	-0.17	1e-06
125	Report_high	Washing_machine	597	-0.18	8.2e-06
118	Report_high	Ironing	588	-0.18	9.2e-06
92	Primary	All_low_price_relative	887	-0.18	3.1e-08
91	Primary	All_low_price	887	-0.19	1.7e-08
87	People	N_of	797	-0.19	8.6e-08
128	Report_high	Weekday_mornings	579	-0.19	3.9e-06
145	Report_low	Weekday_mornings	558	-0.19	5.1e-06
110	Primary	SF_low_wd_18_21	887	-0.19	7.4e-09
94	Primary	Consumption_annual	887	-0.2	1.3e-09
143	Report_low	Weekday_afternoons	557	-0.21	9.7e-07
108	Primary	SF_low_wd_07_10	887	-0.21	1.6e-10
15	Accommodation	Is_house	814	-0.21	7.4e-10
45	Appliances	N_tumble_driers	837	-0.22	5.4e-11
126	Report_high	Weekday_afternoons	583	-0.23	1.8e-08
111	Primary	SF_low_we	887	-0.24	1.4e-13
17	Accommodation	N_rooms	832	-0.25	3.8e-13
16	Accommodation	N_bedrooms	832	-0.26	3.1e-14
112	Report_high	Dishwasher	283	-0.26	8.9e-06
106	Primary	SF_low_wd	887	-0.26	7.8e-16
105	Primary	SF_low_price	887	-0.28	4.1e-18
109	Primary	SF_low_wd_10_17	887	-0.29	8.8e-19
124	Report_high	Tumble_drier	256	-0.31	4.3e-07

Table B.15: Table of significant correlations against root variable SF_high_wd_00_07.

Index	Group	Alias	N	r	p
100	Primary	SF_high_wd_00_07	887	1	0
98	Primary	SF_high_price	887	0.47	7.4e-51
99	Primary	SF_high_wd	887	0.44	2.1e-44
89	Primary	All_high_price	887	0.4	1.2e-35
90	Primary	All_high_price_relative	887	0.37	1.1e-31
104	Primary	SF_high_we	887	0.35	1e-26
88	Primary	All_default_price	887	0.32	3.4e-23
101	Primary	SF_high_wd_07_10	887	0.3	4.9e-20
102	Primary	SF_high_wd_10_17	887	0.3	6.3e-20
96	Primary	Engagement_rank	887	0.3	7.3e-20
103	Primary	SF_high_wd_18_21	887	0.26	4.5e-15
93	Primary	CM_high_price	887	0.25	8.5e-14
111	Primary	SF_low_we	887	-0.15	9.1e-06

Table B.16: Table of significant correlations against root variable SF_high_wd_07_10.

Index	Group	Alias	N	r	p
101	Primary	SF_high_wd_07_10	887	1	0
98	Primary	SF_high_price	887	0.6	6e-96
99	Primary	SF_high_wd	887	0.54	2.8e-73
89	Primary	All_high_price	887	0.52	2.3e-66
104	Primary	SF_high_we	887	0.52	1.6e-65
102	Primary	SF_high_wd_10_17	887	0.47	6.4e-53
90	Primary	All_high_price_relative	887	0.45	7.4e-47
96	Primary	Engagement_rank	887	0.41	3.7e-39
103	Primary	SF_high_wd_18_21	887	0.33	4.6e-25
93	Primary	CM_high_price	887	0.33	6.2e-24
88	Primary	All_default_price	887	0.31	5.9e-22
100	Primary	SF_high_wd_00_07	887	0.3	4.9e-20
27	Appliances	N_dishwashers	837	-0.16	5.4e-06
45	Appliances	N_tumble_driers	837	-0.16	4.9e-06
92	Primary	All_low_price_relative	887	-0.18	4.8e-08
182	Trial_impressions	Often_talked_about_it	614	-0.19	3.2e-06
108	Primary	SF_low_wd_07_10	887	-0.19	1e-08
111	Primary	SF_low_we	887	-0.2	2.7e-09
87	People	N_of	797	-0.2	6.5e-09
109	Primary	SF_low_wd_10_17	887	-0.2	8.2e-10
91	Primary	All_low_price	887	-0.2	6.9e-10
106	Primary	SF_low_wd	887	-0.2	6.4e-10
16	Accommodation	N_bedrooms	832	-0.21	1.2e-09
17	Accommodation	N_rooms	832	-0.22	7.2e-11
105	Primary	SF_low_price	887	-0.23	9.8e-12
94	Primary	Consumption_annual	887	-0.25	2.6e-14

Table B.17: Table of significant correlations against root variable SF_high_wd_10_17.

Index	Group	Alias	N	r	p
102	Primary	SF_high_wd_10_17	887	1	0
98	Primary	SF_high_price	887	0.83	1.2e-268
99	Primary	SF_high_wd	887	0.8	5e-230
89	Primary	All_high_price	887	0.75	4.6e-180
90	Primary	All_high_price_relative	887	0.66	2.2e-121
96	Primary	Engagement_rank	887	0.6	3.9e-96
104	Primary	SF_high_we	887	0.49	9.7e-58
93	Primary	CM_high_price	887	0.48	4.5e-54
101	Primary	SF_high_wd_07_10	887	0.47	6.4e-53
103	Primary	SF_high_wd_18_21	887	0.45	1.4e-47
88	Primary	All_default_price	887	0.4	2.8e-36
100	Primary	SF_high_wd_00_07	887	0.3	6.3e-20
14	Accommodation	Is_flat	814	0.18	1.3e-07
27	Appliances	N_dishwashers	837	-0.16	2.5e-06
110	Primary	SF_low_wd_18_21	887	-0.17	5.8e-07
6	Accommodation	Has_hot_water_tank	658	-0.17	8.8e-06
36	Appliances	N_ovens_electric	837	-0.17	4.6e-07
15	Accommodation	Is_house	814	-0.18	1.3e-07
92	Primary	All_low_price_relative	887	-0.18	3.3e-08
144	Report_low	Weekday_evenings	558	-0.19	7.9e-06
91	Primary	All_low_price	887	-0.19	8.2e-09
128	Report_high	Weekday_mornings	579	-0.2	1.8e-06
108	Primary	SF_low_wd_07_10	887	-0.2	3.3e-09
147	Response_helper	Coordinating_with_household_members	523	-0.21	1.8e-06
87	People	N_of	797	-0.21	3.4e-09
111	Primary	SF_low_we	887	-0.22	2.4e-11
94	Primary	Consumption_annual	887	-0.22	1.9e-11
145	Report_low	Weekday_mornings	558	-0.22	7.1e-08
17	Accommodation	N_rooms	832	-0.23	2.5e-11
16	Accommodation	N_bedrooms	832	-0.23	1.7e-11
143	Report_low	Weekday_afternoons	557	-0.24	4.9e-09
106	Primary	SF_low_wd	887	-0.25	3.8e-14
126	Report_high	Weekday_afternoons	583	-0.25	5.8e-10
109	Primary	SF_low_wd_10_17	887	-0.25	1.3e-14
105	Primary	SF_low_price	887	-0.26	1.5e-15
45	Appliances	N_tumble_driers	837	-0.29	1.6e-17

Table B.18: Table of significant correlations against root variable SF_high_wd_18_21.

Index	Group	Alias	N	r	p
103	Primary	SF_high_wd_18_21	887	1	0
99	Primary	SF_high_wd	887	0.73	3.2e-167
98	Primary	SF_high_price	887	0.72	8.7e-161
89	Primary	All_high_price	887	0.66	6.6e-124
90	Primary	All_high_price_relative	887	0.61	4e-97
96	Primary	Engagement_rank	887	0.54	3.3e-71
93	Primary	CM_high_price	887	0.46	7.2e-49
102	Primary	SF_high_wd_10_17	887	0.45	1.4e-47
104	Primary	SF_high_we	887	0.37	8.1e-31
101	Primary	SF_high_wd_07_10	887	0.33	4.6e-25
88	Primary	All_default_price	887	0.32	2.2e-23
100	Primary	SF_high_wd_00_07	887	0.26	4.5e-15
14	Accommodation	Is_flat	814	0.18	4.3e-07
92	Primary	All_low_price_relative	887	-0.15	3.6e-06
110	Primary	SF_low_wd_18_21	887	-0.16	1.8e-06
22	Appliances	Lighting_N_halogen	837	-0.16	1.9e-06
91	Primary	All_low_price	887	-0.16	9.2e-07
94	Primary	Consumption_annual	887	-0.17	6.1e-07
15	Accommodation	Is_house	814	-0.18	4.3e-07
108	Primary	SF_low_wd_07_10	887	-0.2	3.6e-09
16	Accommodation	N_bedrooms	832	-0.2	7e-09
17	Accommodation	N_rooms	832	-0.2	2.3e-09
109	Primary	SF_low_wd_10_17	887	-0.22	6.8e-11
111	Primary	SF_low_we	887	-0.22	3.9e-11
106	Primary	SF_low_wd	887	-0.23	5.6e-12
105	Primary	SF_low_price	887	-0.24	1e-13
112	Report_high	Dishwasher	283	-0.27	3.7e-06

Table B.19: Table of significant correlations against root variable SF_high_we.

Index	Group	Alias	N	r	p
104	Primary	SF_high_we	887	1	0
98	Primary	SF_high_price	887	0.58	8.7e-88
101	Primary	SF_high_wd_07_10	887	0.52	1.6e-65
89	Primary	All_high_price	887	0.51	5.4e-63
102	Primary	SF_high_wd_10_17	887	0.49	9.7e-58
90	Primary	All_high_price_relative	887	0.45	1.9e-46
96	Primary	Engagement_rank	887	0.43	3.4e-42
99	Primary	SF_high_wd	887	0.4	8.5e-37
103	Primary	SF_high_wd_18_21	887	0.37	8.1e-31
100	Primary	SF_high_wd_00_07	887	0.35	1e-26
93	Primary	CM_high_price	887	0.31	9.3e-22
88	Primary	All_default_price	887	0.27	7.5e-17
87	People	N_of	797	-0.16	3.6e-06
111	Primary	SF_low_we	887	-0.18	1.3e-07
182	Trial_impressions	Often_talked_about_it	614	-0.18	5.9e-06
108	Primary	SF_low_wd_07_10	887	-0.18	4.4e-08
123	Report_high	Sundays	582	-0.18	7.3e-06
122	Report_high	Saturdays	600	-0.18	4.9e-06
16	Accommodation	N_bedrooms	832	-0.19	3.7e-08
128	Report_high	Weekday_mornings	579	-0.19	3.8e-06
139	Report_low	Saturdays	576	-0.2	1.5e-06
17	Accommodation	N_rooms	832	-0.2	7e-09
125	Report_high	Washing_machine	597	-0.2	7.4e-07
45	Appliances	N_tumble_driers	837	-0.21	1.2e-09
140	Report_low	Sundays	564	-0.21	3.3e-07
118	Report_high	Ironing	588	-0.22	6.4e-08
94	Primary	Consumption_annual	887	-0.22	1.4e-11
126	Report_high	Weekday_afternoons	583	-0.22	4.1e-08
92	Primary	All_low_price_relative	887	-0.24	4.2e-13
106	Primary	SF_low_wd	887	-0.24	1.1e-13
109	Primary	SF_low_wd_10_17	887	-0.25	3.5e-14
105	Primary	SF_low_price	887	-0.26	7.8e-15
91	Primary	All_low_price	887	-0.26	8e-16
124	Report_high	Tumble_drier	256	-0.28	5.8e-06

Table B.20: Table of significant correlations against root variable SF_low_price.

Index	Group	Alias	N	r	p
105	Primary	SF_low_price	887	1	0
106	Primary	SF_low_wd	887	0.97	0
109	Primary	SF_low_wd_10_17	887	0.84	1.1e-292
91	Primary	All_low_price	887	0.81	2.1e-241
92	Primary	All_low_price_relative	887	0.79	2.5e-225
110	Primary	SF_low_wd_18_21	887	0.7	2.3e-148
108	Primary	SF_low_wd_07_10	887	0.66	9e-121
111	Primary	SF_low_we	887	0.52	1.9e-65
107	Primary	SF_low_wd_00_07	887	0.44	1.7e-44
124	Report_high	Tumble_drier	256	0.37	6.4e-10
141	Report_low	Tumble_drier	247	0.32	2.6e-07
112	Report_high	Dishwasher	283	0.32	3.9e-08
126	Report_high	Weekday_afternoons	583	0.27	1.7e-11
125	Report_high	Washing_machine	597	0.26	4.4e-11
122	Report_high	Saturdays	600	0.26	4.4e-11
143	Report_low	Weekday_afternoons	557	0.26	3.4e-10
182	Trial_impressions	Often_talked_about_it	614	0.24	1e-09
128	Report_high	Weekday_mornings	579	0.24	3.5e-09
139	Report_low	Saturdays	576	0.24	7.2e-09
123	Report_high	Sundays	582	0.24	7.1e-09
145	Report_low	Weekday_mornings	558	0.23	3.4e-08
140	Report_low	Sundays	564	0.23	4.9e-08
144	Report_low	Weekday_evenings	558	0.22	2.1e-07
142	Report_low	Washing_machine	567	0.2	1.6e-06
118	Report_high	Ironing	588	0.2	1.7e-06
17	Accommodation	N_rooms	832	0.17	3.8e-07
27	Appliances	N_dishwashers	837	0.17	3.9e-07
94	Primary	Consumption_annual	887	0.16	1.3e-06
16	Accommodation	N_bedrooms	832	0.16	4.2e-06
160	Response_limiter	Increaseing_elec_usage	608	-0.19	2.1e-06
164	Response_limiter	Savings_too_small	616	-0.2	9.8e-07
155	Response_limiter	Complexity	613	-0.2	1e-06
165	Trial_impressions	Too_much_hassle	633	-0.2	2.1e-07
101	Primary	SF_high_wd_07_10	887	-0.23	9.8e-12
159	Response_limiter	Forget_price_changes	621	-0.23	7.1e-09
103	Primary	SF_high_wd_18_21	887	-0.24	1e-13
93	Primary	CM_high_price	887	-0.25	1.6e-14
104	Primary	SF_high_we	887	-0.26	7.8e-15
102	Primary	SF_high_wd_10_17	887	-0.26	1.5e-15
90	Primary	All_high_price_relative	887	-0.28	1.7e-17
99	Primary	SF_high_wd	887	-0.28	4.1e-18
98	Primary	SF_high_price	887	-0.3	1.3e-19
89	Primary	All_high_price	887	-0.3	5.2e-20
96	Primary	Engagement_rank	887	-0.53	4.8e-68

Table B.21: Table of significant correlations against root variable SF_low_wd.

Index	Group	Alias	N	r	p
106	Primary	SF_low_wd	887	1	0
105	Primary	SF_low_price	887	0.97	0
109	Primary	SF_low_wd_10_17	887	0.82	3e-256
91	Primary	All_low_price	887	0.8	1.6e-233
92	Primary	All_low_price_relative	887	0.78	1.5e-216
110	Primary	SF_low_wd_18_21	887	0.68	8.5e-133
108	Primary	SF_low_wd_07_10	887	0.66	8.6e-124
107	Primary	SF_low_wd_00_07	887	0.46	4e-49
124	Report_high	Tumble_drier	256	0.37	6.9e-10
111	Primary	SF_low_we	887	0.35	5.8e-28
141	Report_low	Tumble_drier	247	0.3	1.8e-06
112	Report_high	Dishwasher	283	0.29	6.1e-07
126	Report_high	Weekday_afternoons	583	0.26	1.4e-10
125	Report_high	Washing_machine	597	0.24	1.1e-09
143	Report_low	Weekday_afternoons	557	0.23	1.9e-08
182	Trial_impressions	Often_talked_about_it	614	0.23	3.7e-09
128	Report_high	Weekday_mornings	579	0.23	1.9e-08
122	Report_high	Saturdays	600	0.23	1.8e-08
145	Report_low	Weekday_mornings	558	0.21	3.7e-07
123	Report_high	Sundays	582	0.21	4.1e-07
139	Report_low	Saturdays	576	0.2	7.6e-07
140	Report_low	Sundays	564	0.2	2.6e-06
144	Report_low	Weekday_evenings	558	0.19	3.4e-06
118	Report_high	Ironing	588	0.18	8.3e-06
27	Appliances	N_dishwashers	837	0.18	3e-07
94	Primary	Consumption_annual	887	0.16	7.6e-07
17	Accommodation	N_rooms	832	0.16	2.9e-06
88	Primary	All_default_price	887	0.15	7.9e-06
165	Trial_impressions	Too_much_hassle	633	-0.18	8.6e-06
155	Response_limiter	Complexity	613	-0.18	7.2e-06
164	Response_limiter	Savings_too_small	616	-0.19	2.4e-06
101	Primary	SF_high_wd_07_10	887	-0.2	6.4e-10
159	Response_limiter	Forget_price_changes	621	-0.21	1.3e-07
103	Primary	SF_high_wd_18_21	887	-0.23	5.6e-12
104	Primary	SF_high_we	887	-0.24	1.1e-13
93	Primary	CM_high_price	887	-0.25	7.8e-14
102	Primary	SF_high_wd_10_17	887	-0.25	3.8e-14
90	Primary	All_high_price_relative	887	-0.26	1.4e-15
99	Primary	SF_high_wd	887	-0.26	7.8e-16
98	Primary	SF_high_price	887	-0.28	4.4e-17
89	Primary	All_high_price	887	-0.29	2.1e-18
96	Primary	Engagement_rank	887	-0.51	2.1e-62

Table B.22: Table of significant correlations against root variable SF_low_wd_00_07.

Index	Group	Alias	N	r	p
107	Primary	SF_low_wd_00_07	887	1	0
106	Primary	SF_low_wd	887	0.46	4e-49
105	Primary	SF_low_price	887	0.44	1.7e-44
91	Primary	All_low_price	887	0.42	6.9e-40
92	Primary	All_low_price_relative	887	0.41	6.3e-39
110	Primary	SF_low_wd_18_21	887	0.28	4.7e-17
108	Primary	SF_low_wd_07_10	887	0.27	1.6e-16
109	Primary	SF_low_wd_10_17	887	0.25	6.9e-14
125	Report_high	Washing_machine	597	0.19	4.4e-06
88	Primary	All_default_price	887	0.16	1.1e-06
111	Primary	SF_low_we	887	0.16	2.6e-06
96	Primary	Engagement_rank	887	-0.25	1.4e-14

Table B.23: Table of significant correlations against root variable SF_low_wd_07_10.

Index	Group	Alias	N	r	p
108	Primary	SF_low_wd_07_10	887	1	0
106	Primary	SF_low_wd	887	0.66	8.6e-124
105	Primary	SF_low_price	887	0.66	9e-121
91	Primary	All_low_price	887	0.57	3.7e-84
92	Primary	All_low_price_relative	887	0.57	1e-81
109	Primary	SF_low_wd_10_17	887	0.53	8.9e-70
110	Primary	SF_low_wd_18_21	887	0.35	2.7e-27
112	Report_high	Dishwasher	283	0.32	3.8e-08
124	Report_high	Tumble_drier	256	0.31	3e-07
111	Primary	SF_low_we	887	0.3	2.8e-20
107	Primary	SF_low_wd_00_07	887	0.27	1.6e-16
125	Report_high	Washing_machine	597	0.24	1.4e-09
126	Report_high	Weekday_afternoons	583	0.18	7.5e-06
17	Accommodation	N_rooms	832	0.18	9e-08
16	Accommodation	N_bedrooms	832	0.18	1.2e-07
27	Appliances	N_dishwashers	837	0.15	8.6e-06
104	Primary	SF_high_we	887	-0.18	4.4e-08
101	Primary	SF_high_wd_07_10	887	-0.19	1e-08
103	Primary	SF_high_wd_18_21	887	-0.2	3.6e-09
102	Primary	SF_high_wd_10_17	887	-0.2	3.3e-09
99	Primary	SF_high_wd	887	-0.21	1.6e-10
93	Primary	CM_high_price	887	-0.22	1.3e-11
98	Primary	SF_high_price	887	-0.23	9.4e-12
159	Response_limiter	Forget_price_changes	621	-0.23	9.3e-09
90	Primary	All_high_price_relative	887	-0.24	9e-13
89	Primary	All_high_price	887	-0.25	5e-14
96	Primary	Engagement_rank	887	-0.4	3.7e-36

Table B.24: Table of significant correlations against root variable SF_low_wd_10_17.

Index	Group	Alias	N	r	p
109	Primary	SF_low_wd_10_17	887	1	0
105	Primary	SF_low_price	887	0.84	1.1e-292
106	Primary	SF_low_wd	887	0.82	3e-256
91	Primary	All_low_price	887	0.71	4e-152
92	Primary	All_low_price_relative	887	0.7	1.4e-143
108	Primary	SF_low_wd_07_10	887	0.53	8.9e-70
111	Primary	SF_low_we	887	0.45	2.6e-47
110	Primary	SF_low_wd_18_21	887	0.43	2.1e-43
124	Report_high	Tumble_drier	256	0.34	1.8e-08
112	Report_high	Dishwasher	283	0.28	1.7e-06
125	Report_high	Washing_machine	597	0.26	1.3e-10
122	Report_high	Saturdays	600	0.26	1.6e-10
107	Primary	SF_low_wd_00_07	887	0.25	6.9e-14
126	Report_high	Weekday_afternoons	583	0.24	2.7e-09
123	Report_high	Sundays	582	0.23	1.9e-08
128	Report_high	Weekday_mornings	579	0.22	4e-08
143	Report_low	Weekday_afternoons	557	0.22	7.9e-08
182	Trial_impressions	Often_talked_about_it	614	0.22	4e-08
139	Report_low	Saturdays	576	0.21	4.7e-07
118	Report_high	Ironing	588	0.21	4.2e-07
140	Report_low	Sundays	564	0.19	4.4e-06
145	Report_low	Weekday_mornings	558	0.19	5.2e-06
149	Response_helper	In_home_display	634	0.18	2.6e-06
27	Appliances	N_dishwashers	837	0.16	4.5e-06
17	Accommodation	N_rooms	832	0.16	5.2e-06
165	Trial_impressions	Too_much_hassle	633	-0.19	7.9e-07
155	Response_limiter	Complexity	613	-0.2	5.2e-07
101	Primary	SF_high_wd_07_10	887	-0.2	8.2e-10
103	Primary	SF_high_wd_18_21	887	-0.22	6.8e-11
159	Response_limiter	Forget_price_changes	621	-0.22	3.5e-08
104	Primary	SF_high_we	887	-0.25	3.5e-14
102	Primary	SF_high_wd_10_17	887	-0.25	1.3e-14
99	Primary	SF_high_wd	887	-0.29	8.8e-19
98	Primary	SF_high_price	887	-0.3	3.4e-20
93	Primary	CM_high_price	887	-0.3	7.8e-21
90	Primary	All_high_price_relative	887	-0.31	6.2e-21
89	Primary	All_high_price	887	-0.32	1.3e-23
96	Primary	Engagement_rank	887	-0.5	1.6e-60

Table B.25: Table of significant correlations against root variable SF_low_wd_18_21.

Index	Group	Alias	N	r	p
110	Primary	SF_low_wd_18_21	887	1	0
105	Primary	SF_low_price	887	0.7	2.3e-148
106	Primary	SF_low_wd	887	0.68	8.5e-133
92	Primary	All_low_price_relative	887	0.57	4e-83
91	Primary	All_low_price	887	0.56	2.2e-80
109	Primary	SF_low_wd_10_17	887	0.43	2.1e-43
111	Primary	SF_low_we	887	0.43	3.6e-42
108	Primary	SF_low_wd_07_10	887	0.35	2.7e-27
107	Primary	SF_low_wd_00_07	887	0.28	4.7e-17
126	Report_high	Weekday_afternoons	583	0.22	8.7e-08
143	Report_low	Weekday_afternoons	557	0.21	3.5e-07
128	Report_high	Weekday_mornings	579	0.19	2.6e-06
144	Report_low	Weekday_evenings	558	0.19	4.3e-06
125	Report_high	Washing_machine	597	0.18	9e-06
103	Primary	SF_high_wd_18_21	887	-0.16	1.8e-06
102	Primary	SF_high_wd_10_17	887	-0.17	5.8e-07
89	Primary	All_high_price	887	-0.18	1.2e-07
90	Primary	All_high_price_relative	887	-0.19	2.1e-08
98	Primary	SF_high_price	887	-0.19	1.1e-08
99	Primary	SF_high_wd	887	-0.19	7.4e-09
159	Response_limiter	Forget_price_changes	621	-0.2	5.9e-07
96	Primary	Engagement_rank	887	-0.37	2.9e-30

Table B.26: Table of significant correlations against root variable SF_low_we.

Index	Group	Alias	N	r	p
111	Primary	SF_low_we	887	1	0
105	Primary	SF_low_price	887	0.52	1.9e-65
109	Primary	SF_low_wd_10_17	887	0.45	2.6e-47
110	Primary	SF_low_wd_18_21	887	0.43	3.6e-42
92	Primary	All_low_price_relative	887	0.41	1.4e-37
91	Primary	All_low_price	887	0.4	2e-35
106	Primary	SF_low_wd	887	0.35	5.8e-28
108	Primary	SF_low_wd_07_10	887	0.3	2.8e-20
122	Report_high	Saturdays	600	0.27	1.6e-11
112	Report_high	Dishwasher	283	0.26	6.5e-06
139	Report_low	Saturdays	576	0.24	3.6e-09
142	Report_low	Washing_machine	567	0.23	3.6e-08
123	Report_high	Sundays	582	0.23	2.9e-08
140	Report_low	Sundays	564	0.22	7e-08
146	Response_helper	Calendar_or_notepad	578	0.22	1e-07
143	Report_low	Weekday_afternoons	557	0.21	3.1e-07
125	Report_high	Washing_machine	597	0.21	1.4e-07
144	Report_low	Weekday_evenings	558	0.21	4.2e-07
128	Report_high	Weekday_mornings	579	0.21	4.9e-07
126	Report_high	Weekday_afternoons	583	0.2	9.6e-07
180	Trial_impressions	Low_easy	636	0.2	3.6e-07
145	Report_low	Weekday_mornings	558	0.19	6e-06
182	Trial_impressions	Often_talked_about_it	614	0.18	6.5e-06
107	Primary	SF_low_wd_00_07	887	0.16	2.6e-06
100	Primary	SF_high_wd_00_07	887	-0.15	9.1e-06
104	Primary	SF_high_we	887	-0.18	1.3e-07
93	Primary	CM_high_price	887	-0.19	2.4e-08
101	Primary	SF_high_wd_07_10	887	-0.2	2.7e-09
155	Response_limiter	Complexity	613	-0.2	5.6e-07
159	Response_limiter	Forget_price_changes	621	-0.21	2.3e-07
166	Trial_impressions	Convenience_trumps_saving	634	-0.21	7.9e-08
165	Trial_impressions	Too_much_hassle	633	-0.22	3.1e-08
160	Response_limiter	Increaseing_elec_usage	608	-0.22	4.8e-08
103	Primary	SF_high_wd_18_21	887	-0.22	3.9e-11
102	Primary	SF_high_wd_10_17	887	-0.22	2.4e-11
89	Primary	All_high_price	887	-0.24	7.9e-13
99	Primary	SF_high_wd	887	-0.24	1.4e-13
90	Primary	All_high_price_relative	887	-0.25	7.3e-14
98	Primary	SF_high_price	887	-0.25	1.9e-14
96	Primary	Engagement_rank	887	-0.36	2.8e-29