Department for Business, Energy & Industrial Strategy

NEED 2020 Annex A: Comparison with other sources

25 June 2020 **National Statistics**

This document provides comparative analysis and commentary assessing how representative the National Energy Efficiency Data-Framework (NEED) is of the domestic dwelling stock in England and Wales. In all cases, the data are broadly consistent with the other sources available for comparison with NEED.

- The regional distribution of properties in NEED is representative and comparable to estimates of the number of households and the dwelling stock, with London showing the largest variation at 0.3 per cent.

- NEED domestic gas and electricity consumption is aligned with data published in Energy Consumption in the UK (ECUK) and subnational domestic consumption. Mean domestic gas consumption is 13,200 kWh in both NEED and the Subnational Gas statistics, whereas in ECUK it is 13,600 kWh.

- The distribution of Valuation Office Agency (VOA) and Experian property attribute data contained in NEED is aligned with estimates published in the latest English Housing Survey (EHS) statistics. Overall, there is good alignment with some variations between VOA and NEED likely as a result of lower address matching rates for flats.

Contents

Introduction

The domestic National Energy Efficiency Data-Framework (NEED) is based on data from various sources which are linked together using a Unique Property Reference Number (UPRN). This document provides information on the quality assurance of data used in the production of analysis using NEED. The comparison with other sources covers properties in England and Wales but not Scotland, this is due to the availability of comprehensive data sources in England and Wales that allows the range of comparisons included in this document. This publication is updated and reviewed alongside each NEED publication. More information on NEED, including a domestic NEED methodology note and outputs from NEED are available at the following link:

https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

Address Matching
Identifying properties using Unique Property Record Numbers (UPRN) is central to the NEED database. BEIS use its own address matching algorithm to match address information from different datasets to a UPRN. For example, using address information provided by energy suppliers, metered energy consumption is matched to a property using its UPRN. Household attributes such as the number of bedrooms at a property (supplied by Experian) are also linked using the UPRN. With over 25m domestic properties in England and Wales, some properties have conflicting information and it is not always possible to match to a UPRN. In other cases incorrect property information may be matched to a meter, though these are a minority of cases. UPRNs contained in the NEED database underwent a thorough revision prior to the NEED 2020 publication, significantly improving the quality of the address matching used by NEED.

Outputs from NEED are based on all UPRN's contained in the VOA Council Tax Database[1] that have been successfully matched to a database of UPRNs that BEIS has matched to gas and electricity meter data. Figure A.1 shows the distribution of properties in 2020 NEED compared with the latest ONS English regional[2] and Welsh Government[3] household estimates and MHCLG English regional[4] and Welsh Government[5] estimates of the housing stock. The overall distribution of properties in NEED is aligned with ONS household estimates, with the largest discrepancy in London. This is largely due to loss of records from the address matching process. London has higher proportion of flats which have a lower match rate than houses which leads to this loss of records. In comparison with estimates of the dwelling stock the picture is broadly the same, with overall alignment between NEED across regions with the largest discrepancy in London. There is also variation in the proportion of properties that have gas consumption compared with properties with electric consumption. This is due to the lower prevalence of properties with a gas connection, there is also regional variation to the degree that properties have been connected to the gas grid. Users should note that there can be small differences between the number of households and the number of dwellings. In example a

[1] VOA council tax register data can be found here: https://www.gov.uk/government/statistics/council-tax-stock-of-properties-2019
[2] ONS Household Estimates can be found here: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/datasets/householdprojectionsforengland
[3] Welsh household estimates can be found here: https://statswales.gov.wales/Catalogue/Housing/Households/Estimates/households-by-localauthority-year
[4] MHCLG English dwelling stock estimates can be found here: https://www.gov.uk/government/statistical-data-sets/live-tables-on-dwelling-stock-including-vacants
[5] Welsh Government estimates of the housing stock can be found here: https://statswales.gov.wales/Catalogue/Housing/Dwelling-Stock-Estimates/dwellingstockestimates-by-localauthority-tenure

dwelling can contain more than one household, whereas some households may own more than one dwelling in the case of second homes and holiday homes.

Figure A.1: Distribution of NEED compared with Council Tax Database (VOA), household and dwelling stock estimates.

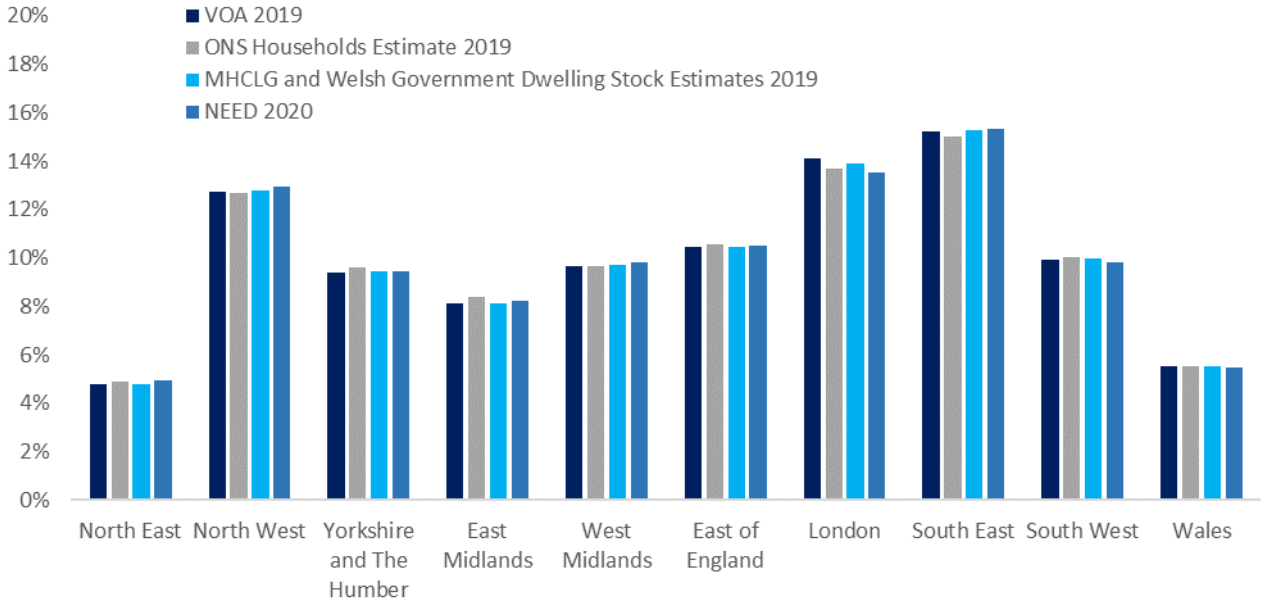


Table A.1 below summarises the strengths and weaknesses of each of the main data sources used for the June 2020 NEED publication. The below strengths and weaknesses should be borne in mind when interpreting the results.

Table A.1: Strengths and weaknesses of data in NEED

Data source	Coverage	Strengths	Weaknesses
Consumption data	• Covers Great Britain	• Good coverage of almost all properties (post matching) • Data provided by energy suppliers • Gas data are weather corrected	• Based on billing data (sometimes estimated) • Gas and electricity years don't cover calendar year (or the same period as each other • Domestic/non-domestic split (by profile class, or consumption threshold)
Valuation Office Agency (VOA)	• Covers every property in England and Wales	• Excellent coverage–more than 99 per cent of properties in the NEED sample for all variables	• No data for Scotland • Some data may not be up to date
Experian	• Data available for each household in the UK	• Best source of data at property level on household characteristics	• Modelled data with varying accuracy at property level
DECC/Ofgem/HEED energy efficiency datasets (including ECO, ECO excess and Green Deal measures)	• Covers households in the UK	• These datasets contain data for measures installed in homes in the UK including the date of installation	• Only covers measures installed through Government schemes; no information on measures installed by households themselves or installed when the property is built • Matching of (converted) flats not reliable • ECO excess dataset has inconsistent address information and almost a third of the dataset could not be matched to AddressBase
Central Feed-in Tariff Register (FiTs)	• Covers every property in Great Britain.	• Excellent coverage–contains detailed property information on all microgeneration installations receiving FiTs	• Excludes microgeneration installations that are not registered for FiTs
Gas Safe Boiler Installations	• Covers every property in Great Britain.	• Excellent coverage–contains information on allregistered boiler installations in Great Britain	

Address Matching

The NEED dataset compiled for the 2020 publication is comprised of data sets covering different time periods: gas and electric consumption data covers 2018, VOA property data covers 2019, ONS UPRN directory data was extracted in February 2020 and Experian data covers 2019 household characteristics.

Alongside this document NEED publish a comprehensive methodology overview that can be accessed here: https://www.gov.uk/government/publications/domestic-national-energy-efficiency-data-framework-need-methodology.

Prior to 2019 a stratified random sample of approximately one in five records was selected from the complete property attribute dataset held by VOA. The resulting sample contained approximately 4 million records. From 2019 the NEED has utilised all properties contained in the VOA Council Tax Database.

The final NEED dataset for England and Wales (combining the UPRNs from both gas and electric data) used in the 2020 publication contains 23.8 million records. This is 90 per cent of the full VOA Council Tax Database. The 10 per cent of records removed during the address matching process occurs when a gas or electric meter can't be linked to a property or a non-typical property type (e.g. annexes). Further records were excluded as a result of invalid or missing consumption values. For 2020, 22.7m (87.2 per cent) had a valid electric consumption value and 18.2m (70.1 per cent) had a valid gas consumption value – this is lower than the valid electric consumption value as not all properties have a gas meter. Combining both the valid gas and valid electricity data sets is considered the full NEED 2020 data set and contains 23.8m records and is used in the comparison with other sources.

To create NEED, address information from a range of property attribute and home energy efficiency installations has been matched to a UPRN. Table A.2 shows the proportion of records on each dataset which were matched to a UPRN. Finally, only properties that match the VOA Council Tax Database were used for analysis. Following improvements in BEIS address matching processes the number of matched records has increased in this publication.

Table A.2: Match rates

Data source	Match rate to UPRN in AddressBase (%)
Electricity Consumption (Domestic)	99
Gas Consumption (Domestic)	96
VOA property attribute data	100
Experian	98
Central Feed-in Tariff Register	94
ECO measures	95
HEED	94
Green Deal measures	90
Gas safe	92

Consumption data

UK Government has collected and published energy consumption data in the *Digest of UK Energy Statistics* since 1948[6]. A time series on how energy has been used, including data back to 1970, is also published in *Energy Consumption in the UK*[7]. Data at individual meter point level (which makes up the consumption part of NEED) were first obtained in 2004 in order to produce local area estimates of consumption—this work was awarded a Royal Statistical Society Award for innovation in 2010. These meter point consumption data cover both gas and electricity consumption for all homes and businesses within England, Scotland and Wales. Property-level data are not available for other heating fuels such as oil, wood or coal. The electricity and gas data are from energy suppliers' administrative systems and cover all domestic and non-domestic meters in England, Wales and Scotland, which is around 31 million electricity meters and 24.5 million gas meters. Consumption data based on these meter level readings are published by BEIS down to postcode level[8]. This section provides more detail on the gas and electricity consumption data used in NEED.

Gas consumption data

Data collection

BEIS obtains annualised consumption estimates for all gas meters in Great Britain. All meter-point data in 2018 came from Xoserve, the data service provider responsible for the collation and aggregation of gas consumption. BEIS is provided with annualised estimates of consumption for all the Meter Point Reference Numbers (MPRNs) in Great Britain based on an Annual Quantity (AQ). The latter is an estimate of annualised consumption using consumption recorded between two meter readings at least six months apart. The estimate is then adjusted using a weather correction factor. The AQ for each MPRN represents consumption relating to the gas period —for 2018 this covers consumption from mid-May 2018 through to mid-May 2019[9].

The data are provided with permission from the owners of the Local Distribution Zones (LDZ) network (i.e. the four major gas transporters in Great Britain – National Grid, Scotia, Wales and West Utilities and Northern Gas Networks) and by agreement by the gas suppliers.

There is currently no reliable way to distinguish between gas used by domestic customers and that used by industry/commerce. The gas industry uses a cut off of 73,200 kWh, with customers using less than this assumed to be domestic. This cut off is therefore also used in BEIS' sub-national consumption publication[10]. It means that there are a significant number of businesses (estimated to be approximately 2 million[11]) misallocated in the sub-national estimates. BEIS is looking to resolve this issue but it does not affect NEED data. NEED uses the allocation of property for council tax to define which customers are domestic.

[6] DUKES can be accessed here: https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes.
[7] ECUK can be accessed here: https://www.gov.uk/government/collections/energy-consumption-in-the-uk.
[8] Sub-national energy consumption statistics can be found in two locations: https://www.gov.uk/government/collections/sub-national-gas-consumption-data (gas) and https://www.gov.uk/government/collections/sub-national-electricity-consumption-data (electricity)
[9] The 2018 "gas period" runs from mid-May 2018 to mid-May 2019.
[10] Available at https://www.gov.uk/government/collections/sub-national-electricity-consumption-data#summary-reports
[11] This figure is an approximation based on the difference between the number of non-domestic meters recorded in electricity and gas data in the Sub-national Consumption publication, accounting for the proportion of properties with electric meters which are off the gas grid

Data validation

Gas consumption in the majority (99.1 per cent) of VOA classified domestic properties is below 50,000 kWh and the relatively small number of properties with consumption greater than this have been excluded. This should reduce the likelihood of including non-domestic properties or domestic properties with invalid consumption in the analysis.

At the lower end of the distribution of consumption figures, there is a cluster of values around 1 kWh to 100 kWh. In 2018, 1.7 per cent of gas consumption records in NEED fell into this category. These have also been excluded from all analysis, as they are likely to be properties with gas supplies which are not used, unoccupied new build properties or where there has previously been an over estimate of the gas consumed at a property.

In addition, meter readings thought to be estimated have been excluded from the data before analysis was undertaken. These estimated meter readings take two forms. For any given year, an estimate is assumed if a property has a gas consumption value identical to the previous year. There are also a small number of cases which are suspected to be estimated readings used by suppliers (i.e., estimates based on typical consumption for certain property types, sizes, etc). Here, a particular value for consumption appeared in the data more often than would be expected given the frequency of similar consumption values.

Comparison with other sources

To check that the sample used for analysis is consistent with the other estimates of domestic consumption published by BEIS and lend confidence in use of the data, the mean consumption for NEED[12] has been compared with the data published by BEIS in ECUK and as sub-national consumption statistics.

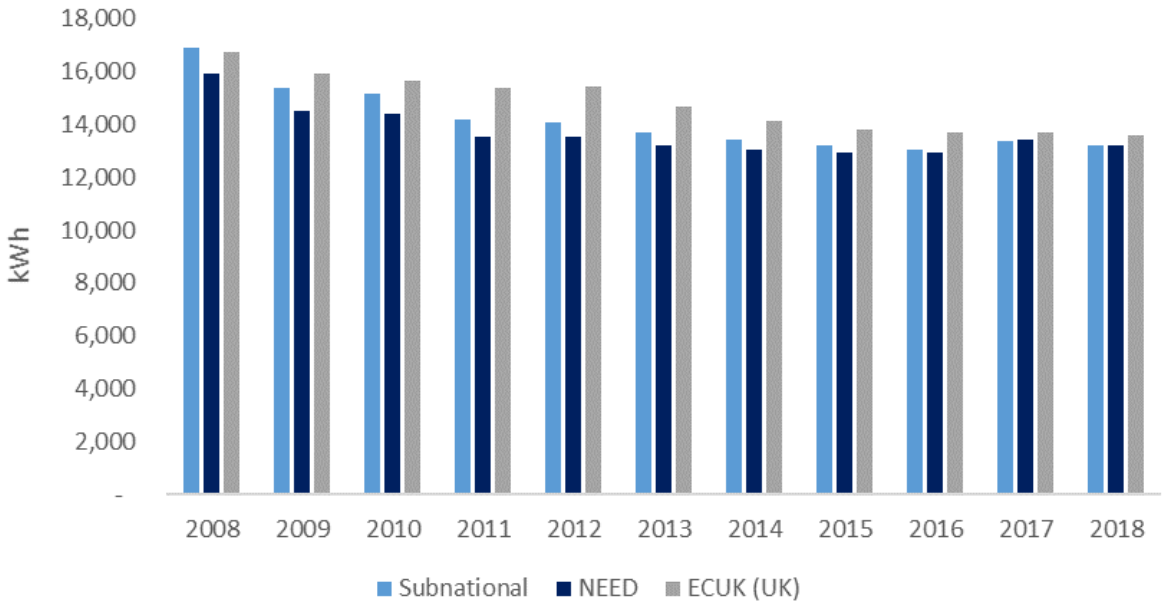
Figure A.2 shows that when looking at gas consumption for 2018, the mean is very similar for the published sub-national statistics and NEED both at 13,200 kWh when rounded to 100 kWh. Since both these sources are based on the same input data and UPRN matching rate is high, it is expected that these values should be similar. Figure A.2 also illustrates that there is more variation when comparing the mean consumption in NEED with that presented in ECUK[13]; the mean consumption presented in ECUK being 400 kWh (2.7 per cent) higher than that of NEED. Gas consumption data in the NEED sample cannot be exactly reconciled with ECUK for a number of reasons:

- the consumption data in ECUK are based on a calendar year whereas the gas consumption data in NEED cover October to September for years 2008 to 2015, July 2016 to July 2017 for 2016, June 2017 to June 2018 for 2017 and May 2018 to May 2019 for 2018.
- there are differences in the weather correction method used for ECUK and that for the meter point consumption data[14]
- the consumption data in ECUK cover the United Kingdom, whereas NEED covers England and Wales

[12] NEED covers England only for 2005 to 2010, and England and Wales from 2011 to 2018.
[13] Source: Energy Consumption in the UK (ECUK), table C9, https://www.gov.uk/government/statistics/energy-consumption-in-the-uk
[14] ECUK reference Energy Trends (June 2011 and September 2011) to read about the methodology of their weather correction: http://webarchive.nationalarchives.gov.uk/20130106091008/http://www.decc.gov.uk/en/content/cms/statistics/publications/trends/trends.aspx

- the different sources of data used for these publications: ECUK estimates are based on aggregate estimates of energy supplied, while NEED is based on end use consumption from gas meters
- ECUK data are based on the number of customers. This differs from the number of meter points since it is possible for a single property to have more than one meter installed

Figure A.2: Comparison of estimates of mean gas consumption (kWh) per property



Electricity consumption data

Data collection

Data are collected with the full co-operation of the electricity industry. Annualised consumption data are generated by the data aggregators who work on behalf of electricity suppliers to collate and aggregate electricity consumption levels for each customer meter or Meter Point Administration Number (MPAN). In addition to this, address information for each meter is obtained from the Electricity Central Online Enquiry Service (ECOES).

The electricity consumption data are generated for both Non-Half Hourly (NHH) meters (domestic and small/medium commercial/industrial customers) and for Half Hourly (HH) meters (larger commercial/industrial customers). There are around 30.7 million NHH meters and approximately 386,000 HH meters in Great Britain. For the NHH data, annualised estimates are based on either an Annualised Advance (AA) or Estimated Annual Consumption (EAC). The AA is an estimate of annualised consumption based on consumption recorded between two meter readings. In comparison, an EAC is used where two meter readings are not available and an estimate of annualised consumption is produced by the energy company using historical information. These data provide a good approximation of annualised consumption, but do not exactly cover the calendar year. For example, 2018 annualised consumption estimates cover the period from 31st January 2018 up to 30th January 2019. For

the half hourly meter consumption estimates, data aggregators produce a report for each MPAN for the relevant calendar year.

BEIS publishes estimates of domestic and non-domestic consumption with aggregate and average consumption figures provided for each local authority. The domestic consumption is based on NHH meters with profiles 1 and 2 (these are the standard domestic and economy 7 meters respectively). Non-domestic consumption is based on NHH meters with profiles 3 to 8 and all HH meters[15]. However, it should be noted that these assumptions differ from those used in NEED, where the additional data available mean it is more appropriate to use a slightly different approach to ensuring a property is domestic and has valid consumption. This is described in more detail in the data validation section below.

Data validation

There are differences in the consumption records included in the sub-national consumption publications and those used in NEED.

Electricity consumption in the majority of VOA classified domestic properties 26.2m (99.4 per cent) is below 25,000 kWh. The relatively small number of properties with consumption greater than this have been excluded from NEED, in order to avoid biasing estimates. This should reduce the likelihood of including non-domestic properties or domestic properties with invalid consumption in the analysis.

At the lower end of the distribution, there is a further cluster of values (1.8%) including negatives with consumption up to 100 kWh. These have been excluded from all analysis, as they are likely to represent properties with electricity supplies which are not used (or new build properties which are not yet occupied). Unlike in sub-national consumption statistics, all negative meter readings are also excluded[16], which raises the overall mean of NEED electric data.

In addition, suspected estimated values have been excluded from the data before analysis was undertaken. These take two forms. For any given year, if a property has a consumption value identical to the previous year it is assumed to be an estimate. There are also a small number of values which are suspected to be estimated readings used by suppliers. As for the gas estimations, these were assumed on the basis of values that appear in the data more often than would be expected given the frequency of similar consumption values.

The impact of removing these records is small. It causes the mean for NEED to be slightly lower than if these filters were not applied, due to the elimination of a relatively small number of records with a high consumption. The median remains almost the same.

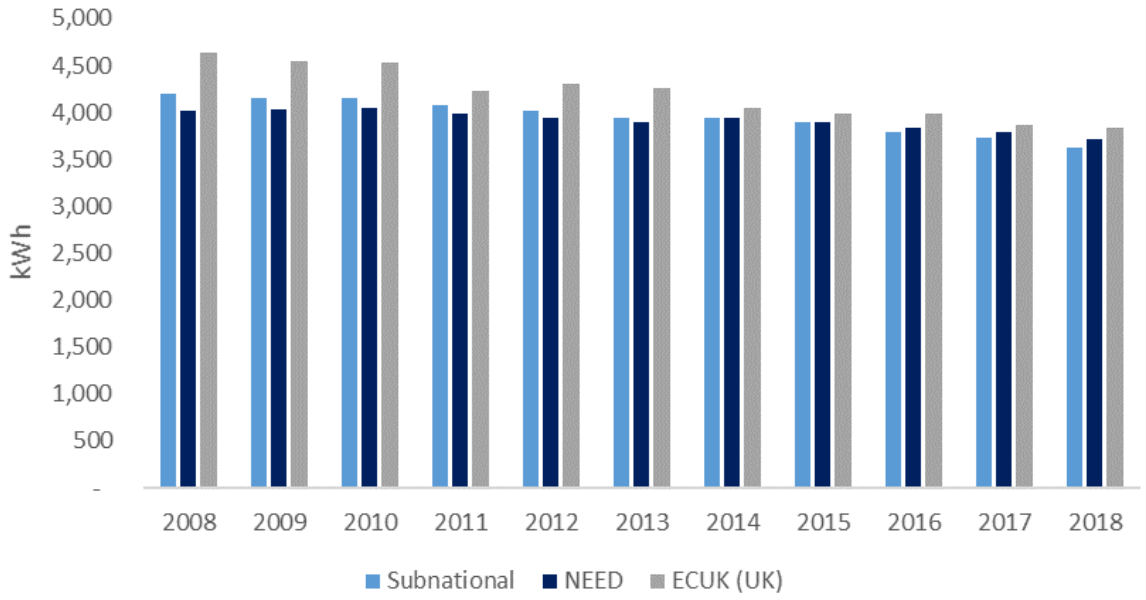
Comparison with other sources

To assess the consistency of the analysis sample with the other estimates of domestic consumption published by BEIS—and therefore increase confidence in use of the data—mean consumption for NEED[17] was compared with the data published by BEIS in ECUK and in sub-national consumption statistics.

[15] Non-domestic consumption also includes any nominally domestic meters with consumption of more than 100,000 kWh in a year or meters with consumption between 50,000 and 100,000 kWh with address information which suggests non-domestic use.
[16] As data are based on billed consumption, it is possible that a negative reading is valid if an estimated reading provided in a previous year was too high, as at the aggregate level the effect of this will cancel out. However, these reading are not considered valid in NEED.
[17] The NEED sample covers England only for 2005 to 2010, and England and Wales between 2011 and 2018.

Figure A.3 below shows that the mean electricity consumption is similar for all three sources being compared. When looking at consumption in 2018, the difference between the mean electricity consumption in NEED and ECUK is 100 kWh (2.9 per cent). The 2018 sub-national electricity consumption mean is just over two and a half per cent higher than that of NEED. This is due to NEED excluding meters where consumption is more than 25,000 and less than 100 kWh, whereas the meters are included in the sub-national analysis.

Figure A.3: Comparison of estimates of mean electricity consumption (kWh) per property



Conclusion

The consumption data are a rich source of data which form the core of NEED. Table A.3 summarises the approaches taken towards using the meter point consumption data in domestic NEED and BEIS' sub-national estimates.

The differences lead to small discrepancies in mean consumption, but provide confidence in the analysis carried out with NEED, including in relation to the impact of installing energy efficiency measures. The comparisons carried out against other data sources confirm that the consumption estimates based on NEED are consistent with other sources.

Table A.3: Differences in domestic consumption data[18,19]

NEED data	Sub-national consumption estimates
The property must be included as a domestic property in the Valuation Office Agency property attribute dataset to be included in domestic NEED analysis.	Domestic properties classified based on consumption for gas (less than 73,200 kWh) and profile class for electricity (profiles 1 and 2 are domestic).
Gas consumption between 100 kWh and 50,000 kWh.	Gas consumption below 73,200 kWh.
Electricity consumption between 100 kWh and 25,000 kWh.	Electricity consumption below 100,000 kWh and profile class 1 or 2 (including negative readings).
Data matched to other sources via unique property reference number (UPRN) at property level.	Data assigned to Lower Layer Super Output Area.
Suspected estimated readings removed.	

[18] Electricity consumption of between 50,000 and 100,000 kWh is reviewed and if it has a likely non-domestic address then it is also excluded from the sub-national domestic estimates.
[19] This means that for the sub-national consumption statistics some properties can be assigned accurately if the street is identified even if the exact property is not known.

Valuation Office Agency Data

Introduction

The Valuation Office Agency (VOA) is the central Government agency responsible for valuing homes in England and Wales for council tax purposes[20]. The VOA has had responsibility for valuing properties for council tax since it was first introduced in 1993 and, before then, for the earlier system of domestic rates. Property attribute data was originally introduced in the 1970s in order to provide a simple system for understanding the main features and attributes of a property.

In order to maintain accurate and fair lists of council tax bandings, the VOA needs to keep the information it holds about properties up to date. It does this in a number of ways, including:

- Obtaining information from the local authority when a home is extended or altered to the extent that planning permission is required
- Using voluntary questionnaires to enable the occupier to confirm information about a property
- Obtaining other sources of freely available and publicly published information. For example, a contract with Calnea Analytics to access the Residata website which contains details of properties marketed through mouseprice.com since 2007
- In addition, the VOA will sometimes ask to visit a property when the information it needs cannot be ascertained from other sources. This can be at the occupier's request; for example when they have challenged the council tax banding of their property and wish the VOA to carry out a review

There are 16 individual property attributes collected, four of which are used in NEED analysis:

- property type
- property age
- floor area (m^2)
- number of bedrooms

Coverage

Table A.4 below shows the categories of data used in the analysis for each of the VOA variables (categories are those published in the English Housing Survey). In most cases, VOA has more detailed data; the VOA categories have been grouped into the categories used for

[20] It does not set the level of council tax nor collect the money, which is the task of local government.

the purposes of the NEED analysis and presentation of results. Full details of the breakdowns included in the VOA dataset are available on the VOA website[21]

Table A.4: VOA property attribute data categories used in NEED

	Property age	Property type	Number of bedrooms	Floor area (m^2)
Categories	Pre 1919	Detached	1	1-50
	1919-44	Semi-detached	2	51-100
	1945-64	End terrace	3	101-150
	1965-82	Mid terrace	4	151-200
	1983-92	Bungalow	5 or more	Greater than 200
	1993-99	Purpose-built flat		
	Post 1999	Converted flat		

Summary of data and comparison with other sources

This section shows how NEED compares with the distribution of the data in the full VOA property attribute database and with the English Housing Survey (EHS)[22].

The EHS will vary compared with the VOA data as it is a sample survey and only covers England, whereas VOA data and NEED cover England and Wales. However it still provides context to validate the NEED data.

Figures A.4 to A.6 show the proportion of properties in each category for three of the VOA variables used in NEED. EHS calculate different floor area categories which make meaningful comparison between sources difficult,

To validate that distribution of dwelling stock is consistent with other sources, the graphs below consider the distribution of properties with “valid” gas and electricity consumption separately to understand that the distribution is representative of the housing stock. This is because that as well as there being less properties that have a gas meter compared to having an electric meter, the number of properties with valid gas and valid electricity values is also different.

The most notable variations between NEED and other data sources are seen for property type of flats and for one and three bedroom properties. There is also variation between fuel type. In example there are less one-bedroom gas properties and more three-bedroom gas properties as a proportion. This is likely to reflect difficulties in address matching flats, but also an indication that flats are more likely to use electric as the main heating fuel.

[21] http://www.voa.gov.uk/corporate/Publications/DwellingHouseCodingGuide/index.html
[22] EHS data are from the English Housing Survey Headline Report 2018-19: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/860076/2018-19_EHS_Headline_Report.pdf

Figure A.4: Comparison of distributions – number of bedrooms

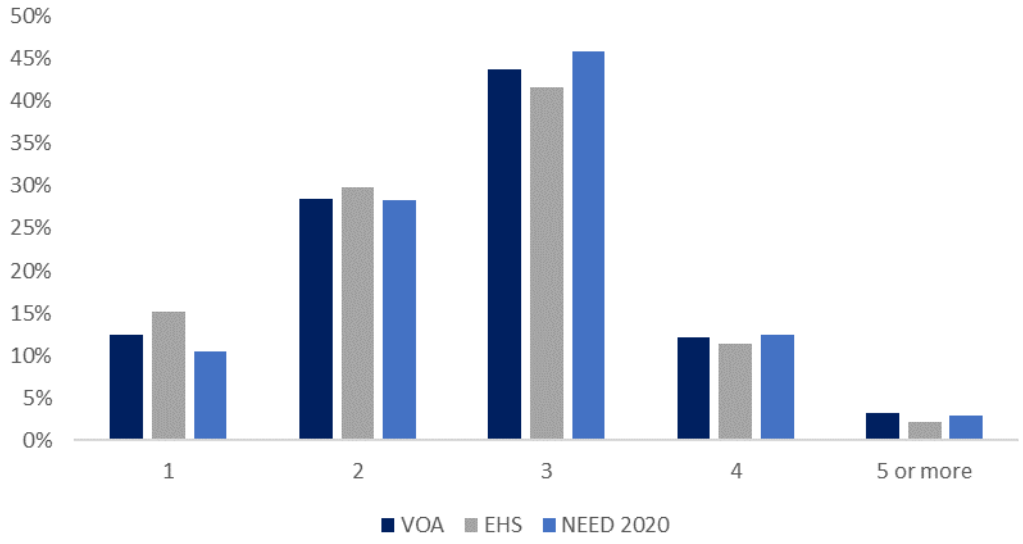


Figure A.5: Comparison of distributions – property type

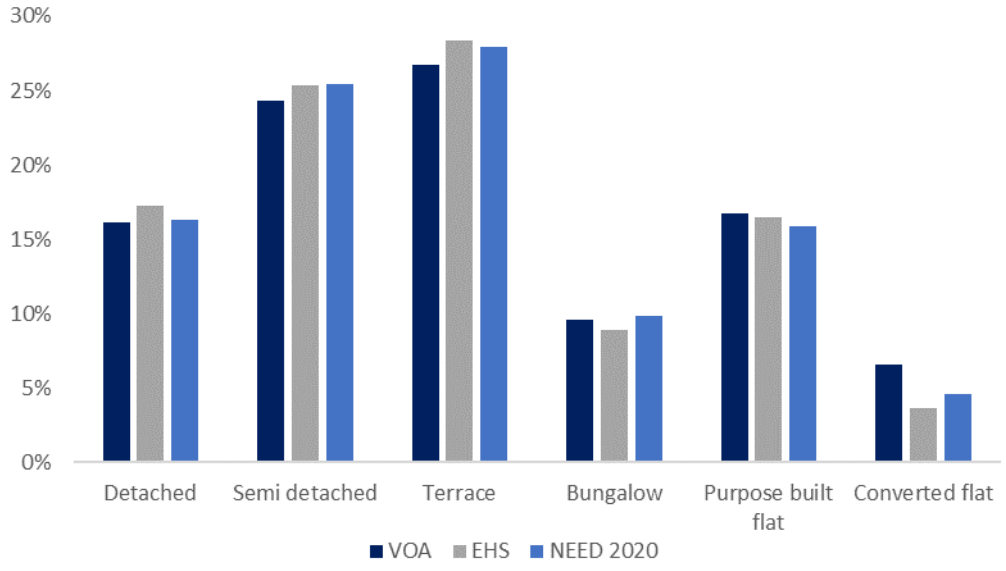
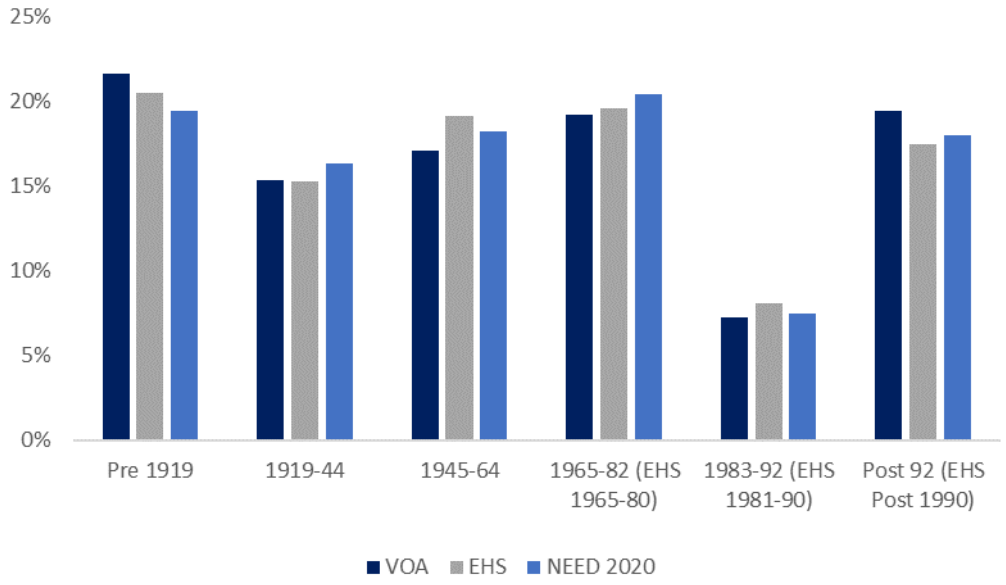


Figure A.6: Comparisons of distributions – property age



Conclusion

The NEED data shows similar distributions of property attributes compared to the VOA administrative dataset and the EHS sample survey data. There are some differences for flats and smaller properties which should be taken into account when interpreting analysis based on NEED.

Experian data

BEIS purchased data from Experian for each property in the UK based on the dwelling stock in 2019. Data are modelled by Experian based on other data sources including Experian surveys and aggregate published data (such as the Census). A unique property reference number could be assigned to 98 per cent of records in the dataset provided by Experian.

Coverage and comparison with other sources

The Experian household characteristics data used in NEED are:

- household income
- tenure
- the number of adult occupants

Household income

The household income variable identifies the likely household income for each property. The data are based on results from responses to Experian's consumer survey, which is then used alongside other predictive data (including Experian's person and household level demographics and Mosaic) to build a model. Household income is available in ten income bands which are set out below in table A.5.

Table A.5: Distribution of income band using the 2019 full Experian dataset

Band	Description	Households (%)
1	Less than £15,000	15.7%
2	£15,000 - £19,999	8.9%
3	£20,000 - £29,999	19.0%
4	£30,000 - £39,999	15.7%
5	£40,000 - £49,999	12.0%
6	£50,000 - £59,999	7.7%
7	£60,000 - £69,999	5.7%
8	£70,000 - £99,999	8.7%
9	£100,000 - £149,999	4.3%
10	£150,000 or more	2.3%

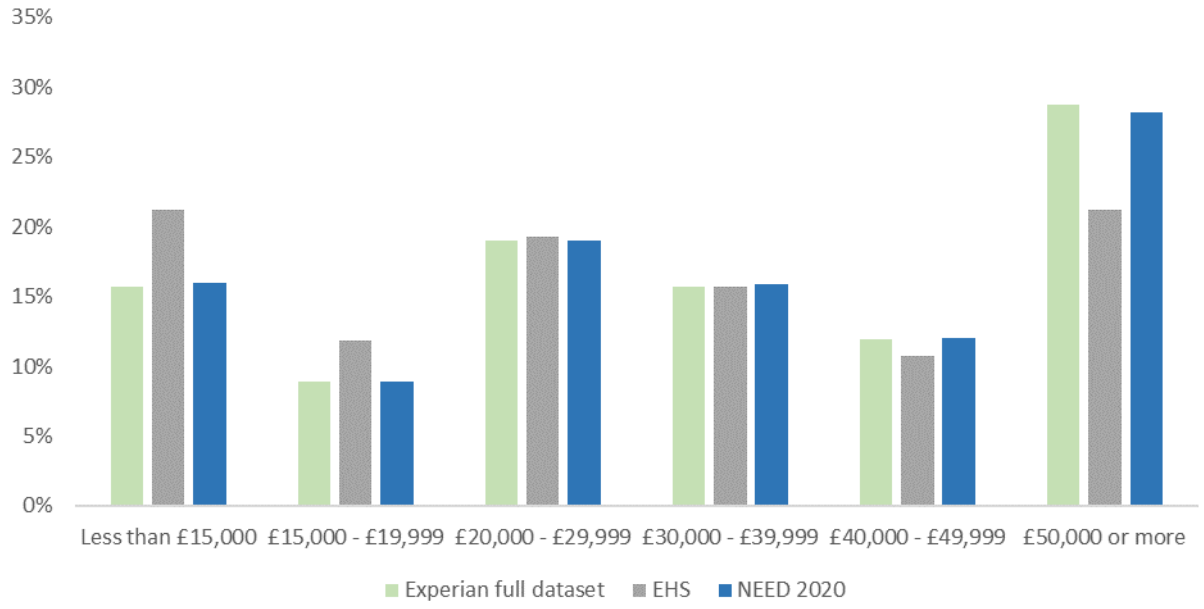
When interpreting any analysis of income in the NEED report it should be noted that data for each property are modelled and therefore are indicative of the income a household is likely to have rather than being an actual value for the current occupant of the property.

Based on Experian's assessment of the data, 25 per cent of properties are in the correct category and 58 per cent of properties are assigned to within one band of the correct

category[23]. Figure A.7 shows how the distribution of income for the Experian dataset and NEED compares with the income reported by the EHS for 2016-17[24]. Note that some of the income categories from the Experian data have been grouped together to allow comparison with the categories used in the EHS.

Figure A.7 shows that Experian appears to be under-assigning properties to the lowest income band and over-assigning them at the higher income bands, compared to the EHS data. This is consistent with BEIS' understanding that the Experian income data is less reliable at the extremes. However, it should also be noted that the EHS is a survey and therefore subject to variation. Income is a self-reported variable and therefore likely to be less reliable compared to the EHS property variables which are based on a physical survey on the property carried out by a trained surveyor.

Figure A.7: Comparison of distributions – household income band



Tenure

Tenure data from Experian allocates each household in the UK to one of three categories; owner occupied, council/housing association or privately rented. The data are based on responses to Experian's lifestyle survey which are then used to predict the status of all households. As with the household income variable, a model is used to predict the tenure for each property.

Figure A.8 shows how the Experian data compares with the NEED gas and electric samples and EHS at the national level[25]. It shows that the proportion of properties assigned to each tenure category is similar for all sources. It appears that the Experian dataset as a whole and NEED allocates too many properties to the owner occupied category and too few to privately

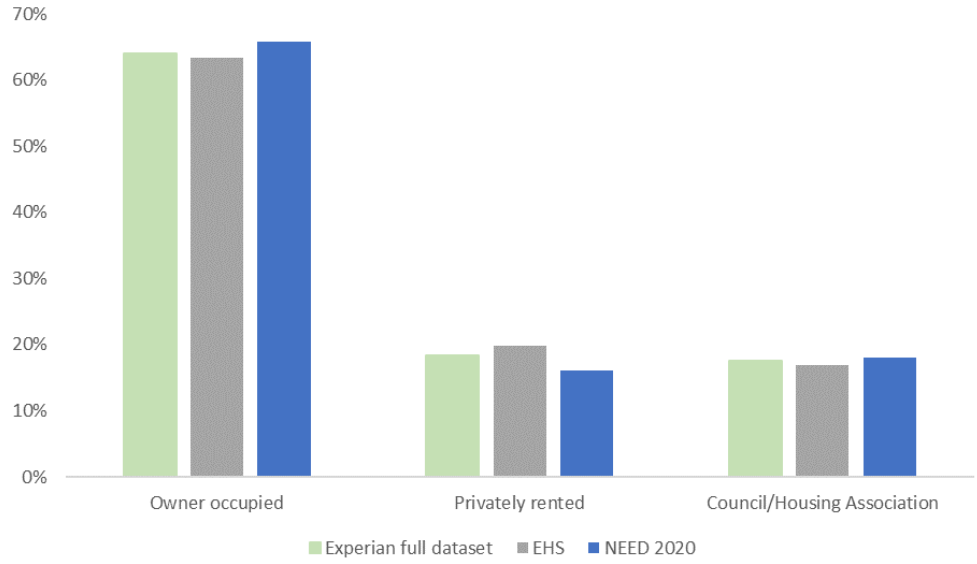
[23] Experian quality assurance of household income classification data: https://www.experianintact.com/content/uk/documents/productSheets/HouseholdIncome.pdf

[24] EHS data are from the English Housing Survey Headline Report 2016-17: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/774820/2017-18_EHS_Headline_Report.pdf

[25] Note that the Experian full dataset covers the UK, the NEED sample covers England and Wales, and the EHS covers England only.

rented. This is likely to be linked to the loss of flats resulting from the address matching process.

Figure A.8: Comparison of distributions – tenure



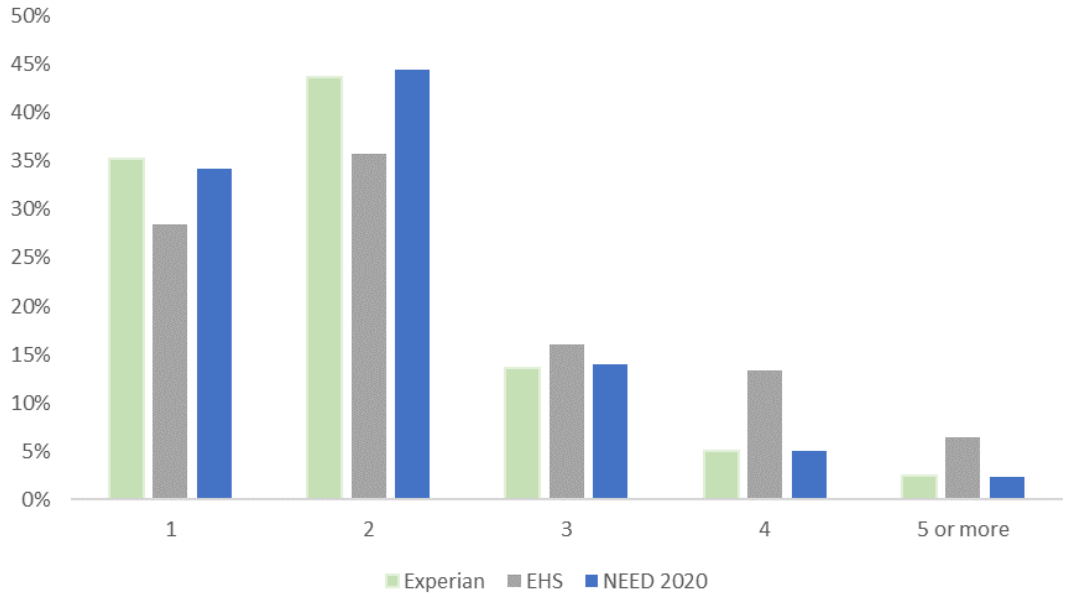
Number of adult occupants

The number of adults variable gives the number of adults over 18 living in a household. Experian takes the number of adults information from its Consumer View database. Figure A.9 shows how the data in the gas and electric NEED samples compare with other sources.

The variation in the distribution is likely to be because the EHS estimates are based on household size whilst the Experian data is based on the number of occupants aged 18 and over. This means a household with two adults and two children would be classified as two in the Experian data and four in the EHS. Therefore, there are more properties with one or two occupants in the Experian data and more properties with three or more in the EHS.

While the Experian data are valuable since they provide an understanding of the properties in NEED and how consumption and impact of energy efficiency measures vary for different types of households, it is important that interpretation of results relating to income, tenure and number of adult occupants is in the context of the limitations of the data.

Figure A.9: Comparison of distributions – number of adults



Energy efficiency measures data

The Energy Company Obligation and Green Deal

The Energy Company Obligation (ECO) and Green Deal (GD) are Government energy efficiency schemes which began operating in 2013[26]. They replaced the previous schemes: Carbon Emissions Reduction Target, Community Energy Saving Programme and Warm Front. Their aim is to encourage the uptake of energy efficiency measures so that the efficiency of the building stock is improved. A unique property reference number could be assigned to 95.2 per cent of records in the dataset provided by Ofgem.

Geographical coverage

The statistics cover various geographies depending on the delivery mechanism:

- GD Plans: Great Britain
- ECO: Great Britain
- GDHIF: England and Wales
- GD Communities: England (selected LAs, see methodology note for full list of participating areas)

[26] For more detail on ECO and GD the household energy efficiency national statistics report here: https://www.gov.uk/government/collections/household-energy-efficiency-national-statistics

Uncertainty and bias

The data sources listed in this document are all subject to a range of data quality checks employed by both data providers and BEIS to ensure that data are as fit for purpose as possible[27]. Energy efficiency measures that are installed outside of Government schemes would not be captured in the estimates published. This can occur when measures are financed through another route such as savings, payment from a landlord, housing association or Local Authority.

Homes Energy Efficiency Database

The Homes Energy Efficiency Database (HEED) is a national database developed by the Energy Savings Trust (EST). It was set up to help monitor and target carbon reduction and fuel poverty work. It contains details of energy efficiency and micro-generation installations such as cavity wall insulation and solar hot water. HEED also includes data about property attributes (such as property age and type) and heating systems. However due to coverage and quality these data are not used in NEED.

Data have been recorded in HEED since 1995 until 2013 including activity reported from Government programmes, such as the Energy Efficiency Commitment (EEC) and the Carbon Emissions Reduction Target (CERT), and activity reported by trade associations such as CORGI (Gas Safe) and FENSA.

Approximately 50 per cent of UK homes have a record in HEED. However there may not be complete information for each of these records. For example, if a measure has been installed through a Government scheme then there may be information on the measure installed but no information on what other energy efficiency measures the property has if they were not installed through a Government scheme. However, there is no information on measures that properties have installed themselves or measures installed at the time the property was built.

Because the majority of measures recorded in HEED are measures installed through Government schemes which are aimed at particular segments of the population, the types of properties receiving measures are not representative of the populace or of the housing stock as a whole.

HEED includes a high proportion of the measures reported by suppliers to Ofgem. As no information is known about the specific properties receiving measures reported by Ofgem it is not possible to determine whether there is any bias in the HEED data, but the good coverage means that any bias should be small.

Coverage of current installations extracted from ECO is comparable to previous schemes such as CERT. Over the four-year period that CERT was active Ofgem[28] report that 55,000 (around 14,000 per year) properties had solid wall insulation installed. This compares with 42,000 installations (around 14,000 per year) recorded in NEED using the ECO data between 2015 and 2017.

When considering the quality of HEED data included in NEED it should also be noted that the installation dates associated with records are of varying quality. This is particularly so for earlier installations of solid wall insulation where it is not possible to distinguish when between 2005 and 2008 measures were installed.

[27] Details on the checks that are carried out are included in the 'Data quality of data sources' section of the methodology note which accompanies the statistical releases. This methodology note can be accessed from: https://www.gov.uk/government/publications/household-energy-efficiency-statistics-methodology-note

[28] Ofgem final CERT report 2008-11: https://www.ofgem.gov.uk/ofgem-publications/58425/certfinalreport2013300413pdf.

Central Feed-in Tariff Register

The Central Feed-in Tariff Register (CFR) is an electronic, web-based system used to manage the Feed-in Tariff (FIT) scheme that Ofgem administers on behalf of BEIS. Extracts from the CFR are provided to BEIS on a monthly basis, and serve as the basis of a number of statistical publications on the FIT scheme, available at https://www.gov.uk/government/collections/feed-in-tariff-statistics. The CFR contains installation-level data on every microgeneration installation that is registered for and receives feed-in tariffs. The scheme ended for new entrants in April 2019 and the data used in this publication are from the March 2019 extract. A unique property reference number could be assigned to 93.6 per cent of records in the dataset provided by Ofgem.

Coverage and comparison with other sources

The FIT installation data recorded in the CFR include:

- geographical location
- technology type (e.g., wind, solar)
- installed capacity
- type of installation (i.e. domestic, community, or commercial)

Gas Safe Boiler Installations

Gas Safe are the organisation that maintains the registration of all legally installed gas appliances and replaced CORGI as the gas registration body in 2009. Boiler installation data is provided to BEIS annually and contains information on the type of boiler installed and property address information that can link the installation to a UPRN. BEIS cross checks new boiler installations at a property to only include the most recent installation at a property. Gas Safe data includes installations up to October 2019, this does not cover the full gas year for 2018 and should be borne in mind when interpreting the impact of measures data for the current year.

Coverage and comparison with other sources

Gas Safe is the most comprehensive register of boiler installations and records all boiler installations in the UK and therefore offers a more complete view of installations than those installed through government schemes. In example Gas Safe boiler installations are registered against around 25 per cent of properties contained in NEED, whereas ECO registered boiler installations reflect only 2 per cent of properties.

This publication is available from: https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

If you need a version of this document in a more accessible format, please email energyefficiency.stats@beis.gov.uk. Please tell us what format you need. It will help us if you say what assistive technology you use.