Who's On Our Wires?

Methodology Report: Low Carbon Technologies

March 2013

Prepared by:

Centre for Sustainable Energy

3 St Peter's Court

Bedminster Parade

Bristol BS3 4AQ

Tel: 0117 9341 400

Fax: 0117 9341 410

Email: info@cse.org.uk

Web: www.cse.org.uk

Registered charity no.298740

Contents

1. Introduction ... 3
2. Project aims and outline ... 3
Project Aims ... 3
Outline ... 3
3. Methodology overview ... 4
4. Housing Assessment Model Methodology ... 5
The Housing Assessment Model ... 5
Developing the dataset ... 6
Assessing baseline energy efficiency levels ... 6
Modelling improvement scenarios ... 6
The importance of SAP ... 7
5. Assumptions ... 7
Solar PV - Assumptions about which households will install ... 7
Solar PV - Assumptions about capacity ... 9
Solar PV - Further notes ... 9
Heat pumps - Assumptions about which households will install ... 10
Heat pumps - Assumptions about demand ... 12
Direct Electric Heating - Assumptions about which households will install ... 13
Direct Electric Heating - Assumptions about capacity ... 13
Electric Vehicles ... 13
References ... 15
6. Matching households to substations ... 16
WPD data ... 16
Experian / Housing Assessment Model ... 16
Matching the two datasets ... 17
Match quality indicators ... 18
Records with coordinate errors ... 19
7. Substation capacity and maximum demands ... 20
8. Existing generation ... 20
9. Bringing it all together: outputs ... 21
10. Appendices ... 23
Appendix 1: Description of fields in main summary table ... 23
Appendix 2: Mosaic Public Sector Groups ... 25

1. Introduction

Western Power Distribution (WPD) commissioned the Centre for Sustainable Energy (CSE) to help to develop a better picture of the households connected to WPD's network. The 'Who's On Our Wires?' project combines WPD's datasets about substations and customer locations with a range of other datasets to obtain an insight into the characteristics of customers attached to specific substations and to use these to make predictions about what they will do in the future. This in turn can inform predictions about changes in load and the priority areas for network reinforcement. This is the subject of the current report.

The work can also help to locate the most vulnerable customers, who can be helped through affordable warmth programmes or added to the Priority Services Register. This is the subject of a second methodology report which is provided alongside the current report.

2. Project aims and outline

Project Aims

The project aims were:

- To create a database of customers (identified by MPAN and postcode only - no personal information has been used) and the substations to which they are attached.
- To use the database to inform the identification of substations likely to experience the highest increases in load over the period to the early 2020s (that is, covering the current price control period and the subsequent one, RIIO-ED1), based on household data and (to a lesser extent) data about businesses. The project does not set out to model load, although where possible installed capacity figures are estimated. The project is focussed on looking at the characteristics of individual households which are attached to each substation, to identify the households most likely to increase demand or on-site generation. This is the subject of the current report.
- To use the database to support targeting of affordable warmth and other social programme initiatives to customers. This is the subject of a second report.

Outline

The core elements of the part of the work relating to load on the network are WPD datasets and outputs from CSE's own Housing Assessment Model.

Key information from WPD includes:

- Customer profile class, energisation status, location and substation to which the customer is connected
- Substation capacity and maximum demands
- Connected generation (where the installer has informed WPD of its existence)

The Housing Assessment Model uses household-level information on property characteristics and socio-demographic type to assess whether the household would take up certain technologies, based on a set of assumptions which will be set out in this document. The technologies considered are:

- Ground source heat pumps
- Air source heat pumps
- Direct electric heating
- Electric Vehicles
- Solar PV

The database also includes information about distributed generation from Ofgem's Feed-In Tariff installation report. WPD's data on generation contains fewer installations than the Feed-In Tariff report. In some areas the difference between the two sources is quite large. The difference between the two sources is likely to be due to timing differences between installers' registrations for the Feed-In Tariff and notification to the DNO, and timing of record updates. However, it should be noted that this difference exists and that is why both sources have been reported in the outputs of this work.

3. Methodology overview

The work was carried out in three stages. First, a test run was carried out on a small area covered by one primary substation, to provide a proof of concept. Next, the methodology was applied to the South West distribution area and following this, to the larger dataset of the South Wales, East Midlands and West Midlands licence areas. The modelling work was done using SQL Server and ArcGIS, with the results tables being output into spreadsheets.

The work involves modelling the potential actions of households over the next decade and the resulting impacts on substations. The main dataset used in modelling household-level actions was the Experian[1] data. Therefore all the results from the modelling of future installations are output at the level of a household as identified in the Experian dataset.

However, the aim of the work was to look at impacts at substation level and so these results had to be aggregated at substation level. The WPD data shows which customers are connected to which substations and so, to aggregate the modelling results to substation level, the households in the Experian dataset had to be linked to domestic customers in the WPD dataset.

Throughout this report, the term 'household' is used when referring to data originating in the Experian dataset, and the term 'customer' is used to mean an MPAN as identified in the WPD data.

In the real world, there should be an exact match between the WPD profile class 1 and 2 customers and the Experian households in the same areas, because they both represent the same real world

[1] Experian is a company which is best known for credit checking but also holds and maintains a large number of datasets and models of both businesses and households. This work uses Experian's data about homes (built form, number of bedrooms, gas connection and tenure), and its Mosaic Classification (public sector version), which is a socio-demographic classification. For more information on Mosaic Public Sector, see http://publicsector.experian.co.uk/Products/~/media/Brochures/MosaicPublicSector_Brochure_final.ashx

entities. However, both datasets are representations of reality with some inaccuracies and there is not an exact match. To match the datasets, a Geographic Information System was used to match on the basis of location, because this was the attribute common to both datasets. When coordinates did not match exactly, a household was matched to the nearest WPD customer record. The household was then given the substation of its matched WPD record. Where there was no match within 20 metres, the household was initially excluded, but where this process left substations with fewer domestic customers than they should have had, households which were within 200m of their customer record were added incrementally until the substation had the right number of households associated with it. The matching process is described in more detail in section 6.

As well as modelled future installations, the outputs from the work include estimates of the remaining capacity on distribution substations, derived from WPD data on substation capacity and WPD's own modelled maximum demands. For more information, see section 7. Existing generation on each substation, as represented by WPD's own data, is shown, and Ofgem's Feed-In Tariff Installation reports are also used to estimate generation on each substation. This is described in section 8.

4. Housing Assessment Model Methodology

CSE's Housing Assessment Model (HAM) is a tool that models the energy performance and energy efficiency (SAP rating[2]) of dwellings at address level. It uses information about individual properties (from Experian) to assess the property's energy use, suitability for improvement measures such as wall insulation, and energy use after improvement measures. The data about individual properties and their inhabitants that the model uses and calculates is used for modelling PV, heat pump, and electric heating uptake, which are largely influenced by property type and this is why the HAM has been used in this project.

This section of the report gives an overview of how the HAM works. The following section describes the assumptions that are applied to data within the HAM in order to make an assessment of each household's likelihood of installing the different technologies.

The Housing Assessment Model

The HAM uses Experian variables, a classification analysis to predict further variables, and assumptions from the 'Reduced data SAP' assessment (RdSAP). The model determines all possible improvement measures applicable to each house that will increase energy efficiency and reduce energy bills and carbon emissions. The model can also calculate the financial implications of each package of measures, including the use of Green Deal and ECO finance (Carbon Reduction and Affordable Warmth). The HAM is normally used to provide analysis for local authorities, social housing providers and community groups which are interested in the improvement potential of the local housing stock. For the On Our Wires project, the HAM was used in a slightly different way: to

[2] The Standard Assessment Procedure (SAP) is the methodology used by the Department of Energy & Climate Change (DECC) to assess and compare the energy and environmental performance of dwellings. For more information see https://www.gov.uk/standard-assessment-procedure

identify types of households which would be likely to install specific technologies, including those which would need improvement to a specific level of energy efficiency before installing.

Developing the dataset

A database of the housing stock in the South West, Midlands and Wales was developed using address level Experian data, which provides key descriptive data about the nature of dwellings and occupants at household -level (i.e. address level), specifically: age of property, tenure, dwelling type, number of bedrooms, mains gas connection and Mosaic group[3] (e.g. see Table 1).

Table 1: Building an address-level dataset using Experian data (data for illustrative purposes only)

Address	Postcode	Experian_ N.Beds	Experian_Property Age	Experian_Dwelling type	Experian_Tenure	Experian_ Mains gas	Experian_Mosaic group
1	BA14 0XX	2	post-1980	Semi-detached	Owner occupied	No	I Lower income workers…
2	BA14 0XX	3	1955-1979	Terraced	Owner occupied	Yes	I Lower income workers…
3	BA14 0XX	3	1920-1945	Terraced	Private rented	Yes	K Residents with sufficient incomes…
4	BA14 0XX	4	1946-1954	Terraced	Owner occupied	Yes	K Residents with sufficient incomes…
5	BA14 0XX	3	1955-1979	Terraced	Private rented	Yes	K Residents with sufficient incomes…
6	BA14 0XX	4	1920-1945	Detached	Owner occupied	No	K Residents with sufficient incomes…

In order to assess the baseline energy efficiency of this housing stock, some key additional information - essential to perform the SAP assessment - is needed about each property, and this must be imputed into the dataset. These additional variables are wall type and insulation levels; loft insulation levels; and heating system for each property in the dataset.

The English Housing Survey (2010)[4] is used to derive a predictive model for each of the additional variables, based on what is known about the properties from the Experian data. The resulting dataset should therefore be viewed as a modelled representation, albeit one based on significantly robust underlying data and predictive models.

The RdSAP methodology[5] is used to complete any other missing data needed for the baseline energy efficiency assessment, for example ventilation details such as the presence of extractor fans and draught lobbies (which are determined by intrinsic characteristics such as dwelling type, age and size). Any missing data was imputed by making assumptions based on the dwelling type and age and using variables from similar properties with complete data.

Assessing baseline energy efficiency levels

The baseline energy assessment uses an energy calculation based on SAP 2009 (the most current SAP methodology) with a complementary algorithm for assessing energy required for cooking. This generates a standard SAP rating, plus energy requirements (kWh) for heating, hot water, lights and appliances and cooking at address level for housing stock databases.

Modelling improvement scenarios

Having assessed baseline energy efficiency levels and energy requirements for every property in a dataset, the HAM then makes an assessment of each property to determine the applicability and

[3] Mosaic is a socio-demographic classification. For more information on Mosaic Public Sector, see http://publicsector.experian.co.uk/Products/~/media/Brochures/MosaicPublicSector_Brochure_final.ashx
[4] https://www.gov.uk/government/organisations/department-for-communities-and-local-government/series/english-housing-survey
[5] https://www.gov.uk/standard-assessment-procedure

potential impact of different improvement measures according to a set of user-defined criteria. For this work, the improvement measures were limited to the main measures of cavity wall insulation, solid wall insulation, loft insulation / top-up, and heating system improvement.

The process of evaluating each package involves a calculation that takes account of: total package costs (including the level of subsidy available from any given policy); individual costs of each measure; the change in SAP rating and associated reduction in household energy requirements, energy bill and CO_2 emissions. These results are used to determine whether the improvement measures are suitable for the property.

The importance of SAP

In the case of PV, the full feed-in tariff will only be paid where the property has SAP D or higher, while heat pumps work best in energy efficient properties. Being able to model the SAP rating of a dwelling, and the possibility of improving the SAP rating where it is below D, is therefore very useful for modelling uptake of these technologies.

The SAP assessment gives a property a numerical score, which is then banded into a rating from A to G (with A being the best). According to data from the 2010 English House Condition Survey (EHS), the distribution of SAP ratings across the housing stock is as shown in Table 2.

Table 2: Distribution of SAP ratings across the housing stock in England, 2010

Energy efficiency rating band (EHS) SAP 2009)	A	B	C	D	E	F	G	Total
Percentage of housing stock	0.00	0.10	11.70	46.90	31.70	7.50	2.20	100.00

5. Assumptions

Solar PV - Assumptions about which households will install

The assumptions for solar PV are:

- Flats will not install PV because they do not have the roof space (although in reality some social housing blocks may install PV).
- All other property forms (terraced, detached, semi-detached, bungalow) are considered eligible for PV.
- Private rented properties will not install PV. Keirstead (2006) found that 97% of solar PV installations were in owner occupied properties, and although this study was carried out before the Feed-In Tariff (FIT) was introduced, it remains a valid assumption because the installation and the associated Feed-In Tariff finances are more complicated when the property is privately rented.
- Socially rented properties are considered suitable for PV (if they are not flats) because social landlords are able to use economies of scale to bulk-buy solar systems for their housing stock. They are also motivated to improve their housing stock and to lower fuel bills for residents.

- It is assumed that all roofs have a suitable orientation for a solar panel. This simplifying assumption is necessary because we do not know the actual orientation of the property.
- Properties must be rated SAP D or above, or they must be capable of improvement to SAP D rating with cavity / solid wall insulation, loft insulation, and heating system upgrade. This is because in order to claim the full Feed-In Tariff, properties must be rated SAP D or above.

The above criteria cover a large proportion of the housing stock. In order to get some more detail, eligible households have been divided into different 'layers'. The layers are ordered from those households considered more likely to install PV, to those which are considered less likely to install. The layers are created from a mixture of what we will call 'SAP status' (see description below), the socio-demographic group of the household, and the tenure of the household.

Keirstead (2006) found that the largest consumer group for solar PV are 45-64 years old (53%), degree educated (77%), earn £50,000-£100,000 (30%) and have a household size of two (46%). Mosaic, Experian's socio-demographic classification system, has been used to identify these types of households, which have been labelled Group 1. This is a rather crude way of identifying these households because age, income and household size are not explicit in the Mosaic groups, but this is the best information available.

There are three categories of SAP status:

- Currently SAP D or above
- Capable of improvement to SAP D (or higher) with cavity wall insulation, loft insulation, and heating system upgrade
- Capable of improvement to SAP D (or higher) with solid wall insulation, loft insulation, and heating system upgrade

In terms of socio-demographic group and tenure, households are split into four groups. The groups are:

Table 3: Socio-demographic and tenure groups for PV layers

Group	Tenure	Mosaic Groups
1	Owner Occupier	A01, A02, B05, B07, C09, C10, C11, D13, D14, D15, D16, E17, E19, F23, G27, G29, L52 (older, most affluent)
2	Social housing	Any
3	Owner Occupier	A03, B06, E18, E20, E21, F22, F24, H35, H38, I39, I40, J46, J47, K48 (affluent, but not as much as group 1, younger than group 1)
4	Owner Occupier	Any not included above

Group and SAP status are then combined into 12 layers, as follows:

Table 4: PV layers

Layer	SAP Status	Group
1	Currently SAP D	1
2	Currently SAP D	2
3	Currently SAP D	3
4	Currently SAP D	4
5	Cavity Wall and Loft Insulation	1
6	Cavity Wall and Loft Insulation	2
7	Cavity Wall and Loft Insulation	3
8	Cavity Wall and Loft Insulation	4
9	Solid Wall and Loft Insulation	1
10	Solid Wall and Loft Insulation	2
11	Solid Wall and Loft Insulation	3
12	Solid Wall and Loft Insulation	4

SAP status is accorded more importance than the group, and so all groups which are currently in SAP D are considered more likely to install solar PV than any group which would require insulation measures to get to SAP D.

Solar PV - Assumptions about capacity

The assumptions about PV capacity are as follows:

- Capacity is 0.125 kW per square metre of panel
- The roof area of the property is calculated from the age, built form, and number of bedrooms of the property, based on property sizes in the English House Condition Survey.
- It is assumed that the roof is pitched, and that one side will have suitable orientation. Therefore the roof area is divided by two to get the suitable roof area.
- It is assumed that 60% of this suitable roof area will be covered in PV panels

Solar PV - Further notes

The calculations for PV do not take into account where households already have PV. Therefore the total predicted figures for PV by licence area and by substation are for the total capacity (currently installed plus future predicted installations), rather than just future predicted installed capacity. The reason for calculating the figures in this way is due to uncertainties about the current installed capacity. As mentioned in section 2, the WPD data on currently installed capacity does not include as much installed capacity as the Feed In Tariff register. We have made an estimate of installed PV per substation based on the FIT installation report[6]; however the estimate is based on pro-rating by postcode sector and so has some inherent inaccuracy. Therefore, in order not to compound the inaccuracy, neither the currently installed PV figures from WPD nor the estimates based on the FIT register have been subtracted from the predicted PV figures.

Currently solar PV is considerably more popular than solar thermal, mainly due to support from the Feed-In Tariff, but also because PV is easier to install, because it can be 'bolted on' whereas solar thermal needs to be integrated into the existing water heating system. Support for solar thermal under the domestic Renewable Heat Incentive, due to start in Summer 2013, may increase the popularity of solar thermal, meaning that some of the roofs assigned to PV in the model may end up being fitted with solar thermal instead. Installation of solar thermal, as a water heating technology, would only have a limited impact on the electricity network, by reducing demand in homes where water is electrically heated.

The predicted PV figures which have resulted from this modelling are large, because the criteria mean that most households are suitable for PV. Therefore in the outputs at substation level, the total number of installations and installed capacity modelled has been reduced to 25% of the original figure. That is, the results show what would happen if a quarter of suitable households in each group would actually install PV.

Heat pumps - Assumptions about which households will install

Ground source heat pumps (GSHPs) and air source heat pumps (ASHPs) are considered in this modelling. The assumptions are applied in several stages:

1. Built form, size, and location characteristics are used to identify the households where heat pumps could be installed - the assumptions we have used here (explained below) result in 59% of the housing stock across the four licence areas being considered suitable for heat pumps.

2. The properties identified in the first stage are further filtered so that only those which are either currently SAP D or above, or could reach SAP D or above with cavity or solid wall insulation, loft insulation, and replacement of the heating system with a heat pump are included. The model considers the cost of improvement within the terms of the Green Deal and ECO so where the energy savings (and, where applicable, Carbon savings[7]) will not pay for the measures they are not installed. When the properties which cannot be improved in this way to SAP D or above are excluded (in addition to those already excluded in step 1 due to built form), 53% of the housing stock in the three licence areas remains.

3. The remaining properties are organised into layers according to their likelihood of installing heat pumps. Likelihood is defined according to SAP status, whether the property has access to the gas network, and whether the tenure is private or social.

[6] This is available from Ofgem and provides a list of all installations; however, the location of each installation is only indicated by postcode sector (e.g. BS3), an area which normally covers several substations.

[7] From the Carbon Saving Obligation, part of the Energy Company Obligation, which for certain households gives carbon saved a value.

Two differences to the solar PV assumptions which the reader may notice are that private rented housing is not excluded, and that there is no division into socio-demographic categories. These were additional requirements added to the solar PV assumptions due to the large number of properties which are suitable for PV. For heat pumps we have the built form criteria which reduces the number of suitable properties, and so further assumptions were not added to refine the criteria.

The built form / size / location criteria are as follows:

- Ground source heat pumps are suitable for larger properties, and so only properties with three bedrooms or more have been considered suitable for GSHPs.
- GSHPs require a reasonably sized garden for the ground loop. We do not have data available on the size of the garden itself, but we have assumed that even large properties in urban areas may not have suitably sized gardens, and so GSHPs are not installed in urban areas[8]. We have also assumed that GSHPs will not be installed in flats, as they may not have a garden.
- Air source heat pumps are suitable for smaller properties and do not require a garden. We have assumed they will be installed in flats, terraces, and bungalows with three bedrooms or less[9].
- The criteria mean that terraces and bungalows with 3 bedrooms in non-urban areas would be suitable for both types of heat pumps. Where this is the case, GSHPs have been given priority.

The assumptions about SAP status (point 2 above) are important because the background research carried out for this study indicates that UK housing is generally too inefficient for heat pumps (Fawcett, 2011; NERA & AEA, 2009). Housing must be well insulated for heat pumps to be effective. Some studies (Fawcett, 2011; Hoggett et al., 2011; Singh et al., 2010) suggest that it is primarily new build houses that are suitable for heat pump installations. However new build housing is beyond the scope of this study, as it is not possible to know where new housing will be located over the three distribution areas in the next decade, nor what type of housing this will be.

Heat pumps are more suitable for properties which are off of the gas network, as they are often more cost-effective than non-gas heating technologies, although not as cost-effective as gas (Fawcett, 2011). In our opinion there are unlikely to be many installations in on-gas areas; however we have included on-gas layers in the modelling (in the higher numbered layers, which are the less likely ones). Caird et al. (2012) found that the majority of heat pump installations were in properties off the gas network.

The Experian dataset shows whether or not a property has a gas connection (this is modelled by Experian). Properties without their own gas connection may be located in areas which have access to the gas network and which over the next decade may connect to the gas network. We believe that over this time period (and indeed, well beyond it) gas is highly unlikely to become a more expensive heating fuel than electricity, and so installing a gas heating system is the more logical financial choice. In addition, gas distributors are extending their networks and are likely to continue doing so, meaning that areas which are off-gas now may have access to gas in a decade's time. Due to these

[8] Our data divides location into: 'Urban', 'Town and Fringe', 'Village', and 'Hamlet & Isolated Dwelling'. The resolution is Census Output Area, which covers an average of 125 houses.

[9] Some research has indicated that ASHPs are not suitable for flats (NERA & AEA, 2009). However there have been some occurrences of successful ASHP installations in flats and apartment blocks, with maintenance access via balconies, such as Thistle Housing Association in partnership with British Gas installing ASHPs in six tower blocks (Cool Planet; NHIC, 2011).

considerations, we have adapted the Experian gas connection variable so that all properties which do not have a gas connection but are within 200 metres of another property with a gas connection are considered to be 'on gas'. The exception to this is high-rise flats which cannot have a gas connection above the fifth floor, due to safety regulations.

The layers include a distinction between private sector (including both owner-occupied and privately rented) and social housing tenures, with social housing being considered slightly more likely to install heat pumps due to social landlords' motivation to improve the quality of their stock and reduce their tenants' heating bills.

The layers for heat pumps are as follows:

Table 5: Heat pump layers

Layer	SAP Status	On-gas	Tenure
1	Currently SAP D	No	Social housing
2	Currently SAP D	No	Private housing
3	Can reach SAP D with cavity and loft insulation	No	Social housing
4	Can reach SAP D with cavity and loft insulation	No	Private housing
5	Can reach SAP D with solid wall and loft insulation	No	Social housing
6	Can reach SAP D with solid wall and loft insulation	No	Private housing
7	Currently SAP D	Yes	Social housing
8	Currently SAP D	Yes	Private housing
9	Can reach SAP D with cavity and loft insulation	Yes	Social housing
10	Can reach SAP D with cavity and loft insulation	Yes	Private housing
11	Can reach SAP D with solid wall and loft insulation	Yes	Social housing
12	Can reach SAP D with solid wall and loft insulation	Yes	Private housing

Heat pumps - Assumptions about demand

In the outputs of this work, predicted heat pump numbers are reported both in terms of number of installations (i.e. number of households installing a heat pump) and maximum demand. The maximum demand is derived from the maximum heating demand of each property assigned a heat pump in the model.

The maximum heating demand for each individual property was calculated in the HAM. It is the heating requirement for a January day with an external temperature of minus 5 °C. The maximum demand figure for a substation is the total of the maximum heating demands for all properties on that substation which were allocated a heat pump in the modelling.

Heat pumps operate with different coefficients of performance (COP) at different times, meaning that for a given heating requirement the demand on the network can vary. For this reason we have made a simplifying assumption here that the COP is 1, meaning that the demand on the network is equal to the heating requirement of the properties. Therefore the maximum demand figure is something of a 'bad-case' scenario: it shows a scenario where all installed heat pumps are operating at once, catering to a high heating requirement and operating with a COP of 1.

Direct Electric Heating - Assumptions about which households will install

As well as heat pumps, the research considers the possibility that households may install electric radiators (direct heating rather than storage heaters). Properties in the dataset have been flagged as suitable for direct electric heating if:

- They are in off-gas areas
- The property does not already have night storage heaters (households from Experian have been matched to substations rather than individual properties, and so it is not possible to use the profile class of a customer record to determine this; instead, it is estimated based on a model CSE uses as part of the HAM)
- They will not be installed in properties if a heat pump is preferable

Essentially, the assumption is that properties which do not have access to gas, which do not already have night storage heaters and which are not suitable for heat pumps will switch to direct electric heating.

Therefore, the figures for direct electric heating are *an estimate of a worst case scenario* where all households which cannot heat using gas and which are deemed unlikely to install a heat pump switch to electric heating. Whereas heat pumps will be supported by the domestic Renewable Heat Incentive, there will be no financial support to those who install direct electric heating and so there is unlikely to be wide take-up.

In the outputs, direct electric heating installations are organised in the same layers as used for heat pumps (see Table 4). However, because the number of direct electric heating installations reduces each time another layer is added to the heat pump assumptions, the layers for direct electric heating should not be read cumulatively (as they are in the case of heat pumps). Instead, a single layer should be chosen, based on the layer heat pumps are being read from. For example, in a scenario where heat pumps are installed by households in layers 1-7, then the direct electric heating capacity in the column for layer 7 is the correct corresponding figure.

Direct Electric Heating - Assumptions about capacity

As for heat pumps, a maximum demand figure is calculated per property using the Housing Assessment Model for a January day with an external temperature of -5 °C.

Electric Vehicles

Electric Vehicles (EVs) are perhaps the most difficult technology to model as their uptake will bear little relation to housing characteristics, which for the other technologies are one of the key determinants.

Forecasts for EV growth are very speculative and vary widely, so have not been used to inform our predictions. They range from a 4.5% uptake by 2025 (Morgan Stanley, 2011), to a 16% uptake by 2020 (CCC, 2012). There are currently (latest figure at the time of writing) 28,813,200 cars (of all propulsion types) licensed in Great Britain (DfT 2012).

Research suggests that early owners of EVs are likely to be “fleet or business users and consumers in urban and suburban locations” (DfT, 2011), as these consumers will be the most able to take advantage of the environmental and cost benefits of EVs such as short journeys, reduced congestion charges and availability of charge points. Beyond 2015 the Department for Transport (DfT, 2011) finds it difficult to forecast EV ownership as it is too dependent on a range of factors, such as the level of consumer acceptance and oil prices, which are difficult to predict.

Those most likely to buy an EV are expected to be affluent people of working age in urban areas, under the assumption that they will drive to work and other short distances, possibly primarily using the EV as a second vehicle, as found in research by Transport for London (2010).

However, the market will need to expand significantly beyond these early adopters for EVs to become as widespread as shown in the Ofgem / DECC scenarios.

Some demographic factors, such as car ownership, are not available within the data set used for this modelling. The assumptions chosen were based on the datasets available.

CSE requested data from the Department for Transport showing the number of EVs registered at the end of 2011 by local authority area. This has been incorporated in the criteria as it is possible that people living in areas where they see EVs on the road may be more likely to consider buying one themselves.

Whether or not the household is located in an urban area is also included in the layers, as urban areas will have more opportunities to use an EV. However, rural areas are not excluded completely (rather, they are assigned to a lower likelihood layer) because people may still have opportunities to use an EV for short journeys locally.

Socio-demographic groups were chosen based first on those which were most likely to include wealthy commuters, followed by groups which are still affluent although not to the extent of the first group. Next come all other groups apart from those which are very unlikely to purchase any new car, still less an EV. Layer 9 is a very large group containing over 1.5 million households across the three licence areas; however the Mosaic groups do not provide enough information to be able to split this group down any further.

The layers are therefore as shown in Table 6.

Table 6: EV layers

Layers	Mosaic	Urban	Areas with EVs already
1	Six mosaic groups representing those likely to be wealthy commuters: ('C09', 'C10', 'C11', 'C12', 'D16', 'F22')	Urban	Yes
2	As above	Urban	No
3	As above	Not urban	Yes
4	As above	Not urban	No
5	Six mosaic groups representing those which are affluent (but not as wealthy as the previous group) and also commuters: ('D16', 'E17', 'E20', 'G26', 'G29', 'G31')	Urban	Yes
6	As above	Urban	No
7	As above	Not urban	Yes
8	As (5)	Not urban	No
9	All other Mosaic groups apart from 17 groups which are unlikely to purchase a new car: the groups excluded are: ('I44', 'I42', 'J45', 'K49', 'K50', 'K51', 'L54', 'M56', 'M57', 'M58', 'M59', 'N60', 'N61', 'N66', 'O67', 'O68', 'O69')	Urban	Yes
10	As above	Urban	No
11	As above	Town and fringe	Both

The figures reported show the predicted number of households owning an EV. No capacity figure has been included for EVs due to uncertainty about the capacity and charge speed which will dominate if EVs become more popular.

References

***Caird, Sally, Roy, Robin & Potter,* Stephen** (2012), "Domestic heat pumps in the UK: user behaviour, satisfaction and performance", *Energy Efficiency*, 5(3), pp. 283–301

CCC (Committee on Climate Change) (2012), "Electric Cars & Vans", http://www.theccc.org.uk/sectors/surface-transport/electric-cars

Cool Planet, "A Guide to Heat Pumps", http://www.cool-planet.co.uk/images/about/downloads/HeatPumpGuidev1.pdf

Department for Transport (DfT) (2011), "Making the Connection: the Plug-In Vehicle Infrastructure Strategy", *Office for Low Emission Vehicles*, London

Department for Transport (DfT) (2012), VEH0101 - Licensed vehicles by body type.

Fawcett, Tina (2011), "The future role of heat pumps in the domestic sector", Proceedings of ECEE 2011 Summer Study

Hoggett, R., Ward, J. & Mitchell, C. (2011), "Heat in Homes: customer choice on fuel and technologies", Study for Scotia Gas Networks

Keirstead, J. (2006), "Behavioural responses to photovoltaic systems in the UK domestic sector", PhD thesis, University of Oxford, UK

Morgan Stanley (2011), "Tesla: Morgan Stanley Turns Bearish On Slower EV Adoption",http://www.forbes.com/sites/ericsavitz/2011/12/08/tesla-morgan-stanley-turns-bearish-from-concept-to-reality/

NERA (Radov, D., Klevnäs, P., Hanif, A.) & AEA (Abu-Ebid, M., Barker, N., Stambaugh, J.) (2009), "The UK Supply Curve for Renewable Heat: Study for the Department of Energy and Climate Change", *Ref: URN 09D/689 (DECC)*

NHIC (National Home Improvement Council) *(2011), "Category Seven, Energy Efficient Improvements Using Renewable Technologies", http://www.nhic.org.uk/awards-2012/awards-2011/category-seven/*

Singh, H., Muetze, A. & Eames, P.C. (2010), "Factors influencing the uptake of heat pump technology by the UK domestic sector", *Renewable Energy* (35), pp. 873-878

Transport for London (2010), "Electric Vehicles market development", 09201, Mayor of London, http://www.tfl.gov.uk/assets/downloads/customer-research/electric-vehicles-maket-development-presentation.pdf

6. Matching households to substations

In order to express the projected installations of new technologies at substation level, the two major datasets used in this project - Experian household data and WPD customer (MPAN) data - needed to be matched together. This section describes the process and associated issues in more detail.

WPD data

WPD provided a dataset with details of MPANs, including XY coordinates, profile class, energisation status, and substation number. The dataset includes all customers in the South West distribution area which are connected to the distribution network (thus high voltage customers are not included). WPD MPAN records are referred to in this report as 'customers'.

Experian / Housing Assessment Model

The HAM uses data from Experian, a company which has developed a database containing information about every household in the UK from a variety of sources and in-house models. This data includes information about the age, built form, tenure and size (number of bedrooms) of the property and whether it has mains gas, as well as a socio-economic classification of the household. Experian data was also used for modelling EV uptake outside of the HAM. In this report the term 'households' is used to refer to the Experian records.

Matching the two datasets

As described in section 3, the HAM has been used to make predictions about which households will take up particular technologies, which will subsequently affect demand on their local substation. Once predictions are made for the household level, the next step is to aggregate these to substation level. This requires the matching of the Experian dataset, which contains the household-level predictions, to the WPD customer dataset, which shows which substation each household is connected to. Inevitably, the two datasets are not a perfect match.

First, the distribution licence areas do not match exactly to any government boundaries, such as region or groups of local authorities. For example, in the South West, the DNO licence area covers most, but not all, of the political South West Region. The main difference is that only parts of Dorset and Gloucestershire are covered by the WPD distribution area, and Wiltshire is not covered at all. Larger population centres of the South West which are excluded include Bournemouth, Poole, Gloucester and Swindon. The Gloucester and Swindon areas are in the political South West, but Gloucester is in the West Midlands licence area while Swindon is the Southern England licence area, which is operated by Scottish and Southern Energy. Around the edges of the licence areas it is common for only part of a local authority area to be covered. For example, in the South East, only half of Aylesbury Vale district local authority area is in the East Midlands licence area, and in Wales only half of Powys is in the South Wales licence area. The boundaries are therefore somewhat blurred, and at the edges of the licence areas there could be many households but not many customers to match them to. There are thus increased risks of error for the matches at the edge of the licence areas.

More importantly, although both datasets include coordinates for each household / customer, these do not always coincide. In the real world, there are households (represented in the Experian data) and each household has a meter (represented by the Meter Point Reference Number, MPAN, in the WPD dataset). Ideally, for every point in the Experian dataset (representing the household), there should be a corresponding point in the WPD customer dataset, and both points should have the same coordinates. However, each point in the Experian dataset does not match exactly to a point in the WPD dataset. Coordinates are the only information we have to link the two datasets together, so this is an important factor to understand when interpreting the results of this work. This issue was anticipated as it would be very unusual for two different datasets produced by different organisations and using different criteria to match perfectly, even though they both represent the same real world entities (or almost the same; there will be situations when in the real world they are not the same, for example a house divided into bedsits will count as several households but may have one shared meter).

The main reason for the difference in coordinates is that the coordinates are given with a resolution of 1 metre. For example, the two points (441000, 292000) and (441000, 292001) are one metre apart and so when matching on X, Y these two points will not match. The plot for a detached house plus garden could measure 30m x 12m and so many different coordinate pairs will fall within its boundaries. If the coordinate pair representing the house can go anywhere within the plot then it is easy to see how the Experian record could be matched to the customer record for the house next door rather than the correct house. In blocks of flats, the coordinates for each flat are the same and there is no way to differentiate different flats.

For this reason, we did not attempt to match households to individual customer records. Instead, we found the nearest customer for each household and then assigned the household to the substation associated with that customer. The process was as follows:

1. The two datasets were viewed in GIS and the boundaries of the WPD distribution area were identified. Experian records within this area, and slightly outside of the area, were selected. Only these Experian records were used in the matching process.
2. Where there was an MPAN on exactly the same point as an Experian record, the Experian record was assigned the substation number of that MPAN.
3. This left a number of Experian records which were not on the same point as an MPAN. A GIS was used to match the remaining Experian records to the nearest MPAN, where there was an MPAN less than 20 metres away.

The match distance of 20 metres is quite short and so is a conservative figure to use. The reason for this is that where properties are on the boundary between two substation areas, even a fairly small distance could mean that the property ends up being matched to the wrong substation. We looked at the difference between the actual number of domestic customers on a substation (as given by the WPD data) and the number of matched households on a substation under different match distances. The error appeared to increase after 20 metres and so this limit was not increased. However, some substations had substantially fewer households than they should have done (the number of households which ought to be connected to a substation being determined by the number of profile class 1 and 2 energised MPANs connected to that substation). To correct this, where substations did not have enough households at a 20 metre distance but had matched households at more than 20 metres, households were incrementally added until the substations had the correct number of matched households, as long as the households had a match distance less than 200 metres.

Of the 145,077 substations which have domestic customers across the four licence areas, 27% have the same number of matched households as the number of domestic energised customers and 40% have a difference or 1 or 2. Overall 89% have a difference of 0-10.

Four percent of domestic energised MPANs across all four licence areas do not have coordinates in the WPD dataset, and could therefore not be used when matching to Experian data.

Some small business customers may be on profile classes 1 or 2; where this is the case there will be an overestimate of the number of domestic customers on a substation.

Match quality indicators

In the substation-level outputs, indicators of the quality of the match between the two datasets are included for each substation. The indicators are:

- **Match difference:** This shows the number of domestic energised customers in the WPD dataset minus the number of households identified in the same area in the Experian dataset. If both datasets were completely accurate this value would always be zero. A positive value here shows that there are more customers in the WPD dataset, while a negative value shows that there are more households in the Experian dataset.

- **Match difference proportion:** This is the absolute value of the match difference divided by the number of domestic energised customers on the substation according to the WPD dataset. This is given in addition to the match difference to give a sense of scale. For example, a match difference of 10 is important if there are only 12 domestic customers on the substation (match difference proportion of 0.83), but less important if there are 400 domestic customers (match difference of 0.03).

Records with coordinate errors

A small number of substations appear to have one or more customers (WPD data) located far from the other customers on the substation. This could be because the customer record is linked to the wrong substation, or because the customer is linked to the correct substation but the customer's coordinates are wrong. When Experian data is matched to WPD data, these coordinate / substation errors can result in the wrong households being matched to the substation. However, although it is possible to tell that there is something wrong with the locations of one or more customers on a substation, it is not always possible to tell which customer record has the error. All we know is that the match will be less accurate in these cases.

In the outputs we have added a flag to the substations which appear to have at least one customer in the wrong location to show where the match with Experian data may have been affected. The relevant flags are:

- **Bad location flag**: This shows substations where some of the coordinates of customer locations appear to be wrong within the WPD dataset. This is identified by looking at the greatest difference between X coordinates and Y coordinates among customers on each substation. If the distance between the X or Y coordinates is more than 2 kilometres, it is likely that at least one customer has the wrong coordinates. This in turn affects the match to Experian. A value of '1' means that there are one or more customers with bad coordinates; a value of '0' means that there are no obvious location errors in the data.
- **Bad locations excluded from match**: Where it is possible to confidently identify which customer points have bad coordinates, these have been excluded from the match with Experian. A value of '1' shows that bad locations have been excluded. It is not always possible to identify which customer points have bad coordinates - in many cases it is clear that the customer's coordinates are too far apart to be connected to the same substation, but they are too evenly spaced to identify one which is clearly wrong.

Where substations are flagged as having a bad location, the figure for modelled installations should be treated with more caution, unless the substation is also flagged to show that bad locations have been excluded from the match.

7. Substation capacity and maximum demands

WPD provided CSE with data for each distribution substation in the South West distribution area, which included information on location, number of transformers, transformer capacity, number of customers in each profile class, and maximum demands. It is important to know the capacity of individual substations and what level of demand they are currently experiencing. This section summarises some data issues encountered when calculating this.

The first issue is that the maximum demands are modelled by WPD based on meter advances. Among these are some values which look incorrect - mostly very high values which, even when divided by 1,000 (to account for the possibility that the unit of measurement is watts rather than kilowatts), are much too large for the substation capacity. Maximum demand values which were more than twice the capacity of the substation were divided by 1,000 to account for the possibility that they were measured in watts. If they were still more than twice the capacity of the substation, they were removed from the dataset. There were five day and five night maximum demands in the dataset, so removing some did not necessarily mean that the substation was left with no maximum demand data. However, it was quite common that if one maximum demand figure was very high, the others were also very high, so in some cases all of a substation's maximum demands were removed. For all substations there remains the possibility that the maximum demand figure is incorrect, as only the easily identifiable errors were removed from the dataset. Where a substation is identified as having a low remaining capacity from this data, further on-the-ground investigation by WPD would be required to ascertain whether these substations are truly overloaded or whether the maximum demand figure is wrong.

The second issue with the data is that, in some cases, the capacity looks very high relative to the number of customers. This could be because these substations have one transformer which is specifically for use by a high voltage customer, and if this is the case, the capacity of the distribution substation should for our purposes be reduced by removing the capacity of that transformer from the total. This is something that WPD could investigate further using data on high voltage customers.

The 'headroom', or unused capacity on the substation, was calculated by subtracting the highest of the ten maximum demand figures (five for day and five for night, where a value was present for all of these) from the capacity. Negative headroom values appear for those substations where maximum demand figures were higher than capacity (but less than twice the capacity).

Where data on capacity is missing or there are no maximum demand figures for a substation, headroom cannot be calculated.

8. Existing generation

While the HAM modelling includes a figure for predicted PV capacity, this chapter looks at estimates of capacity already installed. There are two sources for this. The first is WPD's own data, which shows installations at MPAN level. The second is Ofgem's FIT installation report[10], which shows

[10] See http://www.ofgem.gov.uk/Sustainability/Environment/fits/Pages/fits.aspx

installations at postcode sector level. Postcode sector is the first half of the postcode (e.g. BS3) and the area it represents can contain from 1,000 to 30,000 addresses. The FIT data includes the installed capacity for each technology, whether it is a domestic or non-domestic installation and whether it is grid-connected. For all the grid-connected installations, we have allocated the postcode-sector level installations for each technology to substations based on the number of domestic and non-domestic customers within each postcode sector on each substation (as one substation may have customers within more than one postcode sector). This gives an even distribution of a postcode sector's generation across all substations with customers in that postcode sector, which is the best estimate; however the real distribution across substations will be different - sometimes slightly, sometimes greatly - depending on the degree to which the postcode sector's installations are concentrated in specific areas.

WPD's dataset includes all generation of which WPD had been notified at the time the data was extracted from its database. Not all of this generation will be claiming the FIT. Conversely, not all generation registered for the FIT has yet been picked up in WPD's records (due to timing differences between registration for FITs and notification to the DNO). This can be seen in the fact that, for example, in the South West, installed PV capacity for the region in the FIT installation report is approximately 2.5 times the PV capacity that WPD has recorded.

Where the FIT estimated figure for the substation is higher than the WPD figure, this could be because the installation has not yet been recorded in WPD's data (although it is also possible that the generation in the postcode sector is concentrated on another substation, which would make the estimated figure inaccurate). Where the WPD figure is greater than the FIT figure, it could mean that there is distributed generation which is not claiming the FIT (or that the generation in the postcode sector is concentrated on this substation, making the FIT estimate inaccurate).

9. Bringing it all together: outputs

Licence-area level summaries have been produced which shown predicted heat pumps, solar PV, and EVs at licence area level, with a comparison to Ofgem / DECC targets, with the Ofgem / DECC targets being pro-rated to licence area based on the proportion of GB customers that they contain.

Substation-level spreadsheets have also been produced containing the data brought together by this work with a record for each distribution substation (apart from spreadsheet 5, which has a record for each LSOA). These spreadsheets are:

1. Main Summary Table: This table brings together information about substation capacity and load, with modelled new installations of PV, heat pumps, and direct electric heating at levels approximately equal to DECC / Ofgem's 2023 medium scenario[11]. The column headings are described in the appendix on p.23.

2A. Ground Source Heat Pumps: For each layer for heat pumps described in section 5, the number of installations and the maximum demand of these installations (see p.12 for how this was

[11] Normally to the nearest layer which exceeds the Ofgem / DECC scenario figure, as shown in the licence area summary spreadsheets.

calculated) is given. Indicators of the quality of the match between the customer dataset and the household dataset are also given (these are the same indicators used as in the main summary table - see the sections *Match quality indicators* and *Records with coordinate errors* on p.18 for a description).

2B. Air Source Heat Pumps: The columns in this spreadsheet are the same as for 2A, but covering air source heat pumps.

2C. All Heat Pumps: The columns in this spreadsheet are the same as for 2A, but covering both types of heat pumps.

2D. Direct Electric Heating: Again, the number of installations in each layer and the maximum demand are shown. Note that (unlike the heat pump and PV spreadsheets) the layers must not be summed. Match quality indicators are also shown.

2E. Solar PV: This shows the number of installations and capacity in each layer, along with demand quality indicators.

3. Generation (existing generation): This spreadsheet shows existing generation by technology type and totals. Where the information comes from WPD data, it is prefixed with 'WPD', and where it is derived from the FIT installation report, it is prefixed with 'FIT'.

4. Electric Vehicles: This shows the number of households in each layer, along with match quality indicators.

5. Social Indicators by LSOA: These are described in a separate methodology report.

6. Social Indicators by Substation: These are described in a separate methodology report.

There is a version of each of the above spreadsheets for each licence area.

10. Appendices

Appendix 1: Description of fields in main summary table

A description of the fields in the main summary table (table 1 of the substation-level tables) because the fields in this table require some explanation.

There is a main summary table for each licence area. The columns in the table are as follows:

- **Substation number**
- **Substation name**
- **District**
- **SS Capacity kVA**: This is the substation capacity, based on WPD data. Zero or null values mean there is missing data.
- **Headroom_kVA**: This is the 'unused' capacity, calculated by subtracting the highest recorded maximum demand figure from the capacity.
- **Headroom as a proportion of capacity**: This is the headroom figure divided by the capacity figure. A figure which is small but positive shows little headroom remaining on the substation. A negative figure shows that the maximum demand figure was higher than capacity.
- **WPD DG Data Total kW**: This is the total kW capacity of distributed generation on the substation, according to WPD's data.
- **FIT estimated data**: This is the total kW capacity of distributed generation claiming the feed-in tariff (FIT). It is labelled 'estimated' because the FIT data (provided by Ofgem) is provided at postcode sector (e.g. BS3), and we have distributed it proportionally to substations based on the number of customers.
- **Modelled PV kW (Ofgem):** The figure shown in this column uses assumptions which make the total installed capacity approximately the same as the Ofgem medium scenario for 2023 (at the licence area level, based on an assumed percentage of GB customers in each region), which is why it is labelled 'Ofgem'. Note that this calculation does not take into account where properties already have PV, so this figure overlaps the figure for existing installed capacity. The layers used for each licence area to meet the Ofgem scenario are:
    - South Wales:25% of layers 1-5
    - East Midlands: 25% of layers 1-4
    - West Midlands: 25% of layers 1-5
    - South West: 25% of layers 1-5
- **Modelled HP installations (Ofgem):** The figure shown in this column uses assumptions which make the total number of heat pump installations approximately the same as the Ofgem / DECC medium scenario for 2023 (at the licence area level, based on an assumed percentage of GB customers in each region), which is why it is labelled 'Ofgem'. The layers used for each licence area to meet the Ofgem scenario are:
    - South Wales: Layers 1-6
    - East Midlands: Layers 1-7
    - West Midlands: Layers 1-7
    - South West: Layers 1-4

- **Modelled HP max demand kW (Ofgem):** The figure shown here is the maximum demand from heat pumps with installations at a level which at the aggregate for the South West meets Ofgem's medium scenario for 2023 (layers vary for licence areas, as shown above).
- **Modelled Direct Electric Heating max demand (kW)**: The figure shown here corresponds to the amount of direct electrical heating which might be installed in homes which are not deemed suitable for heat pumps, under the same layer assumptions which give the heat pump figures in the previous columns.
- **Remaining Headroom after modelled additions (kVA)**: This shows what the headroom on the substation would be if all of the potential heat pump and direct electric heating capacity was connected (this is a worst-case assumption that all the peak demand would happen at the same time).
- **Remaining Headroom as proportion of capacity**: This shows the above as a proportion of substation capacity.
- **EVs**: This gives the number of households owning an EV under assumptions based on household socio-demographic type and location, which for each licence area gives a total number of EVs approximately equal to the Ofgem scenario for 2023. The layers used for each licence area to meet the Ofgem scenario are:
    - South Wales: Groups 1-9
    - East Midlands: Groups 1-8, plus half of group 9
    - West Midlands: Groups 1-8, plus half of group 9
    - South West: Groups 1-8, plus half of group 9
- **Domestic MPANs**: This is the number of energised MPANs with profile class 1 or 2 in the substation area.
- **Profile 1**: Number of energised MPANs with profile class 1
- **Profile 2**: Number of energised MPANs with profile class 2
- **Non-domestic MPANs**: Number of energised MPANs with profile classes other than 1 or 2

The summary spreadsheet and other spreadsheets which give information derived from matching the WPD dataset with the Experian dataset include four indicators showing location data quality and the quality of the match for the individual substation. These are as follows, and are described in the sections 'Match quality indicators' and 'Records with coordinate errors' on p.18.

- **Bad location flag**
- **Bad locations excluded from match**
- **Match difference**
- **Match difference proportion**

Appendix 2: Mosaic Public Sector Groups

Table 7 below summarises Experian's Mosaic Public Sector groups. For more information about Mosaic Public Sector and a more detailed of the groups, see http://publicsector.experian.co.uk/Products/~/media/Brochures/MosaicPublicSector_Brochure_final.ashx and the interactive guide: http://guides.business-strategies.co.uk/mosaicpublicsector2009/html/visualisation.htm

Table 7: Mosaic Public Sector Groups

Group			
Group A: Residents of isolated rural communities	A	01	Rural families with high incomes, often from city jobs
	A	02	Retirees electing to settle in environmentally attractive locations
	A	03	Remote communities with poor access to public and commercial services
	A	04	Villagers with few well paid alternatives to agricultural employment
Group B: Residents of small and mid-sized towns with strong local roots	B	05	Better off empty nesters in low density estates on town fringes
	B	06	Self employed trades people living in smaller communities
	B	07	Empty nester owner occupiers making little use of public services
	B	08	Mixed communities with many single people in the centres of small towns
Group C: Wealthy people living in the most sought after neighbourhoods	C	09	Successful older business leaders living in sought-after suburbs
	C	10	Wealthy families in substantial houses with little community involvement
	C	11	Creative professionals seeking involvement in local communities
	C	12	Residents in smart city centre flats who make little use of public services
Group D: Successful professionals living in suburban or semi-rural homes	D	13	Higher income older champions of village communities
	D	14	Older people living in large houses in mature suburbs
	D	15	Well off commuters living in spacious houses in semi rural settings
	D	16	Higher income families concerned with education and careers
Group E: Middle income families living in moderate suburban semis	E	17	Comfortably off suburban families weakly tied to their local community
	E	18	Industrial workers living comfortably in owner occupied semis
	E	19	Self reliant older families in suburban semis in industrial towns
	E	20	Upwardly mobile South Asian families living in inter war suburbs
	E	21	Middle aged families living in less fashionable inter-war suburban semis
Group F: Couples with young children in comfortable modern housing	F	22	Busy executives in town houses in dormitory settlements
	F	23	Early middle aged parents likely to be involved in their children's education
	F	24	Young parents new to their neighbourhood, keen to put down roots
	F	25	Personnel reliant on the Ministry of Defence for public services
Group G: Young, well-educated city dwellers	G	26	Well educated singles living in purpose built flats
	G	27	City dwellers owning houses in older neighbourhoods
	G	28	Singles and sharers occupying converted Victorian houses
	G	29	Young professional families settling in better quality older terraces
	G	30	Diverse communities of well educated singles living in smart, small flats
	G	31	Owners in smart purpose built flats in prestige locations, many newly

Group			
			built
	G	32	Students and other transient singles in multi-let houses
	G	33	Transient singles, poorly supported by family and neighbours
	G	34	Students involved in university and college communities
Group H: Couples and young singles in small modern starter homes	H	35	Childless new owner occupiers in cramped new homes
	H	36	Young singles and sharers renting small purpose built flats
	H	37	Young owners and rented developments of mixed tenure
	H	38	People living in brand new residential developments
Group I: Lower income workers in urban terraces in often diverse areas	I	39	Young owners and private renters in inner city terraces
	I	40	Multi-ethnic communities in newer suburbs away from the inner city
	I	41	Renters of older terraces in ethnically diverse communities
	I	42	South Asian communities experiencing social deprivation
	I	43	Older town centre terraces with transient, single populations
	I	44	Low income families occupying poor quality older terraces
Group J: Owner occupiers in older-style housing in ex-industrial areas	J	45	Low income communities reliant on low skill industrial jobs
	J	46	Residents in blue collar communities revitalised by commuters
	J	47	Comfortably off industrial workers owning their own homes
Group K: Residents with sufficient incomes in right-to-buy social houses	K	48	Middle aged couples and families in right-to-buy homes
	K	49	Low income older couples long established in former council estates
	K	50	Older families in low value housing in traditional industrial areas
	K	51	Often indebted families living in low rise estates
Group L: Active elderly people living in pleasant retirement locations	L	52	Communities of wealthy older people living in large seaside houses
	L	53	Residents in retirement, second home and tourist communities
	L	54	Retired people of modest means commonly living in seaside bungalows
	L	55	Capable older people leasing / owning flats in purpose built blocks
Group M: Elderly people reliant on state support	M	56	Older people living on social housing estates with limited budgets
	M	57	Old people in flats subsisting on welfare payments
	M	58	Less mobile older people requiring a degree of care
	M	59	People living in social accommodation designed for older people
Group N: Young people renting flats in high density social housing	N	60	Tenants in social housing flats on estates at risk of serious social problems
	N	61	Childless tenants in social housing with modest social needs
	N	62	Young renters in flats with a cosmopolitan mix
	N	63	Multicultural tenants renting flats in areas of social housing
	N	64	Diverse homesharers renting small flats in densely populated areas
	N	65	Young singles in multi-ethnic communities, many in high rise flats
	N	66	Childless, low income tenants in high rise flats
Group O: Families in low-rise social housing with high levels of benefit need	O	67	Older tenants in low rise social housing estates where jobs are scarce
	O	68	Families with varied structures living in low rise social housing estates
	O	69	Vulnerable young parents needing substantial state support